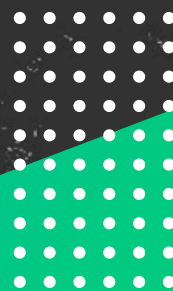




MINICURSOS

XXIV Simpósio Brasileiro de Computação Aplicada à Saúde

25 a 28 de junho de 2024 | Goiânia - GO



INF
INSTITUTO DE
INFORMÁTICA



UFG
UNIVERSIDADE
FEDERAL DE GOIÁS



24° Simpósio Brasileiro de Computação Aplicada à Saúde
Goiânia, Goiás, 25 a 28 de junho de 2024

Minicursos do XXIV Simpósio Brasileiro de Computação Aplicada à Saúde

Organização do Livro

Dianne Scherly Varela de Medeiros
Eduardo Simões de Albuquerque
Carlos Henrique Rorato Souza

Capa

Wellington Santos Martins (fotografias – @gynsualinda)
Daniella Fernandes do Nascimento (arte)

Coordenação Geral do Simpósio

Sérgio Teixeira de Carvalho

Realização

Universidade Federal de Goiás – UFG
Sociedade Brasileira de Computação – SBC

Porto Alegre
Sociedade Brasileira de Computação – SBC
2024

Dados Internacionais de Catalogação na Publicação (CIP)

S612 Simpósio Brasileiro de Computação Aplicada à Saúde (24. : 25 – 28 junho 2024 : Goiânia)
Minicursos do SBCAS 2024 [recurso eletrônico] / organização: Dianne Scherky Varela de Medeiros ... [et al.]. – Dados eletrônicos. – Porto Alegre: Sociedade Brasileira de Computação, 2024.
291 p. : il. : PDF ; 41 MB

Modo de acesso: World Wide Web.
ISBN 978-85-7669-590-5 (e-book)

1. Computação – Brasil – Evento. 2. Saúde aplicada. 3. Tecnologia em saúde. I. Medeiros, Dianne Scherly Varela de. II. Albuquerque, Eduardo Simões de. III. Souza, Carlos Henrique Rorato. IV. Sociedade Brasileira de Computação. V. Título.

CDU 004.6(063)

Ficha catalográfica elaborada por Annie Casali – CRB-10/2339
Biblioteca Digital da SBC – SBC OpenLib

Índices para catálogo sistemático:

1. Ciência e tecnologia dos computadores : Informática : Dados – Publicação de conferências, congressos e simpósios etc. ... 004.6(063)

Sociedade Brasileira de Computação – SBC

Presidência

Thais Vasconcelos Batista (UFRN), Presidente

Cristiano Maciel (UFMT), Vice-Presidente

Diretorias

Renata de Matos Galante (UFRGS), Diretora Administrativa

Lisandro Zambenedetti Granville (UFRGS), Diretor de Finanças

Denis Lima do Rosário (UFPA), Diretor de Eventos e Comissões Especiais

Claudia Lage Rebello da Motta (UFRJ), Diretora de Educação

José Viterbo Filho (UFF), Diretor de Publicações

André Luís de Medeiros Santos (UFPE), Diretor de Planejamento e Programas Especiais

Eunice Pereira dos Santos Nunes (UFMT), Diretora de Secretarias Regionais

Alírio Santos de Sá (UFBA), Diretor de Comunicação

Tanara Lauschner (UFAM), Diretora de Relações Profissionais

Carlos Eduardo Ferreira (USP), Diretor de Competições Científicas

Ronaldo Alves Ferreira (UFMS), Diretor de Cooperação com Sociedades Científicas

Michelle Silva Wangham (UNIVALI), Diretora de Inovação

Leila Ribeiro (UFRGS), Diretora de Computação na Educação Básica

Contato

Av. Bento Gonçalves, 9500

Setor 4 - Prédio 43.412 - Sala 219

Bairro Agronomia

91.509-900 – Porto Alegre (RS)

CNPJ: 29.532.264/0001-78

<http://www.sbrc.org.br>

Mensagem da Coordenação Geral do SBCAS 2024

O XXIV Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2024), organizado pelo Instituto de Informática (INF) da Universidade Federal de Goiás (UFG) e realizado no período de 25 a 28 de junho de 2024 na cidade de Goiânia–GO, reuniu um público especializado, oriundo de diversos segmentos das áreas de Computação, Engenharia e Saúde. Profissionais, professores, pesquisadores, estudantes de graduação e de pós-graduação puderam encontrar no SBCAS 2024 a oportunidade de adquirir um conjunto atualizado de informações científicas e tecnológicas, e de debater temas e soluções com especialistas nacionais e internacionais nas diversas áreas de computação aplicada à saúde.

É uma grande satisfação poder reunir as comunidades de computação e de saúde em um evento multidisciplinar, que promove e incentiva a realização de pesquisas e trabalhos inovadores com potencial de trazer benefícios diretos à nossa sociedade. Com o propósito de disseminar conhecimento técnico-científico produzido em nosso país sobre tecnologia, informática, computação e saúde, a programação técnica do SBCAS 2024 envolveu sessões técnicas, palestras, painéis, apresentação de ferramentas e aplicações, além dos concursos de teses e dissertações, e de trabalhos de iniciação científica.

Este Livro de Minicursos, além dos próprios Anais com toda a produção científica gerada no SBCAS 2024, refletem a qualidade da programação e o resultado do empenho de diversas pessoas. Portanto, o nosso profundo agradecimento às pessoas que integraram o comitê da trilha de minicursos, Dianne Scherly Varela Medeiros (UFF) e Eduardo Simões de Albuquerque (UFG), e ao coordenador de publicação dos Anais e deste Livro de Minicursos, Carlos Henrique Rorato Souza (UFG).

Nossos agradecimentos também aos coordenadores da Trilha Principal do SBCAS 2024, Rodrigo da Rosa Righi (UNISINOS) e Paulo Eduardo Ambrósio (UESC), às coordenadoras da trilha de Ferramentas e Aplicações, Lina Maria Garcés Rodriguez (USP) e Renata Dutra Braga (UFG), aos coordenadores do Concurso de Teses e Dissertações (CTD), Debora Christina Muchaluat Saade (UFF) e Iwens Gervasio Sene Junior (UFG), às coordenadoras do Concurso de Trabalhos de Iniciação Científica (CTIC), Andrea Campos Gomes Bianchi (UFOP) e Luciana de Oliveira Berretta (UFG), e aos coordenadores da trilha Journal First, Márcia Ito (Fatec/SP - CEETEPS) e Renato de Freitas Bucão Neto (UFG).

Agradecemos, ainda, o apoio financeiro da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), do CEIA (Centro de Excelência em Inteligência Artificial) da UFG, CGI.br (Comitê Gestor da Internet no Brasil), do CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), e da FAPEG (Fundação de Amaro à Pesquisa do Estado de Goiás), o qual tornou viável a realização do SBCAS 2024. Nosso muito obrigado também ao apoio da SBC (Sociedade Brasileira de Computação), de sua Comissão Especial de Computação Aplicada à Saúde (CECAS), e de sua Secretaria Regional de Goiás, na realização do evento. Agradecemos também à Faculdade de Medicina e à Faculdade de Enfermagem, ambas da UFG, e à Comissão de Governança da Informação em Saúde (CGIS) da UFG, pela parceria na organização do evento. Agradecemos o apoio local do Instituto de Informática da UFG, e da Reitoria da UFG. Por fim, fazemos um agradecimento muito especial ao grupo de estudantes de graduação e de pós-graduação da UFG, que atuaram como voluntários durante a realização do SBCAS 2024. De forma espontânea e leve, a colaboração desse grupo tornou viável a realização do nosso evento.

Em nome da comissão organizadora do SBCAS 2024, desejamos uma excelente leitura deste livro e que seja tão enriquecedora em conhecimentos quanto foram os dias de troca de experiências e de agradável convívio em Goiânia.

Sérgio Teixeira de Carvalho (UFG)

Coordenador Geral

XXIV Simpósio Brasileiro de Computação Aplicada à Saúde

Mensagem da Coordenação de Minicursos do SBCAS 2024

O Livro de Minicursos do XXIV Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2024) reúne os textos dos minicursos selecionados e apresentados nesta edição do evento, sediada em Goiânia – Goiás de 25 a 28 de junho de 2024. Foram submetidas 11 propostas, das quais 6 foram selecionadas para publicação e apresentação, representando uma taxa de aceitação de 54,5%. O Comitê de Programa foi composto por 12 pesquisadores para elaboração dos pareceres e cada proposta recebeu pelo menos 3 revisões, gerando ao todo 34 revisões. Aproveitamos para agradecer aos membros do Comitê de Programa pela dedicação na contribuição voluntária no processo de avaliação e de seleção dos minicursos, fornecendo revisões valiosas para os autores.

As questões abordadas são de grande relevância para a computação aplicada à saúde, tanto no contexto nacional, quanto no internacional, considerando temas como processamento de sinais de eletroencefalografia para desenvolvimento de aplicações de interface cérebro-computador (Capítulo 1), simuladores imersivos baseados em realidade virtual (Capítulo 2), avaliação de software para a área de saúde digital (Capítulo 3), aprendizado de máquina para produção de estruturas de índices para otimizar recuperação e acesso a registros de uma base de dados (Capítulo 4), padronização de vocabulários de dados e ferramentas para mapeamento de vocabulários locais (Capítulo 5) e aplicações de inteligência artificial para apoio ao diagnóstico, tratamento e prognóstico de transtornos mentais (Capítulo 6). Nesse contexto, os minicursos trazem uma oportunidade de aprendizado e atualização de conhecimentos para a comunidade, inspirando novos trabalhos e áreas de pesquisa em níveis de graduação e pós-graduação.

Gostaríamos de agradecer a todos os autores que submeteram suas propostas de minicursos ao SBCAS 2024, nos motivando a realizar anualmente este evento de interesse, visibilidade e sucesso crescentes. Em especial, agradecemos aos autores dos minicursos aceitos que prepararam os capítulos deste livro.

Expressamos também nossos agradecimentos à coordenação geral do SBCAS 2024, realizada pelo Prof. Sérgio Teixeira de Carvalho da Universidade Federal de Goiás (UFG), por todo o seu apoio e suporte à realização dos minicursos.

Desejamos que todos os leitores aproveitem o conteúdo desse livro, preparado com muito zelo e qualidade técnica pelos autores dos minicursos.

Dianne Scherly Varela de Medeiros (UFF)

Eduardo Simões de Albuquerque (UFG)

Coordenadores dos Minicursos

XXIV Simpósio Brasileiro de Computação Aplicada à Saúde

Comitê de Programa de Minicursos

Debora Muchaluat-Saade, UFF
Edelberto Franco Silva, UFJF
Igor Monteiro Moraes, UFF
Lucas Ferrari de Oliveira, UFPA
Marcia Ito, PPG-GTSP-CEETEPS
Nicollas Rodrigues de Oliveira, UFF
Paulo Ambrósio, UESC
Robespierre Pita, UFBA
Taiane Ramos, UFF

Comissão Especial de Computação Aplicada à Saúde (CE-CAS) - SBC

Cristiano André da Costa (Unisinos), Coordenador
Rodrigo de Melo Souza Veras (UFPI), Vice-Coordenador
Débora Christina Muchaluat Saade (UFF)
Lina Maria Garcés Rodriguez (UNIFEI)
Lucas Ferrari de Oliveira (UFPR)
Marcia Ito (FATEC/SP)
Natália Castro Fernandes (UFF)
Paulo Eduardo Ambrósio (UESC)
Sérgio Teixeira de Carvalho (UFG)

Sumário

Mensagem da Coordenação Geral	iv
Mensagem da Coordenação de Minicursos	v
Comitês	vi
Capítulo 1 - Uma Abordagem Prática Voltada a Aplicações de Interface Cérebro-Computador Utilizando Sinais de EEG	1
<i>Álfe C. Macedo, Murilo B. Silva, Augusto C. S. Rocha, Hugo A. D. do Nascimento, Diego Resende Faria</i>	
Capítulo 2 - Simuladores Imersivos para Ensino em Saúde: Futuro ou Realidade?	43
<i>Luciana Nedel, Matheus Negrão, Wesley Ferreira, Carla M.D.S. Freitas, Anderson Maciel, Betânia Bohrer, Ana Paula Oliveira Rosses e Rafael Piccin Torchelsen</i>	
Capítulo 3 - Avaliação de Tecnologia em Saúde Digital: Fundamentos e Aplicações Práticas	82
<i>Thais Maria Yomoto Ferauche, Jean Jose Silva, Wilian Hisatugu e Márcia Ito</i>	
Capítulo 4 - Learned Index: Perspectivas e Desafios na Gestão de Dados de Saúde	131
<i>Guilherme Waldschmidt, Alberto Sironi, Lais Sacramento, Mirlei Moura da Silva, Maíra Lima Souza, Nívea Bispo da Silva e Robespierre Dantas da Rocha Pita</i>	
Capítulo 5 - Harmonização Global de Dados de Saúde: O Papel dos Vocabulários Padronizados OHDSI	183
<i>Maria Tereza Fernandes Abrahão, Pablo Jorge Madril e Mateus de Lima Freitas</i>	
Capítulo 6 - Applications of Artificial Intelligence to Support the Diagnosis, Treatment, and Prognosis of Mental Disorders	236
<i>Paulo Mann, Elton H. Matsushima e Aline Paes</i>	

Capítulo

1

Uma abordagem prática voltada a aplicações de Interface Cérebro-Computador utilizando sinais de EEG

Álefe C. Macedo (UFG), Murilo B. Silva (IFG), Augusto C. S. Rocha (UFG), Hugo A. D. do Nascimento (UFG), Diego Resende Faria (UH – UK)

Abstract

This chapter addresses the processing of Electroencephalography (EEG) signals with a focus on the development of Brain-Computer Interface (BCI) applications. The objective is to stimulate interest in the subject and contribute to the construction of new research in the area. A historical overview of the use of EEG, low-cost BCI devices, computational tools for signal processing, and related public databases are discussed. Some practical case studies are also discussed, involving programming codes and databases for the treatment and classification of EEG signals.

Resumo

Este capítulo aborda o processamento de sinais de Eletro-Encefalografia (EEG) com foco no desenvolvimento de aplicações de Interface Cérebro-Computador (Brain-Computer Interface, BCI). O objetivo é estimular o interesse sobre o assunto e contribuir para a construção de novas pesquisas na área. São discutidos um panorama histórico do uso de EEG, dispositivos de BCI de baixo custo, ferramentas computacionais para o processamento dos sinais e bases de dados públicas relacionadas. Também são comentados alguns estudos de casos práticos, envolvendo códigos de programação e bases de dados para o tratamento e a classificação de sinais de EEG.

1.1. Introdução

Nos últimos vinte anos, percebe-se que o uso de tecnologias que aplicam os sinais biológicos do cérebro para interagir com dispositivos eletrônicos, ou interfaces cérebro-computador (ou BCI, do inglês *Brain-Computer Interface*) [Abbaspour et al., 2020], está

cada vez mais presente. Esses sinais cerebrais podem ser capturados, direta ou indiretamente, de muitas maneiras e por meio de equipamentos variados como dispositivos de Eletroencefalografia (EEG) e de ressonância magnética funcional (fMRI – *functional MRI*).

As primeiras referências na literatura acerca da possibilidade de se capturar os sinais cerebrais de seres vivos data do final do século 19, quando Richard Caton verificou a existência de correntes elétricas no córtex cerebral de coelhos e macacos. Contudo, existem poucos detalhes sobre tal experimento, já que a única referência seria a demonstração feita em um congresso de fisiologistas ingleses, que estavam reunidos em Edimburgo [of Birkenhead, 1959]. Ainda no século 19, outros pesquisadores replicaram o experimento de Caton por meio da leitura dos sinais cerebrais em diversos animais, como Fleichl von Marxow em 1883, Francis Gotch e Victor Horsley em 1889, Adolf Beck em 1890 e Vasily Danilewsky em 1891 [Silva Junior, 1945].

Em 1902, Hans Berger iniciou seus primeiros experimentos com a captura de sinais cerebrais. Em 1924, Berger obteve o primeiro eletroencefalograma humano, com o galvanômetro de dupla bobina de Siemens e Halske. Em 1929, ele fez sua primeira publicação intitulada “Über das Elektrenkephalogramm des Menschen”, na qual apresentou os resultados obtidos [Ince et al., 2021, Pazzanese and Pupo, 1948].

É possível classificar os métodos de coleta de sinais de EEG em duas categorias principais: invasivas e não invasivas. As técnicas de coleta invasivas necessitam de um canal de comunicação direto com o corpo humano para obter os sinais desejados. Em geral, isso ocorre na forma de um grupo de eletrodos inseridos no cérebro e resulta em informações mais assertivas, com uma resolução espacial maior e com interferências mínimas ou inexistentes causadas por outros sinais, como movimentos musculares e outros tipos de artefatos. Contudo, essa metodologia traz um risco maior à saúde do usuário, uma vez que é inserido um elemento estranho diretamente no corpo da pessoa via processos cirúrgicos [Nam et al., 2018, Kawala-Sterniuk et al., 2021].

Já as metodologias não invasivas captam sinais por meio de equipamentos que dispensam procedimentos cirúrgicos e que, portanto, são muito mais seguros ao usuário. Por outro lado, tais equipamentos possuem, comumente, uma menor resolução temporal das informações e originam sinais mais propensos a interferências externas como correntes elétricas produzidas por outros artefatos e movimentos musculares [Nam et al., 2018, Kawala-Sterniuk et al., 2021].

Alguns fatores influenciam a disseminação das tecnologias baseadas em EEG. Entre eles, destaca-se o desenvolvimento de novas técnicas de processamento de dados, o avanço da miniaturização dos componentes eletrônicos e o desenvolvimento de processadores cada vez mais rápidos. O emprego de técnicas de aprendizado de máquina (AM) na classificação de sinais de EEG também vem contribuindo para tanto.

Consequentemente, a popularização dos dispositivos de EEG fez com que se intensificassem certas linhas de pesquisa, bem como surgissem novas aplicações que antes eram difíceis de serem exploradas. Dentre os temas mais populares relacionados ao uso de EEG, podemos citar a meditação guiada [Hunkin et al., 2021], o *neuromarketing* [Golnar-Nik et al., 2019], a ergonomia [Hou et al., 2015] e a neurociência [Lo-

pes da Silva, 2013], além de estudos psiquiátricos [Koo et al., 2017] e de psicologia cognitiva [Kawala-Sterniuk et al., 2021]. Muitos outros campos de aplicação existem e se constituem oportunidades interessantes de pesquisa.

O presente capítulo aborda as atividades de coleta, processamento e classificação de sinais de EEG, com foco no desenvolvimento de aplicações de Interface Cérebro-Computador (em inglês, *Brain-Computer Interface*, BCI). O objetivo é proporcionar conhecimentos básicos, mas abrangentes, que possibilitem iniciar os estudos sobre o assunto, estimulando o interesse de novos pesquisadores e contribuir para o início de novas pesquisas científicas na área.

O restante do capítulo está organizado da seguinte forma: na Seção 1.2, são apresentados os principais conceitos necessários para compreender os temas e experimentos apresentados nas seções seguintes; a Seção 1.3 apresenta o processo de planejamento dos experimentos a serem executados, como a aplicação e o objetivo do estudo e um resumo dos diversos equipamentos de EEG disponíveis no mercado para a pesquisa com BCIs; por sua vez, a Seção 1.4 apresenta os diversos sistemas computacionais de apoio utilizadas pelos pesquisadores ao longo do desenvolvimento de seus trabalhos, bem como algumas soluções implementadas por eles; na Seção 1.5, são descritas os principais métodos usados no pré-processamento, extração de características e divisão dos dados dos sinais de EEG, além de um relato sobre o treinamento dos modelos de aprendizado de máquina para a classificação dos dados e a avaliação dos resultados; a Seção 1.6 apresenta algumas das bases de dados disponíveis como *benchmarks* e uma base construída pelos autores; a Seção 1.7 descreve dois estudos de caso da área de BCI; e, por fim, na Seção 1.8, são sumarizadas as principais conclusões e sugeridas bibliografias para aprofundar os estudos na área.

1.2. Conceitos básicos

Nesta seção, abordamos brevemente a estrutura e o funcionamento do cérebro, bem como padrões em sinais de EEG. Os conceitos teóricos apresentados são baseados, prioritariamente, nos trabalhos de Mark Bear *et al.* [Bear et al., 2017], Daniel Fuentes *et al.* [Fuentes et al., 2014], Robert Lent [Lent, 2010], e Saeid Sanei e Jonathan Chambers [Sanei and Chambers, 2007].

O sistema nervoso central humano consiste no encéfalo e na medula espinhal. O encéfalo, por sua vez, compreende as estruturas desse sistema que estão no crânio, sendo, portanto, o cérebro, o cerebelo, o tronco encefálico e suas conexões¹.

O cérebro é o componente principal do sistema nervoso central, responsável por

¹Outras classificações dividem o encéfalo em três partes: Prosencéfalo (ou Cérebro Anterior, formado pelo Telencéfalo e Diencéfalo), Mesencéfalo (Cérebro Médio, contendo o Teto e o Tegmento) e Rombencéfalo (Cérebro Posterior, composto pelo Metencéfalo e o Mielencéfalo). O Telencéfalo contém o Córtex Cerebral, os Gânglios da Base e o Sistema Límbico (este inclui estruturas como o hipocampo e a amígdala). O Diencéfalo contém o Tálamo (que atua como um centro de retransmissão para a maioria das informações sensoriais que chegam ao córtex cerebral) e o Hipotálamo (que regula funções vitais como temperatura corporal, fome, sede e liberação de hormônios). O Metencéfalo contém a Ponte (realiza a comunicação entre várias partes do encéfalo e a medula espinhal, além de participar na regulação da respiração) e o Cerebelo. O Mielencéfalo é formado pelo Bulbo (que controla as funções vitais autônomas, como a respiração, o batimento cardíaco e a pressão arterial).

diversas funções cognitivas, como pensamento, memória, emoção, controle motor e percepção sensorial. O cérebro humano é normalmente dividido em seções para facilitar o seu estudo. São considerados, assim, a presença de dois hemisférios cerebrais, o hemisfério esquerdo (associado com habilidades linguísticas, lógica, matemática e análise detalhada) e o hemisfério direito (associado com habilidades espaciais, reconhecimento de rostos, música e interpretação do contexto). Também assume-se uma separação do cérebro em uma camada externa de substância cinzenta, chamada de *córtex cerebral*, e uma massa interna de substância branca. Já o cerebelo é uma estrutura localizada na parte de trás do cérebro, abaixo do córtex cerebral, que tem um papel importante no controle motor fino, na coordenação de movimentos e no equilíbrio. O tronco encefálico conecta o cérebro à medula espinhal. Esse tronco compõe-se do bulbo raquidiano, que controla as funções autonômicas vitais, e da ponte e do mesencéfalo, que desempenham papéis importantes no processamento sensorial e motor. O encéfalo também possui uma rede de conexões neurais que realizam a comunicação entre diferentes áreas cerebrais e a coordenação de atividades.

Concentraremos o estudo agora nos componentes do cérebro. O córtex cerebral possui cerca de 2mm a 4mm em um cérebro adulto, e é responsável pelo processamento de informações, tomada de decisões, controle motor e percepção sensorial. Ele é formado por dobras com elevações (referenciadas como *giros*) e depressões (conhecidas como *sulcos*). Já a massa branca está associada às funções de transmissão de sinais entre diferentes partes do cérebro e à integração de funções motoras, sensoriais e cognitivas.

O córtex cerebral, em cada um de seus hemisférios, é comumente dividido em lobos, nos quais se verificam funções cognitivas e/ou de percepção dos sentidos:

- Lobo Frontal – localizado na parte frontal do cérebro, está associado a funções executivas, ao controle motor voluntário, ao planejamento e à tomada de decisões. A porção mais anterior do lobo frontal (região do cérebro localizada na parte frontal da cabeça, atrás da testa) é chamada de área *Pré-frontal*, sendo crucial para o planejamento, a resolução de problemas, a tomada de decisões, a adaptação e o controle emocional.
- Lobo Parietal – localizado na parte superior do cérebro, é responsável pelo processamento sensorial e pela integração, percepção espacial e manipulação de objetos.
- Lobo Temporal – localizado nas laterais do cérebro, está associado ao processamento auditivo, à memória e ao reconhecimento de padrões e faces.
- Lobo Occipital – localizado na parte posterior do cérebro, é onde se concentra o processamento visual.

Além desses quatro lobos, destaca-se a *ínsula* (ou *lobo insular*), uma região do córtex cerebral localizada profundamente nas dobras do cérebro, entre o lobo temporal e o lobo frontal. A ínsula está associada a funções importantes, como o processamento sensorial (por exemplo, a percepção do gosto e de sensações como dor e desconforto), as funções emocionais e sociais, as funções autonômicas e a integração multissensorial.

Existem ainda estruturas subcorticais do cérebro, a saber: Gânglios da Base (conjunto de núcleos envolvidos no controle motor e na coordenação de movimentos), Hipocampo (crucial para a formação de novas memórias) e Amígdala (envolvida em respostas emocionais e processamento emocional).

Deve ser notado, contudo, que as regiões do cérebro não são propriamente sedes de funções cognitivas, mas sim locais em que há maior convergência de circuitos neurais responsáveis por elas [Fuentes et al., 2014].

Algumas outras formas de subdivisão do cérebro enfatizam a existência de mais de uma região dedicada a um mesmo tipo de processamento. Luria (1973) [Luria, 1976], por exemplo, propõe a existência de unidades de *recepção* de informação e de *execução* no cérebro. Além disso, o autor assume que essas unidades possuem três áreas hierárquicas – primárias, secundária ou terciárias –, por meio das quais as informações são geradas indo de níveis mais básicos de percepção até um nível mais alto e abstrato. Mesulam [Mesulam, 2000], por sua vez, apresenta uma proposta de subdivisão funcional do cérebro em córtex de projeção, córtex de associação unimodal, córtex de associação multimodal, córtex de associação supramodal e córtex límbico, com essas áreas associadas a graus diferentes de abstração.

O sentido da visão, como caso ilustrativo, apoia-se em porções do lobo occipital para um processamento inicial das informações visuais. Dessa região, saem conexões em direção aos lobos parietal e frontal, para pontos que determinam “onde” a informação está. Também saem conexões para o lobo temporal, onde ocorre o reconhecimento do que é a imagem, com a interpretação de objetos e rostos.

O processamento da linguagem humana é outro caso representativo do envolvimento de mais de um lobo na sua realização, uma vez que contempla processos que ocorrem nos lobos frontal, temporal e parietal.

A Figura 1.1 mostra o mapa citoarquitetônico cortical de Brodmann [GUILLERY, 2000] adaptado para destacar, em cores, os lobos do córtex. As regiões numeradas estão associadas a atividades cerebrais específicas, sendo que a Tabela 1.1 desta algumas dessas vinculações. Informações sobre a relação entre outras atividades e as regiões no mapa de Brodmann podem ser obtidas em [GUILLERY, 2000, wik, 2022].

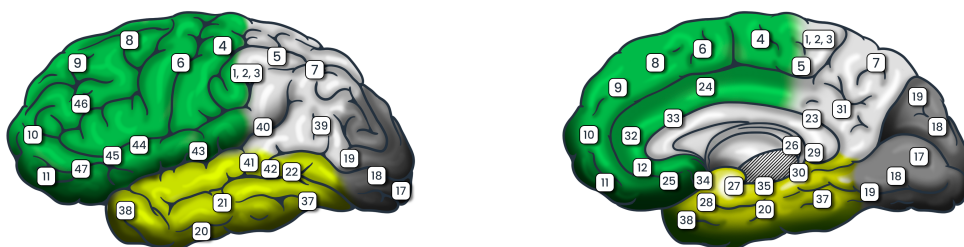


Figura 1.1. Mapa citoarquitetônico cortical de Brodmann adaptado para destacar, em cores, os lobos do córtex. Os lobos apresentados são: Frontal (Verde-escuro), Parietal (Branco), Temporal (Verde-claro) e Occipital (Cinza). Fonte: adaptação de figuras em domínio público disponibilizadas em https://pt.wikipedia.org/wiki/\%C3\%81rea_de_Brodmann.

Tabela 1.1. Regiões no mapa citoarquitetônico cortical de Brodmann associadas a algumas funções cognitivas. Fonte: Adaptado de <https://www.fmriconsulting.com/brodmann/functions.html>.

Grupo	Função	Áreas de Brodmann
Motor	Motor primário	4, 1, 2, 3
	Motor secundário	6, 8
	Planejamento motor	6, 13–16, 24, 32–33, 40
	Imaginação motora	5, 7, 4, 6, 8, 24, 32–33
	Aprendizado motor	4, 1-3, 6, 8, 23, 26, 29-31
	Movimentos sacádicos	4, 5, 7, 6, 8, 17, 18, 19, 46
	Inibição do piscar	4
Audição	Processamento básico	41, 42
	Processamento de sons complexo	21, 22
	Imaginação auditiva	8, 9, 10
	Vozes familiares	38
Visão	Intensidade / padrões de luz	17, 18, 19
	Discriminação de cor	17
	Integração visual	20
	Processamento de movimento visual	37
Atenção	Visual	17, 18, 37
	Visiomotor	5, 7, 6, 8
	Visioespacial	6, 8, 39, 24, 32-33, 45
	Seletiva a sons	6, 9, 10, 24, 32–33
	À fala	20, 22; 23, 26, 29-31, 38, 47

Em um nível mais elementar, o funcionamento do encéfalo humano se baseia em suas células e na forma como elas se interligam [Lent, 2010, Bear et al., 2017]. O encéfalo possui cerca de 86 bilhões de neurônios e quase a mesma quantidade de células gliais, além de apresentar outros tipos de células, como as endoteliais, ependimárias e células-tronco neurais. Os neurônios são, geralmente, as células que recebem maior atenção dos pesquisadores, por serem as responsáveis pela transmissão de sinais elétricos e pela comunicação no cérebro. A maior concentração de neurônios no encéfalo se encontra no cerebelo, o qual possui cerca de 69 bilhões dessas células. Os neurônios são normalmente classificados de acordo com sua morfologia, função, localização e conexões. As células gliais, por sua vez, são de vários tipos e desempenham importante papel de suporte, proteção e manutenção das funções neuronais.

A Figura 1.2 ilustra a estrutura de um neurônio, a qual é formada por três partes: o soma (também referenciado como corpo celular e pericário), os dendritos e o axônio. O *soma* é a parte central do neurônio e tem formato aproximadamente esférico. O axônio e os dendritos, classificados como *neuritos*, são tubos finos ligados ao soma. Os dendritos possuem pequena dimensão e responsáveis pela recepção de sinais químicos advindos de outros neurônios. A quantidade de dendritos ligados ao soma varia segundo o tipo de neurônio, podendo ser de poucas unidades a várias dezenas ou centenas. O axônio é o componente responsável por transmitir um sinal elétrico do soma por meio de seu duto e, no final, por uma terminação axonal. Ele é único em cada neurônio e pode ter

ramificações nas suas terminações, as quais se conectam a pontos receptores de dendritos de outros neurônios. Enquanto os dendritos passam raramente de $2mm$, o comprimento de um axônio pode variar de alguns milímetros a vários centímetros, conforme o tipo de neurônio a que pertence.

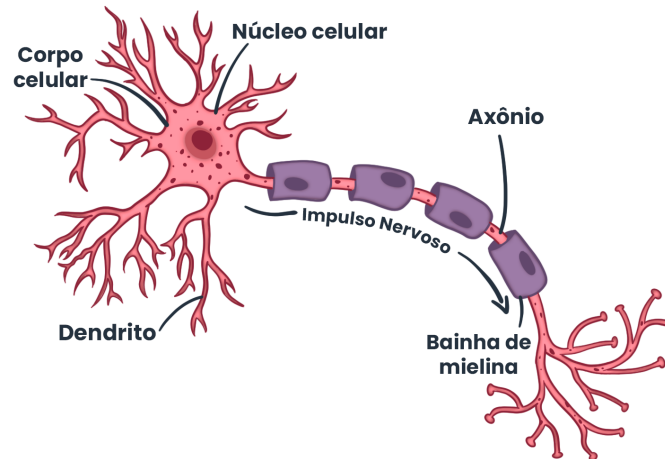


Figura 1.2. Estrutura de um neurônio.

O processo de transmissão de sinal de um neurônio para outros chama-se de *sinapse*. Nesse processo, sinais elétricos recebidos pelos dendritos de um neurônio se propagam até seu corpo celular. Se a agregação desses sinais alcançar um limiar crítico, um potencial de ação é gerado e transmitido na forma de impulso pelo axônio. Ao chegar nos terminais axonais, o impulso nervoso libera neurotransmissores que atravessam o que é chamado de fenda sináptica e se ligam aos receptores nos dendritos dos demais neurônios.

Sinapses também são possíveis entre dendritos de diferentes neurônios, embora sejam menos comuns do que aquelas entre axônios e dendritos. Essas conexões, chamadas de sinapses dendrodendríticas ou sinapses dendríticas, são encontradas em áreas específicas do sistema nervoso, como no córtex cerebral e no bulbo olfatório.

Quando os neurônios são ativados, as correntes elétricas decorrentes das sinapses geram um campo magnético mensurável por máquinas de eletromiograma (EMG) e um campo elétrico secundário sobre o couro cabeludo que pode ser medido por sistemas de EEG [Sanei and Chambers, 2007]. Entretanto, os sinais elétricos sofrem atenuação quando atravessam as estruturas do corpo humano (o crânio, por exemplo, produz uma atenuação que é cerca de cem vezes maior do que aquela causada pelos tecidos moles) e são contaminados por ruídos produzidos no cérebro (ruído interno) ou sobre o couro cabeludo (ruído do sistema ou ruído externo). Assim, para que esses sinais possam efetivamente ser utilizados, eles devem advir de uma grande quantidade de neurônios ativos para gerar um potencial elétrico suficiente, e terem passados por um processo de amplificação. Também deve ser considerada a localização dos sensores de EEG, uma vez que eles resultam em sinais mais claros se posicionados próximos às regiões do encéfalo onde se concentram as funções cerebrais a serem monitoradas.

O *Sistema 10-20* é um padrão internacional amplamente utilizado para o posi-

cionamento de sensores de EEG no escalpo. Esse sistema é ilustrado na Figura 1.3, mostrando uma configuração com 21 eletrodos (à esquerda) e outra com 71 eletrodos. O sistema baseia-se em quatro posições de referência principais: o *nasion* — o qual é o ponto de depressão na junção do nariz com a testa, o *inion* — representado pela protuberância óssea na parte de trás do crânio, e as orelhas esquerda e direita. A cabeça é então medida longitudinalmente (do nasion ao inion) e lateralmente (de uma orelha à outra), passando pelo ponto mais alto. Em seguida, a posição dos eletrodos é definida em intervalos de 10% ou 20% dessas distâncias, garantindo uma cobertura completa e simétrica do couro cabeludo.

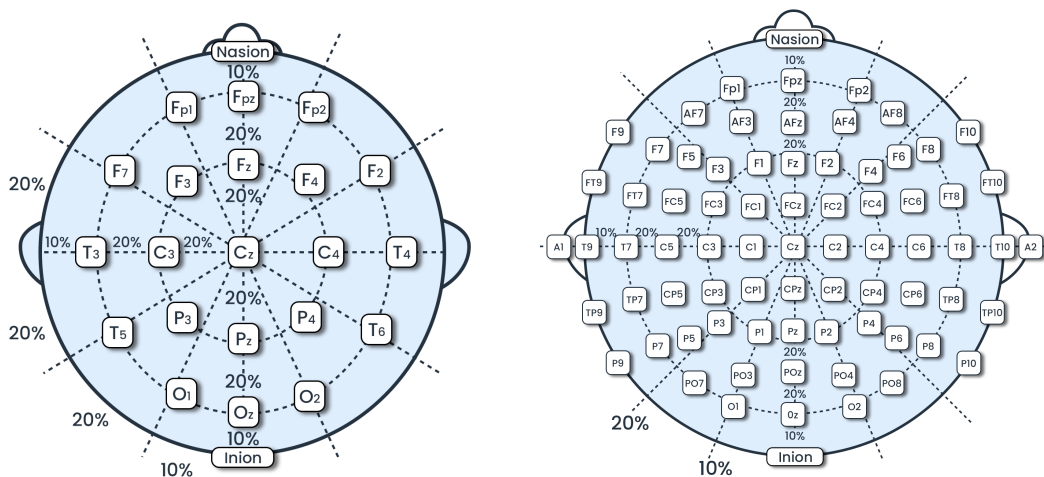


Figura 1.3. Sistema 10-20, com 21 posições para eletrodos (à esquerda) e 71 posições (à direita).

Na imagem à esquerda, os espaçamentos são de 10% na vizinhança do nasion e do inion em ambas as direções (longitudinal e lateral), e de 20% nas demais regiões. Na imagem da direita, um novo posicionamento é definido para cada intervalo de 20% de espaçamento, elevando assim a quantidade possível de eletrodos.

Cada posicionamento no Sistema 10-20 recebe um identificador próprio. Esse identificador começa por um termo que indica a região longitudinal da cabeça, sendo **Fp** para Frontal Polar, **F** para Frontal, **C** para Central, **P** para Parietal, **O** para Occipital e **T** para Temporal. Quando a posição dos lóbulos das orelhas são considerados, usa-se a letra **A** para marcar suas posições. O identificador finaliza com um número, indicando a posição lateral. São empregados, neste caso, números ímpares (1, 3, 5, 7) para as posições no hemisfério esquerdo, números pares (2, 4, 6, 8) para o hemisfério direito, e a letra **z** (indicando o zero) para a linha média central.

1.2.1. Padrões mentais

As ondas cerebrais brutas adquiridas pelos equipamentos de EEG podem ser caracterizadas em cinco tipos bastante comuns, conforme suas faixas de frequência: delta (δ), teta (θ), alfa (α), beta (β) e gama (γ). Dado que há pequenas variações quanto às frequências inicial e final que delimitam cada tipo de onda, apresentamos aqui os seus intervalos mais comuns.

As ondas delta, com frequências de 0,5 a 4Hz, exibem altas amplitudes e são predominantes durante o sono profundo, intensificando-se quando o indivíduo está privado de sono. Elas estão ligadas ao sonho e ao aprendizado. As ondas teta, que variam de 4 a 8Hz, compartilham alta amplitude, ocorrendo em estados de sono leve e associadas à reminiscência. As ondas alfas, com uma frequência de 8,0 a 13,0Hz e alta amplitude, são proeminentes durante a vigília, descanso e relaxamento, particularmente associadas ao córtex visual. Por outro lado, as ondas betas, com uma frequência de 13,0 a 30,0Hz e menor amplitude, são predominantes durante a atividade intelectual e o estado de alerta. Por último, as ondas gama, que variam de 30,0 a 140Hz, estão associadas a tarefas cognitivamente exigentes que requerem atenção, memória e percepção.

Em termos de localização, as ondas alfa e beta se manifestam principalmente na região frontal do cérebro, enquanto as ondas teta ocorrem na região temporal, e as ondas delta e gama são globais. A Tabela 1.2 sintetiza as principais características dessas cinco faixas de frequência.

Tabela 1.2. Ondas cerebrais em faixas de frequência específicas.

Onda	Faixa de Freq.	Região Predominante	Associação
Delta	0,5 – 4Hz	Global	Sono profundo e aprendizado
Teta	4 – 8Hz	Temporal	Estados de relaxamento profundo, meditação, sono leve
Alfa	8 – 13Hz	Frontal	Relaxamento, estado de alerta tranquilo
Beta	13 – 30Hz	Frontal	Atenção ativa, resolução de problemas, pensamento crítico
Gama	> 30Hz	Global	Processamento de informações de alta complexidade, funções cognitivas elevadas

Outros padrões de EEG bem conhecidos incluem os *steady-state evoked potentials* (SSEPs) e o P300, os quais são sinais decorrentes da percepção de determinados estímulos externos [Tanaka and Arvaneh, 2018]. Os SSEPs são gerados quando o indivíduo percebe um estímulo periódico, como uma imagem piscante ou um som repetitivo. Eles aparecem na região do cérebro onde o estímulo é processado e possuem frequências correspondentes às da estimulação. Aplicações que exploram esse padrão piscam, por exemplo, diferentes áreas de uma tela de computador, cada uma em uma frequência distinta. Quando o usuário da aplicação foca a sua atenção em uma dessas áreas, é possível, pela frequência do sinal capturado por meio de EEG, saber onde ele está olhando. Já o P300 é um sinal que aparece principalmente na lobo parietal cerca de 300 ms após a percepção de um estímulo relevante mas pouco frequente. Uma forma típica de empregar o P300 consiste em apresentar diversos ícones (representando opções a serem escolhidas) distribuídos em uma tela e solicitar ao usuário para prestar atenção no ícone de seu interesse. Repetidamente e de forma aleatória, um ícone por vez é selecionado e destacado visualmente. Quando a seleção recai no ícone de interesse do usuário, o sinal P300 emerge.

Há mais padrões de ondas cerebrais que são perceptíveis quando determinados

eventos sensoriais ou cognitivos acontecem, ou que são característicos de movimentos musculares. Alguns estudos também têm utilizado EEG para detectar transtornos mentais. Para uma visão mais detalhada de alguns desses padrões, sugere-se a leitura de Sanei e Chambers [Sanei and Chambers, 2007], Nam *et al.* [Nam et al., 2018] e Tanaka e Arvaneh [Tanaka and Arvaneh, 2018].

1.3. Planejamento de experimentos de EEG

A criação de um experimento envolvendo dispositivos de EEG consiste em estabelecer inicialmente o que se pretende com o experimento e, a partir deste escopo definido, determinar qual será o ambiente de coleta, quais serão os estímulos (visuais, táteis, auditivos dentre outros) e como se dará a atividade à qual os participantes serão expostos (quantidade de sessões, repetições, tempo de exposição, intervalo entre sessões, dentre outros), bem como a coleta propriamente dita e a análise dos dados coletados. A escolha dos participantes da pesquisa deverá considerar a idade, a escolaridade, o sexo, o conhecimento prévio do domínio da aplicação e a existência de limitações físicas e mentais, entre outros aspectos.

1.3.1. Definindo a aplicação e o objeto de estudo

Os projetos de pesquisa com EEG de baixo custo abrangem uma ampla gama de aplicações, desde a análise da carga cognitiva de um condutor de veículos até a identificação de estados mentais durante a realização de atividades específicas. Teo *et al.* [Teo et al., 2018, Teo and Chia, 2018] exploraram o reconhecimento passivo de preferências e emoções humanas em resposta a estímulos visuais 3D, simulando uma montanha-russa. Bird *et al.* [Bird et al., 2018, 2019, Ashford et al., 2020] buscaram identificar características em sinais EEG para classificar estados mentais (relaxamento, neutro e concentração) durante diferentes atividades. No estado de relaxamento, os participantes ouviam música calma e sons meditativos. No estado neutro, o teste era semelhante ao de relaxamento, mas sem estímulos. Para o estado de concentração, os participantes jogavam um jogo de esconder uma bolinha em um copo. Sakalle *et al.* [Sakalle et al., 2021] submeteram os participantes a trechos de filmes e classificaram quatro classes negativas de emoções (tristeza, nojo, raiva e surpresa), além de três classes básicas de emoções (positivas, negativas e neutras). O trabalho de Almogbel *et al.* [Almogbel et al., 2018, 2019] combinou EEG com um simulador de direção para investigar a carga cognitiva durante a condução. O estudo envolveu dois tipos de sessões de coleta. O primeiro tipo tinha como objetivo simular um ambiente com alta carga cognitiva e, para isto, o trânsito era similar ao de uma metrópole. Já o segundo sugeria baixa carga cognitiva, sendo representado por um ambiente monótono de uma rodovia vazia, não necessitando de muita atenção ao dirigir. O experimento envolveu um único participante do sexo masculino com 29 anos, que já exercia a profissão de motorista há 11 anos. Foram realizadas 24 sessões (12 para cada tipo de condição) no período de um mês. Cada sessão teve duração de 15 a 30 minutos. Por serem randomizados, os ambientes utilizados nas sessões poderiam variar entre noturno e diurno e com diferentes rotas.

Um dos programas existentes para a criação de experimentos ligados à área de

Psicologia é o Psychology Experiment Building Language (PEBL)². Com ele, é possível utilizar experimentos já disponíveis ou desenhar novos, por meio de uma linguagem de programação simples. Trata-se de um software livre, licenciado sob licença GPL, disponibilizado para várias plataformas de sistemas operacionais. Cho *et al.* [Cho et al., 2024] utilizaram, na pesquisa, experimentos relativos à atividade cerebral em Ciência, Tecnologia, Engenharia e Matemática (STEM). Eles avaliaram a memória de trabalho, o funcionamento aritmético, a flexibilidade cognitiva, a flexibilidade mental e o planejamento para uma tarefa. As atividades envolvidas em cada proposta foram: memorizar e combinar 9 imagens, respostas sim e não ao reconhecer problemas matemáticos, a combinação de cartões de estímulos e mudança de regras, alternando alfabeto e números com a conexão ordenada entre eles e empilhamento de objetos com o mínimo de movimentos. O PEBL foi empregado para desenvolver as interfaces computacionais. O trabalho avaliou ainda a correlação das tarefas com as regiões do cérebro.

Uma vez estabelecidas as características do experimento e dos participantes, faz-se necessário verificar a forma em que os dados serão coletados, transmitidos e armazenados. Em geral, cada equipamento de EEG possui uma forma própria de transmitir os dados coletados para algum outro dispositivo computacional, a fim deles serem armazenados e processados.

Por exemplo, no caso de estudos com o equipamento Muse 2, os sinais coletados geralmente são transmitidos para um smartphone ou tablet em tempo real a uma taxa de amostragem de 220 a 256Hz via Bluetooth. Em versões anteriores desse equipamento, a empresa criadora, a Interaxon, oferecia um *framework* para desenvolvimento de aplicações, sendo possível a construção de interfaces computacionais para receber os dados coletados. No entanto, com a versão atual do Muse, esse *framework* não é mais disponibilizado. Uma empresa terceirizada comercializa o aplicativo denominado Mind Monitor, o qual permite receber os sinais do Muse 2 e retransmití-los, via OSC³, para qualquer outra aplicação. A Figura 1.4 demonstra o fluxo de coleta até a gravação em alguma base de dados. Alternativamente, o Mind Monitor pode, ao final da coleta, gerar um arquivo padrão CSV e transmití-lo para uma conta do Dropbox, previamente configurada. Na figura, uma aplicação apresenta estímulos visuais ao usuário ao mesmo tempo em que envia informações sobre eles a outra aplicação, destacada como Coletor. O Coletor recebe as informações sobre os estímulos e as associa aos dados de EEG correspondentes recebidos em tempo real do Mind Monitor.

1.3.2. Escolhendo o dispositivo e os sinais dos sensores

Os equipamentos de EEG diferem quanto a algumas características como a quantidade e a localização de sensores no crânio, o tipo de sensor utilizado, a taxa de amostragem, o tipo de aplicação para o qual foi projetado e a plataforma de software disponível para seu uso, além do próprio valor de aquisição. O padrão 10-20 [Aznan et al., 2018a] tem sido adotado para definir as posições dos sensores desses equipamentos no couro cabeludo.

²Disponível gratuitamente no site <https://pebl.sourceforge.net>.

³O Open Sound Control (OSC) é um protocolo de comunicação entre computadores e sintetizadores de som, dentre outros dispositivos, que atua como padrão aberto e independente, baseado em mensagens [Wright et al., 1997].

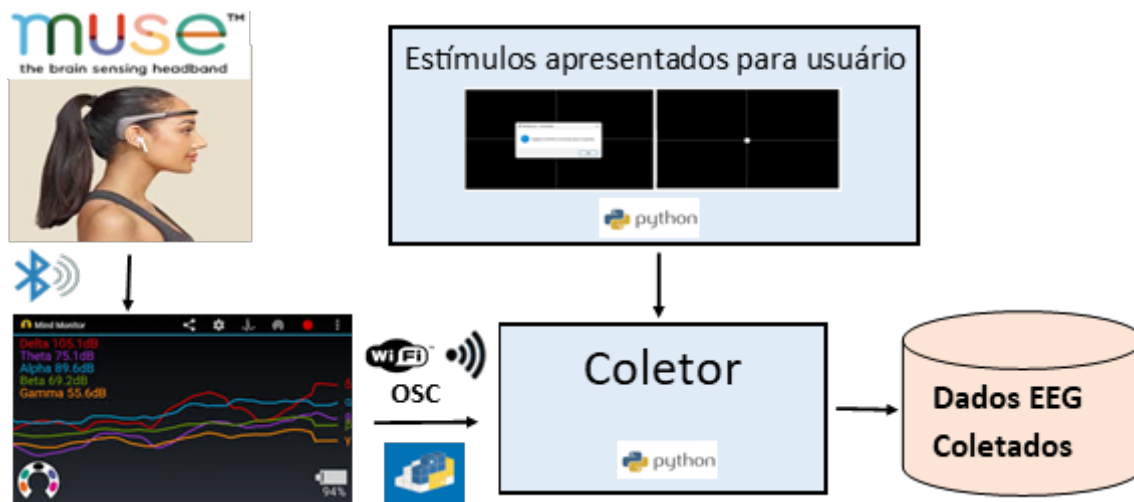


Figura 1.4. Fluxo de coleta para Muse 2.

Roy *et al.* [Roy et al., 2019] realizaram uma extensa revisão sistemática sobre o aprendizado profundo baseado em equipamentos de eletroencefalografia. Os relatos são de que a quantidade de sensores utilizados nas pesquisas apresentam uma melhoria substancial ao sair de 2 para 6 sensores, além de que, acima de 22 sensores, é possível conseguir melhores resultados. Metade dos trabalhos estudados utilizaram taxas de 250Hz ou menos e as maiores taxas chegaram a 5000Hz. As bases de dados privados representaram 42% da origem dos dados que constam dos trabalhos avaliados nessa revisão. Houve dificuldade em comparar dados originários de vários neurologistas, já que foram utilizadas diferentes localizações no crânio. O estudo também trouxe informações sobre alguns equipamentos de baixo custo, a saber: o *Emotiv EPOC* com 14 sensores, o *Emotiv Insight* com 5 sensores, *OpenBCI* configurado com 8 ou 16 sensores, o Muse 2 com 5 sensores e Neurosky com 1 sensor. Os equipamentos de baixo custo são o foco do presente capítulo, e, por isso, os mais populares deles descritos a seguir.

Dispositivos de EEG

Diversos dispositivos de coleta de sinais de EEG de baixo custo vêm sendo empregados em pesquisas com BCI, como o EPOC+ da Emotiv, o Muse 2 da Interaxon, e o *OpenBCI*. Esses e outros equipamentos possuem diferentes capacidades e funcionalidades próprias e, portanto, são voltados a aplicações distintas. No entanto, é crucial ter conhecimento das características de cada equipamento para determinar o mais adequado à pesquisa em andamento. O objetivo desta seção é descrever alguns dispositivos de baixo custo que um pesquisador tem à sua disposição durante a realização de uma pesquisa na área de interfaces cérebro-computador, bem como fornecer os conhecimentos necessários para analisá-los segundo as suas necessidades.

- EPOC X – Emotiv

Um dos primeiros dispositivos dessa natureza é o EPOC, lançado em 2009 pela empresa Emotiv Systems na Austrália. Outra empresa, a Emotiv Inc., criada nos

Estados Unidos por um dos fundadores da empresa anterior, produziu em seguida dois novos dispositivos entre 2013 e 2014: o EPOC+ e o Emotiv Insight. O Emotiv EPOC+ (<https://www.emotiv.com/epoc>) possui 16 sensores gerando 14 canais de EEG na largura de banda de 0,16 a 43Hz, sendo dois sensores de referência, além de acelerômetro e magnetômetro. O dispositivo é compatível com Windows, MAC, iOS e Android e pode ser utilizado totalmente sem fio via bluetooth (cerca de 6 horas de autonomia da bateria, neste caso) ou por meio de um receptor USB proprietário (12 horas de autonomia). Para o seu uso, é necessário embeber os sensores em uma solução salina disponibilizada junto ao equipamento. Já o Emotiv Insight possui apenas 7 sensores, com 5 canais de EEG e bateria com capacidade de até 8 horas. Em 2020, a Emotiv lançou o EPOC X, também com 14 sensores de EEG e mais 2 sensores de referência.

- Insight – Emotiv

Insight é um dispositivo de 5 sensores semi secos, que não necessita de uma solução salina ou gel para captação, mas retiram a umidade necessária do próprio ambiente. Possui uma bateria que dura até 20 horas, conectividade Bluetooth Low Energy aprimorada, antena atualizada com suporte para Bluetooth 5, taxa de amostragem de 128Hz por canal, 5 canais posicionados nos referenciais AF3, AF4, T7, T8 e Pz do padrão 10-20.

A detecção conta com comandos mentais, com até 4 itens pré-treinados por perfil de treinamento, mais métricas de desempenho (Excitação, Engajamento, Relaxamento, Interesse, Estresse, Foco), expressões faciais (piscar, piscar para a esquerda/direita, surpresa, franzir a testa, sorrir, apertar maxilar), entretanto o acesso está sujeito a licença e EEG bruto também disponível.

Este dispositivo foi pensado para ser versátil, removendo interferências e artefatos na captura e transmissão de sinais, mantendo a mobilidade com um equipamento confiável, simples e prático, tanto no transporte como na preparação e montagem para uso.

- FLEX 2 Saline ou Gel – Emotiv

O FLEX é um sistema sem fio com os sensores acoplados a uma touca, variando sua configuração de até 32 sensores que podem ser posicionados em qualquer local em um touca com 74 orifícios seguindo o padrão internacional 10-20. Há uma variedade de tamanhos de toucas flexíveis, proporcionando conforto ao usuário e conformidade para dados de EEG ideais.

O Flex disponibiliza os dados brutos e dados de banda de frequência adquirindo uma licença PRO para uso. Ele capta sinais de EEG, ECG, EMG e EOG e possui uma conectividade Bluetooth 5.2 de alta fidelidade. Ele foi projetado para uma taxa de amostragem de 2048Hz interno, no entanto, faz uma redução para 256 amostras por segundo, sendo essa a taxa de entrega dos dados. Seus sensores precisam ser umedecidos com uma solução salina ou com gel, dependendo da versão escolhida.

O Flex tem uma bateria com duração de até 6 horas e conexão USB-C para um carregamento mais rápido.

- MN8 – Emotiv

O MN8 é um fone de ouvido que, além de receber o som do dispositivo conectado, também monitora as atividades cerebrais do usuário. Ele conta com apenas dois sensores dedicados para EEG que ficam posicionados atrás de cada uma das orelhas, limitando a quantidade de ondas a serem captadas e que podem ser estudadas. O MN8 possui um foco na detecção da cognitividade do usuário, auxiliando-o em atividades de meditação e concentração, com ou sem áudio.

- X-trodes – Emotiv

X-trodes é uma versão que faz uso de fitas adesivas condutivas, em pontos onde seriam colocados os sensores convencionais, aderindo à pele e adaptando-se de forma discreta, pensado para obter o uso do dispositivo por maior período de coleta atingindo até 10 horas. Desenvolvido pela Emotiv em parceria com X-trodes, este dispositivo é voltado principalmente para pesquisas. O dispositivo X-trodes possui 3 variações de montagem, podendo ser adquiridas em kits com os multi-eletrodos.

Outro ponto abordado é a necessidade de aquisição da licença da versão PRO do software da Emotiv para obtenção dos dados do dispositivo, além da compatibilidade com poucos sistemas operacionais, o que limita a possibilidade de implementação e desenvolvimento com o dispositivo.

- Muse 2 – Interaxon

O equipamento Muse 2 é fabricado pela Interaxon e comercializado principalmente no Canadá, possuindo um dos menores preços para aquisição. Ele tem, por finalidade, auxiliar na meditação, podendo essa ser guiada com um aplicativo instalado em um Smartphone. Composto por cinco sensores de EEG, o Muse 2 consegue fornecer tanto dados brutos, com taxa amostral entre 220Hz e 256Hz, quanto ondas separadas (delta, teta, alfa, beta e gama) a uma taxa amostral de 10Hz. O equipamento possui ainda giroscópio, acelerômetro e detector de batimento cardíaco.

Os sinais capturados por esse dispositivo advêm de quatro sensores (posicionados em TP9, TP10, FP1 e FP2) e são captados em microvolts com valores entre 0,0 e 1682,815uV. Há também um sensor auxiliar (AUX), cujos sinais servem de referência. Não necessita de uma solução salina ou gel para captação dos dados.

A forma principal de obtenção e transferência dos dados coletados, como explicado previamente, é por meio de um software denominado Mind Monitor, disponível nas lojas de aplicativos móveis. O software fornece comunicação via Bluetooth entre o Muse 2 e o dispositivo em que é instalado, seja ele um *smartphone* ou tablet. O Mind Monitor possibilita diversas configurações da sua interface gráfica com o usuário e de pré-processamentos, como, por exemplo, a aplicação de Transformada de Fourier sobre os dados coletados e a escolha da forma como esses dados serão transferidos ou salvos.

O Muse 2 conta com conectividade Bluetooth 4.2, uma porta Micro USB para carregamento e uma bateria íon de lítio recarregável com faixa de tempo operante de 10 horas ou 12 sessões.

- Muse S – Interaxon

O Muse S é semelhante ao Muse 2, mas com uma estrutura física mais robusta e com mais sensores de referência. Ele tem, por finalidade, auxiliar na preparação para o sono, na meditação e na concentração, de forma simples, com esses processos também guiados pelo uso de um aplicativo instalado em um Smartphone. O Muse S possui 7 sensores de EEG, sendo 2 na testa (posições AF7 e AF8), 2 atrás das orelhas (posições TP9 e TP10) e 3 sensores de referência (na posição frontal). De forma similar ao Muse 2, ele conta com um sensor PPG, que mede a frequência cardíaca, e um acelerômetro e um giroscópio que medem o ritmo da respiração e o movimento do corpo. Não necessita de uma solução salina ou gel para captação dos dados.

O Muse S mede de 48 à 63cm e conta com conectividade Bluetooth 4.2, uma porta Micro USB para carregamento e uma bateria íon de lítio recarregável com faixa de tempo operante de 10 horas ou 12 sessões. A taxa de amostragem dos seus sensores de EEG é de 256Hz.

- OpenBCI

O OpenBCI (<https://openbci.com>) é uma plataforma que oferece hardware e software abertos para estudo de EEG.

Existem diversas configurações e modelos de hardware para o OpenBCI, tais como placas de 4, 8 e 16 canais, todas elas compatíveis com outros dispositivos microcontroladores. Algumas dessas configurações utilizam eletrodos secos, enquanto outros adotam uma solução condutiva, como um *HeadBand* ou um capacete. As suas placas de circuito EEG são compatíveis com sistemas operacionais Mac OS X, Windows 7+ e Linux (Ubuntu/Debian ou Fedora/RHEL). Os dados das placas podem ser transmitidos por Bluetooth ou *dongle RFDuino*. Alguns dos sistemas utilizam pilhas AA, mas é possível trocá-las. Um kit vendido na loja *OpenBCI* pode custar entre US\$ 829 e US\$ 3699. Além disso, é possível imprimir alguns componentes do capacete em impressora 3D e criar a própria placa de circuito. O site da *OpenBCI* oferece, além da descrição dos componentes de hardware, um software gráfico para o estudo e o processamento dos sinais de EEG, um SDK para o desenvolvimento de aplicações de EEG e um sistema desenvolvido pela comunidade da *OpenBCI* para a realização de testes com o equipamento.

A Tabela 1.3 traz um resumo dos equipamentos descritos acima com detalhes para o preços, os sensores que compõe o equipamento e sua finalidade.

1.4. Software de apoio

Várias ferramentas de software auxiliam na utilização dos equipamentos de EEG mencionados na seção anterior. Existem aplicações proprietárias desenvolvidas pelas próprias empresas fabricantes dos equipamentos, algumas pagas e outras gratuitas, além de ferramentas de software livre desenvolvidas pela comunidade. As ferramentas mais utilizadas são descritas a seguir [Alves et al., 2019].

Para acessar os dados do equipamento Emotiv EPOC+, é necessário o uso do SDK disponibilizado em <https://www.emotiv.com/pages/developer> para

Tabela 1.3. Equipamentos de EEG de baixo custo.

Equipamento	Preço	Sensores	Quantidade de sensores	Finalidade
Emotiv EPOC X	U\$ 999,00	AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4; 2 referências: referências CMS/DRL em P3/P4; alternativa ao processo mastóide esqueda/direita	14	Uso pessoal e em pesquisa
Emotiv Insight	U\$ 499,00	AF3, AF4, T7, T8, Pz; 2 referências: referências CMS/DRL no processo mastóide esqueda	5	Uso pessoal e em pesquisa
Emotiv FLEX 2 Saline ou Gel	U\$ 1999,00	Posições configuráveis seguindo o padrão 10-20; CMS/DRL configurável em qualquer local 10-20 ou nas orelhas	32	Uso geral e ênfase em pesquisa
Emotiv MN8	U\$ 399,00	Atrás da orelha esquerda e direita (TP9 e TP10)	2 (+ 4 referências)	Uso pessoal e em pesquisa
Emotiv X-trodes	U\$ 10000,00	Três configurações de EEG; 16 canais (AF7, AF8, FP1, FP2, L1, L3, L4, L5, L8, R1, R3, R4, R5, R8, E1, E2), 12 canais (AF7, AF8, L1, L3, L4, L5, L8, R1, R3, R4, R5, R8) e outra com 2 canais	2, 12 e 16	Uso pessoal e em pesquisa
Muse 2	U\$ 250,73	TP9, AF7, AF8, TP10, Aux	5	Uso para auxílio em meditação
Muse S (Gen2)	U\$ 378,23	TP9, AF7, AF8, TP10	4	Uso para auxílio em Meditação e sessões para indução do sono
OpenBCI	U\$ 2.399,00	8-channel: Fp1, Fp2, C3, C4, P7, P8, O1, O2 16-channel: Fp1, Fp2, F3, F4, F7, F8, T7, T8, C3, C4, P7, P8, P3, P4, O1, O2	8, 16	Uso pessoal e em pesquisa

Mac OS e Windows. Existem três tipos de licença. A primeira, denominada Básica Gratuita, fornece APIs para recursos referenciados como Comandos Mentais, Frequência de Banda, Expressões Faciais, Dados de Movimentação e Métricas de Desempenho de Baixa Resolução. A segunda, chamada Não-comercial PRO, dá acesso aos dados brutos do dispositivo, a todas as APIs da licença Básica Gratuita. Esse software permite a visualização em tempo real dos dados do dispositivo Emotiv, acesso a dados brutos de EEG, dados de movimentação, qualidade de contato, gravação e armazenamento em nuvem.

Além do SDK de pesquisa, a Emotiv disponibiliza o EmotivBCI, software proprietário gratuito disponível para Windows e Mac OS. Este software oferece recursos de detecção de comandos mentais, reconhecimento de expressões faciais, métricas de desempenho e sensores de movimentação. Essas funções podem ser utilizadas em conjunto com o Node-RED Toolbox, uma ferramenta de código aberto desenvolvida pela IBM para conexão de dispositivos de IoT (*Internet of Things*).

Os desenvolvedores de aplicações utilizando o Muse, na sua primeira versão, podiam utilizar um SDK Developer com versões para Windows, Android, Unity e IOS, contudo esta ferramenta foi descontinuada. Há também ferramentas para os casos em que não se deseja escrever códigos, mas sim utilizar recursos prontos. Tais ferramentas funcionam nas plataformas Windows, Linux e Mac OS. O SDK Developer era de uso gratuito.

Para o *OpenBCI*, existe o *OpenBCI GUI* (destinado à visualização dos dados EEG em tempo real), os *SDKs OpenBCI Cyton* e *OpenBCI Ganglion* (para conexão de um bio-sensor a um computador). Há também um firmware a ser instalado no bio-sensor escolhido. Esses programas estão disponíveis para download gratuito no portal <https://openbci.com/> e são de código aberto. Há versões dos programas para Mac OS X, Windows 64-bit e Linux 64-bit e os SDKs permitem o desenvolvimento em Java, Python e Node.js. No site do *OpenBCI*, é possível baixar ainda o *The Ultracortex*, um modelo de capote para instalar os sensores pode ser impresso em 3D.

Além das ferramentas padrão disponibilizadas pelas empresas, existem algumas *toolboxes* que auxiliam na coleta e no processamento de dados de EEG. Uma delas é a plataforma livre e gratuita *OpenViBE*⁴, compatível com vários dispositivos de captura EEG e sistemas operacionais Windows e Linux. Ela permite a criação de diferentes tipos de interfaces cérebro-computador e a visualização de dados de EEG em tempo real, além de possuir algoritmos de processamento de sinais, funções de aprendizado de máquina e suporte a *scripts*.

Outra ferramenta livre e gratuita é a PyEEG⁵. O foco da PyEEG é ser uma biblioteca de funções em Python para a extração de características de dados de EEG, incluindo a possibilidade de exportação dessas informações para ferramentas de aprendizado de máquina.

Algumas ferramentas livres e gratuitas foram desenvolvidas para estender as funcionalidades do MATLAB. Entre elas se destacam a BCILAB, a FieldTrip e a EEGLAB.

⁴<http://openvibe.inria.fr/>

⁵<https://code.google.com/archive/p/pyeeg/>

A BCILAB⁶ é uma *toolbox* que adiciona ao MATLAB recursos para projetar, prototipar, experimentar e avaliar interfaces cérebro-computador. A FieldTrip⁷, usada para análise de dados de EEG, possui funções como métodos de processamento e análise avançados, e análise de tempo-frequência. Já a EEGLAB⁸ é uma *toolbox* que possui uma interface de usuário gráfica (GUI), permite processar dados EEG contínuos de alta densidade e oferece um ambiente de programação estruturado para armazenar, acessar, medir, manipular e visualizar dados de EEG relacionados a eventos. Alguns dos recursos da EEGLAB funcionam também com o software livre Octave, similar ao MATLAB.

Por fim, há a biblioteca de rotinas Biosig⁹, desenvolvida como software livre e gratuito para as áreas de Neuroinformática, Interfaces Cérebro-Computador, Neurofisiologia, Psicologia, processamento de sistemas cardiovasculares e pesquisa do sono. Ela possui funções para aquisição e processamento de dados, extração de características, classificação, modelagem e visualização de dados, entre outras tarefas. Há versões da Biosig para MATLAB, Octave, C/C++ e Python.

Na Tabela 1.4, é possível encontrar a relação dos programas de apoio com as informações do dispositivo alvo, a descrição, a linguagem de programação, o sistema operacional e o ano de atualização.

1.5. Fluxo de classificação de dados de EEG

O surgimento dos equipamentos de EEG abriu uma ampla área de pesquisa, não apenas na área da saúde, mas também na área da computação. Na computação, algumas etapas da pesquisa foram se consolidando, permitindo a obtenção de resultados palpáveis. A execução é composta por sete etapas: a definição da aplicação e dos objetivos de estudo; a escolha do dispositivo e dos sinais a serem utilizados; a configuração do experimento e aquisição dos dados; o pré-processamento; a extração de características; o uso e a separação dos dados; e, por fim, a utilização de métodos de classificação. As próximas sessões procuram explicar cada uma das etapas descritas.

A Figura 1.5, trata das etapas específicas do processamento da aplicação, após a coleta. Podem ainda operar sobre sinais de EEG de um tipo (bruto ou onda) previamente selecionado.

1.5.1. Pré-processamento

No pré-processamento dos dados de EEG a remoção de ruídos e de artefatos indesejáveis é uma etapa importante. Isto garantirá que a análise posterior irá se concentrar nas informações relevantes obtidas da atividade cerebral. Diversas técnicas podem ser utilizadas nessa etapa, como, por exemplo, a aplicação de filtros passa-baixa, passa-alta, transformadas específicas, dentre outras. A transformada de Fourier é bastante utilizada em aplicações de EEG, devido a sua capacidade de decompor o sinal em suas frequências constituintes e com isto permite a identificação e remoção de componentes indesejados com maior precisão.

⁶<https://scn.ucsd.edu/wiki/BCILAB>

⁷<http://www.fieldtriptoolbox.org/>

⁸<https://scn.ucsd.edu/eeglab/index.php>

⁹<http://biosig.sourceforge.net/>

Tabela 1.4. Software de apoio a pesquisas com EEG.

Software	Dispositivos Alvo	Descrição	Linguagem de Programação	Sistema Operacional	Ano de Atualização
Emotiv Development Kit(SDK)	Emotiv EPOC+, Emotiv Insight, Emotiv EPOC Flex	Fornecer acesso à API para diferentes fluxos de dados do EMOTIV. (Software proprietário, três níveis de licença)	Java, C#, C++, Python, Ruby, NodeJS, PHP	Mac, Windows e iOS	2024
Muse SDK	Muse na sua primeira versão	Oferecia conectividade entre o dispositivo Muse e diferentes plataformas operacionais. (Software proprietário e gratuito)	C++	iOS, Android, Windows e Unity	Descontinuado
OpenBCI GUI	OpenBCI	O OpenBCI GUI é a ferramenta de software padrão do OpenBCI para visualizar, gravar e transmitir dados dos OpenBCI Boards. (Software Livre e gratuito)	Python, Node.js, Java	Mac, Windows e Linux	2023
OpenViBE	Emotiv EPOC+, Emotiv Insight, Emotiv EPOC Flex, OpenBCI, Muse	O OpenViBE é uma plataforma de software que permite projetar, testar e usar Interfaces Cérebro-Computador (BCIs). Também pode ser usado como um sistema genérico de aquisição, processamento e visualização de EEG em tempo real. (Software Livre e gratuito)	Python, MATLAB, C++, Lua	Windows e Linux	2024
PYEEG	Emotiv EPOC+, Emotiv Insight, Emotiv EPOC Flex, OpenBCI, Muse	Módulo Open Source para extração de características EEG/MEG. (Software Livre e gratuito)	Python	Mac, Windows, Linux	2018
BCILAB	Emotiv EPOC+, Emotiv Insight, Emotiv EPOC Flex, OpenBCI, Muse	Toolbox MATLAB para pesquisa de interface cérebro-computador. (Software Livre e gratuito)	MATLAB	Mac, Windows, Linux	2020
FieldTrip	Emotiv EPOC+, Emotiv Insight, Emotiv EPOC Flex, OpenBCI, Muse	Toolbox MATLAB para análise de MEG, EEG, iEEG e NIRS. (Software Livre e Gratuito)	MATLAB	Mac, Windows, Linux	2023
EEGLAB	Emotiv EPOC+, Emotiv Insight, Emotiv EPOC Flex, OpenBCI, Muse	Toolbox para MATLAB interativa capaz de processar dados EEGs contínuos, relacionados a eventos. (Software Livre e gratuito)	MATLAB	Mac, Windows, Unix, Linux	2024
BioSig	Emotiv EPOC+, Emotiv Insight, Emotiv EPOC Flex, OpenBCI, Muse	Biblioteca de software Open Source para processamento de sinais biomédicos. (Software Livre e gratuito)	MATLAB, C/C++, Python	Mac, Windows, Linux	2024

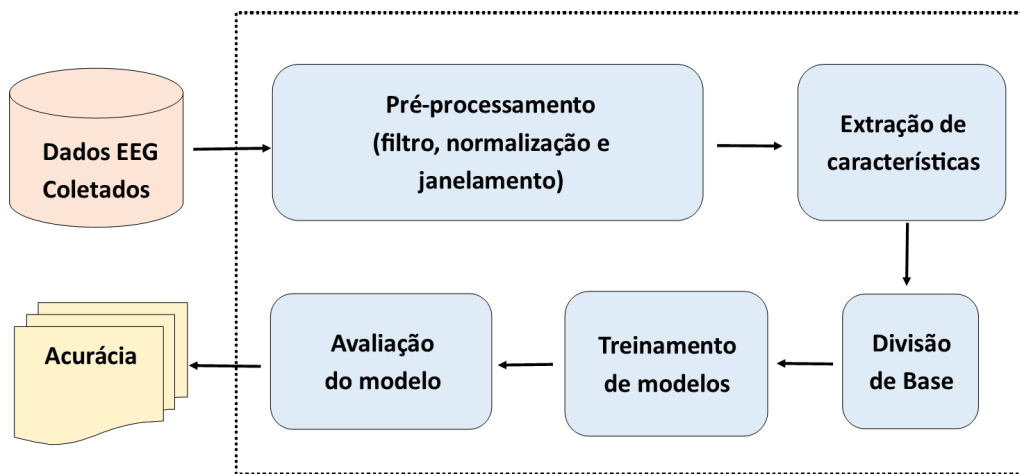
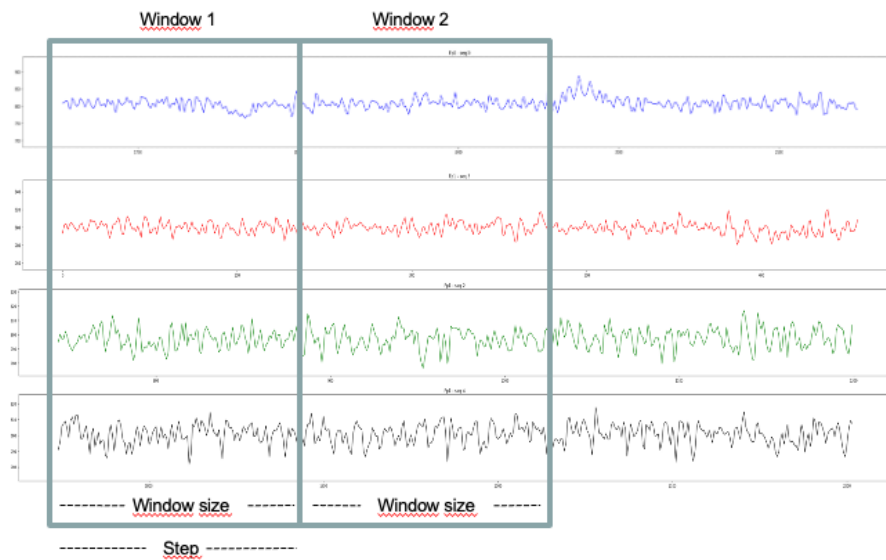


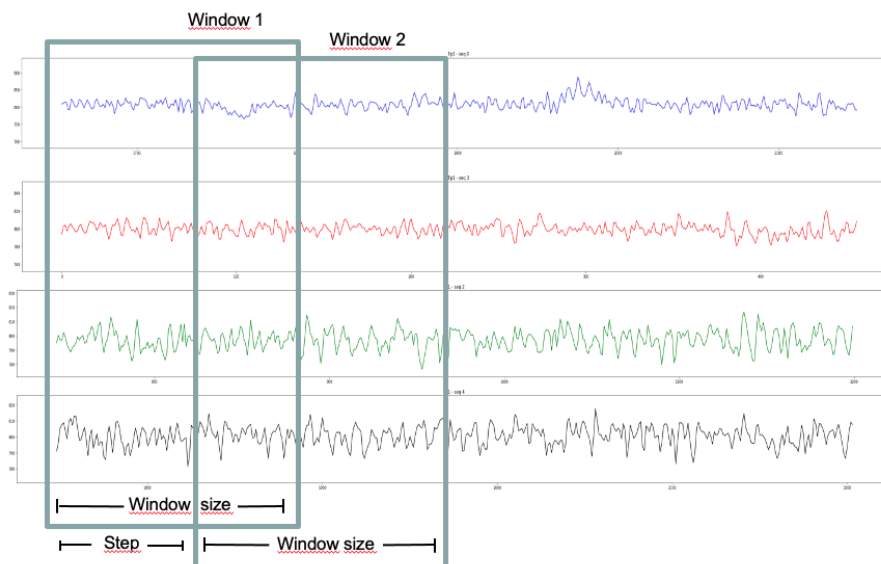
Figura 1.5. Fluxo de processamento de sinais de EEG.

Apesar da possibilidade de uso do pré-processamento, esta etapa demanda tempo de processamento. Recentemente, alguns trabalhos de pesquisa baseados em aprendizado profundo para classificação de sinais de EEG, como os de Almogbel [Almogbel et al., 2018] e Khadijah [Aznan et al., 2018b], têm dispensado tal etapa, deixando para que o próprio método de aprendizado de máquina implicitamente assumira essa função.

O janelamento é outra etapa importante no pré-processamento dos dados EEG. Essa técnica divide o sinal EEG contínuo em pedaços menores, com tamanhos fixos, denominados de janelas, conforme ilustrado na Figura 1.6. Cada janela representa um período curto da atividade cerebral, permitindo uma análise mais detalhada do sinal. Encontrar o tamanho ideal de janela é o ponto crucial para a qualidade da análise dos dados. Se as janelas forem muito pequenas podem conter pouca informação para a análise dos dados, por outro lado, se for muito grande, podem conter mais de uma informação por vez, atrapalhando a classificação dos dados. Outro fator vinculado ao janelamento é o passo de deslocamento, sendo o quanto a próxima janela estará distante da janela atual. Passos muito pequenos geram sobreposição dos dados, e caso sejam muito pequenos, podem gerar sobreposição excessiva, caso sejam muito grandes, podem levar à perda de dados. A definição ideal dos parâmetros de janelamento depende da aplicação específica, da taxa de amostragem do EEG e do tipo de análise desejada. Em geral, a escolha do tamanho da janela e do passo ocorre de forma empírica seguida de testagens e ajustes. Também é utilizado o termo janelas deslizantes. Um exemplo das janelas deslizantes sem sobreposição de dados pode ser visto na Figura 1.6(a) e janelas deslizantes com sobreposição de dados na Figura 1.6(b).



(a)



(b)

Figura 1.6. Janelamento sem sobreposição (a) e com sobreposição (b).

Na revisão realizada por Roy *et al.* [Roy et al., 2019] sobre o uso de DL em EEG, a escolha de parâmetros de janelamento apresentou grande variação entre os estudos. As variações envolviam desde o tamanho da janela em minutos, o tamanho do passo para a próxima janela gerando assim sobreposição de dados nas janelas, a não utilização de sobreposição. A utilização de janelas deslizantes com sobreposição é comum quando a base de dados coletada é pequena e não há como realizar novas coletas. Esse tema também foi estudado por Almogbel [Almogbel et al., 2018], Bird [Bird et al., 2018], Teo [Teo et al., 2018], Almogbel [Almogbel et al., 2019], Bird [Bird et al., 2019], Ju [Ashford et al., 2020] e Sakelle [Sakalle et al., 2021], dentre outros.

1.5.2. Extração de características

A extração de características é usada com frequência em alguns trabalhos ligados a BCI [Bird et al., 2021, Ashford et al., 2020]. O objetivo é transformar os dados coletados em informações mais discriminativas para auxiliar na classificação, buscando identificar características relevantes nos dados brutos e de ondas de EEG, contribuindo na diferenciação e facilitando consequentemente a classificação. Diversas técnicas são utilizadas para extrair características, dentre elas os cálculos de medidas de estatísticas como a média, mínimo, máximo, desvio padrão, entropia, curtose, dentre outros. Podem ser usadas isoladamente ou em conjunto, podendo ser incorporadas juntamente com os dados originais de EEG. Em contraste com os métodos tradicionais de extração de características, algumas abordagens de aprendizado profundo, como as redes neurais convolucionais (CNNs) e redes recorrentes (RNNs), têm demonstrado sucesso na classificação de sinais de EEG diretamente a partir dos dados brutos, sem a necessidade de uma fase preliminar de extração de características. Estudos como os de O'Shea [OrShea et al., 2018] e Zhou [Zhou et al., 2021] demonstram a capacidade dessas redes de aprender automaticamente características relevantes dos sinais de EEG, obtendo resultados comparáveis ou até superiores aos métodos tradicionais que dependem da extração manual de características.

1.5.3. Divisão de Dados em Conjuntos de Treino e Teste

Campos *et al.* Campos et al. [2014] discutem a importância da separação dos dados de treino e teste com o intuito de garantir a confiabilidade dos resultados em modelos de aprendizado de máquina. Essa separação é crucial quando se trabalha com bases de dados *off-line*, uma vez que impede que o modelo seja treinado com dados que também serão utilizados para avaliar. Quando um modelo de aprendizado de máquina é treinado com sobreposição de dados de treino e teste, isso pode levar a resultados artificiais e enviesados. Os autores estabeleceram ainda questionamentos sobre a base de dados quanto à ordem de classificação dos conjuntos de treino e teste, sobre qual método de avaliação cruzada deve ser utilizado, a quantidade de classificações e quais itens são relevantes para cada usuário. A definição do conjunto de treino e teste pode ser considerada de várias formas, como, por exemplo: a divisão percentual da base de dados; pela quantidade de usuários contida na base de dados (em que os dados de alguns usuários são empregados para treino e os dados dos demais ficam para teste); estabelecendo uma data e hora como divisão do conjunto de treino e teste; ou pela validação cruzada, como, por exemplo, a *k-fold*, a qual pode aumentar a generalização por meio de re-amostragem.

Nam *et al.* [Nam et al., 2018] alertam para a importância da submissão correta dos dados para classificação. Em algumas aplicações, ocorre da coleta e classificação ocorrerem em momentos distintos, nestes casos é fundamental a separação dos dados em conjuntos de treino e teste. O treino do modelo de classificação se dá com o conjunto de dados para treino, após esta fase, é possível avaliar o desempenho em um conjunto que não foi ainda utilizado. Isso permite verificar se o modelo conseguiu generalizar o conhecimento aplicado a novos dados, ou se apenas memorizou os rótulos de classe do conjunto de treino. A memorização ocorre normalmente com os do conjunto de treino são apresentados com o conjunto de teste, gerando dados artificialmente altos, desempenho irreal, levando a conclusões incorretas sobre a efetividade do modelo, dificultando a comparação com outros métodos.

Para gerar conjuntos de treinamento e teste em BCI, os artigos pesquisados utilizaram principalmente dois métodos:

1. Validação cruzada: essa técnica divide os dados em várias dobras, utilizando cada dobra como conjunto de teste em um momento diferente. Isso garante que todos os dados sejam usados para treinamento e teste em diferentes combinações. Bibliotecas como *sklearn.model_selection.KFold* no Python automatizam esse processo.
2. Divisão aleatória: nesse método, os dados são divididos aleatoriamente em conjuntos de treinamento e teste em proporções predefinidas. A biblioteca que facilita esta divisão no Python é a *sklearn.model_selection.train_test_split*.

A escolha do método de validação pode depender da ordem temporal dos dados. Se a ordem for crucial para a aplicação, métodos como a validação cruzada que preservam a ordem temporal devem ser utilizados, uma vez que a biblioteca K-Fold do Python possui um parâmetro de embaralhamento, que por padrão é Falso. No caso do *train_test_split* do Python, existe um fator preocupante, pois esta biblioteca também possui o parâmetro de embaralhamento, que, por padrão, é Verdadeiro. Caso o padrão seja mantido neste parâmetro, a ordem temporal será alterada, gerando dados embaralhados e enviesando possivelmente os dados de classificação.

Luu e Chau [Luu and Chau, 2009] realizaram um estudo onde nove participantes escolheram sua bebida preferida entre duas opções. Sinais biológicos foram coletados por um dispositivo de espectroscopia de infravermelho próximo (NIRS) com 48 canais. Para avaliar um método de análise discriminante linear ajustado para classificar esses sinais, os autores utilizaram validação cruzada, obtendo uma precisão de 80%.

No ano seguinte, Dominguez [Dominguez, 2009] reavaliou o estudo, questionando a falta de avaliação independente e validação robusta dos resultados. Através da simulação de séries temporais com dados aleatórios, Dominguez obteve resultados estatisticamente semelhantes ao estudo original. Em resposta, Chau e Damouras [Chau and Damouras, 2009] reformularam o estudo, aprimorando a separação dos conjuntos de dados de treinamento e teste. Essa reformulação resultou em uma queda na precisão média geral para 53%.

Este estudo demonstra a importância de avaliações rigorosas e da consideração da ordem temporal dos dados em BCI. A escolha de métodos de validação adequados e a consideração de fatores como a ordem temporal dos dados podem influenciar significativamente os resultados da classificação.

1.5.4. Treinamento de modelos de aprendizado de máquina

Considerando que o problema computacional que se apresenta é o de classificação, diversos métodos podem ser utilizados para abordá-lo. Isso inclui, por exemplo, verificar diretamente se os atributos extraídos das ondas cerebrais possuem certos padrões, nos casos deles serem bem conhecidos e fáceis de identificar. Como raramente essa é a situação, a forma mais comum de tratar o problema tem sido empregar métodos de aprendizado de máquina, uma subárea da Inteligência Artificial [Skansi, 2018], em particular, os métodos supervisionados.

Algoritmos clássicos de aprendizado supervisionado, como Support Vector Machine, K-Nearest Neighbors, Decision Tree, Random Forest e Extreme Gradient Boosting, podem ser empregados e apresentam bons resultados quando há um conjunto reduzido de atributos extraídos dos dados que ajudam a discriminar as classes desejadas. Por outro lado, o uso de Redes Neurais Profundas [Skansi, 2018] no contexto dos sistemas de BCI (Brain-Computer Interfaces) vem ganhando maior atenção, dado o seu sucesso em superar o humano em diversas tarefas. De fato, a natureza complexa dos sinais cerebrais, que apresentam variadas formas e intensidades de onda, torna o aprendizado profundo bastante promissor para o problema de classificação desses dados.

Um estudo de Zhang *et al.* Zhang et al. [2020] analisaram mais de 200 trabalhos que aplicaram aprendizado profundo em BCI nos últimos 5 anos. Entre os métodos utilizados, as redes neurais convolucionais (CNNs) se destacaram, com presença em 37% das pesquisas. O uso híbrido de CNNs com outras técnicas também foi frequente (22,6%), enquanto Deep Belief Networks (DBNs), Multi-Layer Perceptron (MLP), Deep Autoencoders (D-AE), Recurrent Neural Networks (RNNs) e redes Generativas (Generative) tiveram menor prevalência (14,5%, 9,8%, 7,2%, 5,1% e 3,8%, respectivamente).

Outro levantamento, realizado por Roy *et al.* [Roy et al., 2019], corrobora a popularidade das CNNs em BCI, com 40,3% dos estudos as utilizando. Autoencoders (AEs), RNNs, DBNs, CNNs+RNNs, redes Fully Connected e outras técnicas também foram empregadas, em proporções decrescentes (13%, 13%, 7,1%, 7,1%, 6,5% e 4,3%, respectivamente). Vale ressaltar que 3,9% dos trabalhos não mencionaram o método específico e 2,6% utilizaram RBM (Restricted Boltzmann Machine) e 1,9% GAN (Generative Adversarial Network).

Em relação ao tipo de dado utilizado nas CNNs, Zhang *et al.* Zhang et al. [2020] observou que metade dos estudos empregou dados brutos de EEG, 22% utilizou dados no domínio de frequência e o restante, 28%, optou por outros domínios.

O crescente interesse em redes neurais convolucionais (CNNs) para aplicações BCI com EEG se intensificou desde 2016. Essa preferência se deve à capacidade das CNNs em processar dados brutos de EEG com eficiência, dispensando etapas intermediárias de pré-processamento e viabilizando aplicações em tempo real.

A popularidade das CNNs em BCI não se limita apenas a essa área. Elas também são amplamente utilizadas no reconhecimento de imagens [Aggarwal, 2018]. Um exemplo disso é o trabalho de Ashford *et al.* [Ashford et al., 2020], que propõe a classificação de sinais de EEG através da representação de imagens associadas a características estatísticas. As ondas de EEG são transformadas em vetores no tempo, cálculos estatísticos são aplicados, os dados são normalizados e redimensionados em uma matriz de valores fixos, similar a uma imagem em tons de cinza. Essa representação visual é então submetida à classificação por uma rede CNN, obtendo resultados de acurácia superiores a outros métodos.

As aplicações de CNNs em BCI se estendem para além da classificação binária de estados mentais (positivo/negativo). Diversos estudos utilizam CNNs e Deep Belief Networks-Restricted Boltzmann Machine (DBN-RBM) para classificar estados mentais em três categorias: positivo, neutro e negativo. A capacidade das CNNs em extrair carac-

terísticas relevantes dos sinais EEG as torna ferramentas poderosas para essa tarefa.

A conversão de sinais EEG unidimensionais para formatos bidimensionais como imagens facilita a análise por redes CNN. Essa abordagem, combinada com a robustez e expressividade das CNNs, as torna uma escolha popular para aplicações BCI com EEG, como evidenciado por Zhang *et al.* [Zhang et al., 2020].

1.5.5. Avaliação dos Resultados

A maioria das aplicações de BCI trabalha com o princípio de classificar os sinais de EEG coletados, visando usar os resultados dessas classificações em comandos de iteração para um participante e o computador ou dispositivo a ser controlado.

Para avaliar o desempenho dos modelos em tarefas de classificação, é necessário estabelecer uma métrica que sirva de referência para comparações. A acurácia é a métrica mais utilizada, que indica o quão bom um modelo de classificação é através da divisão do número de classificações corretas pelo número de classificações feitas [Skansi, 2018]. No entanto, essa métrica só é eficiente se a base de dados estiver balanceada. Em uma base onde não há balanceamento entre as classes, a acurácia balanceada seria a mais adequada, uma vez que considerará tal fator para ser mais precisa. Além das acurácias, pode-se avaliar os resultados de um modelo conforme suas métricas de precisão (obtida pela divisão do total de classificações corretas sobre o total de classificações feitas) e o F1-Score [Skansi, 2018].

No entanto, apesar da escolha de uma métrica adequada, é necessário considerar outros aspectos dos sinais de EEG, tais como a sua temporariedade. A temporariedade de um sinal de EEG tem um grande impacto nas aplicações de BCI, pois os sinais tendem a mudar com o tempo à medida que uma pessoa se adapta à tarefa executada. Os dados mais recentes coletados tendem a ter um impacto mais significativo no treinamento dos modelos AM treinados para classificar sinais de EEG. No entanto, assegurar que o modelo esteja atualizado com as informações mais recentes é extremamente difícil, sobretudo no contexto de uma aplicação *online*. Em uma aplicação *offline*, todos os dados já foram coletados, sendo necessário, portanto, separar os dados mais recentes para a avaliação dos modelos e os dados mais antigos para o treinamento deles. Em um ambiente *online*, os dados aumentam com o decorrer do tempo, tornando necessário retreinar periodicamente para manter o modelo atualizado.

1.6. Bases de Dados

Este tópico abordará algumas bases de dados públicas na área, as quais podem ser utilizadas na avaliação comparativa de métodos computacionais e de dispositivos de hardware para o processamento e a classificação de sinais de EEG. Também será apresentada uma base de dados coletada pelos autores, durante a execução de 3 trabalhos, um de graduação, um de mestrado e outro de doutorado, utilizada nos experimentos discutidos no tópico dos estudos de caso 1.7.

Foram utilizadas duas bases de dados na execução deste trabalho. A primeira foi uma base de dados pública, chamada *BCI Competition 2a*¹⁰, que contém os dados de nove

¹⁰<https://www.bci.de/competition/iv/>

peças que executaram atividades de *Motor Imagery – MI*. A segunda base, denominada *Mindwork Dataset*, foi criada como parte desta e de outra pesquisa¹¹ sobre EEG. Essa base traz dados de nove participantes que realizaram tarefas de imaginação de um cursor se movendo em uma interface gráfica, através da utilização do equipamento de baixo custo Muse 2.

1.6.1. BCI Competition IV 2a

O *BCI Competition IV 2a (BCICIV)*, descrito com mais detalhes no trabalho de [Brunner et al., 2008/1], apresenta os dados coletados por um equipamento de EEG com 25 sensores (22 sensores EEG e 3 EOG) na frequência de 250Hz . Esta base contém as informações de 9 indivíduos para quatro classes de movimentos, tais como: a movimentação da mão direita (*right_hand*), da mão esquerda (*left_hand*), da língua (*tongue*) e de ambos os pés (*feet*). Foram executadas 12 sessões de coleta em dois dias diferentes. Neste caso, uma sessão foi composta por 12 coletas (denominadas *tentativas* ou, em inglês, *trials*) para cada uma das classes, totalizando 48 coletas. O experimento inteiro envolveu 576 coletadas por participante.

Cada tentativa iniciou com um efeito sonoro no instante 0, seguido de uma orientação representada visualmente na forma de uma seta indicando o movimento a ser imaginado. Desse modo, temos: seta para a direita igual à mão direita; seta para a esquerda igual à mão esquerda; seta para cima igual à língua; e seta para baixo igual a ambos os pés. A dica permanece na tela por um período de 1.25s e, logo após, desaparece para o participante poder imaginar o movimento. A coleta continua por 3s e, em seguida, o indivíduo é orientado a descansar por 1.5s até a próxima tentativa. A Figura 1.7 demonstra o protocolo.

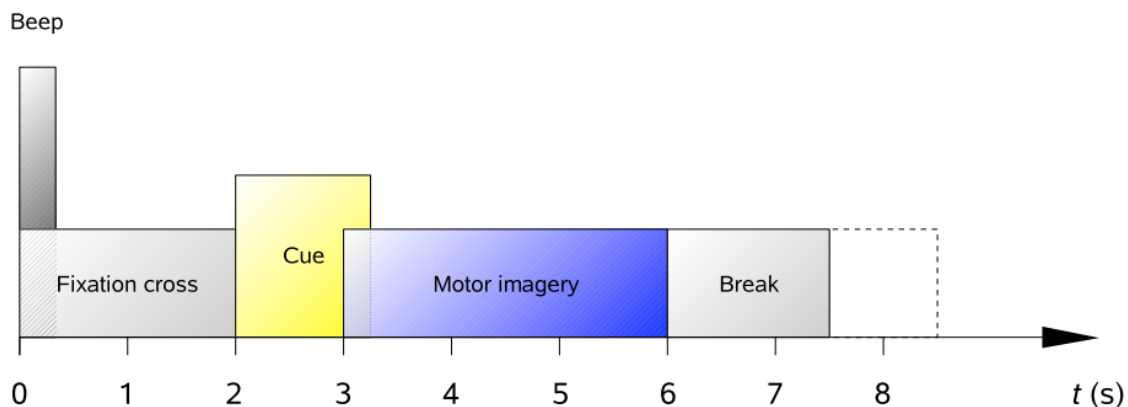


Figura 1.7. Protocolo de coleta da base BCI Competition IV 2a.

Fonte: [Brunner et al., 2008/1]

1.6.2. Base de Dados *MindBigData*

A base de dados *MindBigData* foi criada inicialmente com gravações de sinais de EEG de um único indivíduo imaginando um dígito entre 0 e 9, e gravações em que essa mesma

¹¹A base Mindwork integra dados coletados para apoiar esta pesquisa de mestrado e uma pesquisa de doutorado de Murilo Borges Silva, ambos do Programa de Pós-graduação do INF/UFMG.

pessoa não pensava em qualquer dígito, rotuladas como -1 . Na sua versão 1.03, a base contém 1.207.293 sinais cerebrais com 2 segundos de duração cada, gravados entre 2014 e 2015 usando os equipamentos *NeuroSky MindWave*, *Emotiv EPOC*, *Interaxon Muse* e *Emotiv Insight*. Hoje, a base também possui dados para outros cenários de experimentação e está disponível em <https://mindbigdata.com/>.

1.6.3. Base de Dados *Kaggle*

Além das bases de dados públicas citadas acima, existe uma outra opção bastante utilizada: a plataforma de aprendizado de ciência de dados chamada Kaggle. A Kaggle é também uma comunidade que promove competições para desenvolvimento de soluções para determinados problemas, disponibilização de modelos de aprendizado de máquina, exemplo de códigos de diversas aplicações, fóruns de discussão e disponibilização de bases de dados. É possível pesquisar por bases de dados específicas na Kaggle como, por exemplo, dados de EEG. O site oficial da plataforma pode ser acessado pelo link <https://www.kaggle.com>.

1.6.4. MindWork

Apesar de existirem algumas bases de dados disponíveis, as características dos dados com as quais se planeja trabalhar podem não estar disponíveis e neste caso será necessário criar sua própria base de dados. Para algumas situações deve ser possível gerar bases de dados artificialmente, mas em outros, somente uma base de dados específica permitirá o desenvolvimento das análises características e direcionadas ao que se aspira descobrir.

Um dos pontos fundamentais na coleta com humanos é que se faz necessário a autorização de um comitê de ética, no Brasil, representado pela “Plataforma Brasil” que avale e características do experimento proposto e autorize o início da coleta. O acesso ao site oficial pode ser feito pelo site <https://plataformabrasil.saude.gov.br/>. Para isto algumas informações devem ser disponibilizadas, alguns documentos devem ser produzidos a fim de deixar claro aos participantes o objetivo da pesquisa, ao que o participante será submetido, quais são as consequências da coleta, se ele irá autorizar a utilização dos dados coletados, das imagens geradas durante a coleta, se permitirá divulgar os seus dados nas publicações e também se seus dados poderão ficar disponíveis para futuras pesquisas, dentre outras informações. O processo leva tempo para ser escrito, para ser submetido, para ser avaliado e adequar as possíveis solicitações do comitê, por isso, inclua no seu cronograma geral da pesquisa um tempo razoável para que este documento seja elaborado, submetido e aprovado, algo em torno de três a seis meses. Após a aprovação do projeto e conforme o cronograma informado será possível dar continuidade ao trabalho e iniciar as coletas. Na Seção 1.7 detalharemos com mais propriedade e de forma contextualizada as bases de dados utilizada em nossas pesquisas.

1.7. Estudos de caso

Dois estudos de casos serão discutidos na presente seção, cada um deles tratando de uma aplicação de classificação de sinais de EEG usando uma base de dados pública ou criada pelos autores. Os estudos de caso apresentam abordagens práticas, que visam demonstrar o processamento e a classificação de sinais de EEG em aplicação reais e, então, dar uma

noção de como os conhecimentos apresentados nas seções anteriores podem ser utilizados na prática.

1.7.1. *Brainoscope*

O *Brainoscope* é um projeto artístico desenvolvido por parte dos autores deste documento [da Silva Rocha et al., 2022], que envolve a apresentação de uma imagem dinâmica de um caleidoscópio controlada em tempo real por um *iterator* humano utilizando um dispositivo de EEG. Ao *iterator*, é dada a capacidade de girar virtualmente o caleidoscópio, acelerar sua rotação até certo ponto e desacelerar o movimento até o paralisar. A principal dificuldade em fazer funcionar o sistema reside na correta classificação dos sinais cerebrais do *iterator* entre padrões de aceleração e desaceleração de movimento. Estudos anteriores demonstram que a classificação de padrões cerebrais é sensível à presença de ruídos nos sinais de EEG, varia de pessoa para pessoa e não é consistente ao longo do tempo [Nam et al., 2018]. Parte dessas questões é mitigada pelo uso de um processo contínuo de calibração e avaliação do sistema.

A arquitetura geral do funcionamento do Brainoscope está ilustrada na figura abaixo e envolve cinco elementos:

- um software simulador de caleidoscópio¹², que produz uma imagem caleidoscópica a partir de uma imagem 2D qualquer;
- um *iterator* humano;
- um equipamento de EEG, acoplado à cabeça do *iterator*, que coleta continuamente sinais eletromagnéticos;
- um teclado comum, pelo qual o *iterator* informa diretamente suas intenções de pensamento (acelerar ou desacelerar a movimentação do caleidoscópio); e
- um software de Controle Central, encarregado de interpretar os sinais de EEG e determinar a movimentação do caleidoscópio. Esse elemento também é responsável por ajustar parâmetros internos com base em eventos de teclado, para melhorar a interpretação dos sinais cerebrais.

O simulador de caleidoscópio é baseado em um software livre já existente¹³, escrito em Javascript, que emprega a tecnologia Electron¹⁴ para apresentação na Web. Esse software livre produz uma imagem caleidoscópica a partir de uma imagem 2D comum recebida como entrada, pela repetição de padrões poligonais extraídos dela considerando uma posição central (x,y) . O software original permitia deslocar as coordenadas centrais (x,y) por meio do *mouse*, gerando a ilusão do caleidoscópio estar girando. Para o projeto do Brainoscope, a interação por *mouse* foi desativada, já que a movimentação passou a ser controlada pelo software de Controle Central. Além disso, optou-se por empregar apenas

¹²Utiliza-se o termo “simulador” para indicar um software que simula o funcionamento de um caleidoscópio físico real.

¹³<http://coldhead.github.io/kaleidos/>

¹⁴<https://www.electronjs.org/>

a coordenada x central para criar a impressão de movimento do caleidoscópio, mantendo-se a coordenada y constante. A aplicação de caleidoscópio foi modificada para iniciar com um valor x padrão e foi integrada a um código de comunicação via rede escrito em Python que espera, do Controle Central, a posição x para redesenho do caleidoscópio.

O equipamento de EEG adotado neste projeto foi o OpenBCI, modelo Cyton, acrescido da placa expansora Daisy, somando 16 canais de EEG ao todo. Dezesesseis sensores foram acoplados a um capacete nas posições FP1, FP2, C3, C4, P7, P8, O1, O2, F7, F8, F3, F4, T7, T8, P3 e P4, conforme o padrão 10–20. Quando ativo, os dados coletados de todos os sensores eram enviados a 250Hz por um *dongle wireless* para um computador rodando o software de Controle Central.

O Controle Central teve como função principal interpretar continuamente os sinais de EEG quanto às duas possíveis intenções do *interator* (iniciar/acelerar o movimento do caleidoscópio e desacelerar/parar tal movimento) e transformar essas interpretações em alterações de deslocamento no eixo x , as quais eram enviadas ao simulador de caleidoscópio. Para tanto, utilizou-se a API BrainFlow¹⁵ para capturar e processar os dados do OpenBCI de modo contínuo. O Controle Central capturava apenas dois valores por amostragem de EEG, correspondentes às médias das intensidades das ondas alfa e beta extraídas dos 16 sensores em uma janela de tempo. O Controle então armazenava esses pares alfa e beta instantâneos e calculava as suas médias $alfa_{mdio}$ e $beta_{mdio}$ considerando outra janela com as últimas 256 amostragens desses dados. A tupla $v = (alfa_{mdio}, beta_{mdio})$ era então comparada, em termos de distância, às duas tuplas padrão_médio_aceleração e padrão_médio_desaceleração, que representavam respectivamente os valores médios dos dois tipos de intenção de pensamento. Se a tupla v estivesse próxima dessas tuplas de valores médios e muito mais perto de uma delas do que da outra, então uma ação correspondente era executada para alterar (aumentar ou reduzir) a velocidade de modificação da coordenada x da imagem caleidoscópica. Caso contrário, a velocidade de modificação de x era mantida como estava. Esse controle de movimentação foi operacionalizado mantendo internamente uma variável x que representava a posição horizontal da imagem 2D original a ser usada como centro do caleidoscópio e duas outras variáveis que indicavam a velocidade e a aceleração (positiva ou negativa) do incremento de x .

Inicialmente, as tuplas padrão_médio_aceleração e padrão_médio_desaceleração eram definidas com valores nulos, o que fazia com que as intenções de pensamento do *interator* fossem interpretadas indiscriminadamente como sendo de um ou de outro tipo, ou como nenhum deles. No entanto, o *interator* podia pressionar a tecla “a” ou “z” do teclado para indicar explicitamente que está pensando em uma ação de aceleração ou de desaceleração, respectivamente. Isso fazia com que a tupla de padrão correspondente, padrão_médio_aceleração ou padrão_médio_desaceleração, fosse imediatamente atualizada na direção da tupla v que acabara de ser calculada. A ideia desse procedimento era possibilitar ao *interator* recalibrar o sistema para o reconhecimento dos tipos de intenções sempre que ele se mostrassem pouco efetivo.

Um experimento foi realizado com sete participantes, em que cada um deles tentou calibrar o Brainoscope para perceber a sua intenção de girar e de parar o caleidoscópio virtual. Ao preencher um questionário final sobre o uso do sistema, cinco dos sete partici-

¹⁵<https://brainflow.readthedocs.io/en/stable/UserAPI.html>

pantes disseram que conseguiram efetivamente controlar o movimento do caleidoscópio. Um aspecto adicional observado nesse trabalho é que a interligação dos elementos de hardware e de software por meio de uma rede de comunicação ofereceu maior liberdade de deslocamento físico ao *interator* no espaço ocupado pelo projeto artístico.

1.7.2. Mindwork com Múltiplos Parâmetros

O estudo Mindwork com múltiplos parâmetros se propôs a classificar os sinais cerebrais de 10 participantes, coletados a partir de um equipamento de EEG de baixo custo, o Muse 2 da Interaxon. Foram exploradas muitas combinações de parâmetros a fim de ser avaliar a acurácia de classificação, como aplicação ou não de filtro, tamanho de janela e passo da janela com tamanhos variados, extraindo dados característicos por meio de cálculos estatísticos, variando o tamanho de conjunto de treino e teste, aplicando alguns métodos de aprendizado de máquina, além de processar sobre dados brutos e de ondas separadamente.

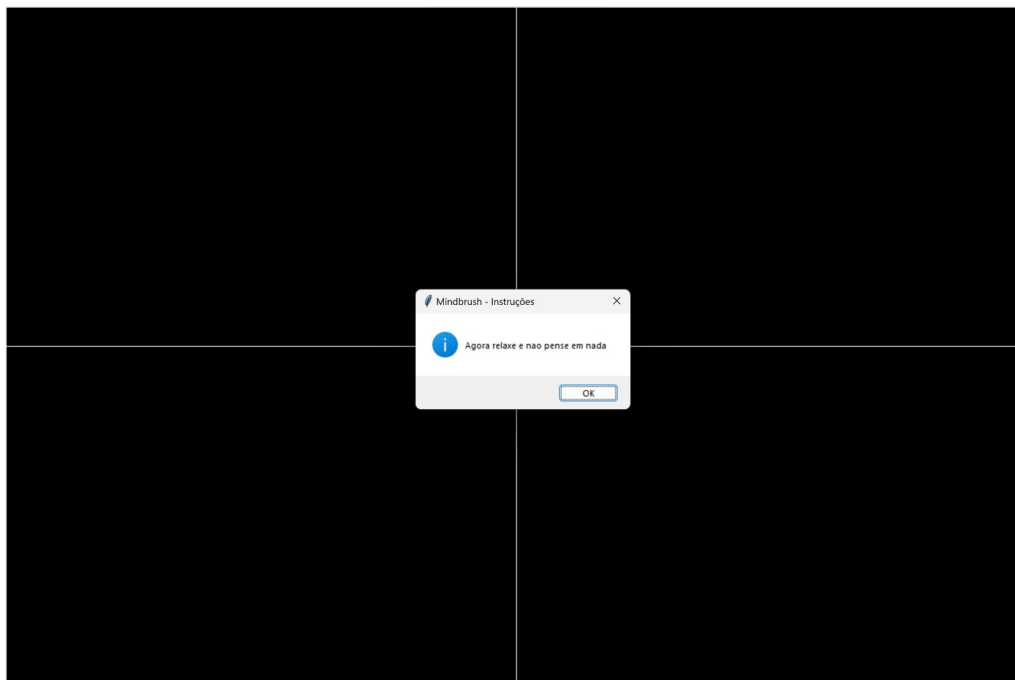
A coleta de dados

Foi utilizado para coleta de dados o Muse 2. Após a escolha do tipo de equipamento para coleta, definimos a base tecnológica capaz de transmitir, armazenar os dados coletados. Foram coletadas ondas alfa, beta, delta, gama e teta separadamente, dados brutos (RAW), juntamente com o respectivo *Timestamp*, etiqueta de identificação da situação de observação do objeto, dentre outras informações disponibilizadas pelo Muse 2. Foi definida uma aplicação computacional que permitisse a observação da movimentação de objetos na tela pelos participantes, a captação das ondas cerebrais e o armazenamento dos dados em uma estrutura de banco de dados, possibilitando o posterior processamento. O ambiente de coleta foi controlado em sala com tratamento acústico, climatizada, sem estímulos externos, possuía cadeira com assento de espuma, revestida em tecido. Sete homens e três mulheres participaram da coleta, composta por vinte e quatro sessões, cada sessão com cinco situações de observações, cada situação com dezoito segundos cada. A cada seis sessões de coleta houve uma paralisação de cinco minutos para descanso. As cinco situações de observação e coleta envolvem observar um círculo pequeno se movendo na tela para a esquerda, para a direita, para cima, para baixo ou se mantendo parado no centro da tela, conforme pode ser visto na Figura 1.9.

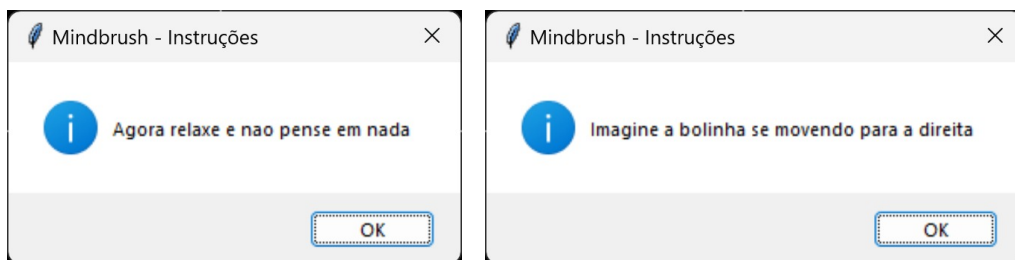
Foram indicadas pelas seguintes mensagens: “*imagine à bolinha se movendo para a esquerda*”, “*imagine à bolinha se movendo para a direita*”, “*imagine à bolinha se movendo para baixo*”, “*imagine à bolinha se movendo para cima*” e “*agora relaxe e não pense em nada*”. A gravação dos dados e consequente transmissão só ocorreram após o pressionar do botão OK, presente nas mensagens de orientação. Após o período de 18 segundos da execução da situação de observação, a gravação e transmissão eram encerrados. Na Figura 1.8 é possível ver a tela de mensagem bem como a tela de movimentação da bolinha.

A comunicação entre Muse 2 ocorreu via bluetooth com um iPad Pro (11 pol. de 3ª geração) com iPadOS 16.6 através do APP Mind Monitor. O envio dos dados para um notebook, que continha a base de dados, ocorreu por meio do protocolo OSC juntamente

com TCP-IP. Foi utilizado um notebook Dell XPS 13 9315 2-in-1, processador i7 de 12^a geração, 16 GB de RAM, SSD 936 GB com Windows 11. A aplicação, desenvolvida em Python 3.8, exibia na tela as mensagens de orientação e recebia os dados advindos do Muse 2 juntamente com a indicação de qual situação de coleta estava ocorrendo naquele momento. As informações coletadas foram armazenadas em um banco de dados não-relacional orientado a documentos denominado MongoDB. Seis arquivos CSVs com os dados coletados eram gerados para cada participante, sendo divididos, um para cada tipo de onda (RAW, alfa, beta, delta, gama e teta). Os dados contidos em cada arquivo eram: data e hora, o nome do participante, a sessão, qual das 5 situações de observação ocorria naquele instante, o registro dos sensores (TP9, TP10, FP1, FP2 e AUX), dentre outros.



(a) Mensagem de orientação de ação sobreposta à bolinha



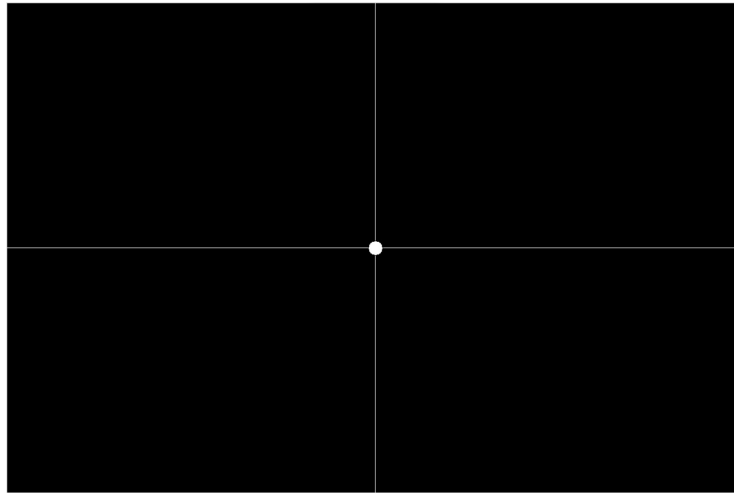
(b) Mensagem de bolinha parada

(c) Mensagem de bolinha movimentando-se para a direita

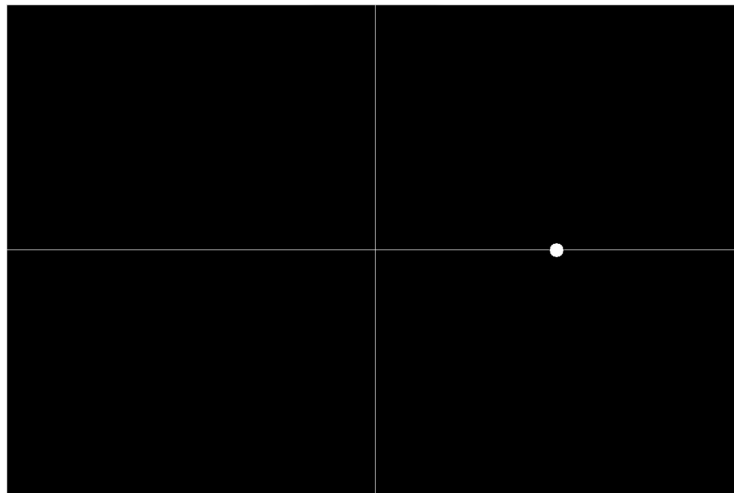
Figura 1.8. Mensagem de orientação sobre a movimentação ou não da bolinha.

Fluxo utilizado para processamento dos dados coletados

Os dados coletados de cada participante foram submetidos a 5 (cinco) etapas visando classificar a sua forma pensamento, explorando um conjunto de parâmetros. O fluxo



(a) Tela de coleta com bolinha parada



(b) Tela de coleta com bolinha movimentando para a direita

Figura 1.9. Telas com a bolinha e sua movimentação.

utilizado para o processamento de dados contemplou as seguintes etapas: aplicação ou não de filtros, janelamento dos dados, extração de características, divisão de dados em treino, validação e teste, e aplicação dos métodos de aprendizado de máquina. A Figura 1.5, demonstrada na Seção 1.2, ilustra as etapas citadas. O processo foi executado para cada participante em seus dados brutos, ondas alfa, beta, delta, gama e teta. Outro fator importante é que, nos dados brutos, não foi utilizada a coluna AUX. Para cada combinação de parâmetros foram gerados arquivos CSVs para que fossem processados à posteriori.

Os parâmetros relativos a filtro se dividiram em quatro características. O primeiro se manteve com dados inalterado, denominado “Sem filtro”. O segundo, denominado “Filtro A”, e terceiro, denominado “Filtro B”, foram submetidos a uma transformada de Fourier, onde foram convertidos ao domínio da frequência, aplicados cortes nas maiores e menores frequências e convertidos de volta ao domínio do tempo. “O Filtro A” teve corte na primeira posição à esquerda, considerada a de menor frequência. O “Filtro B” passou por corte à esquerda em suas doze primeiras posições, consideradas as de menores

frequências, e à direita nas suas últimas posições, consideradas as de maiores frequências. O quarto e último filtro, denominado “Filtro C”, cortou as frequências menores que 4Hz e as maiores que 38Hz.

Com os filtros aplicados, foi possível separar os dados de cada participantes em janelas temporais divididas em pedaços menores. As janelas foram organizadas em formato de matriz, com as colunas representadas pelos sensores, ou dados oriundos deles e linhas conforme a sequência temporal. Para os dados brutos, foram considerados linhas de tamanho 125, 250, 500, 750, 1000 e sobreposição. Para os dados de ondas, adotou-se linhas de tamanho 5, 10, 15, 20, 44, 88 e 112. As janelas foram compostas com e sem sobreposição. As janelas com sobreposição avançavam a metade do tamanho da janela original. Os pares ordenados gerados pelo tamanho da janela e o tamanho do passo nos dados brutos foram: [[125, 62], [125, 125], [250, 125], [250, 250], [500, 250], [500, 500], [750, 375], [750, 750], [1000, 500], [1000, 1000]]. Para as ondas alfa, beta, delta, gama e teta, os pares ordenados foram: [[5,5], [10,5], [10,10], [15,7], [15,15], [20,10], [20,20], [44,10], [44,44], [88,10], [112,10], [112,112]]. Nos casos onde não houve sobreposição de janelas, os pares ordenados eram iguais, por exemplo: [500, 500].

Três novos conjuntos foram considerados, como parte dos dados característicos: a manutenção dos dados originais, os dados originais acrescidos dos dados estatísticos e somente os dados estatísticos. O cálculo dos dados estatísticos se deu em cada uma das janelas. Foram executados cálculos estatísticos como a média, o desvio padrão, o mínimo e o máximo para todos os conjuntos, bem como realizados cálculos de assimetria (*Skewness*), curtose e covariância. Para cada janela, foi feito um cálculo dividido em três etapas: a primeira etapa foi composta da janela completa; a segunda separou o mesmo conjunto inicial em duas partes iguais; a terceira etapa dividiu o conjunto inicial em 4 partes iguais.

A Figura 1.10 mostra uma representação visual da combinação das etapas anteriores. Os dados brutos e de ondas estão representados com cores diferentes na etapa de dados e sinais, bem como na etapa de janela e passo. Nessas duas etapas, os dados brutos e ondas foram combinados única e exclusivamente com as janelas de mesma cor.

A variação do conjunto de treino, validação e teste é outro fator que permitiu avaliar a evolução das acurácias. O conjunto de treino foi sendo acrescido de novos dados, enquanto o conjunto de validação e teste sofreu um deslocamento proporcional no tempo, mantendo o seu tamanho com três subconjuntos de coleta. Foram oito variações do conjunto de treino. O primeiro conjunto de treino foi composto pela sessão 0 e conjunto de validação de 1 a 3. O segundo conjunto de treino foi composto pelas sessões de 0 a 1 com conjunto de validação de 2 a 4. O terceiro conjunto de treino era composto pelas sessões de 0 a 3 com conjunto de validação de 4 a 6. Daí por diante, cada conjunto de treino e validação foram somados de duas unidades. As sessões posteriores ao conjunto de validação foram utilizadas para teste. A Tabela 1.5 relaciona o código da evolução das acurácias e os respectivos conjuntos de treino e validação.

Após termos utilizado todas as combinações entre dados brutos e ondas, filtros e não filtros, janelamento com e sem sobreposição e a extração de características, submetemos cada conjunto combinado aos métodos de aprendizado de máquina, no intuito de classificar os padrões de pensamento. Foram utilizadas as fases de treinamento, validação

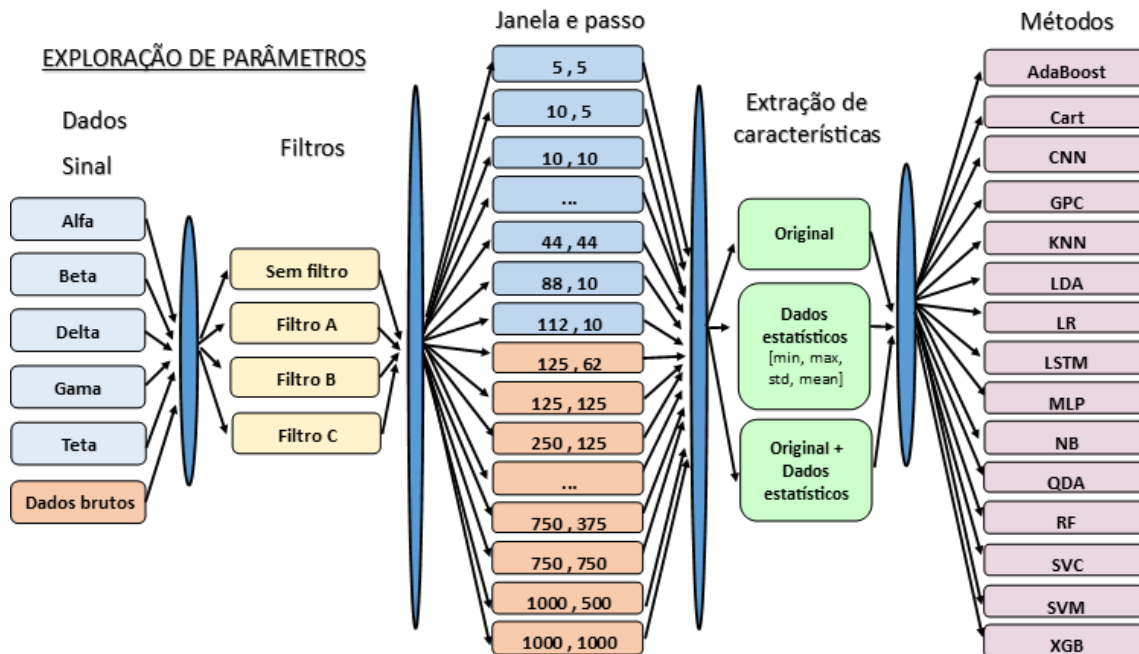


Figura 1.10. Combinações das etapas da coleta final.

Tabela 1.5. Conjunto de sessões de treino e validação.

Código da Evolução das acurácias	conjunto de treino	conjunto de validação
0	com 0	do 1 ao 3
1	do 0 ao 1	do 2 ao 4
2	do 0 ao 3	do 4 ao 6
3	do 0 ao 5	do 6 ao 8
4	do 0 ao 7	do 8 ao 10
5	do 0 ao 9	do 10 ao 12
6	do 0 ao 11	do 12 ao 14
7	do 0 ao 13	do 14 ao 16

e teste. Os Métodos de aprendizado de máquina utilizados foram: *Ada Boost Classifier* (AdaBoost), *Classification and Regression Trees* (CART), *Convolutional Neural Network* (CNN), *Gaussian Process Classifier* (GPC), *K Nearest Neighbors* (KNN), *Linear Discriminant Analysis* (LDA), *Logistic Regression* (LR), *Long Short Term Memory* (LSTM), *Multi Layer Perceptron Classifier* (MLP), *Gaussian Naive Bayes* (NB), *Quadratic Discriminant Analysis* (QDA), *Random Forest Classifier* (RF), *Support Vector Classification* (SVC), *Support Vector Machine* (SVM) e *eXtreme Gradient Boosting* (XGB).

Ao final, foi possível comparar os resultados das várias combinações com suas respectivas acurácias. Observamos que os dados de ondas obtiveram uma acurácia um pouco superior aos dados brutos. As janelas com tamanho médio de 2 segundos se sobressaíram dentre as maiores acurácias em comparação a janelas menores. Cada participante obteve um conjunto diferente de parâmetros com melhores acurácias, sendo que três deles se destacaram com maior quantidade de combinações com alta acurácia.

1.8. Considerações finais

Como explicado neste capítulo, a eletroencefalografia vem se popularizando com aplicações nas mais diversas áreas do conhecimento. A evolução e o barateamento dos equipamentos de EEG, bem como o surgimento de novas técnicas de classificação baseadas em aprendizado profundo (*deep learning*) têm contribuído para esse fenômeno.

Ainda assim, a coleta e o processamento de sinais de EEG são atividades que exigem atenção e planejamento. No presente trabalho, foram explicadas as etapas envolvidas nessas atividades e apresentadas orientações para maximizar as suas chances de sucesso. Adicionalmente, a análise de estudos de caso permitiu demonstrar na prática aspectos que, em um primeiro momento, foram apresentados apenas de forma teórica.

Para a obtenção de mais informações sobre a aplicação de EEG, recomendamos a consulta às bibliografias abaixo, as quais incluem livros e *surveys* na área.

- Literatura sobre Neurociência:
 - Robert Lent. *Cem bilhões de neurônios: conceitos fundamentais de neurociência*. Editora Atheneu, 2 edition, 2010. ISBN 9788538801023.
 - Daniel Fuentes, Leandro F Malloy-Diniz, Candida Helena Pires de Camargo, and Ramon M Cosenza. *Neuropsicologia: Teoria e Prática*. Artmed Editora, 2014.
 - Mark F. Bear, Barry W. Connors, Michael A. Paradiso, Carla Dalmaç, Jorge Alberto Quillfeldt, Maria Elisa Calcagnotto, Carlos Alberto Saraiva Gonçalves, Carmem Juracy Silveira Gottfried, Denise Maria Zancan, Renata Menezes Rosat, and Tadeu Mello e Souza. *Neurociências: Desvendando o Sistema Nervoso*. Artmed, 4ª edição edition, July 2017. ISBN 978-85-8271-432-4.
- Literatura sobre EEG e BCI:
 - Saeid Sanei and Jonathan Chambers. *EEG Signal Processing*. John Wiley & Sons Ltd, 2007. ISBN 9780470025819. URL <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470511923>.

- Chang S. Nam, Anton Nijholt, and Fabien Lotte, editors. *Brain-computer interfaces handbook: technological and theoretical advances*. Taylor & Francis, CRC Press, Boca Raton, 2018. ISBN 978-1-4987-7343-0.
- Toshihisa Tanaka and Mahnaz Arvaneh, editors. *Signal Processing and Machine Learning for Brain-Machine Interfaces*. Institution of Engineering and Technology, September 2018. ISBN 978-1-78561-398-2 978-1-78561-399-9. doi: 10.1049/PBCE114E. URL <https://digital-library.theiet.org/content/books/ce/pbce114e>.
- Rabie A. Ramadan and Athanasios V. Vasilakos. Brain computer interface: control signals review. *Neurocomputing*, 223:26–44, February 2017. ISSN 09252312. doi: 10.1016/j.neucom.2016.10.024. URL <https://linkinghub.elsevier.com/retrieve/pii/S0925231216312152>.
- F Lotte, L Bougrain, A Cichocki, M Clerc, M Congedo, A Rakotomamonjy, and F Yger. A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *Journal of Neural Engineering*, 15(3):031005, June 2018. ISSN 1741-2560, 1741-2552. doi: 10.1088/1741-2552/aab2f2. URL <https://iopscience.iop.org/article/10.1088/1741-2552/aab2f2>.
- Annushree Bablani, Damodar Reddy Edla, Diwakar Tripathi, and Ramalingaswamy Cheruku. Survey on Brain-Computer Interface: An Emerging Computational Intelligence Paradigm. *ACM Computing Surveys*, 52(1):1–32, February 2019. ISSN 0360-0300, 1557-7341. doi: 10.1145/3297713. URL <https://dl.acm.org/doi/10.1145/3297713>.
- Xiang Zhang, Lina Yao, Xianzhi Wang, Jessica Monaghan, David Mcalpine, and Yu Zhang. A survey on deep learning-based non-invasive brain signals: Recent advances and new frontiers. arXiv preprint arXiv:1905.04149, 2020.
- Yueying Zhou, Shuo Huang, Ziming Xu, Pengpai Wang, Xia Wu, and Daoqiang Zhang. Cognitive Workload Recognition Using EEG Signals and Machine Learning: A Review. *IEEE Transactions on Cognitive and Developmental Systems*, pages 1–1, 2021. ISSN 2379-8920, 2379-8939. doi: 10.1109/TCDS.2021.3090217. URL <https://ieeexplore.ieee.org/document/9458953/>.
- Samaa Abdulwahab, Hussain Khleaf, and Manal Jasim. A Survey in Implementation and Applications of Electroencephalograph (EEG)-Based Brain-Computer Interface. *Engineering and Technology Journal*, 39(7):1117–1132, July 2021. ISSN 2412-0758. doi: 10.30684/etj.v39i7.1854. URL https://etj.uotechnology.edu.iq/article_169379.html.

Agradecimentos

Agradecemos ao Laboratório Multiusuário de Computação de Alto Desempenho (LaM-CAD) da UFG, pela disponibilização de recursos computacionais usados na experimentação dos casos de uso, e à Psicóloga Heloisa Bueno Machado, pelas sugestões e revisões no texto. Agradecemos também ao Instituto Federal de Goiás – Campus Senador Canedo,

pelo apoio à pesquisa de doutorado de Murilo Borges Silva, cujos resultados culminaram em partes deste texto.

Referências Bibliográficas

Área de brodmann, May 2022. URL https://pt.wikipedia.org/w/index.php?title=%C3%81rea_de_Brodmann&oldid=63517878. Page Version ID: 63517878.

Hamidreza Abbaspour, Nasser Mehrshad, and Seyyed Mohammad Razavi. Identifying motor imagery activities in brain computer interfaces based on the intelligent selection of most informative timeframe. *SN Applied Sciences*, 2(2):232, January 2020. ISSN 2523-3971. doi: 10.1007/s42452-020-2020-0. URL <https://doi.org/10.1007/s42452-020-2020-0>.

Samaa Abdulwahab, Hussain Khleaf, and Manal Jasim. A Survey in Implementation and Applications of Electroencephalograph (EEG)-Based Brain-Computer Interface. *Engineering and Technology Journal*, 39(7):1117–1132, July 2021. ISSN 2412-0758. doi: 10.30684/etj.v39i7.1854. URL https://etj.uotechnology.edu.iq/article_169379.html.

Charu C. Aggarwal. *Neural Networks and Deep Learning: A Textbook*. Springer International Publishing, Cham, 2018. ISBN 978-3-319-94462-3 978-3-319-94463-0. doi: 10.1007/978-3-319-94463-0. URL <http://link.springer.com/10.1007/978-3-319-94463-0>.

Mohammad A Almogbel, Anh H Dang, and Wataru Kameyama. EEG-Signals Based Cognitive Workload Detection of Vehicle Driver using Deep Learning. *International Conference on Advanced Communications Technology (ICACT)*, page 4, 2018.

Mohammad A. Almogbel, Anh H. Dang, and Wataru Kameyama. Cognitive Workload Detection from Raw EEG-Signals of Vehicle Driver using Deep Learning. In *2019 21st International Conference on Advanced Communication Technology (ICACT)*, pages 1–6, PyeongChang Kwangwoon_Do, Korea (South), February 2019. IEEE. ISBN 979-11-88428-02-1. doi: 10.23919/ICACT.2019.8702048. URL <https://ieeexplore.ieee.org/document/8702048/>.

João Bosco Ferreira Alves, Hugo A. D. do Nascimento, and Murilo Borges Silva. Uma revisão dos estudos de detecção e reconhecimento de sinais cerebrais para interação humano-computador com equipamentos de baixo custo. In *Anais do VI Simpósio Internacional de Inovação em Mídias Interativas*, pages 369–381, Buenos Aires, Argentina, Maio 2019. UBA & UFG. doi: <https://doi.org/10.5281/zenodo.5787257>. URL <https://siimi.medialab.ufg.br/p/29801-anais-2019>.

Jodie Ashford, Jordan J. Bird, Felipe Campelo, and Diego R. Faria. Classification of EEG Signals Based on Image Representation of Statistical Features. In Zhaojie Ju, Longzhi Yang, Chenguang Yang, Alexander Gegov, and Dalin Zhou, editors, *Advances in Computational Intelligence Systems*, volume 1043, pages 449–460. Springer International Publishing, Cham, 2020. ISBN 978-3-030-29932-3 978-3-030-29933-0.

doi: 10.1007/978-3-030-29933-0_37. URL http://link.springer.com/10.1007/978-3-030-29933-0_37. Series Title: Advances in Intelligent Systems and Computing.

Nik Khadijah Nik Aznan, Stephen Bonner, Jason Connolly, Noura Al Moubayed, and Toby Breckon. On the classification of ssvep-based dry-eeeg signals via convolutional neural networks. In *2018 IEEE international conference on systems, man, and cybernetics (SMC)*, pages 3726–3731. IEEE, 2018a.

Nik Khadijah Nik Aznan, Stephen Bonner, Jason D. Connolly, Noura Al Moubayed, and Toby P. Breckon. On the Classification of SSVEP-Based Dry-EEG Signals via Convolutional Neural Networks. *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3726–3731, October 2018b. doi: 10.1109/SMC.2018.00631. URL <http://arxiv.org/abs/1805.04157>. arXiv: 1805.04157.

Annushree Bablani, Damodar Reddy Edla, Diwakar Tripathi, and Ramalingaswamy Cherkuru. Survey on Brain-Computer Interface: An Emerging Computational Intelligence Paradigm. *ACM Computing Surveys*, 52(1):1–32, February 2019. ISSN 0360-0300, 1557-7341. doi: 10.1145/3297713. URL <https://dl.acm.org/doi/10.1145/3297713>.

Mark F. Bear, Barry W. Connors, Michael A. Paradiso, Carla Dalmaz, Jorge Alberto Quillfeldt, Maria Elisa Calcagnotto, Carlos Alberto Saraiva Gonçalves, Carmem Juracy Silveira Gottfried, Denise Maria Zancan, Renata Menezes Rosat, and Tadeu Mello e Souza. *Neurociências: Desvendando o Sistema Nervoso*. Artmed, 4^a edição edition, July 2017. ISBN 978-85-8271-432-4.

Jordan J. Bird, Luis J. Manso, Eduardo P. Ribeiro, Aniko Ekart, and Diego R. Faria. A Study on Mental State Classification using EEG-based Brain-Machine Interface. In *2018 International Conference on Intelligent Systems (IS)*, pages 795–800, Funchal - Madeira, Portugal, September 2018. IEEE. ISBN 978-1-5386-7097-2. doi: 10.1109/IS.2018.8710576. URL <https://ieeexplore.ieee.org/document/8710576/>.

Jordan J. Bird, Diego R. Faria, Luis J. Manso, Anikó Ekárt, and Christopher D. Buckingham. A Deep Evolutionary Approach to Bioinspired Classifier Optimisation for Brain-Machine Interaction. *Complexity*, 2019:1–14, March 2019. ISSN 1076-2787, 1099-0526. doi: 10.1155/2019/4316548. URL <https://www.hindawi.com/journals/complexity/2019/4316548/>.

Jordan J Bird, Michael Pritchard, Antonio Fratini, Anikó Ekárt, and Diego R Faria. Synthetic biological signals machine-generated by gpt-2 improve the classification of eeg and emg through data augmentation. *IEEE Robotics and Automation Letters*, 6(2): 3498–3504, 2021.

Clemens Brunner, Robert Leeb, Gernot Müller-Putz, Alois Schlögl, and Gert Pfurtscheller. Bci competition 2008–graz data set a. *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology*, 16:1–6, 2008/1.

- Pedro G. Campos, Fernando Díez, and Iván Cantador. Time-aware recommender systems: a comprehensive survey and analysis of existing evaluation protocols. *User Modeling and User-Adapted Interaction*, 24(1-2):67–119, February 2014. ISSN 0924-1868, 1573-1391. doi: 10.1007/s11257-012-9136-x. URL <http://link.springer.com/10.1007/s11257-012-9136-x>.
- Tom Chau and Sotirios Damouras. Reply to ‘On the risk of extracting relevant information from random data’. *Journal of Neural Engineering*, 6(5):058002, October 2009. ISSN 1741-2560, 1741-2552. doi: 10.1088/1741-2560/6/5/058002. URL <https://iopscience.iop.org/article/10.1088/1741-2560/6/5/058002>.
- Ryan Cho, Mobasshira Zaman, Kyu Taek Cho, and Jaejin Hwang. Analyzing brain activity during learning tasks with eeg and machine learning. *arXiv preprint arXiv:2401.10285*, 2024.
- Augusto César da Silva Rocha, Álefe Costa Macedo, Murilo Borges da Silva, Hugo A. D. do Nascimento, Kliger Kissinger Fernandes Rocha, and João Bosco Ferreira Alves. Brainoscope. In *Anais do IX Simpósio Internacional de Inovação em Mídias Interativas*, pages 1–14, Santiago, Chile, Novembro 2022. UBA & UFG. 4-5/11/2022.
- Luis Garcia Dominguez. On the risk of extracting relevant information from random data. *Journal of Neural Engineering*, 6(5):058001, October 2009. ISSN 1741-2560, 1741-2552. doi: 10.1088/1741-2560/6/5/058001. URL <https://iopscience.iop.org/article/10.1088/1741-2560/6/5/058001>.
- Daniel Fuentes, Leandro F Malloy-Diniz, Candida Helena Pires de Camargo, and Ramon M Cosenza. *Neuropsicologia: Teoria e Prática*. Artmed Editora, 2014.
- Parnaz Golnar-Nik, Sajjad Farashi, and Mir-Shahram Safari. The application of EEG power for the prediction and interpretation of consumer decision-making: A neuromarketing study. *Physiology & Behavior*, 207:90–98, August 2019. ISSN 00319384. doi: 10.1016/j.physbeh.2019.04.025. URL <https://linkinghub.elsevier.com/retrieve/pii/S0031938418306723>.
- R. W. GUILLERY. Brodmann’s ‘localisation in the cerebral cortex’. translated and edited by Laurence J. Garey. (pp. xviii+300; illustrated; £28 hardback; isbn 1 86094 176 1.) london: Imperial college press. 1999. *Journal of Anatomy*, 196(3):493–496, April 2000. ISSN 1469-7580. doi: 10.1046/j.1469-7580.2000.196304931.x. URL <http://dx.doi.org/10.1046/j.1469-7580.2000.196304931.x>.
- Xiyuan Hou, Yisi Liu, Olga Sourina, and Wolfgang Mueller-Wittig. CogniMeter: EEG-based Emotion, Mental Workload and Stress Visual Monitoring. In *2015 International Conference on Cyberworlds (CW)*, pages 153–160, October 2015. doi: 10.1109/CW.2015.58.
- Hugh Hunkin, Daniel L. King, and Ian T. Zajac. EEG Neurofeedback During Focused Attention Meditation: Effects on State Mindfulness and Meditation Experiences. *Mindfulness*, 12(4):841–851, April 2021. ISSN 1868-

8535. doi: 10.1007/s12671-020-01541-0. URL <https://doi.org/10.1007/s12671-020-01541-0>.

Aleksandra Kawala-Sterniuk, Natalia Browarska, Amir Al-Bakri, Mariusz Pelc, Jaroslaw Zygarlicki, Michaela Sidikova, Radek Martinek, and Edward Jacek Gorzelanczyk. Summary of over fifty years with brain-computer interfaces—a review. *Brain Sciences*, 11(1), 2021. ISSN 2076-3425. doi: 10.3390/brainsci11010043. URL <https://www.mdpi.com/2076-3425/11/1/43>.

Ping Chai Koo, Johannes Thome, Christoph Berger, Paul Foley, and Jacqueline Hoepfner. Current source density analysis of resting state EEG in depression: a review. *Journal of Neural Transmission*, 124(S1):109–118, February 2017. ISSN 0300-9564, 1435-1463. doi: 10.1007/s00702-015-1432-2. URL <http://link.springer.com/10.1007/s00702-015-1432-2>.

Robert Lent. *Cem bilhões de neurônios: conceitos fundamentais de neurociência*. Editora Atheneu, 2 edition, 2010. ISBN 9788538801023.

Fernando Lopes da Silva. EEG and MEG: Relevance to Neuroscience. *Neuron*, 80(5):1112–1128, December 2013. ISSN 08966273. doi: 10.1016/j.neuron.2013.10.017. URL <https://linkinghub.elsevier.com/retrieve/pii/S0896627313009203>.

F Lotte, L Bougrain, A Cichocki, M Clerc, M Congedo, A Rakotomamonjy, and F Yger. A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. *Journal of Neural Engineering*, 15(3):031005, June 2018. ISSN 1741-2560, 1741-2552. doi: 10.1088/1741-2552/aab2f2. URL <https://iopscience.iop.org/article/10.1088/1741-2552/aab2f2>.

Aleksandr Romanovich Luria. *The working brain: An introduction to neuropsychology*. Basic Books, 1976.

Sheena Luu and Tom Chau. Decoding subjective preference from single-trial near-infrared spectroscopy signals. *Journal of Neural Engineering*, 6(1):016003, February 2009. ISSN 1741-2560, 1741-2552. doi: 10.1088/1741-2560/6/1/016003. URL <https://iopscience.iop.org/article/10.1088/1741-2560/6/1/016003>.

M-Marsel Mesulam. *Principles of behavioral and cognitive neurology*. Oxford University Press, 2000.

Chang S. Nam, Anton Nijholt, and Fabien Lotte, editors. *Brain-computer interfaces handbook: technological and theoretical advances*. Taylor & Francis, CRC Press, Boca Raton, 2018. ISBN 978-1-4987-7343-0.

The Rt. Hon. Lord Cohen of Birkenhead. Richard caton (1842-1926): Pioneer electrophysiologist. In *Proceedings of the Royal Society of Medicine, Section of the History of Medicine*, volume 52, page 21, Liverpool, 1959. Royal Society of Medicine. Meeting: March 4, 1959. President: Sir Weldon Dalrymple-Champneys, Bt., C.B., M.A., D.M., F.R.C.P.

- Alison OrShea, Gordon Lightbody, Geraldine Boylan, and Andriy Temko. Investigating the Impact of CNN Depth on Neonatal Seizure Detection Performance. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5862–5865, Honolulu, HI, July 2018. IEEE. ISBN 978-1-5386-3646-6. doi: 10.1109/EMBC.2018.8513617. URL <https://ieeexplore.ieee.org/document/8513617/>.
- Olavo Pazzanese and Paulo Pinto Pupo. A eletrencefalografia na epilepsia: Considerações sobre 271 casos. *Arquivos de Neuro-Psiquiatria*, 6(2):107–150, June 1948. ISSN 0004-282X. doi: 10.1590/s0004-282x1948000200001. URL <http://dx.doi.org/10.1590/s0004-282x1948000200001>.
- Rabie A. Ramadan and Athanasios V. Vasilakos. Brain computer interface: control signals review. *Neurocomputing*, 223:26–44, February 2017. ISSN 09252312. doi: 10.1016/j.neucom.2016.10.024. URL <https://linkinghub.elsevier.com/retrieve/pii/S0925231216312152>.
- Yannick Roy, Hubert Banville, Isabela Albuquerque, Alexandre Gramfort, Tiago H Falk, and Jocelyn Faubert. Deep learning-based electroencephalography analysis: a systematic review. *Journal of Neural Engineering*, 16(5):051001, August 2019. ISSN 1741-2552. doi: 10.1088/1741-2552/ab260c. URL <https://iopscience.iop.org/article/10.1088/1741-2552/ab260c>.
- Aditi Sakalle, Pradeep Tomar, Harshit Bhardwaj, Divya Acharya, and Arpit Bhardwaj. A LSTM based deep learning network for recognizing emotions using wireless brainwave driven system. *Expert Systems with Applications*, 173:114516, July 2021. ISSN 09574174. doi: 10.1016/j.eswa.2020.114516. URL <https://linkinghub.elsevier.com/retrieve/pii/S095741742031160X>.
- Saeid Sanei and Jonathan Chambers. *EEG Signal Processing*. John Wiley & Sons Ltd, 2007. ISBN 9780470025819. URL <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470511923>.
- J. A. Caetano Da Silva Junior. Idéias gerais sôbre a eletrencefalografia. *Arquivos de Neuro-Psiquiatria*, 3(2):170–176, jun 1945. ISSN 0004-282X. doi: 10.1590/s0004-282x1945000200007. URL <http://dx.doi.org/10.1590/s0004-282x1945000200007>.
- Sandro Skansi. *Introduction to Deep Learning: From Logical Calculus to Artificial Intelligence*. Undergraduate Topics in Computer Science. Springer International Publishing, Cham, 2018. ISBN 978-3-319-73003-5 978-3-319-73004-2. doi: 10.1007/978-3-319-73004-2. URL <http://link.springer.com/10.1007/978-3-319-73004-2>.
- Toshihisa Tanaka and Mahnaz Arvaneh, editors. *Signal Processing and Machine Learning for Brain-Machine Interfaces*. Institution of Engineering and Technology, September 2018. ISBN 978-1-78561-398-2 978-1-78561-399-9. doi: 10.1049/PBCE114E. URL <https://digital-library.theiet.org/content/books/ce/pbce114e>.

Jason Teo and Jia Tian Chia. Deep Neural Classifiers For Eeg-Based Emotion Recognition In Immersive Environments. In *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*, pages 1–6, Shah Alam, July 2018. IEEE. ISBN 978-1-5386-4836-0 978-1-5386-4838-4. doi: 10.1109/ICSCEE.2018.8538382. URL <https://ieeexplore.ieee.org/document/8538382/>.

Jason Teo, Lin Hou, Jia Tian, and James Mountstephens. Classification of Affective States via EEG and Deep Learning. *International Journal of Advanced Computer Science and Applications*, 9(5), 2018. ISSN 21565570, 2158107X. doi: 10.14569/IJACSA.2018.090517. URL <http://thesai.org/Publications/ViewPaper?Volume=9&Issue=5&Code=ijacsa&SerialNo=17>.

Matthew Wright, Adrian Freed, et al. Open soundcontrol: A new protocol for communicating with sound synthesizers. In *ICMC*, 1997.

Xiang Zhang, Lina Yao, Xianzhi Wang, Jessica Monaghan, David Mcalpine, and Yu Zhang. A survey on deep learning-based non-invasive brain signals: Recent advances and new frontiers. arXiv preprint arXiv:1905.04149, 2020.

Yueying Zhou, Shuo Huang, Ziming Xu, Pengpai Wang, Xia Wu, and Daoqiang Zhang. Cognitive Workload Recognition Using EEG Signals and Machine Learning: A Review. *IEEE Transactions on Cognitive and Developmental Systems*, pages 1–1, 2021. ISSN 2379-8920, 2379-8939. doi: 10.1109/TCDS.2021.3090217. URL <https://ieeexplore.ieee.org/document/9458953/>.

Rümeysa İnce, Saliha Seda Adanır, and Fatma Sevmez. The inventor of electroencephalography (EEG): Hans Berger (1873–1941). *Child's Nervous System*, 37(9):2723–2724, September 2021. ISSN 1433-0350. doi: 10.1007/s00381-020-04564-z. URL <https://doi.org/10.1007/s00381-020-04564-z>.

Capítulo

2

Simuladores imersivos para ensino em saúde: futuro ou realidade?

Luciana Nedel (UFRGS), Carla M.D.S. Freitas (UFRGS), Matheus D. Negrão (UFRGS), Wesley Ferreira (UFRGS), Anderson Maciel (IST, UFRGS), Ana Paula Rosses (UFPeI), Betânia Bohrer (UFRGS), Rafael Torchelsen (UFPeI), Lucas Alviene (UFRGS), Leonardo Guths (UFRGS), Zaira Teston (UFRGS)

Abstract

Virtual reality (VR) training has been used for decades, starting with flight simulators in the 1970s. VR has also been applied to physical and mental health, including surgical and rehabilitation applications. However, its use in other healthcare areas remains limited. Virtual simulations offer personalized training for different profiles, from novices to specialists. Recently, advancements and availability in VR hardware have improved the quality of experiences, making simulators more comfortable and efficient. This chapter discusses the use of immersive simulators in teaching technical and behavioral skills in healthcare, highlighting the evolution of the technology, current applications, and future challenges.

Resumo

Treinamento em realidade virtual (RV) vem sendo usado há décadas, tendo iniciado com os simuladores de voo na década de 1970. Realidade virtual também tem sido usada aplicada à saúde física e mental, incluindo aplicações em cirurgias e reabilitação. No entanto, o uso em outros cuidados de saúde ainda é limitado. Simulações virtuais oferecem treinamento personalizado para diferentes perfis, desde novatos até especialistas. recentemente, avanços e disponibilidade no hardware de RV melhoraram a qualidade das experiências, tornando os simuladores mais confortáveis e eficientes. Este capítulo discute o uso de simuladores imersivos no ensino de habilidades técnicas e comportamentais na saúde, destacando a evolução da tecnologia, as aplicações atuais e os desafios futuros.

2.1. Introdução

Simulações utilizando realidade virtual têm sido aplicadas para treinamento em diversos contextos há décadas (Zyda, 2005). Os simuladores de voo começaram na década de 1970 e desde então vários simuladores têm sido usados tanto na indústria da aviação civil como no treinamento militar.

Nas últimas décadas, temos testemunhado ainda o uso crescente da tecnologia de realidade virtual (RV) e realidade aumentada (RA) em diversas aplicações relacionadas à saúde física e mental humana. Uma das áreas onde se tem observado o uso de tecnologias de RV no contexto da saúde é a dos procedimentos cirúrgicos. Há décadas, simuladores virtuais são utilizados para praticar desde técnicas minimamente invasivas até cirurgias maxilofaciais e neurológicas complexas (Lungu et al., 2021). Na saúde mental, encontramos aplicações que visam o tratamento de fobias e reabilitação de situações graves. Um exemplo de aplicação para reabilitação é apresentado por Raya et al. (2023), que relata um aplicativo para uso em UTI pediátrica para tratar e reabilitar crianças em situação de delirium. O aplicativo fornece estímulos externos e representa uma potencial terapia não farmacológica.

Entretanto, a utilização de tecnologia semelhante em outros ambientes de cuidado de saúde, como atendimento inicial de emergência, consultas de pacientes ou enfermarias, tem sido limitada pela necessidade de um ambiente adequado e flexível. Isso se deve ao fato de que, num cenário de ensino prático convencional de *role-playing* de consulta médica, por exemplo, além de fornecer o diagnóstico correto, é importante avaliar a lógica por trás dele (Libin et al., 2010). Isso significa que elementos como as perguntas feitas e as não feitas, a escolha dos exames físicos e a condução da conversa são igualmente relevantes para a formação do profissional (Caldwell, 2019).

A demanda por profissionais de saúde tem aumentado rapidamente e estima-se que cresça ainda nas próximas décadas, mas a capacidade das Universidades para formar médicos, enfermeiros e outros profissionais de saúde não aumenta no mesmo ritmo. Isto se deve ao uso de técnicas tradicionais de ensino-aprendizagem que exigem um especialista dedicado para treinar e avaliar cada estudante de forma individualizada. Embora em muitas outras áreas profissionais o uso da Web, o ensino remoto e os sistemas de auto-aprendizagem tenham obtido bons resultados no aumento das taxas aluno-professor, no ensino na área da saúde isso é mais desafiador, pois requer habilidades práticas que devem ser aprendidas e exercitadas em um cenário prático de *role-playing* sob supervisão de um profissional mais experiente (Joyner and Young, 2009, Nestel and Tierney, 2007), o qual geralmente dispõe de tempo limitado para tais atividades.

Entre os benefícios já consolidados da simulação virtual para o ensino de competências na área da saúde estão a capacidade de personalizar o treinamento e atender os mais diversos perfis de alunos, desde iniciantes até especialistas. Entre os novatos, pode haver um ganho de confiança devido ao treino repetido de procedimentos. Já os especialistas podem dominar algumas situações complexas e raras que não apresentam oportunidades suficientes para o treinamento frequente e, até mesmo, para a manutenção de uma competência conquistada.

Um aspecto importante que facilita essa evolução nas aplicações de ensino em

saúde é a qualidade atual do hardware de realidade virtual. Até recentemente, a tecnologia de RV não atendia aos padrões de qualidade para simular adequadamente cenários em que pudessem ser avaliados tanto aspectos comportamentais como técnicos dos procedimentos médicos. Esse panorama mudou a partir dos avanços recentes, pois o hardware se tornou mais difundido por um custo menor, além de mais confortável e leve. Algumas novas funcionalidades têm contribuído para uma melhor experiência do usuário, por exemplo, o rastreamento da movimentação do corpo e mãos sem fio e sem a necessidade de infraestrutura externa instalada; e o campo de visão dos *head-mounted displays* (HMD), que mais que dobrou. Estes avanços permitem focar em duas frentes de pesquisa e desenvolvimento. A primeira é como projetar e desenvolver simuladores imersivos que se adaptem à necessidade de treinamento de cada situação. Já a segunda é definir como as habilidades a serem aprendidas e treinadas devem ser replicadas e avaliadas no ambiente imersivo.

Dado esse contexto geral, o objetivo deste capítulo é discutir o uso de simuladores imersivos no ensino de habilidades técnicas (*hard skills*) e não técnicas (*soft skills*) na área da saúde, mais especificamente, no ensino médico. Apresentaremos uma breve revisão dos conceitos principais associados à área de realidade virtual e a evolução e o cenário atual dos simuladores utilizando essa tecnologia na área da saúde. Discutiremos ainda como tem sido feita a transposição das habilidades técnicas e comportamentais para a interação tridimensional (3D) nos ambientes imersivos. Apresentaremos o posicionamento mercadológico dos atuais simuladores e seus usos na formação de novos profissionais, permitindo o treinamento prático sem expor os pacientes e de forma autônoma, sem a necessidade de acompanhamento individual com um especialista. Finalmente, discutiremos os desafios atuais e perspectivas desse cenário para o futuro próximo.

2.2. Evolução da área

O conceito de realidade virtual foi introduzido pela primeira vez por Ivan Sutherland em seu trabalho seminal “The Ultimate Display”, de 1965. Naquela época, Sutherland já sugeria que um *display* poderia ser construído para fornecer imagens geradas por computador tão realistas que seriam indistinguíveis das coisas reais. Três anos depois, em 1968, ele produziu o primeiro capacete de realidade virtual acoplado a um computador. Esse capacete era composto por dois pequenos monitores CRT (*cathode ray tube*) montados em uma bandana e rastreava a posição da cabeça do usuário. A ideia por trás desse “novo” produto é reproduzida no parágrafo a seguir:

Não pense nisso como uma tela, pense nisso como uma janela, uma janela através da qual alguém olha para um mundo virtual. O desafio da computação gráfica é fazer com que esse mundo virtual tenha aparência real, soe real, se mova, dê respostas em tempo real, e até mesmo cause sensações reais.¹

2.2.1. Conceitos

Durante a última década, o termo “realidade virtual” tornou-se popular e tem sido usado indiscriminadamente para caracterizar desde simples aplicativos gráficos interativos até

¹tradução livre do original (Sutherland, 1965)

experiências totalmente imersivas, às vezes misturando objetos virtuais e reais. Percebendo a falta de uma taxonomia para distinguir tais aplicações interativas, Milgram and Kishino (1994) propuseram o *continuum* da virtualidade, que está relacionado com a mistura de classes de objetos apresentados em qualquer situação de exibição. Ambientes reais são mostrados em uma extremidade do continuum, enquanto ambientes virtuais estão na extremidade oposta. Nesse contexto, esses autores propuseram um conceito claro e simples para a realidade virtual. Para eles,

Um ambiente de Realidade Virtual (RV) é aquele em que o participante-observador está totalmente imerso e é capaz de interagir com um mundo completamente sintético.

Em outras palavras, podemos dizer que a realidade virtual é um meio composto por simulações computadorizadas e interativas que detectam a posição e as ações do participante e substituem ou aumentam o feedback para um ou mais sentidos, proporcionando a sensação de estar mentalmente imerso ou presente na simulação (um mundo virtual).

Quatro elementos-chave são necessários para uma experiência de realidade virtual: mundo virtual, sensação de imersão, feedback sensorial e interatividade.

Enquanto o conceito de **mundo virtual** é bastante óbvio e se refere a um espaço imaginário composto por um conjunto de objetos representados em 3D e acrescido de regras que governam esses objetos, a **imersão** nesse mundo pode ser tanto física quanto mental. O estado de imersão mental, ou seja, estar profundamente envolvido, é frequentemente referido como ter um “senso de presença” em um ambiente, sendo um desafio na área medir esse “senso de presença” (Nedel et al., 2016). A imersão física, por outro lado, requer o uso de estímulos sintéticos por meio de tecnologias que ajudem o corpo a sentir o ambiente virtual.

O **feedback sensorial** é uma característica essencial das aplicações de realidade virtual. Um sistema de RV fornece feedback sensorial direto aos participantes com base em sua posição física. Na maioria dos casos, o feedback é visual, embora ambientes de realidade virtual ideais devam estimular todos os sentidos humanos (audição, visão, tato, olfato e paladar). Para gerar a saída sensorial do sistema de RV na posição do participante, o sistema deve rastrear seus movimentos. Um sistema de RV típico rastreará a cabeça dos participantes e pelo menos as mãos ou objetos segurados pelas mãos.

Para que a realidade virtual pareça autêntica, ela deve responder às ações do usuário, ou seja, ser responsiva e, mais do que isso, **interativa**. Geralmente, os computadores permitem isso, mas alguns requisitos tecnológicos devem ser considerados: hardware de geração de imagens 3D em tempo real e geração de som estéreo de alta qualidade; dispositivos de entrada e saída específicos que simulam e estimulam os sentidos humanos; e software para simular ambientes virtuais, muitas vezes com alto nível de realismo, que permitem uma resposta imediata às ações do usuário. Esses requisitos tecnológicos compõem os elementos centrais das interfaces humano-computador em RV.

O termo **metaverso** se refere a um espaço virtual tridimensional coletivo e compartilhado, que tenta replicar ou simular a realidade através de dispositivos digitais. Ele representa a possibilidade de uma espécie de realidade paralela onde as pessoas podem

interagir, criar, explorar e socializar (Wikipedia, 2023). A ideia central de metaverso é criar um ambiente digital imersivo e expansível, semelhante a um universo virtual, onde os usuários podem se conectar e interagir como se estivessem fisicamente presentes. Esses ambientes podem variar em termos de realismo, desde mundos virtuais altamente detalhados até espaços mais estilizados e simplificados. Em outras palavras, seria a Internet 3D populada por pessoas reais representadas pelos seus avatares.

Desta forma, ainda que o conceito esteja largamente difundido e o termo venha sendo empregado exaustivamente, na prática o metaverso ainda não existe. Deverá ser uma construção coletiva, a exemplo do que aconteceu anteriormente com a Web. No caso do metaverso, entretanto, são consideradas 7 camadas (ver Figura 2.1).



Figura 2.1. As sete camadas do metaverso e as empresas-chave que atuam em cada uma. Fonte: Wikipedia (2023).

A primeira camada é a **infraestrutura**, a base técnica que dará suporte para todo o projeto, incluindo as tecnologias 5G (e 6G), já que velocidade, processamento e armazenamento em nuvem são essenciais. A segunda camada é a **interface** necessária para acessar o metaverso e inclui os óculos de realidade virtual e aumentada, smartphones e toda a tecnologia necessária para conectar os avatares digitais à experiência sensorial física dos usuários. O terceiro nível é a **descentralização** e envolve a tecnologia necessária para garantir a liberdade para que todas as pessoas naveguem de um ambiente a outro. É esperado que *blockchain*, NFTs e inteligência artificial tenham um papel importante neste nível.

A quarta camada está sendo chamada de **computação espacial** e visa garantir a integração do mundo virtual com o mundo real. Ela faz uso de realidade virtual, realidade aumentada, sensores, dispositivos e técnicas de interação 3D para fazer o mapeamento entre os dois mundos. A quinta camada é chamada de **economia de criação** e envolve uma infinidade de ferramentas de *design*, fluxos de trabalho personalizados, mercados de

ativos, *assets*, etc. Existe a expectativa de que essa camada envolva a colaboração entre as empresas da área.

A sexta camada, **descoberta**, permitirá que as empresas monetizem com os usuários através de publicidade, reviews, lojas, etc. Finalmente, a sétima e última camada é denominada **experiência**. É nesta camada que o público é atraído e cativado através de um espaço de entretenimento e gamificação (ver Figura 2.2). A camada de experiência envolve o desenvolvimento e a disponibilização de conteúdo através das diversas aplicações. Estima-se que educação, entretenimento, trabalho colaborativo, e jogos são as áreas que mais se beneficiarão do metaverso.



Figura 2.2. Espaço do INF-UFRGS no metaverso, criado com a ferramenta Mozilla Hubs.

Para que o metaverso seja uma realidade, pesquisadores, empresas de tecnologia e desenvolvedores independentes estão explorando ativamente esse conceito e trabalhando separadamente nas sete camadas identificadas (ver Figura 2.1).

2.2.2. Evolução dos simuladores em realidade virtual

Sistemas de realidade virtual têm sido amplamente utilizados para treinar profissionais em áreas tão diversas como medicina, indústria, combate a incêndios, educação e saúde. Essas aplicações são apresentadas como jogos com um propósito sério, simplesmente *jogos sérios* – *serious games* (SGs) – ou ainda *jogos aplicados*. Sawyer (2007) os define como “qualquer uso significativo de recursos de jogos informatizados ou da indústria de jogos cuja missão principal não seja o entretenimento”. Zyda (2005) descreve como sendo “um

desafio mental, jogado com um computador de acordo com regras específicas, que usa o entretenimento para promover objetivos de treinamento governamentais ou corporativos, educação, saúde, políticas públicas e comunicação estratégica”. Finalmente, de acordo com Michael and Chen (2005), são “jogos que não têm entretenimento ou diversão como seu principal objetivo”.

Assim como os jogos de vídeo, os SGs envolvem jogabilidade, desafio, interação e objetivo, enquanto as aplicações *gamificadas* incorporam apenas elementos de jogos (Deterding et al., 2011). Segundo esses autores, *gamificação* é definida como “o uso de elementos de design de jogos em contextos não relacionados a jogos”. Esses elementos podem estar relacionados a:

- componentes de *design* de interação e soluções de *design* para um problema conhecido em um contexto, como crachá, classificação e nível;
- componentes de jogabilidade, como limite de tempo, recursos limitados e rodadas;
- diretrizes avaliativas, como um jogo duradouro e objetivos claros;
- modelos conceituais de componentes de jogos, como desafio, fantasia e curiosidade; e
- práticas e processos, como testes de jogo e *design* centrado no jogo.

Além disso, os SGs são construídos sobre estruturas pedagógicas e educacionais, que definem a relação entre aprendizado e mecanismos de jogo, garantindo uma combinação bem sucedida desses fatores para alcançar seu propósito sério. O *framework* apresentado por Ibanez et al. (2011) abrange seis facetas do desenvolvimento de SGs:

- objetivos de aprendizado, que definem um quadro de referência do domínio a ser ensinado;
- simulação de domínio, que define um modelo formal estabelecendo as bases da simulação;
- interação com a simulação de domínio, que é o coração da metáfora;
- problemas e progressão, que definem o nível de dificuldade e a progressão de habilidades;
- decoração, que descreve como entreter e envolver o jogador; e
- implantação, que descreve as condições de uso para preservar as qualidades de aprendizado do jogo.

Técnicas de *design* de jogos podem ser úteis para criar aplicativos de RV mais envolventes e que sejam, ao mesmo tempo, mais atrativos para o público. Essas aplicações, combinadas com a crescente disponibilidade de óculos de realidade virtual (HMDs) para

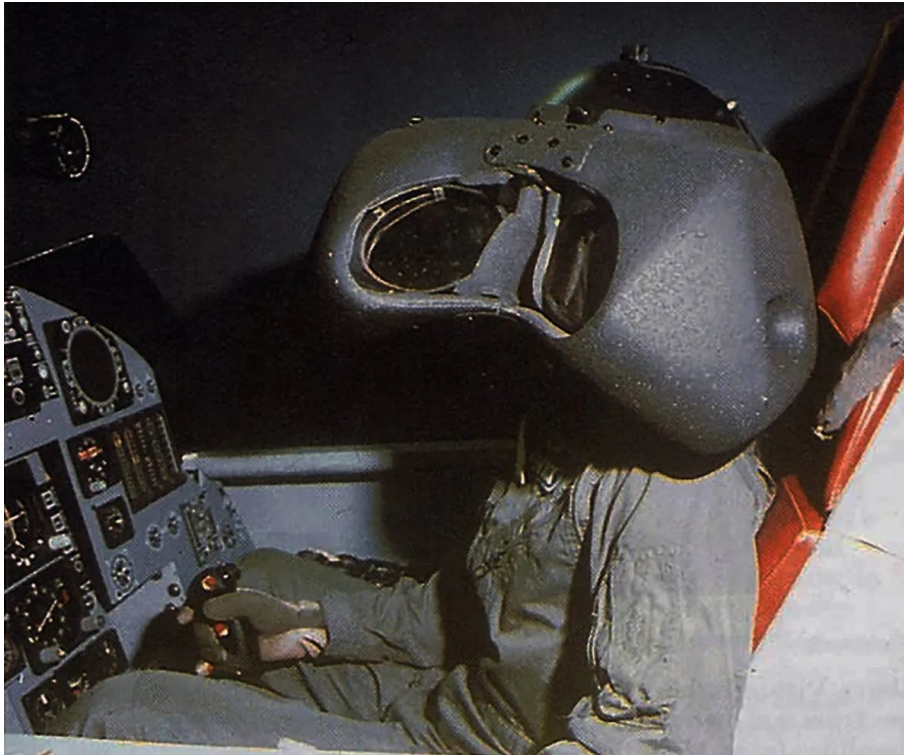


Figura 2.3. Protótipo do Super Cockpit ³

consumidores, podem levar a um futuro em que as pessoas treinem e aprendam usando a RV imersiva. Níveis mais elevados de imersão têm apresentado efeitos positivos em julgamentos espaciais em pequena escala e memorização. A combinação desses fatores leva à memorização de procedimentos complexos, o que permite que o ambiente virtual produza resultados de treinamento e aprendizado. Bowman et al. (2009) fornecem evidências empíricas de que um alto nível de imersão também pode produzir uma melhoria mensurável no desempenho de uma atividade mental abstrata. Chalmers and Debattista (2009) afirmam que um alto nível de realismo é necessário para garantir que o treinamento e a aprendizagem em ambientes virtuais sejam equivalentes ao mundo real. Eles também devem ser capazes de simular todos os sentidos humanos simultaneamente.

A RV imersiva combinada com técnicas usadas em jogos leva ao desenvolvimento de simulações imersivas com aspectos de jogo que permitem o envolvimento em atividades de aprendizado que de outra forma seriam caras ou muito perigosas, difíceis ou impraticáveis de implementar em sala de aula. Seu uso ajuda a mudar o relacionamento das pessoas com as informações, incentivando a visualização, experimentação e criatividade. Elas são flexíveis e complexas o suficiente para atender a diferentes estilos de aprendizado e ampliar a exposição de diferentes pessoas e perspectivas, incentivando a colaboração e apoiando discussões significativas após o jogo.

A evolução notável dos simuladores em realidade virtual alcança o treinamento em saúde, impulsionada tanto pelos avanços tecnológicos já mencionados como por pesquisas científicas dedicadas. A aviação foi parte importante na evolução de simuladores

³<https://garc.space/history/tom-furness-cockpit-vr/>

imersivos, sendo uma das primeiras áreas profissionais a aplicar treinamentos usando *head-mounted displays*, como projeto do Super Cockpit da força aérea dos Estados Unidos (ver fig. 2.3), um dos primeiros protótipos para treinamento usando realidade virtual Furness III (1986). A constatação dessa evolução pode ser observada a partir dos primeiros simuladores imersivos surgidos no final do século XX, como os simuladores para treinamento em procedimentos de laparoscopia (Satava, 1993), até os dias atuais, onde a tecnologia de realidade virtual está integrada com feedback háptico e inteligência artificial.

Nos anos 2000, estudos demonstraram aplicações em realidade virtual que se propunham a expor terapeuticamente militares em situação de combate, oferecendo uma abordagem para o tratamento de estresse pós-traumático (Rizzo et al., 2009). Mais recentemente, com a evolução dos óculos de realidade virtual *All-in-one*, esses dispositivos estão se tornando mais leves e baratos. Car et al. (2022) mapeou e revisou uma série de trabalhos que utilizam Realidade Virtual, Aumentada e Mista na graduação médica, buscando efetividade no uso de instrumentos de medida com evidências de validade em ensaios clínicos randomizados (ECR). Os trabalhos revisados mostram que os resultados menos comuns incluíram as atitudes dos participantes, a satisfação, a carga cognitiva ou mental, a eficácia da aprendizagem, o estado emocional, etc. Kyaw et al. (2019) também apresentaram uma análise que mostra uma ligeira melhoria na eficácia da RV na aprendizagem em comparação com os métodos tradicionais. Por fim, Aliwi et al. (2023) também revisaram o papel da RV e RA imersivas na saúde, demonstrando que a maioria dos estudos com essas tecnologias melhoram a comunicação médica.

2.3. Uso de simuladores na formação de profissionais da saúde

A primeira menção oficial ao ensino baseado em simulação na educação médica surgiu no relatório Flexner, de 1910 (Flexner, 1972), sendo utilizado o uso de simulações para o ensino de obstetrícia. Há décadas, portanto, a simulação é considerada uma das principais ferramentas de ensino-aprendizagem para formação de profissionais de saúde, sendo utilizada no aprimoramento de conhecimentos, habilidades e atitudes.

O aumento da necessidade de horas de treinamento pelo avanço na área da saúde, aumento nas vagas das universidades no campo da saúde, número limitado de encontros com pacientes e a necessidade de proporcionar maior segurança ao paciente corroboraram com o avanço de novas tecnologias e formas de proporcionar um currículo padronizado através de simulações.

A segurança do paciente é uma preocupação da Organização Mundial da Saúde (OMS) que, em 2009, definiu o conceito de *Segurança do Paciente* como “a redução dos riscos de danos desnecessários associados à assistência em saúde até um mínimo aceitável”. Em 2011, a OMS publicou um documento (OMS, 2016) no qual recomenda a simulação como uma ferramenta muito útil para proporcionar maior segurança aos pacientes em substituição à abordagem tradicional “leia, veja e faça”, com a justificativa de que ambientes de simulação proporcionam um cenário seguro e controlado, onde os estudantes podem praticar habilidades técnicas e não técnicas, tomar decisões e enfrentar desafios sem expor pacientes reais a riscos.

Além do aprimoramento técnico e segurança do paciente, a utilização de simula-

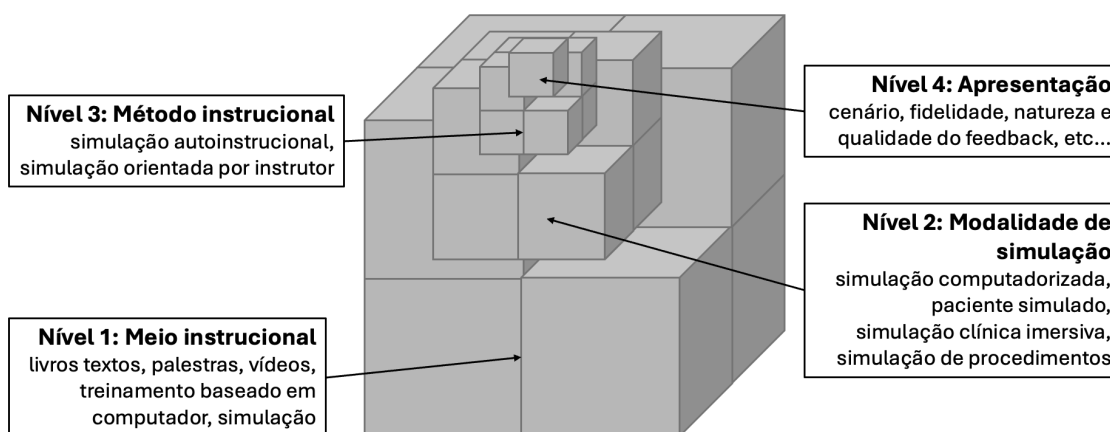


Figura 2.4. Níveis de *design* instrucional para a construção de simuladores para o ensino em saúde. Fonte: Adaptada de Chiniara et al. (2013)

ção nos cuidados da saúde também pode ter a finalidade de gestão de pacientes e trabalho em equipe, sendo essa uma utilização mais recente que vem se tornando cada vez mais significativa.

Existem diferentes modalidades de simulação na saúde e cada uma se ajusta melhor às necessidades de formação, ou seja, não há um método uniforme e sim uma variedade de metodologias que irão proporcionar experiências educacionais muito diferentes.

Na tentativa de melhor descrever e sistematizar o uso de simuladores no ensino em saúde, pesquisadores da *Canadian Network for Simulation in Healthcare Guidelines Working Group* (Chiniara et al., 2013) propuseram um *framework* baseado em quatro níveis do que eles chamaram de *design* instrucional: meio instrucional, modalidade de simulação, método instrucional e apresentação.

O objetivo foi fornecer uma base sólida sobre a qual construir as características da simulação para uma determinada intervenção educacional. Cada nível abrange um conjunto definido de características que correspondem a um nível de detalhes da atividade de simulação. Dito de outra forma, uma atividade envolvendo simulação consiste em diversas características agrupadas em níveis distintos com base no seu impacto específico, na qualidade global e na concepção da intervenção educacional. As escolhas para qualquer característica num determinado nível de detalhe geralmente dependem das escolhas feitas nos níveis anteriores. Em cada nível, as escolhas feitas dependem das necessidades reais de aprendizagem e dos objetivos da atividade. A Figura 2.4 mostra os níveis propostos.

O nível **meio instrucional** abrange os principais meios de ensino, como livros didáticos, palestras, vídeos, treinamento baseado em computador e simulação. A decisão de usar uma simulação deve considerar duas características: relevância (gravidade do evento e seu impacto) e frequência (frequência do evento ou probabilidade de descoberta de problemas). Essas duas características definem uma matriz de eventos em quatro quadrantes (ver Figura 2.5): (1) Alta relevância e baixa frequência (ARBF), por exemplo, hipertermia maligna na sala de cirurgia; (2) Alta relevância e alta frequência (ARAF), por exemplo, hemorragia pós-parto; (3) Baixa relevância e baixa frequência (ARBF), por exemplo, desimpactação manual de massa fecal e (4) Baixa relevância e alta frequência

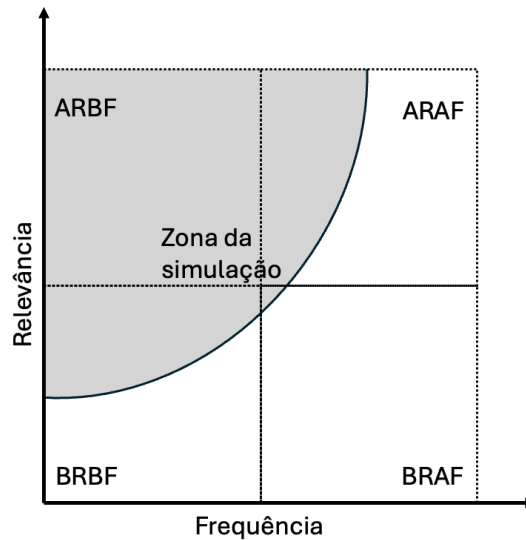


Figura 2.5. Matriz de eventos, onde a “zona de simulação” abrange todas as situações ARBF (alta relevância e baixa frequência) e, quando viável, a maioria das situações ARAF (alta relevância e alta frequência) e BRBF (baixa relevância e baixa frequência). Para situações BRAF (baixa relevância e alta frequência), a simulação pode não ser o método mais eficiente. Fonte: Adaptada de Chiniara et al. (2013)

(BRAF), por exemplo, indução descomplicada de anestesia.

No nível **modalidade da simulação**, a classificação é realizada segundo o tipo de simulação (ver Figura 2.6). Cada modalidade possui características distintas que influenciam os resultados educacionais e são categorizadas para facilitar sua adaptação a objetivos específicos de aprendizagem.

A simulação de procedimentos equivale ao treinamento de habilidades cujo foco principal é a aquisição de habilidades psicomotoras, com necessidade de *checklists* padronizadas e repetição. Trata-se de uma simulação simples e focada num objetivo específico

Na simulação de pacientes padronizados há utilização de atores ou membros da comunidade que representam pacientes reais com o objetivo de treinar as competências clínicas, habilidades de comunicação e comportamento dos treinandos. Pacientes padronizados atuam em um cenário com anamnese, sinais clínicos ou reação emocional predefinida, como ocorre no método OSCE (*Objective Structured Clinical Examination*).

A simulação clínica por imersão difere da simulação de pacientes padronizados por contar com cenários, ou seja, o ambiente profissional é recriado com fidelidade e os participantes exercitam habilidades, atitudes, tomadas de decisão e outros elementos da competência clínica de maneira mais completa. Essas simulações podem ser realizadas com realidade virtual, manequins simuladores de pacientes ou com atores. É normalmente utilizada para treinamento em gerenciamento de pacientes, diagnóstico clínico e competências de segurança do paciente como nas avaliações clínicas da prova prática do Exame Nacional de Revalidação de diplomas médicos expedidos por instituição de Educação Superior Estrangeira, conhecido como Revalida.

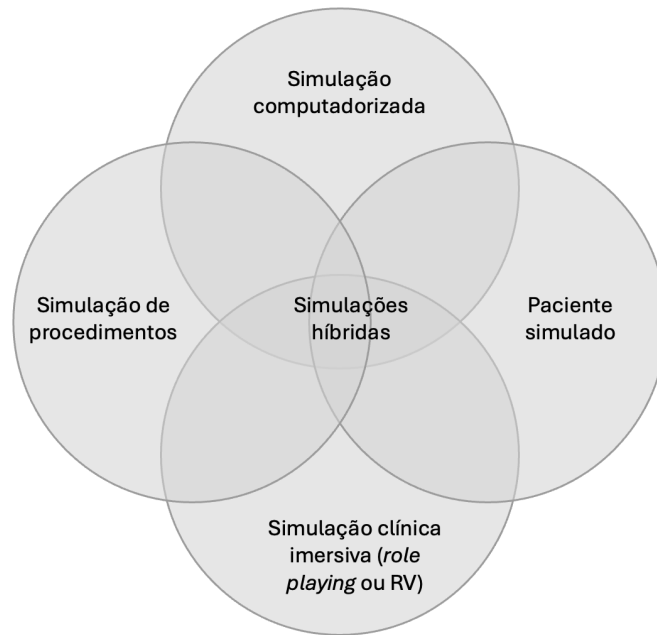


Figura 2.6. Modalidades de simulação. Fonte: Adaptada de Chiniara et al. (2013)

Na simulação computadorizada, a interface de interação com os participantes é uma tela de computador, com graus variáveis de fidelidade. Pode ser destinada ao treinamento de habilidades psicomotoras como a destreza em procedimentos videolaparoscópicos (uso de realidade virtual) ou a tomada de decisão em cenário virtual.

Em algumas situações, para uma avaliação mais completa podemos utilizar diferentes modalidades de simulação ao mesmo tempo, o que pode ser descrito como simulação híbrida (ou combinada). Tais simulações permitem, por exemplo, o treinamento em habilidades técnicas combinadas com proficiência em comunicação como são as simulações em RV utilizando HMDs. Na Tabela 2.3, temos a descrição e exemplos de algumas modalidades de simulação.

Em relação ao nível **método instrucional**, a simulação pode ser autoinstrucional ou orientada por um instrutor. Na autoinstrução os alunos controlam o tempo e o ritmo da aprendizagem, escolhendo seus objetivos. Essa é a forma adequada para simulações de procedimentos e simulação computadorizada. Já a aprendizagem orientada por instrutor é a mais comum na área da saúde e envolve graus diferentes de participação do professor/preceptor desde esclarecimento até a participação ativa nas sessões de treinamento, como o que ocorre nas práticas de *role play*, por exemplo.

Já o nível **apresentação** inclui características que definem exatamente como a atividade de simulação é pensada e projetada. Embora geralmente envolvam pequenas diferenças dentro de cada método de ensino, podem ter um impacto significativo na eficácia real da aprendizagem. Várias características são essenciais na experiência de simulação, entre as quais estão a natureza e a qualidade do feedback, a fidelidade da simulação, o tipo de simulador, o cenário e a composição da equipe, o tempo e periodicidade da simulação e a presença ou não de *debriefing* ao final da sessão.

Tipo de simulação	Descrição	Exemplo(s)
Simulação para treinamento de habilidades ou tarefas parciais	Simulador sintético que replica componentes específicos de um paciente ou sistema para treinamento de habilidades possibilitando que a mesma técnica ou procedimento seja repetido diversas vezes, desenvolvendo competências de menor nível de complexidade.	Simuladores de partes do corpo, como um braço (para técnica de punção venosa) ou um tórax SAM II ¹
Simulação clínica com uso de simulador de pacientes	Uma das formas mais utilizadas, pode ser realizada com simuladores de baixa ou média tecnologia ou com simuladores de alta tecnologia. A construção de casos clínicos contribui para essa estratégia de simulação, seja ela de baixa, média ou alta tecnologia.	Manequim de simulação de partos Noelle ² , simulador clínico Istan ³
A simulação clínica com paciente simulado (humano)	Utiliza pessoas para representar sintomas ou problemas clínicos. Obrigatório ter um caso clínico que deverá ser interpretado pelo participante simulado, seguindo um roteiro de encenação (script) e favorecendo uma maior interação durante a aplicação do cenário. As pessoas podem representar um escopo expandido de funções (por exemplo, clientes, familiares, profissionais de saúde).	Exame Clínico Objetivo Estruturado (OSCE)
O <i>role-play</i> , ou “troca de papéis”	Estudantes são convidados a assumir o papel de outras pessoas por meio de dramatização, com objetivo de compreenderem um fenômeno partindo de uma perspectiva diferente da sua. Esse recurso é utilizado na formação de diferentes profissionais da saúde para sensibilizar estudantes em relação a uma temática ou situação.	Colega que simula o papel de paciente
Simulação com material orgânico	Utilização de material orgânico para treinamento de habilidades, geralmente utilizado para treinamento de habilidades específicas.	Sutura em estômago bovino ou traqueostomia em cães
Simulação híbrida	Combinação de mais de uma modalidade de simulação em um único treinamento, como a associação de um paciente simulado com um simulador de qualquer nível de tecnologia. Essa estratégia permite o desenvolvimento de habilidades processuais e de comunicação com uma pessoa, trazendo uma sensação de realismo à atividade que pode não ser alcançada usando atores ou simuladores isoladamente.	
Realidade ou simulação virtual (mundo virtual)	Envolve a criação de cenários de simulação baseados em técnica computacional que permite ao aluno mergulhar, através de tela ou óculos de RV na recriação de um ambiente ou cenário. Geralmente interage com o meio digital por meio de uma “persona digital” ou avatar cujo participante exerce um papel central no cumprimento de tarefas específicas e no desenvolvimento de habilidades de interação, tomada de decisão e comunicação no atendimento a pacientes virtuais criados, a partir de uma variedade de configurações clínicas.	Metahealth: Plataforma de simulação de casos clínicos com uso de óculos de RV ⁴ , Muiiraquitã: Jogo com simulação de visita domiciliar ⁵ , Gamificação ALICE (Chon et al., 2018), VOSCE (Virtual OSCE)

¹ <https://laerdal.com/>

² <https://gaumard.mogiglass.com.br/obstetricos/noelle-s551-250-simulador-de-paciente-para-parto-com-omni-2/>

³ <http://www.bertas.com.tr/patient-simulation-istan.48.en.html>

⁴ <https://www.inf.ufrgs.br/metahealth/>

⁵ <https://unasus.ufcspa.edu.br/cidadesvirtuais/NovaMqta/story.html5.html>

Tabela 2.1. Descrição e exemplos de modalidades de simulação.

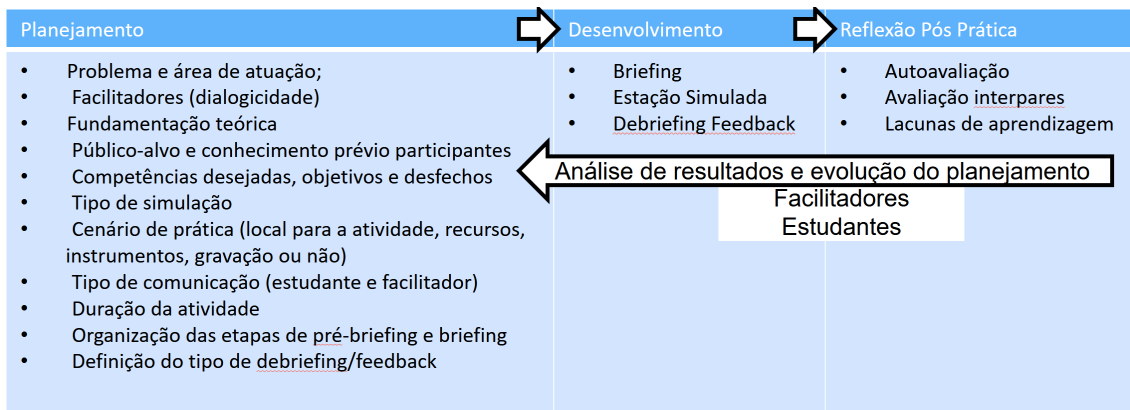


Figura 2.7. Simulação clínica como método de ensino. Adaptada de Júnior and Guedes (2022)

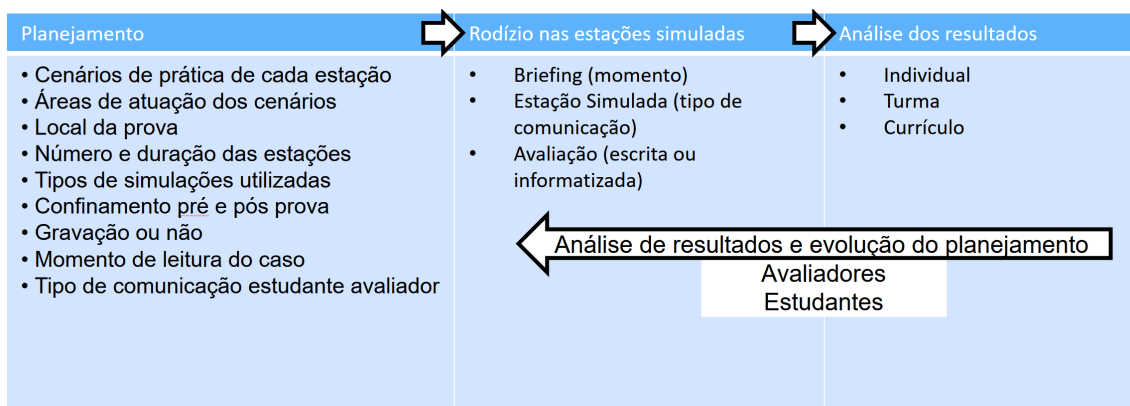


Figura 2.8. Simulação clínica como método de avaliação. Adaptada de Júnior and Guedes (2022)

Como já mencionado, a simulação clínica é uma potencial estratégia tanto de ensino como de avaliação pois permite conquistar e validar competências clínicas, além de otimizar métodos de avaliação na formação profissional. Esta abordagem também desenvolve atributos interpessoais, capacidade de raciocínio clínico, tomada de decisão, habilidades técnicas, entre outros aspectos (Montiel et al., 2012). Porém, existem diferenças no processo de elaboração e desenvolvimento das estações simuladas que precisam ser conhecidas, conforme mostram as Figuras 2.7 e 2.8.

Vale a pena lembrar que no *briefing* há definição de quais informações serão oferecidas aos participantes para a atividade como sinais, sintomas, história médica progressiva, queixa clínica, alterações no exame físico, além de particularidades do paciente como personalidade (calmo, agressivo, prolixo), nível de escolaridade, personalidade, etc. Parâmetros e características que podem facilitar ou dificultar a anamnese, diagnóstico e tratamento de acordo com o planejamento do caso também devem ser definidos no *briefing*. Em resumo, no *briefing* devem ser repassadas todas as orientações específicas quanto ao cenário simulado que será desenvolvido, com a apresentação do problema e dos passos relativos à tarefa a ser realizada, regras e limites.

Já o *debriefing* é uma fase planejada e voltada para a promoção do pensamento

reflexivo e aperfeiçoamento do desempenho do participante. É um momento após aplicação da simulação, no qual se faz uma retomada da competência que foi adquirida, ou não, de forma construtiva e não crítica. Após o desenvolvimento da atividade ainda podemos fazer a realização de autoavaliação, avaliação interpares e identificação de lacunas para o aprimoramento da utilização da simulação num momento futuro.

O uso de simulação no ensino em saúde não é algo novo e tem evoluído ao longo do tempo com a incorporação de novas tecnologias, sendo inclusive recomendado pela OMS. Atualmente, a simulação é amplamente utilizada no ensino na área da saúde, tanto para o treinamento de *hard skills* quanto de *soft skills*. As seções que seguem abordarão o uso de simuladores nestas duas categorias, bem como indicarão o seu uso comercial.

2.4. Simuladores para ensino e treinamento de *hard skills*

Simuladores para ensino e treinamento de *hard skills* na área da saúde concentram-se nas habilidades técnicas necessárias para executar procedimentos. Pela ótica da cartografia do trabalho vivo de Merhy (2005), esta seção concentra-se nas tecnologias duras para os cuidados em saúde. As tecnologias leves e leve-duras, são abordadas na Seção 2.5.

Neste contexto, as técnicas cirúrgicas são as mais exploradas (Lungu et al., 2021). Simulações convincentes do teatro de operações visam melhorar as habilidades psicomotoras, como coordenação entre a visão e as habilidades manuais, orientação espacial e manipulações, e têm como alvo principalmente cirurgias minimamente invasivas, como laparoscopia (Grantcharov et al., 2004, Huber et al., 2018b) e colonoscopia (Wen et al., 2018). Nesses cenários, elementos como o realismo da sala de operações e dos comportamentos do corpo humano simulado são usados como fundamentos para aguçar a percepção visual do treinando, automatizar sua destreza motora e refinar o seu processo decisório.

Simuladores virtuais para treinamento de procedimentos cirúrgicos surgiram naturalmente num ambiente onde outros tipos de procedimentos complexos já vinham sendo treinados com simuladores (Satava and Jones, 1997). O exemplo mais emblemático é o dos simuladores de voo, nascidos ainda nos anos 1970, e seguido de simuladores para treinar a operação de diversos tipos de veículos e máquinas. A tecnologia de computação gráfica necessária para reproduzir virtualmente um espaço aéreo, uma estrada e juntas de rotação é muito mais simples, no entanto, do que a que é necessária para reproduzir virtualmente um organismo vivo. Além disso, a possibilidade de implementar um *cockpit* e controles físicos para os simuladores de veículos se torna inexistente quando o objeto é composto de tecidos moles e varia significativamente consoante as individualidades e patologias.

Fica evidente, portanto, que da gama de ambientes de cuidados médicos, o ambiente cirúrgico é mais controlável e delimitável para uma implementação simulada do que consultas ou exames físicos. É notável, ainda, que operações fortemente centradas em um mediador físico, como é o caso das laparoscopias, são menos complexas de simular (Basdogan et al., 2007). O uso de longos instrumentos rígidos coloca um distanciamento confortável para que algoritmos transformem razoavelmente bem as ações realizadas sobre um comando físico em ações sobre um único órgão simulado. Da mesma forma, procedimentos baseados em câmera, como as endoscopias, onde os tecidos não são vi-

sualizados diretamente pelo olho do cirurgião, mas sim através de um monitor de vídeo, tornam o problema do *rendering* gráfico dos tecidos mais tratável (ver Figura 2.9). Em oposição, se houvesse interações diretas das mãos do cirurgião com o tecido vivo, o desafio para computar interações precisas seria imenso.



Figura 2.9. Exemplos de simuladores para cirurgias minimamente invasivas.

Pela soma desses fatores, simuladores cirúrgicos cada vez mais efetivos foram surgindo ao longo das últimas 3 décadas, acompanhando o desenvolvimento da tecnologia de computação gráfica e de realidade virtual (Hong et al., 2021, Liu et al., 2003). Os diversos componentes necessários na construção de um simulador cirúrgico, como modelos dos tecidos moles (Cotin et al., 1999) – que precisam ter as características dos tecidos reais (Zhang et al., 2018) – foram os primeiros a receberem soluções aceitáveis. Em seguida, são necessários também, algoritmos para detectar e resolver os contatos entre os tecidos e os instrumentos (Maciel and De, 2008, Raghupathi et al., 2004), e *displays*, tanto para o retorno visual (Maciel et al., 2008) quanto para o retorno tátil (Burdea, 1996). Por fim, ainda era necessário desenvolver programas de atividades (Kim-Fine and Brennan, 2016, Tay et al., 2014) que conduzissem o *trainee* no percurso gradual por todas as componentes de habilidades a serem adquiridas combinado com um mecanismo para avaliar a aquisição de competências (Ribeiro et al., 2018).

Simuladores cirúrgicos comerciais foram desenvolvidos e adotados por todo o

mundo (mais detalhes na Seção 2.6). Apenas mais recentemente, entretanto, surgiram iniciativas relevantes que aproveitam a tecnologia imersiva para treinar procedimentos médicos Mao et al. (2021). Esta transição para o imersivo só se tornou possível com o salto qualitativo dos *displays* de realidade virtual que se iniciou há cerca de 10 anos com os Oculus Rift. Era preciso um *display* montado na cabeça (HMD) com baixa latência, amplo campo de visão, não muito desconfortável de usar, e rastreamento em tempo real dos movimentos da cabeça e das mãos. Hoje, isso está disponível e o desenvolvimento mudou o foco para ambientes e interações adequadas ao treino de outras atividades de cuidados à saúde diferentes da cirurgia.

Um simulador imersivo precisa de todos os elementos tecnológicos dos simuladores cirúrgicos e mais alguns, como técnicas de navegação nos espaços virtuais, comunicação e colaboração com outros atores como assistentes, enfermeiros, colegas, pacientes, uso extensivo de gestos e linguagem não-verbal, entre outros. Um ponto que representa ainda uma limitação tecnológica é o retorno tátil (Rangarajan et al., 2020). Enquanto o uso de retorno tátil tem mostrado um efeito positivo na fidelidade do treino em RV, especialmente em simulação de cirurgia, os sistemas atuais são extremamente custosos e se encontram em estágio inicial de maturidade. Ainda requerem muita pesquisa antes de se poder confirmar como um elemento imprescindível. Isso não impede que outras áreas se desenvolvam, no entanto.

O fator mais significativo que difere o treinamento com simuladores não-imersivos dos imersivos é a elicitación da sensação de presença. Um usuário imerso tende a desconectar-se mentalmente do mundo real e é capaz de focar muito mais numa tarefa virtual por sentir que de fato se encontra naquele local realizando aquela tarefa ao invés de sentir como se assistisse a um filme numa tela. É como se estivesse realizando a ação real e o aprendizado, memorização, retenção da habilidade esperados é comparável ao de realizar a prática real da atividade (Ekstrand et al., 2018, Siu et al., 2016)

Um aspecto interessante que pode ser alcançado com simulações de RV é relatado por Ganni et al. (2020). Os autores afirmam que, diferentemente dos simuladores de cirurgia convencionais baseados em uma única estação dedicada e sem distrações, os simuladores de RV permitem replicar o ambiente movimentado e muitas vezes caótico da sala de cirurgia. Um trabalho anterior (Sankaranarayanan et al., 2016) já havia avaliado o efeito de interrupções e distrações no desempenho do usuário em (1) um cenário simulado simples apresentado na tela de um computador, (2) um usuário interagindo com o cenário simulado exibido em um HMD, sem distrações e interrupções, e um usuário interagindo com o cenário simulado, com distrações e interrupções.

O realismo visual do cenário da sala cirúrgica também pode influenciar no desempenho e na experiência do usuário em procedimentos laparoscópicos virtuais. Huber et al. (2018a) apresentam um estudo piloto comparando o desempenho e as preferências dos usuários em um cenário imersivo baseado em RV com duas condições: uma sala cirúrgica virtual e uma sala cirúrgica baseada em vídeo 360°. Num outro estudo, o procedimento técnico de inserção de um catéter Port-o-cath para quimioterapia foi avaliado com 30 estudantes de um programa de enfermagem (Jung and Park, 2022). Um efeito muito significativo no conhecimento pós intervenção ($p = 0.001$) foi detectado, entre outros efeitos, na atenção, satisfação e motivação em comparação com um grupo controle.

Cabe notar que apesar dos bons resultados, o simulador utilizado não permitia interações complexas, apenas visualização imersiva (ver Figura 2.10).

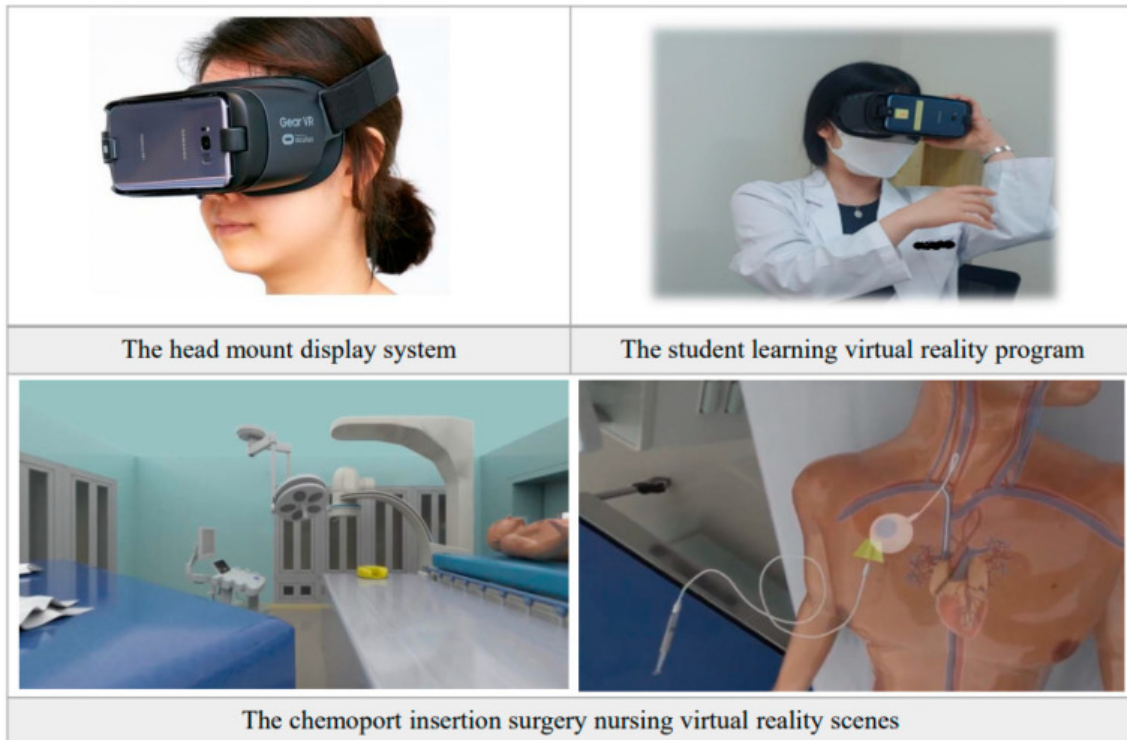


Figura 2.10. Simulador imersivo de inserção de catéter Port-o-cath. Fonte: Jung and Park (2022).

Dispositivos hápticos também podem ser usados em cenários imersivos. Alaraj et al. (2015) apresentaram um protótipo de simulação de procedimentos em aneurisma usando a plataforma Immersive Touch, que contém módulos de treinamento em RV com feedback háptico de força contínuo, incluindo ventriculoplastia, perfuração óssea, procedimentos na coluna vertebral, e punção lombar. Nesse trabalho, o ITACS (Immersive Touch Aneurysm Clipping Simulation) foi acrescido de novos módulos para simular a deformação do volume do aneurisma e vasos sanguíneos e a clipagem do aneurisma (ver Figura 2.11). Outro exemplo utilizando a plataforma Immersive Touch é relatado por Lemole Jr et al. (2007), descrevendo seu uso em simulação da tarefa de colocação de cateter em ventriculostomia como prova de conceito.

Também encontramos diversos dispositivos hápticos com propósitos específicos. Sutherland et al. (2013) apresentam um protótipo de simulação para treinamento de inserção de agulha espinhal utilizando realidade aumentada (ver Figura 2.12). Com o uso do rastreamento óptico MicronTracker2, o sistema permite que os usuários realizem inserções simuladas de agulha com o dispositivo háptico PHANToM em um manequim físico.

Libertados da estação de trabalho, os simuladores imersivos alargaram seu escopo para outros cenários de cuidados à saúde, como cenários de clínica médica, emergência, etc. Casos para os quais normalmente se usam manequins, podem ser implementados nesses ambientes virtuais imersivos, ampliando o acesso a mais estudantes e reduzindo



Figura 2.11. Immersive Touch Aneurysm Clipping Simulation. Fonte: Alaraj et al. (2015).

os custos e a complexidade das sessões. Uma revisão sistemática aponta que em breve o foco deve se dirigir também para estudos a longo prazo capazes de examinar habilidades não-técnicas e multidisciplinaridade das equipes (Pfandler et al., 2017).

É possível também simular uma conversa com um paciente virtual que expõe suas dores usando linguagem natural tanto oral como corporal, gestos, etc. É possível realizar exames físicos no paciente virtual que executará reações programadas de acordo com o caso para prática de diagnóstico e prescrição de tratamento. Com o avanço nessa área, o treino de habilidades duras passa a interagir sinergicamente com o treino de habilidades comportamentais mais ligadas às tecnologias leves.

A RV atinge, portanto, um grau elevado de importância ao conceder horas de voo a novos médicos sem expor pacientes reais a profissionais inexperientes. Aumenta as possibilidades de praticar também para profissionais experientes ao simular casos raros que estes não teriam chance de presenciar frequentemente. Por fim, melhora a qualidade do serviço prestado à população ao auxiliar na formação de mais e melhores profissionais. Há um consenso entre os autores (Bjerrum et al., 2018, Stefanidis et al., 2015) de que o treinamento em simuladores virtuais deve ser parte do currículo, incluindo habilidades técnicas, não técnicas e teóricas, e deve ser fundamentado nas avaliações das necessidades mais relevantes. Esta atividade deve seguir os princípios baseados em evidências para medir os efeitos e otimizar os treinamentos.

Há ainda um grande potencial a ser realizado para o treinamento virtual imersivo à medida que os simuladores evoluem, que mais programas baseados em evidências sejam implementados e que a relação custo-benefício e o impacto na segurança dos pacientes

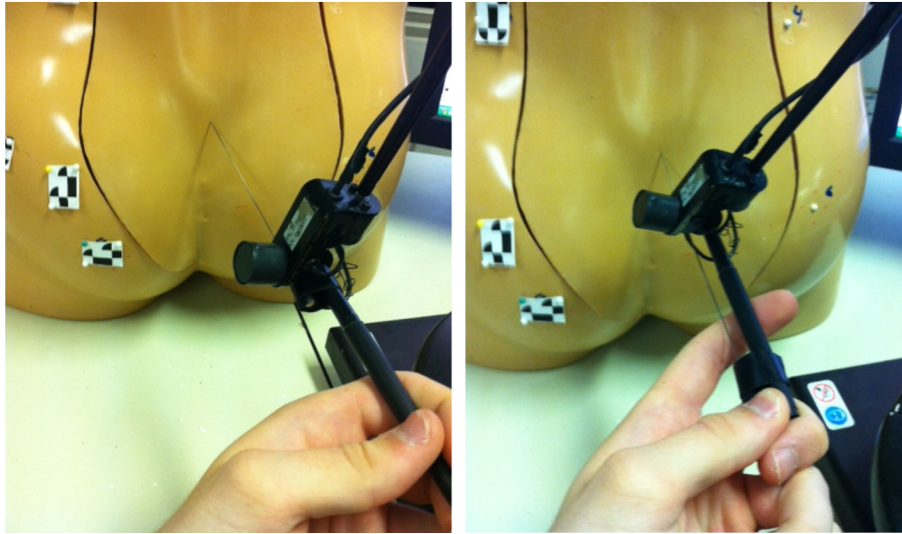


Figura 2.12. Aparato háptico para treinamento de inserção de agulha especial.
 Fonte: Sutherland et al. (2013).

sejam claramente demonstrados (Rizzetto et al., 2020).

2.5. Simuladores para ensino e treinamento de *soft skills*

Diversos procedimentos médicos não cirúrgicos têm se beneficiado das vantagens proporcionadas pelos simuladores em RV (Bianchi et al., 2020, Gerup et al., 2020, Reznik et al., 2002). Esses procedimentos não cirúrgicos são geralmente traduzidos como encontros paciente-profissional de saúde, onde a empatia e as habilidades de comunicação desempenham um papel importante para proporcionar uma boa experiência ao paciente (Caldwell, 2019). Tais habilidades envolvem o uso das tecnologias conhecidas como “leves” ou “leve-duras”.

Estima-se que os profissionais de saúde empreguem de 85% a 90% do seu tempo de trabalho se comunicando. Em especial, a qualidade da comunicação entre profissionais de saúde e pacientes tem um impacto direto na adesão ao tratamento, no prognóstico, na satisfação e na redução de denúncias e processos contra profissionais de saúde (Stewart et al., 2017).

Diante da importância da comunicação na saúde, o treinamento de habilidades de comunicação, conhecidas como *soft skills*, é fundamental para melhorar a qualidade da atenção e do cuidado das pessoas. Essas habilidades incluem a capacidade de ouvir ativamente, mostrar empatia, comunicar-se de forma clara e concisa, e resolver conflitos de forma eficaz.

Nos últimos anos, os simuladores em realidade virtual (RV) têm surgido como uma ferramenta útil para o treinamento e aprimoramento de habilidades de comunicação (Chaby et al., 2022). Esse tipo de treinamento já vem sendo amplamente utilizado na medicina, enfermagem e odontologia, e tem se mostrado eficaz para melhorar a qualidade do atendimento prestado aos pacientes e fortalecer a relação entre profissionais e paciente (Jimenez, 2022, Kleinheksel, 2014, Patel et al., 2023). A Figura 2.13 exemplifica duas situações com personagens virtuais em ambiente médico.



Figura 2.13. Personagens virtuais utilizados nas simulações de Kleinheksel (2014), à esquerda e de Patel et al. (2023), à direita.

A falha na comunicação e a ausência de empatia podem ter origem em vieses e preconceitos. Simuladores de RV podem ser usados para expor os alunos a situações que podem desencadear esses sentimentos. Altmiller et al. (2022) relatam um estudo com estudantes de enfermagem onde eles tiveram que tratar um paciente virtual com vírus da imunodeficiência humana (HIV). No estudo, ao interagir com o paciente virtual, os alunos tiveram a oportunidade de compreender as necessidades médicas desses pacientes e reconhecer o impacto dos cuidados de enfermagem em suas vidas.

Segundo a OMS, uma grande proporção dos eventos adversos que ocorrem nos serviços de saúde em todo o mundo são devido a falhas de comunicação (WHO) e isso não envolve apenas a comunicação profissional de saúde e paciente. Habilidades de comunicação eficiente são necessárias também entre os membros da equipe na rotina clínica diária.

A educação médica/enfermagem convencional tem contado com uma ferramenta conhecida como SBAR (*Situation, Background, Assessment, and Recommendation*, ou seja, situação, antecedentes, avaliação, recomendação) e seus derivados, que são amplamente utilizados em diferentes unidades de saúde como comunicação interprofissional e transferência de pacientes. A Tabela 2.2 apresenta uma breve descrição e um exemplo da ferramenta. Embora não haja fortes evidências de que tal protocolo de comunicação melhore a segurança do paciente (Müller et al., 2018), simuladores de RV recentes proporcionaram aos alunos oportunidades de desenvolver suas habilidades de comunicação usando o SBAR (Streuber et al., 2020, Stuart et al., 2021).

Para que as simulações sejam eficazes no treinamento de profissionais da saúde, os pacientes virtuais devem simular de forma realista as interações com pacientes reais. Isso inclui a comunicação verbal e não verbal, como expressões faciais, tom de voz e respostas verbais coerentes. Além disso, os pacientes virtuais devem reagir de forma adequada às ações do profissional, fornecendo *feedback* positivo ou negativo e adaptando-se às intervenções. Para tornar a simulação ainda mais realista, é importante que cada paciente virtual tenha um histórico médico detalhado, que haja diversidade de personalidades e que os casos simulados abranjam uma variedade de situações clínicas (Consorti et al., 2012, Cook and Triola, 2009).

Além de pacientes virtuais com comunicação realística, outro ponto importante

	Questões	Descrição	Exemplo
Situação	O que está acontecendo com o paciente? Qual a situação que está sendo comunicada?	Primeiro, o locutor apresenta a situação, identificando-se, informando o nome do paciente e descrevendo brevemente o problema.	Dr. Preston, estou ligando para falar do Sr. Lakewood, que está com dificuldade para respirar.
Antecedentes	Qual o histórico ou o contexto do paciente?	O locutor então fornece os antecedentes, como o diagnóstico do paciente ou o motivo da internação, o estado médico e o histórico relevante. O prontuário do paciente é revisado e as dúvidas que o outro prestador de cuidados possa ter são antecipadas.	Ele é um homem de 54 anos com doença pulmonar crônica que está piorando e agora está gravemente pior.
Avaliação	Qual o problema?	Em seguida, são fornecidas informações específicas sobre sinais vitais, exames de laboratórios recentes e outros dados quantitativos ou qualitativos relacionados ao estado atual do paciente. Esta sessão pode incluir um diagnóstico provisório ou uma impressão clínica.	Não ouço nenhum som respiratório no seu peito direito. Eu acho que ele tem um pneumotórax.
Recomendação	Qual é o próximo passo no gerenciamento do paciente?	Uma sugestão informada para o cuidado continuado do paciente deve ser feita pelo locutor. A necessidade imediata é explicada de forma clara e específica, incluindo o que é necessário para resolver o problema.	Preciso que você o veja agora. Acho que ele precisa de um dreno torácico.

Tabela 2.2. Descrição da ferramenta de comunicação SBAR. Traduzida livremente de Müller et al. (2018).

que deve ser levado em consideração no desenvolvimento de simuladores de *soft skills* para treinamento médico é o ambiente. Situações de alto nível de estresse, com tempo restrito, onde a tomada de decisão tem que ser feita rápida e assertivamente fazem parte do cotidiano de cuidado em saúde. A importância de preparar os estudantes para lidar com tais cenários é destacada por Prashar et al. (2022), que enfatiza que a prática sob condições simuladas pode ajudar a desenvolver as habilidades necessárias para enfrentar esses desafios na vida real.

Para garantir que o treinamento seja o mais similar possível à realidade, é fundamental que o simulador recrie fielmente o ambiente de clínica ou hospitalar. Isso inclui não apenas a aparência física do ambiente, mas também os sons, a interferência na comunicação por causa de ruídos típicos e até mesmo as emoções que estão presentes nesses ambientes. Segundo Hafferty (1998), a inclusão do “currículo oculto” – habilidades de comunicação, resiliência, empatia e compaixão – no treinamento simulado é crucial para formar médicos capazes de tomar decisões centradas no paciente em situações de alta pressão.

Os simuladores dedicados ao treinamento de *soft skills* também devem ser capazes de simular as diferentes situações de estresse que os médicos podem enfrentar no dia a

dia. Isso pode incluir situações como pacientes agressivos ou familiares desesperados. O simulador deve permitir que o médico pratique como lidar com essas situações de forma profissional e sensível. O estudo de Prashar et al. (2022) reforça a necessidade de praticar essas habilidades em um ambiente simulado, pois a exposição reduzida a situações clínicas reais pode impactar negativamente no desenvolvimento das habilidades clínicas e práticas dos estudantes.

Kleven et al. (2014), em seu artigo, abordam a importância das habilidades interpessoais na educação e na prática da saúde. O artigo concentra-se no desenvolvimento de habilidades de comunicação, colaboração, empatia e interação com pacientes entre profissionais de saúde e participantes não médicos. Por meio de cenários de *role-playing* e simulações interativas em uma sala de cirurgia virtual, os participantes têm a oportunidade de praticar e aprimorar suas habilidades interpessoais em um ambiente imersivo. O estudo visa melhorar as habilidades dos participantes para se comunicarem efetivamente com os pacientes e trabalhar de forma colaborativa dentro de uma equipe de saúde. A Figura~2.14 ilustra uma das cenas usadas neste simulador.



Figura 2.14. Ambiente e personagens virtuais utilizados nas simulações de Kleven et al. (2014).

Em suas investigações, Lohre et al. (2020) exploraram a eficácia do treinamento em realidade virtual imersiva na melhoria da aquisição de habilidades não técnicas. Os resultados indicaram que o grupo que utilizou a realidade virtual imersiva completou as tarefas consideravelmente mais rápido em comparação com o grupo de controle. Impressionantemente, houve uma diminuição de 570% no tempo de aprendizagem. O treinamento em realidade virtual imersiva demonstrou uma melhoria substancial na aquisição de habilidades não técnicas, sugerindo o seu potencial como uma ferramenta educativa eficaz.

Embora ambos tipos de simuladores, tanto os para treinamento de *hard skills*, quanto para treinamento de *soft skills* sejam importantes para a formação médica, um ambiente real não simulado inclui habilidades dos dois tipos. O MetaHealth, projeto em desenvolvimento pelos autores desse capítulo, aborda a experiência completa de trei-



Figura 2.15. Paciente durante a anamnese (esquerda) e visualização da pia e dos instrumentos médicos disponíveis (direita).

namento em saúde, incluindo os principais pontos técnicos da realização de exames e pensamento clínico, como a realização da anamnese e condução da consulta (*soft skills*).

Existem diversos aspectos científicos por trás do uso de simuladores para aquisição, manutenção e avaliação de habilidades. Porém, cada novo simulador tem particularidades que podem necessitar ser avaliadas experimentalmente. O MetaHealth é pautado por verificar cientificamente os efeitos do seu uso antes de disponibilizar um simulador.

Um desses casos ocorreu no desenvolvimento de uma estação OSCE envolvendo uma clínica pediátrica simulada. Nesse caso foram apresentadas tarefas normalmente exigidas durante uma consulta médica, o que ajuda os alunos a treinar habilidades básicas antes do contato com pacientes reais. Tais tarefas incluem procedimentos de biossegurança e anamnese por meio da coleta de detalhes sobre a doença, sinais, sintomas e histórico médico do paciente. O usuário também devia ser capaz de realizar exames físicos, interpretar os achados e compilar todas as informações para determinar o diagnóstico correto e o tratamento adequado (ver Figura 2.15).



Figura 2.16. Ajuste de percepção visual e dessensibilização ao toque manual realizados no hall de entrada.

Apesar de ser um projeto ainda em desenvolvimento, até o momento foram realizadas várias sessões com usuários para medir a eficácia e a experiência de uso do

simulador (Negrão et al., 2024). No princípio, o usuário deve investigar corretamente o histórico médico do paciente, os sintomas que ele apresenta e a queixa principal através das questões da anamnese, aspectos que estão englobados nas *soft skills*. Em seguida, com base na anamnese, deverá realizar os exames necessários para confirmar os achados. As duas últimas metas são o diagnóstico correto da doença da criança e o tratamento recomendado. As informações do caso e do paciente apresentadas em nosso cenário de estudo foram desenvolvidas e revisadas por um pediatra.

Um dos aspectos mais importantes abordado no projeto MetaHealth é a adaptação do usuário ao simulador e aos ambientes virtuais. Antes de ser capaz de aplicar seus conhecimentos em saúde e tomar decisões com precisão, o usuário deve se sentir adaptado ao ambiente virtual e confiante ao utilizar os controles do aplicativo para navegar dentro da clínica virtual. A interface deve ser transparente, não uma barreira.



Figura 2.17. Ambiente virtual inicial do MetaHealth, onde ocorre a dessensibilização.

Portanto, o curso de ação esperado no simulador envolve primeiro uma dessensibilização no ambiente de realidade virtual, para a seguir partir para um treinamento ou uma avaliação de habilidades clínicas. A fase de dessensibilização é proposta para auxiliar o usuário no ajuste do óculos de VR, treinar os primeiros passos com os controladores e as interações virtuais, como tocar, agarrar, segurar e suas combinações (ver Figura 2.16 e 2.17). O usuário também realiza outras tarefas de adaptação, como testar as técnicas de navegação, por exemplo. Por fim, um usuário já dessensibilizado não precisa voltar a dessensibilizar antes da próxima sessão.

Embora a combinação do treinamento e aprendizado de *soft skills* e *hard skills* seja completamente benéfica ao aluno, elencar e padronizar os principais dados para avaliação ainda é um desafio, como tangibilizar as habilidades não-técnicas de forma semelhante as habilidades técnicas para criar um retorno de aprendizagem e avaliação justo e padronizado.

2.6. Aplicações comerciais de simulação imersiva em saúde

No mercado de simulação, também encontramos empresas que disponibilizam simuladores para o ensino e treinamento de profissionais de saúde, como a Laerdal⁴, uma fornecedora global de sistemas de simulação física e virtual para médicos e enfermeiros. A empresa é líder global de mercado, contando com uma ampla variedade de simuladores físicos e estações de treinamento práticos. Dentre os produtos oferecidos estão manequins de alta fidelidade para procedimentos como: reanimação cardiopulmonar, intubação, situações de emergência, etc. A Figura 2.18 apresenta o SimMan, um dos manequins disponíveis no leque da Laerdal. Seu objetivo é simular uma ampla gama de cenários realísticos em trauma e emergência, incluindo tomada de decisão em situações críticas e em simulações práticas, como desfibrilação.



Figura 2.18. Manequim de Simulação SimMan® 3G Plus⁶

Além dos simuladores físicos, a empresa também investe em soluções imersivas para treinamento, oferecendo cenários para prática de cuidados na enfermagem com o produto vSim, que teve sua interface inicialmente desenvolvida em 3D tradicional e, posteriormente, a adaptação de alguns componentes para 3D imersivo. Recentemente, no ano de 2023, a Laerdal passou a oferecer outra variedade de simuladores imersivos, através de uma parceria com a SimX⁷, uma empresa voltada completamente para treinamentos imersivos para saúde.

A SimX⁸ oferece uma das maiores plataformas de simulação em realidade virtual presentes no mercado, abrangendo em seu currículo não só simulações em cenários de emergência, mas também no cuidado do dia-a-dia do paciente. Embora o foco seja, inicialmente, em treinamentos de enfermagem, as soluções apresentadas não se limitam apenas a essa área, incluindo opções que exploram as habilidades de comunicação, tomada

⁴<https://laerdal.com/us/>

⁶<https://laerdal.com/br/products/simulation-training/emergency-care-trauma/simman-3g/>

⁷<https://laerdal.com/br/products/courses-learning/virtual-simulation/simx/>

⁸<https://www.simxvr.com/>

de decisão e pensamento clínico do usuário. Um dos principais diferenciais ofertados pela empresa é a personalização dos pacientes virtuais, como apresentado na Figura 2.19. A plataforma proporciona diferentes perfis e condições de saúde, bem como a personalização do diálogo, sinais vitais e ferramentas do ambiente médico virtual.

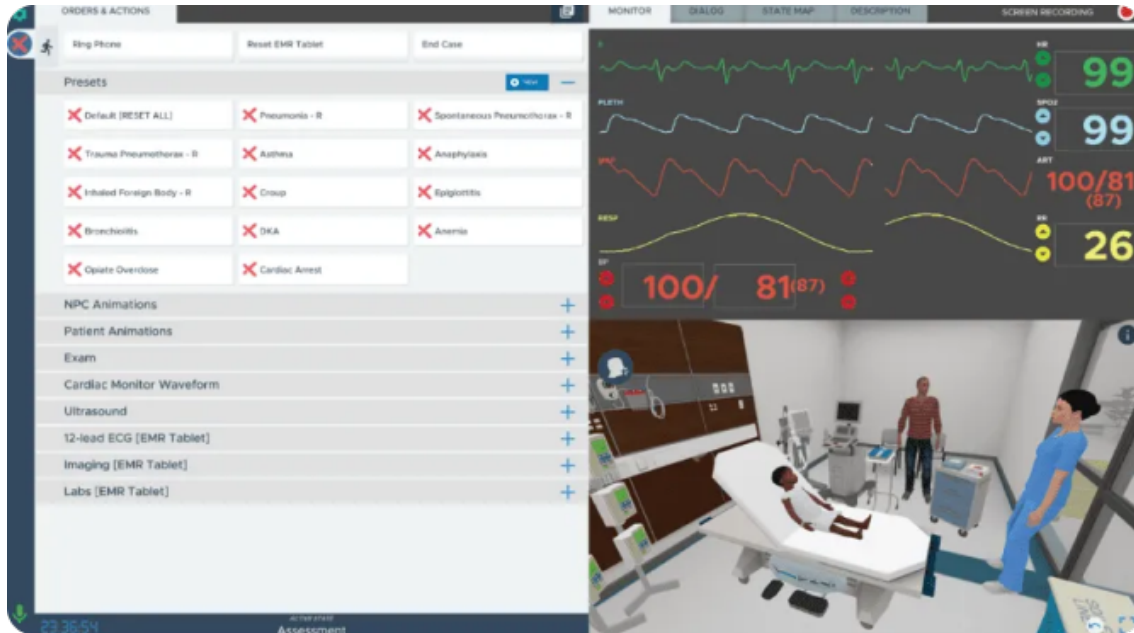


Figura 2.19. Personalização do paciente virtual na plataforma da SimX¹⁰

Através de parcerias firmadas com entidades acadêmicas, como as Universidades de Delaware e do Estado do Arizona, editoras acadêmicas, como a Elsevier e centros de simulação médicas na América do Norte, a SimX torna viável a comercialização de diferentes pacotes de simulações médicas. Na plataforma de *marketplace* (ver Figura 2.20), gerada através dessas parcerias, é possível encontrar simulações desenvolvidas pelas equipes parceiras e filtrar pelos cenários virtuais (e.g. Ambulância, Casa, Hospital), perfil do paciente (e.g. Neonatal, Adulto, Mulher), especialidade médica (e.g. Cardiologia, Neurologia, Psiquiatria), e público alvo (e.g. Residente, Estudante, Enfermeiro). Como uma solução completa que se apresenta, o SimX presa pelo conteúdo apresentado e pela personalização dos casos, além do feedback apresentado. O último, por sua vez, auxilia na identificação dos erros cometidos nos treinamentos, melhorando o aprendizado dos usuários.

No âmbito de simulações imersivas focadas em especialidades e procedimentos complexos, encontramos soluções como as da PrecisionOS¹³, que tem seu foco em procedimentos ortopédicos. Através de um design modular, a empresa tem foco no treinamento e na precisão de habilidades técnicas para cirurgiões ortopédicos. A plataforma permite a prática em um ambiente virtual seguro de sala de cirurgia, com feedback em tempo real. Na mesma especialidade, podemos encontrar as soluções da OssoVR¹⁴, que oferece

¹⁰<https://www.simxvr.com/customizable-medical-simulation-solutions/>

¹²<https://www.simxvr.com/marketplace/>

¹³<https://www.precisionostech.com/>

¹⁴<https://www.ossovr.com/>

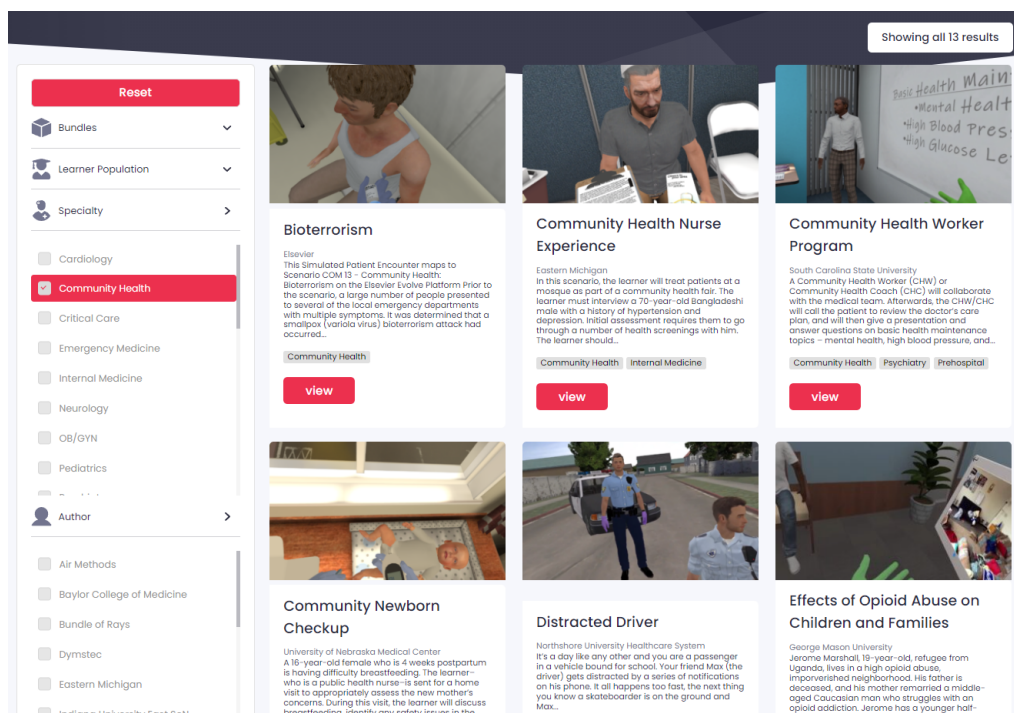


Figura 2.20. Simulações oferecidas no Marketplace da SimX¹²

simulações voltadas para o treinamento ortopédico, incluindo procedimentos diferentes e módulos de treinamento, como trauma, pediatria, coluna vertebral, etc. Na Figura 2.21, é demonstrado um dos diferenciais apresentados pela empresa, a possibilidade de simulação de procedimentos em equipe nos ambientes imersivos em realidade virtual.

Outras iniciativas estão presentes no mercado, sendo importante ressaltar algumas bem consolidadas como a Oxford Medical Simulation¹⁵, da própria Universidade de Oxford, que foca em simulações em escala para instituições acadêmicas e sistemas de saúde, oferecendo treinamentos clínicos e procedimentos médicos como um todo. Também a HealthScholars¹⁶ é uma referência e oferece treinamentos modularizados em cursos que englobam cenários de cirurgia e emergência, além de simulações de treinamento na maternidade.

A empresa MetaMedicsVR¹⁷ oferece, assim como outras mencionadas anteriormente, treinamentos para enfermagem, cenários de emergência e casos clínicos simulados. Porém, o que se destaca é uma das funcionalidades que integra os manequins físicos de manobra de reanimação cardiopulmonar do mundo real com o ambiente virtual para treinamento. O usuário do simulador aprende não só a realizar o procedimento visualmente no ambiente virtual, como tem também o retorno tátil da sua ação.

A funcionalidade não é exclusiva do produto dessa empresa e também está presente em um dos simuladores comercializados pela Surgiscience¹⁸, fornecedora global

¹⁵<https://oxfordmedicalsimmulation.com/>

¹⁶<https://www.healthscholars.com/>

¹⁷<https://metamedicsvr.com/>

¹⁸<https://surgiscience.com/simulators/all-products-gallery/>



Figura 2.21. Simulador RV de procedimento em equipe da Osso VR¹⁴

de simuladores com retorno tátil para treinamento de procedimentos minimamente invasivos. O TraumaVR, um simulador mais robusto que o apresentado pela MetaMedicsVR, é um simulador imersivo de realidade mista, ou seja, elementos do mundo real são utilizados no ambiente virtual (ver Figura 2.22).

O ambiente imersivo realístico, combinado ao retorno tátil simulado dos procedimentos, com a fisiologia e anatomia das cavidades orais, com retorno tátil real dos instrumentos, oferece ao estudante ainda mais ganho durante os treinamentos, incluindo o manejo correto dos instrumentos e do paciente.

Além das empresas citadas acima, muitas outras também estão presentes no mercado com soluções para treinamento de médicos, enfermeiros e outros profissionais, nas mais diversas áreas da saúde. Não é objetivo deste curso fazer uma varredura exaustiva do mercado. No entanto, dentre elas, podemos citar ainda:

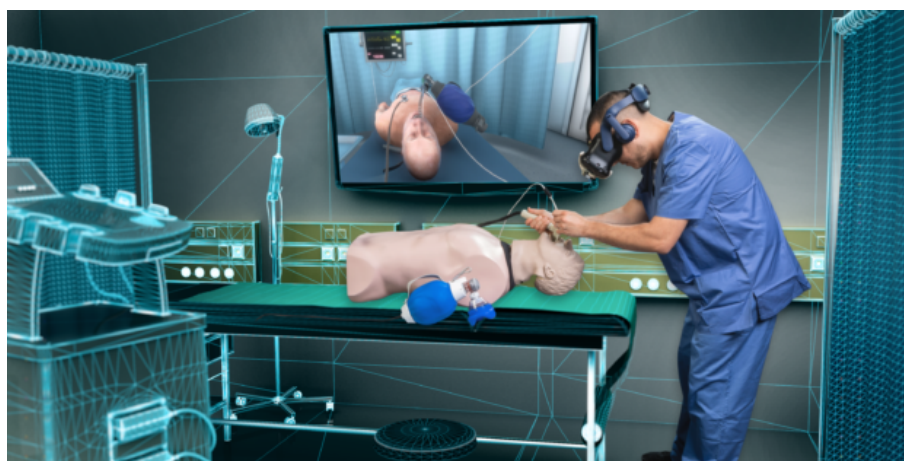


Figura 2.22. Demonstração da aplicação com manequim da MetaMedicsVR¹⁷

- STRIVR¹⁹
- FundamentalVR²⁰
- AppliedVR²¹
- XRHealth²²
- ImmersiveTouch²³

Quando olhamos para o mercado nacional, encontramos iniciativas ainda tímidas em comparação com as apresentadas anteriormente. Uma das empresas que oferece aprendizado em realidade virtual na área médica é a MedRoom, que oferece como seu carro-chefe uma aplicação de aprendizado de anatomia usando experiências imersivas (ver Figura 2.23), além de projetar soluções para treinamentos de casos clínicos.



Figura 2.23. Aplicação Atrium para estudos de anatomia da MedRoom ²⁵

Embora a Laerdal possua atuação no Brasil e esteja presente em centros de simulações mais avançados em universidades, não há informações sobre a comercialização dos simuladores da SimX no país. Por outro lado, *startups* com foco nos simuladores imersivos, com parceria e fomento de órgãos públicos tem ganhado cada vez mais força. Um exemplo é a Simula Health²⁶, que disponibiliza treinamentos através de um hospital virtual onde profissionais e estudantes podem desenvolver e aprimorar habilidades técnicas e comportamentais através de experiências imersivas e desktop.

¹⁹<https://www.strivr.com/use-cases/healthcare>

²⁰<https://fundamentalsurgery.com/>

²¹<https://www.appliedvr.io/>

²²<https://www.xr.health/>

²³<https://www.immersivetouch.com/>

²⁵<https://www.medroom.com.br/atrium-laboratorio>

²⁶<https://simulahealth.com.br>

2.7. Considerações finais e possibilidades futuras

Considerando os resultados positivos obtidos, fica evidente que a realidade virtual se destaca como uma ferramenta eficaz e promissora. Ao oferecer experiências imersivas e interativas, ela proporciona oportunidades únicas de aprendizado prático e aprofundado tanto para estudantes quanto para profissionais da área.

Desde simulações de procedimentos médicos até treinamentos em situações sociais e comportamentais, a realidade virtual está revolucionando a forma como o conhecimento é adquirido e aplicado no contexto da saúde. Com isso, podemos afirmar com confiança que a realidade virtual já se consolidou como uma realidade no ensino em saúde, capacitando os alunos a desenvolver habilidades essenciais de maneira mais envolvente e eficaz.

Cabe ressaltar que a adoção de simulação em saúde emerge como uma prática relativamente nova e que ainda não foi plenamente integrada, sobretudo no que tange ao desenvolvimento de habilidades não técnicas. Mesmo diante de evidências científicas que comprovam sua viabilidade, escalabilidade e eficácia, a implementação dessas tecnologias ainda permanece limitada (Bracq et al., 2019, Zackoff et al., 2023).

Apesar dos avanços tecnológicos e das ofertas cada vez mais diversificadas de simuladores para o treinamento de profissionais de saúde, ainda existem desafios significativos que limitam sua adoção generalizada em larga escala. Estas restrições incluem questões de integração, flexibilidade e validação dos simuladores. Esses obstáculos precisam ser superados para que a utilização de simuladores de realidade virtual se torne mais disseminada e impactante no campo da saúde.

Ainda há um vasto campo para melhorias, especialmente em aspectos técnicos, como a precisão na reprodução de estímulos físicos, o rendering realista, e o desenvolvimento de comportamentos mais realistas em pacientes virtuais.

Testes e experimentos controlados com usuários representativos do público-alvo dos simuladores precisam ser conduzidos a fim de validar as formas de interação e visualização escolhidas, bem como a efetividade dos simuladores em relação ao aprendizado dos estudantes. Ainda que vários estudos neste sentido já tenham sido conduzidos com resultados positivos, o desenvolvimento de novas ferramentas de hardware e software (e.g. novos HMDs, integração de IA e háptica), exige que novos estudos sejam conduzidos continuamente.

Finalmente, ainda que equipamentos de realidade virtual de alta qualidade estejam atualmente disponíveis no mercado a preços razoáveis, a adaptação de estudantes e profissionais ao uso contínuo destes equipamentos ainda não foi plenamente avaliada. Poucos estudos abordam os efeitos do uso de HMDs por longos e recorrentes períodos, em especial envolvendo profissionais da área médica.

Essas questões são fundamentais para avançar em direção a uma utilização ainda mais eficaz das tecnologias de realidade virtual no ensino em saúde.

Agradecimentos

Este estudo vem sendo parcialmente financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código Financeiro 001, CNPq (Universal 421962/2023-2), FAPERGS (PqG 17/2551-0001192-9 e RITE 22/2551-0000390-7) e Rede Nacional de Ensino e Pesquisa (RNP).

Referências

- Ali Alaraj, Cristian J Luciano, Daniel P Bailey, Abdussalam Elsenousi, Ben Z Roitberg, Antonio Bernardo, P Pat Banerjee, and Fady T Charbel. Virtual reality cerebral aneurysm clipping simulation with real-time haptic feedback. *Operative Neurosurgery*, 11(1):52–58, 2015.
- Ibrahim Aliwi, Vico Schot, Michele Carrabba, Phuoc Duong, Silvia Shievano, Massimo Caputo, Jo Wray, Adelaide de Vecchi, and Giovanni Biglino. The role of immersive virtual reality and augmented reality in medical communication: A scoping review. *J. of Patient Experience*, 10, 1 2023. ISSN 23743743. doi: 10.1177/23743735231171562.
- Gerry Altmiller, Francisco Jimenez, Jack Wharton, Cheryl Wilson, and Natalie Wright. HIV and Contact Tracing: Impact of a Virtual Patient Simulation Activity. *Clinical Simulation in Nursing*, 64:58–66, 2022. ISSN 18761399. doi: 10.1016/j.ecns.2021.12.005.
- Cagatay Basdogan, Mert Sedef, Matthias Harders, and Stefan Wesarg. Vr-based simulators for training in minimally invasive surgery. *IEEE Computer Graphics and Applications*, 27(2):54–66, 2007. doi: 10.1109/MCG.2007.51.
- Isabela Bianchi, Alexandre Lazaretti Zanatta, and Rafael Rieder. Augmented reality in medical teaching-learning process content: A systematic review. In *2020 22nd Symposium on Virtual and Augmented Reality (SVR)*, pages 129–133, 2020. doi: 10.1109/SVR51698.2020.00032.
- Flemming Bjerrum, Ann Sofia Skou Thomsen, Leizl Joy Nayahangan, and Lars Konge. Surgical simulation: current practices and future perspectives for technical skills training. *Medical teacher*, 40(7):668–675, 2018.
- D. A. Bowman et al. Higher levels of immersion improve procedure memorization performance. In *Proceeding JVRC'09 Proceedings of the 15th Joint Virtual Reality Eurographics Conference on Virtual Environments*, 2009.
- Marie-Stéphanie Bracq, Estelle Michinov, and Pierre Jannin. Virtual reality simulation in nontechnical skills training for healthcare professionals: a systematic review. *Simulation in Healthcare*, 14(3):188–194, 2019.
- Grigore C Burdea. *Force and touch feedback for virtual reality*. John Wiley & Sons, Inc., 1996.

- Gordon Caldwell. The process of clinical consultation is crucial to patient outcomes and safety: 10 quality indicators. *Clinical Medicine*, 19(6):503–506, 2019. ISSN 1470-2118. doi: 10.7861/clinmed.2019-0263. URL <https://www.rcpjournals.org/content/19/6/503>.
- Lorraine Tudor Car, Bhone Myint Kyaw, Andrew Teo, Tatiana Erlikh Fox, Sunitha Vimalasvaran, Christian Apfelbacher, Sandra Kemp, and Niels Chavannes. Outcomes, measurement instruments, and their validity evidence in randomized controlled trials on virtual, augmented, and mixed reality in undergraduate medical education: Systematic mapping review, 4 2022. ISSN 22919279.
- Laurent Chaby, Abdelkader Benamara, Marco Pino, Etienne Prigent, Brian Ravenet, Jean-Claude Martin, Hubert Vanderstichel, Anne-Sophie Rigaud, and Mohamed Chetouani. Embodied virtual patients as a simulation-based framework for training clinician-patient communication skills: An overview of their use in psychiatric and geriatric care. *Frontiers in Virtual Reality*, 3:827312, 2022. doi: 10.3389/frvir.2022.827312.
- A. Chalmers and K. Debattista. Level of realism for serious games. In *Games and Virtual Worlds for Serious Applications, 2009. VS-GAMES '09. Conference in*, 2009.
- Gilles Chiniara, Gary Cole, Ken Brisbin, Dan Huffman, Betty Cragg, Mike Lamacchia, Dianne Norman, and Canadian Network For Simulation in Healthcare Guidelines Working Group. Simulation in healthcare: A taxonomy and a conceptual framework for instructional design and media selection. *Medical Teacher*, 35(8):e1380–e1395, 2013. doi: 10.3109/0142159X.2012.733451. URL <https://doi.org/10.3109/0142159X.2012.733451>. PMID: 23121247;.
- Seung-Hun Chon, Sabrina Hilgers, Ferdinand Timmermann, Thomas Dratsch, Patrick Sven Plum, Felix Berlth, Rabi Datta, Hakan Alakus, Hans Anton Schlößer, Christoph Schramm, et al. Web-based immersive patient simulator as a curricular tool for objective structured clinical examination preparation in surgery: development and evaluation. *JMIR Serious Games*, 6(3):e10693, 2018.
- Fabrizio Consorti, Rosaria Mancuso, Martina Nocioni, and Annalisa Piccolo. Efficacy of virtual patients in medical education: a meta-analysis of randomized studies. *Computers & Education*, 59(3):1001–1008, 2012.
- David A Cook and Marc M Triola. Virtual patients: a critical literature review and proposed next steps. *Medical education*, 43(4):303–311, 2009.
- S. Cotin, H. Delingette, and N. Ayache. Real-time elastic deformations of soft tissues for surgery simulation. *IEEE Transactions on Visualization and Computer Graphics*, 5(1):62–73, 1999. doi: 10.1109/2945.764872.
- S. Deterding et al. From game design elements to gamefulness: Defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 2011.

- Chelsea Ekstrand, Ali Jamal, Ron Nguyen, Annalise Kudryk, Jennifer Mann, and Ivar Mendez. Immersive and interactive virtual reality to improve learning and retention of neuroanatomy in medical students: a randomized controlled study. *Canadian Medical Association Open Access Journal*, 6(1):E103–E109, 2018. doi: 10.9778/cmajo.20170110. URL <https://www.cmajopen.ca/content/6/1/E103>.
- Abraham Flexner. *Medical Education in the United States and Canada: A Report to the Carnegie Foundation for the Advancement of Teaching*. Arno Press, New York, 1972. Originally published in 1910 with the Carnegie Foundation for the Advancement of Teaching.
- Thomas A Furness III. The super cockpit and its human factors challenges. In *Proceedings of the human factors society annual meeting*, volume 30, pages 48–52. SAGE Publications Sage CA: Los Angeles, CA, 1986.
- Sandeep Ganni, Meng Li, Sanne MBI Botden, Samir Ranjan Nayak, Bhaskar Rao Ganni, Anne-Francoise Rutkowski, Richard HM Goossens, and Jack Jakimowicz. Virtual operating room simulation setup (vorss) for procedural training in minimally invasive surgery—a pilot study. *Indian Journal of Surgery*, 82:810–816, 2020.
- J. Gerup, C.B. Soerensen, and P. Dieckmann. Augmented reality and mixed reality for healthcare education beyond surgery: an integrative review. *Int J Med Educ*, 11:1–18, 2020. doi: 10.5116/ijme.5e01.eb1a. URL <http://www.ijme.net/archive/11/augmented-reality-for-healthcare-education/>.
- T. P. Grantcharov, V. B. Kristiansen, J. Bendix, L. Bardram, J. Rosenberg, and P. Funch-Jensen. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *British Journal of Surgery*, 91(2):146–150, 2004. ISSN 00071323. doi: 10.1002/bjs.4407.
- Frederic W Hafferty. Beyond curriculum reform: confronting medicine’s hidden curriculum. *Academic medicine*, 73(4):403–7, 1998.
- Minsik Hong, Jerzy W Rozenblit, and Allan J Hamilton. Simulation-based surgical training systems in laparoscopic surgery: a current review. *Virtual Reality*, 25(2):491–510, 2021.
- Tobias Huber, Markus Paschold, Christian Hansen, Hauke Lang, and Werner Kneist. Artificial versus video-based immersive virtual surroundings: Analysis of performance and user’s preference. *Surgical Innovation*, 25(3):280–285, 2018a. doi: 10.1177/1553350618761756. URL <https://doi.org/10.1177/1553350618761756>. PMID: 29504470.
- Tobias Huber, Tom Wunderling, Markus Paschold, Hauke Lang, Werner Kneist, and Christian Hansen. Highly immersive virtual reality laparoscopy simulation: development and future aspects. *International journal of computer assisted radiology and surgery*, 13:281–290, 2018b.

- B.C. Ibanez, B. Marne, and J.M. Labat. Conceptual and technical frameworks for serious games. In *Proceedings of the 5th European Conference on Games Based Learning*, pages 81–87, 2011.
- Francisco A Jimenez. Can Virtual Patient Simulation Be Used in Substitution of Traditional Clinical Hours in Undergraduate Nursing Education? A Review of the Evidence. *White Paper*, page 15, 2022.
- Beres Joyner and Dr Louise Young. Teaching medical students using role play: Twelve tips for successful role plays. *Medical Teacher*, 28(3):225–229, 2009. doi: 10.1080/01421590600711252. URL <https://doi.org/10.1080/01421590600711252>. PMID: 16753719.
- Ae-Ri Jung and Eun-A Park. The effectiveness of learning to use hmd-based vr technologies on nursing students: chemoport insertion surgery. *International Journal of Environmental Research and Public Health*, 19(8):4823, 2022.
- Gerson Alves Pereira Júnior and Hermila Tavares Vilar Guedes. *Simulação Clínica: ensino e avaliação nas diferentes áreas da Medicina e Enfermagem Cirurgia Geral, Clínica Médica, Emergência, Ginecologia e Obstetrícia, Pediatria, Saúde Coletiva, Saúde da Família e Comunidade e Saúde Mental*. Associação Brasileira de Educação Médica, 2022. ISBN 978-65-86406-04-7. URL <https://website.abem-educmed.org.br/publicacoes/ebooks/>.
- Shunaha Kim-Fine and Erin A Brennand. Surgical simulation and competency. *Obstetrics and Gynecology Clinics*, 43(3):575–590, 2016.
- A. J. Kleinheksel. Transformative learning through virtual patient simulations: Predicting critical student reflections. *Clinical Simulation in Nursing*, 10(6):e301–e308, 2014. ISSN 18761399. doi: 10.1016/j.ecns.2014.02.001. URL <http://dx.doi.org/10.1016/j.ecns.2014.02.001>.
- Nils Fredrik Kleven, Ekaterina Prasolova-Førland, Mikhail Fominykh, Arne Hansen, Guri Rasmussen, Lisa Millgård Sagberg, and Frank Lindseth. Training nurses and educating the public using a virtual operating room with oculus rift. In *2014 International Conference on Virtual Systems & Multimedia (VSMM)*, pages 206–213. IEEE, 2014.
- Bhone Myint Kyaw, Nakul Saxena, Pawel Posadzki, Jitka Vseteckova, Charoula Konstantia Nikolaou, Pradeep Paul George, Ushashree Divakar, Italo Masiello, Andrzej A. Kononowicz, Nabil Zary, and Lorainne Tudor Car. Virtual reality for health professions education: Systematic review and meta-analysis by the digital health education collaboration, 1 2019. ISSN 14388871.
- G Michael Lemole Jr, P Pat Banerjee, Cristian Luciano, Sergey Neckrysh, and Fady T Charbel. Virtual reality in neurosurgical education: part-task ventriculostomy simulation with dynamic visual and haptic feedback. *Neurosurgery*, 61(1):142–149, 2007.
- Alexander Libin, Manon Lauderdale, Yuri Millo, Christine Shamloo, Rachel Spencer, Brad Green, Joyce Donnellan, Christine Wellesley, and Suzanne Groah. Role-playing

- simulation as an educational tool for health care personnel: Developing an embedded assessment framework. *Cyberpsychology, Behavior, and Social Networking*, 13(2):217–224, 2010. URL <http://doi.org/10.1089/cyber.2009.0040>.
- Alan Liu, Frank Tendick, Kevin Cleary, and Christoph Kaufmann. A Survey of Surgical Simulation: Applications, Technology, and Education. *Presence: Teleoperators and Virtual Environments*, 12(6):599–614, 12 2003. doi: 10.1162/105474603322955905. URL <https://doi.org/10.1162/105474603322955905>.
- Ryan Lohre, Aaron J Bois, George S Athwal, Danny P Goel, et al. Improved complex skill acquisition by immersive virtual reality training: a randomized controlled trial. *JBJS*, 102(6):e26, 2020.
- Abel J. Lungu, Wout Swinkels, Luc Claesen, Puxun Tu, Jan Egger, and Xiaojun Chen. A review on the applications of virtual reality, augmented reality and mixed reality in surgical simulation: an extension to different kinds of surgery. *Expert Review of Medical Devices*, 18:47–62, 2021. ISSN 17452422. doi: 10.1080/17434440.2021.1860750.
- Anderson Maciel and Suvranu De. An efficient dynamic point algorithm for line-based collision detection in real time virtual environments involving haptics. *Computer animation and virtual worlds*, 19(2):151–163, 2008.
- Anderson Maciel, Youquan Liu, Woojin Ahn, T Paul Singh, Ward Dunnican, and Suvranu De. Development of the vblast™: a virtual basic laparoscopic skill trainer. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 4(2):131–138, 2008.
- Randi Q Mao, Lucy Lan, Jeffrey Kay, Ryan Lohre, Olufemi R Ayeni, Danny P Goel, et al. Immersive virtual reality for surgical training: a systematic review. *Journal of Surgical Research*, 268:40–58, 2021.
- Emerson Elias Merhy. Saúde: a cartografia do trabalho vivo. In *Saúde: a cartografia do trabalho vivo*, pages 189–189. 2005.
- D.R. Michael and S.L. Chen. *Serious Games: Games that Educate, Train, and Inform*. Muska & Lipman/Premier-Trade, 2005.
- Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, E77-D(12):1321–1329, 1994.
- I. D. Montiel et al. *Evaluación de competencias en ciencias de la salud*. Unam, Editorial Médica Panamericana, México, 2012.
- Martin Müller, Jonas Jürgens, Marcus Redaelli, Karsten Klingberg, Wolf E Hautz, and Stephanie Stock. Impact of the communication and patient hand-off tool sbar on patient safety: a systematic review. *BMJ open*, 8(8):e022202, 2018.
- M. Müller, J. Jürgens, M. Redaelli, K. Klingberg, W.E. Hautz, and S. Stock. Impact of the communication and patient hand-off tool sbar on patient safety: a systematic review. *BMJ Open*, 8(8):e022202, Aug 2018. doi: 10.1136/bmjopen-2018-022202.

- Luciana Nedel, Vinicius Costa de Souza, Aline Menin, Lucia Sebben, Jackson Oliveira, Frederico Faria, and Anderson Maciel. Using immersive virtual reality to reduce work accidents in developing countries. *IEEE Computer Graphics and Applications*, 36(2): 36–46, 2016. doi: 10.1109/MCG.2016.19.
- Matheus Dias Negrão, Wesley Ferreira, Betania Bohrer, Carla M.D.S. Freitas, Anderson Maciel, and Luciana Nedel. Design and think-aloud study of an immersive interface for training health professionals in clinical skills. In *Proceedings of the 25th Symposium on Virtual and Augmented Reality, SVR '23*, page 157–165, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400709432. doi: 10.1145/3625008.3625037. URL <https://doi.org/10.1145/3625008.3625037>.
- Debra Nestel and Tanya Tierney. Role-play for medical students learning about communication: Guidelines for maximising benefits. *BMC Medical Education*, 7(1): 3–, 2007. doi: 10.1186/1472-6920-7-3. URL <https://doi.org/10.1186/1472-6920-7-3>.
- OMS. *Guia curricular de segurança do paciente da Organização Mundial da Saúde: edição multiprofissional*. Autografia, Rio de Janeiro, 2016. ISBN 978-85-5526-850-2. URL <https://iris.who.int/bitstream/handle/10665/44641/9788555268502-por.pdf?sequence=32>.
- Dixit Bharatkumar Patel, Yong Pei, Mitesh Vasoya, and Paul J. Hershberger. Computer-Supported Experiential Learning-Based Tool for Healthcare Skills. *IEEE Computer Graphics and Applications*, 43(2):57–68, 2023. ISSN 15581756. doi: 10.1109/MCG.2023.3242921.
- Michael Pfandler, Marc Lazarovici, Philipp Stefan, Patrick Wucherer, and Matthias Weigl. Virtual reality-based simulators for spine surgery: a systematic review. *The Spine Journal*, 17(9):1352–1363, 2017. ISSN 1529-9430. doi: <https://doi.org/10.1016/j.spinee.2017.05.016>. URL <https://www.sciencedirect.com/science/article/pii/S1529943017302085>.
- Jai Prashar, Chavini Ranasinghe, and Chirag Bussa Rao. Twelve tips for medical students to enhance clinical skills learning during disrupted placements. *Medical Teacher*, 44(6):596–600, 2022. doi: 10.1080/0142159X.2021.1910644. URL <https://doi.org/10.1080/0142159X.2021.1910644>. PMID: 33856946.
- L. Raghupathi, L. Grisoni, F. Faure, D. Marchal, M.-P. Cani, and C. Chaillou. An intestinal surgery simulator: real-time collision processing and visualization. *IEEE Transactions on Visualization and Computer Graphics*, 10(6):708–718, 2004. doi: 10.1109/TVCG.2004.36.
- Karan Rangarajan, Heather Davis, and Philip H. Pucher. Systematic review of virtual haptics in surgical simulation: A valid educational tool? *Journal of Surgical Education*, 77(2):337–347, 2020. ISSN 1931-7204. doi: <https://doi.org/10.1016/j.jsurg.2019.09.006>. URL <https://www.sciencedirect.com/science/article/pii/S1931720419306609>.

- Laura Raya, Juan Jesus Ruiz, Marc Fabian, Adrian Ron, Javier Garcia, and Cristina Verdu. Development of a Virtual Reality Tool for the Treatment of Pediatric Patients in the ICU. *IEEE Computer Graphics and Applications*, 43(2):69–77, 2023. ISSN 15581756. doi: 10.1109/MCG.2023.3239676.
- M. Reznick, P. Harter, and T. Krummel. Virtual reality and simulation: Training the future emergency physician. *Academic Emergency Medicine*, 9(1):78–87, 2002. ISSN 10696563. doi: 10.1197/aemj.9.1.78.
- Igo B Ribeiro, Janet MC Ngu, Buu-Khanh Lam, and Roger A Edwards. Simulation-based skill training for trainees in cardiac surgery: a systematic review. *The Annals of Thoracic Surgery*, 105(3):972–982, 2018.
- Francesco Rizzetto, Alessandro Bernareggi, Sofia Rantas, Angelo Vanzulli, and Maurizio Vertemati. Immersive virtual reality in surgery and medical education: Diving into the future. *The American Journal of Surgery*, 220(4):856–857, 2020.
- Albert A Rizzo, Joann Difede, Barbara O Rothbaum, Scott Johnston, Robert N McLAY, Greg Reger, Gregory A Gahm, Thomas D Parsons, Ken Graap, and Jarrell Pair. Vr ptsd exposure therapy results with active duty oif/oef combatants. In *MMVR*, pages 277–282, 2009.
- Ganesh Sankaranarayanan, Baichun Li, Kelly Manser, Stephanie B Jones, Daniel B Jones, Steven Schwaitzberg, Caroline GL Cao, and Suvranu De. Face and construct validation of a next generation virtual reality (gen2-vr©) surgical simulator. *Surgical endoscopy*, 30:979–985, 2016.
- Richard M Satava. Virtual reality surgical simulator: the first steps. *Surgical endoscopy*, 7:203–205, 1993.
- Richard M Satava and Shaun B Jones. Virtual environments for medical training and education. *Presence: Teleoperators & Virtual Environments*, 6(2):139–146, 1997.
- B. Sawyer. The "serious games" landscape. In *Instructional & Research Technology Symposium for Arts, Humanities, and Social Sciences*, 2007.
- Ka-Chun Siu, Bradley J Best, Jong Wook Kim, Dmitry Oleynikov, and Frank E Ritter. Adaptive virtual reality training to optimize military medical skills acquisition and retention. *Military medicine*, 181(suppl_5):214–220, 2016.
- Dimitrios Stefanidis, Nick Sevdalis, John Paige, Boris Zevin, Rajesh Aggarwal, Teodor Grantcharov, Daniel B Jones, Association for Surgical Education Simulation Committee, et al. Simulation in surgery: what's needed next? *Annals of surgery*, 261(5): 846–853, 2015.
- Moira Stewart, Judith Belle Brown, W Wayne Weston, Ian R McWhinney, Carol L McWilliam, and Thomas R Freeman. *Medicina centrada na pessoa: transformando o método clínico*. Artmed Editora, 2017.

- Sebastian Streuber, Patrick Saalfeld, Katja Podulski, Florentine Hüttl, Tobias Huber, Holger Buggenhagen, Christian Boedecker, Bernhard Preim, and Christian Hansen. Training of patient handover in virtual reality. *Current Directions in Biomedical Engineering*, 6(1):20200040, 2020. URL <https://doi.org/10.1515/cdbme-2020-0040>.
- Jacob Stuart, Karen Aul, Michael Bumbach, Anita Stephen, and Benjamin Lok. Building a handoff communication virtual experience for nursing students using virtual humans. *CIN: Computers, Informatics, Nursing*, 39(12):1017–1026, December 2021. doi: 10.1097/CIN.0000000000000760.
- Colin Sutherland, Keyvan Hashttrudi-Zaad, Rick Sellens, Purang Abolmaesumi, and Parvin Mousavi. An augmented reality haptic training simulator for spinal needle procedures. *IEEE Transactions on Biomedical Engineering*, 60(11):3009–3018, Nov 2013. ISSN 1558-2531. doi: 10.1109/TBME.2012.2236091.
- Ivan E. Sutherland. The ultimate display. In *Proceedings of the Congress of the International Federation of Information Processing (IFIP)*, volume volume 2, pages 506–508, 1965.
- Charison Tay, Ankur Khajuria, and Chinmay Gupte. Simulation training: a systematic review of simulation in arthroscopy and proposal of a new competency-based training framework. *International Journal of Surgery*, 12(6):626–633, 2014.
- Tuo Wen, David Medveczky, Jie Wu, and et al. Colonoscopy procedure simulation: Virtual reality training based on a real time computational approach. *BioMedical Engineering OnLine*, 17:9, 2018. doi: 10.1186/s12938-018-0433-4. URL <https://doi.org/10.1186/s12938-018-0433-4>.
- WHO. World health organization global patient safety action plan 2021–2030: Towards eliminating avoidable harm in health care. URL <https://www.who.int/teams/integrated-health-services/patient-safety/policy/global-patient-safety-action-plan>. Accessed: May 2024.
- Wikipedia. Metaverso, 2023. URL <https://pt.wikipedia.org/wiki/Metaverso>. Accessed: June 2024.
- Matthew W Zackoff, David Davis, Michele Rios, Rashmi D Sahay, Bin Zhang, Ian Anderson, Matthew NeCamp, Ingrid Rogue, Stephanie Boyd, Aimee Gardner, et al. Tolerability and acceptability of autonomous immersive virtual reality incorporating digital twin technology for mass training in healthcare. Technical report, LWW, 2023.
- Jinao Zhang, Yongmin Zhong, and Chengfan Gu. Deformable models for surgical simulation: A survey. *IEEE Reviews in Biomedical Engineering*, 11:143–164, 2018. doi: 10.1109/RBME.2017.2773521.
- Michael Zyda. From visual simulation to virtual reality to games. *Computer*, 38(9): 25–32, 2005. ISSN 00189162. doi: 10.1109/MC.2005.297.

Capítulo

3

Avaliação de Tecnologia em Saúde Digital: Fundamentos e Aplicações Práticas

Thais Maria Yomoto Ferauche (UEPEP/CEETEPS), Jean José Silva (Santa Casa de Misericórdia de Passos), Wilian Hisatugu (UFES), Márcia Ito (UEPEP/CEETEPS)

Abstract

The use of digital technologies in healthcare has grown significantly, both in support of public and private healthcare. These technologies have shown considerable potential for improving healthcare services. In this scenario, it is necessary that the assessment of health technologies incorporate specific processes to assess the use of these digital technologies. The purpose of this course is to present the relevance, fundamentals and applications of technology assessment in digital health, with a particular focus on the assessment of software for the healthcare. The particularities of applications in the healthcare will be addressed, as well as specific protocols and processes for evaluating digital health technologies.

Resumo

O emprego crescente de tecnologias digitais na área da saúde tem sido expressivo, tanto no apoio aos serviços de saúde pública quanto privada. Estas ferramentas têm evidenciado um considerável potencial para aprimorar o suporte oferecido pelos serviços de saúde. Nesse cenário, é necessário que a avaliação de tecnologias em saúde incorpore processos específicos para avaliar as tecnologias digitais. O propósito deste minicurso é apresentar a relevância, os fundamentos e as aplicações da avaliação de tecnologia em saúde digital, com um foco particular na avaliação de softwares para a área da saúde. Serão abordadas as particularidades dos aplicativos na área da saúde, bem como os protocolos e processos específicos para a avaliação de tecnologias digitais na saúde.

3.1 Introdução

A evolução da tecnologia da informação e comunicação (TIC) aliada a inteligência artificial, ao IoT (Internet das coisas) e a robótica fizeram com que os processos se tornassem integrados e inteligentes desde a produção da matéria prima até o produto, permitindo aumentar a produtividade, a criação de valor e o bem-estar social. Na Indústria, a transformação digital, alterou a estrutura organizacional, fazendo com ela seja mais ágil e responsiva. Assim surge a Indústria 4.0, na qual criam-se papéis que encorajam a inovação e a mudança de cultura aumentando a colaboração e a experimentação entre desenvolvedores e clientes. A mudança é tão significativa que altera o modo de interação dos clientes com a organização de serviços por meio de canais digitais. O impacto da transformação digital nas indústrias inclui o aumento da produtividade na manufatura, aumento da qualidade e eficiência na assistência e indústria de saúde, melhora o acesso à educação na medida que novos meios de disseminação do conhecimento são desenvolvidos, melhoram as condições de trabalhos, facilitam a tomada de decisão pelas organizações e por fim estimulam o crescimento econômico por meio das cidades inteligentes [ABIODUN; RAMPERSAD; BRINKWORTH, 2023].

Portanto, estamos numa nova era denominada, Era Digital, na qual tecnologia, máquinas e seres humanos estão conectados e interagindo em tempo real e em muitas vezes sendo difícil perceber quem é quem na execução do serviço. Na Era digital a globalização e a rápida evolução das tecnologias digitais estão trazendo mudanças diversas e complexas para a sociedade. Ao mesmo tempo, o mundo está se deparando com desafios em escala global, como o esgotamento dos recursos naturais, o aquecimento global, o terrorismo entre outros. Em setembro de 2015 a ONU estabeleceu a agenda 2030 para o desenvolvimento sustentável, o desafio é estabelecer um sistema em que todas as nações trabalhem juntas com a finalidade de ter um mundo mais sustentável, resolvendo tanto o desenvolvimento econômico quanto as questões sociais. A chave para isto é que pessoas e o planeta tenham paz e prosperidade, com uma inclusão para todos sem exceção. Foi assim que em janeiro de 2016, o Japão define a Sociedade 5.0, a sociedade centrada no humano, na qual tanto o desenvolvimento social quanto a solução dos desafios sociais são alcançados, e todos possam desfrutar de uma boa qualidade de vida sem diferenças socioeconômicas. Para isto é preciso equilibrar o avanço econômico com a resolução de problemas sociais na qual computadores e hiper conexão resultem em um modo de vida mais inteligente, eficiente e sustentável [SHEPHERD, 2004] [FUKUYAMA, 2018].

As transformações que a indústria 4.0 e as metas que se deseja alcançar na sociedade 5.0, afetam todas as áreas inclusive a da saúde e com isso faz surgir o conceito de Saúde Digital que expande o conceito da e-Saúde ao incluir os consumidores digitais que possuem os mais diversos dispositivos conectados e inteligentes (WHO, 2021).

Em anos mais recentes, notou-se que para melhorar a qualidade de vida e por consequência a saúde do cidadão, era preciso cuidar não somente da doença, mas do paciente como um todo, seu estado psicossocial, econômico, entre outros. A este tipo de assistência é denominado de cuidado centrado no paciente e a indústria de saúde passa a ter um “consumidor de serviço de saúde” que deseja e tem o controle sobre a gestão da sua saúde. É assim que a indústria de saúde passa a se preocupar com as necessidades e expectativas com este novo “consumidor de serviço de saúde” que busca por experiências personalizadas, confortáveis, rapidez e eficiência na prestação do serviço. A saúde digital

além de melhorar os serviços de saúde trazendo agilidade e o aumento de produtividade, possibilita fornecer ao consumidor de serviço de saúde a experiência inovadora e o controle e gestão sobre sua saúde [ITO, 2020] [STOUMPOS; KITSIOS; TALIAS, 2023]. Assim, surge o consumidor de saúde digital que são os cidadãos, pacientes, cuidadores e familiares que usam os serviços e ferramentas da saúde digital. Os serviços e ferramentas da saúde digital são os que permitem o consumidor de saúde digital ter um papel ativo em sua saúde e na prevenção de doenças de forma segura. A diferença com relação à telessaúde é que neste caso o usuário obrigatoriamente deve ter um papel ativo na condução de sua saúde. Outro ponto é que ele não pode ser uma solução para uso somente do profissional de saúde [WETTER, 2012].

Por outro lado, com a erradicação das doenças infecciosas, houve um aumento das doenças crônicas. Há diferença de abordagem entre os dois tipos de doença, enquanto na doença infecciosa a dependência da participação do paciente em relação ao seu tratamento é menor (não é nulo, pois decidir se vai tomar ou não o medicamento é do paciente), na doença crônica é totalmente dependente do doente, pois a decisão de parar de fumar ou mudar os hábitos alimentares é dele e essa decisão implicará num bom desfecho clínico ou não. Assim, um dos fatores de sucesso no tratamento de doenças é a participação ativa do paciente em relação ao seu estado de saúde. Neste sentido, a saúde digital contribui ao facilitar a participação dos pacientes no processo de tomada de decisão quanto ao seu estado de saúde. O paciente ao buscar por informações seja pela internet ou por aplicativos de saúde digital, por exemplo, aplicativos em celulares, fazem a diferença na tomada de decisão por parte deles. Pode-se então dizer que os consumidores de saúde digital se beneficiam com serviços e ferramentas digitais que possuem informações e funcionalidades que os ajudam na tomada de decisão sobre sua saúde, trazendo um empoderamento para estes cidadãos [ITO, 2020] [ITO; PEREIRA, 2015] [STOUMPOS; KITSIOS; TALIAS, 2023].

De acordo com a Organização Mundial de Saúde, *mobile health (mHealth)* é definido como “a prática clínica ou de saúde pública apoiada por dispositivos móveis, como telefones celulares, dispositivos para o monitoramento de pacientes, assistentes digitais pessoais e outros dispositivos sem fio”¹. Por extensão à definição de cuidado centrado no paciente, pode-se dizer que tecnologias *mHealth* centrado no paciente são aquelas apoiadas por dispositivos móveis projetados para promover o cuidado centrado no paciente e, portanto, um dos seus principais usuários são os consumidores de saúde digital [BRUCE et al., 2020].

Ao mesmo tempo que os aplicativos *mHealth* trazem benefícios aos consumidores, eles podem trazer riscos, se eles forem imprecisos e não confiáveis, principalmente como vimos anteriormente esses usuários usam as informações para a tomada de decisão. Na revisão realizado por Akbar et. al [AKBAR; COIERA.; MAGRABI, 2020] foram encontrados os seguintes problemas na qualidade do conteúdo de alguns aplicativos de saúde:

¹ Tradução livre feito pelos autores

- a. Informação incorreta,
- b. Informação incompleta,
- c. variação no conteúdo entre aplicativos de domínios semelhantes,
- d. resultados incorretos, após cálculos e análises de respostas dos usuários,
- e. respostas inapropriadas às necessidades dos consumidores,

por exemplo, foram encontrados aplicativos que em caso de suspeita de suicídio não indicavam a busca de uma ajuda profissional nos momentos críticos.

Nesta mesma revisão foram encontrados estudos com falhas na funcionalidade do software:

- a. funcionalidades inadequadas de acordo com a tarefa, por exemplo, não permitir a entrada da dosagem em gramas ao invés de mililitros,
- b. falta de validação na entrada dos dados,
- c. processamento demorado que podem trazer risco à segurança do paciente, por exemplo, a demora no monitoramento de sinais vitais,
- d. resposta que trazem perigo à saúde,
- e. falha no alarme.

Além disso, foram encontrados vários aplicativos que não tinham o envolvimento de um especialista, ou em que o aplicativo não foi baseado em evidências clínicas e científicas e o mais preocupante deles é que não houve uma validação formais do aplicativo que pudessem trazer confiabilidade no seu conteúdo ou uso. Com isso fica evidente a necessidade de meios para que os profissionais de saúde e consumidores de saúde digital avaliem os aplicativos e assim possam escolher o aplicativo que mais seja confiável e adequado às suas necessidades [AKBAR; COIERA,; MAGRABI, 2020].

3.2 Avaliação de software

Um componente importante para assegurar a qualidade do software é uma avaliação de software. Incluindo o desenvolvimento de uma avaliação de qualidade de teste onde abrange a conformidade, estabilidade, adequação e eficácia [LU et al., 2022]. A qualidade de um software se refere à conformidade com os requisitos funcionais e de desempenho explicitamente declarados, a padrões de desenvolvimento claramente documentados e às características implícitas que são esperadas de todo software [MAXIM; PRESSMAN, 2021]. Os modelos de qualidade permitem avaliar softwares de acordo com diferentes aspectos e comumente representam as características desejáveis a um software em uma estrutura hierárquica [YANG, 2012].

Em outras palavras, a avaliação de software é um processo importante que visa analisar e medir a qualidade, eficácia e desempenho de um sistema ou aplicativo. Esta é uma prática ampla que visa garantir que o software satisfaça os requisitos estabelecidos e atinja os objetivos propostos. Existem várias dimensões e critérios envolvidos nesse processo, entre eles a usabilidade, confiabilidade, eficiência, manutenibilidade e portabilidade. A usabilidade, por exemplo, refere-se à facilidade de uso e compreensão do software pelo usuário final. Avaliações nessa dimensão consideram a interface do usuário, a intuitividade das operações e a experiência geral de interação. Já a confiabilidade está relacionada à capacidade do software de desempenhar suas funções de maneira consistente e livre de falhas, enquanto a eficiência diz respeito ao desempenho do software em termos de consumo de recursos e tempo de resposta. [ENTRINGER; FERREIRA, 2019].

A avaliação de software também abrange a manutenibilidade, que avalia a facilidade com que o software pode ser modificado ou corrigido, e a portabilidade, que se refere à capacidade do software de ser transferido entre diferentes ambientes. Além desses, outros critérios específicos podem ser considerados dependendo do contexto, como segurança, escalabilidade e interoperabilidade.

Existem diversas metodologias e ferramentas para a avaliação de software, desde testes manuais até abordagens mais automatizadas, como testes de unidade, integração e aceitação. A escolha da abordagem dependerá dos objetivos da avaliação, do tipo de software em questão e das características do projeto.

No campo da arquitetura de software, por exemplo, uma estrutura foi proposta para ajudar arquitetos de software iniciantes a avaliar os requisitos de qualidade de um sistema. Essa estrutura usa técnicas de processamento de imagem e aprendizado de máquina para inferir atributos e táticas de qualidade com base em padrões arquitetônicos [KAPUR et al., 2021].

Outro exemplo, é um método de avaliação da qualidade estática de software baseado no método AHP, que avalia objetiva e quantitativamente a qualidade do software e fornece orientação para a manutenção do software [MINGLIANG et al., 2021].

Porém, a avaliação de software, concentra-se especificamente na análise de programas de computador, aplicativos ou sistemas de software. Tem como principal fim analisar e determinar a qualidade, eficiência e adequação de um programa de software específico para cumprir seus objetivos.

Com algumas limitações, está mais restrita à análise das características e funcionalidades específicas do software, avaliando sua capacidade de realizar tarefas específicas. Por isso, faz uso de métodos mais específicos, como testes de software, revisões de código, avaliações de usabilidade e conformidade com requisitos.

3.3 Qualidade de software

Existem muitas definições de qualidade de software dispostas na literatura, sob diferentes pontos de vista de vários autores:

- A qualidade de um produto de software é medida com base em vários fatores para garantir que ele atenda às necessidades e expectativas do cliente, conforme discutido usando técnicas de sistema de inferência difusa no padrão ISO [BARZEGAR, 2021].
- A qualidade do software é determinada atendendo às necessidades do usuário por meio de processos e métricas bem definidos, como métricas de projeto, produto e processo, conforme discutido no artigo sobre melhoria da qualidade de software [KARUNA; DIVYA; MANGALA, 2018].
- Qualidade de software é a conformidade a requisitos funcionais e de desempenho que foram explicitamente declarados, a padrões de desenvolvimento claramente documentados, e a características implícitas que são esperadas de todo software desenvolvido por profissionais [MAXIM; PRESSMAN, 2021].

O conceito de qualidade de software engloba três fatores importantes, são eles os requisitos de software, os padrões e os requisitos implícitos. Diante disso a ISO/IEC 9126 afirma que qualidade é a totalidade de características e critérios de um produto ou serviço

que exercem suas habilidades para satisfazer às necessidades declaradas ou envolvidas. Baseado nessa fundamentação, a base para qualidade de software é descrita na FIGURA 3.1:

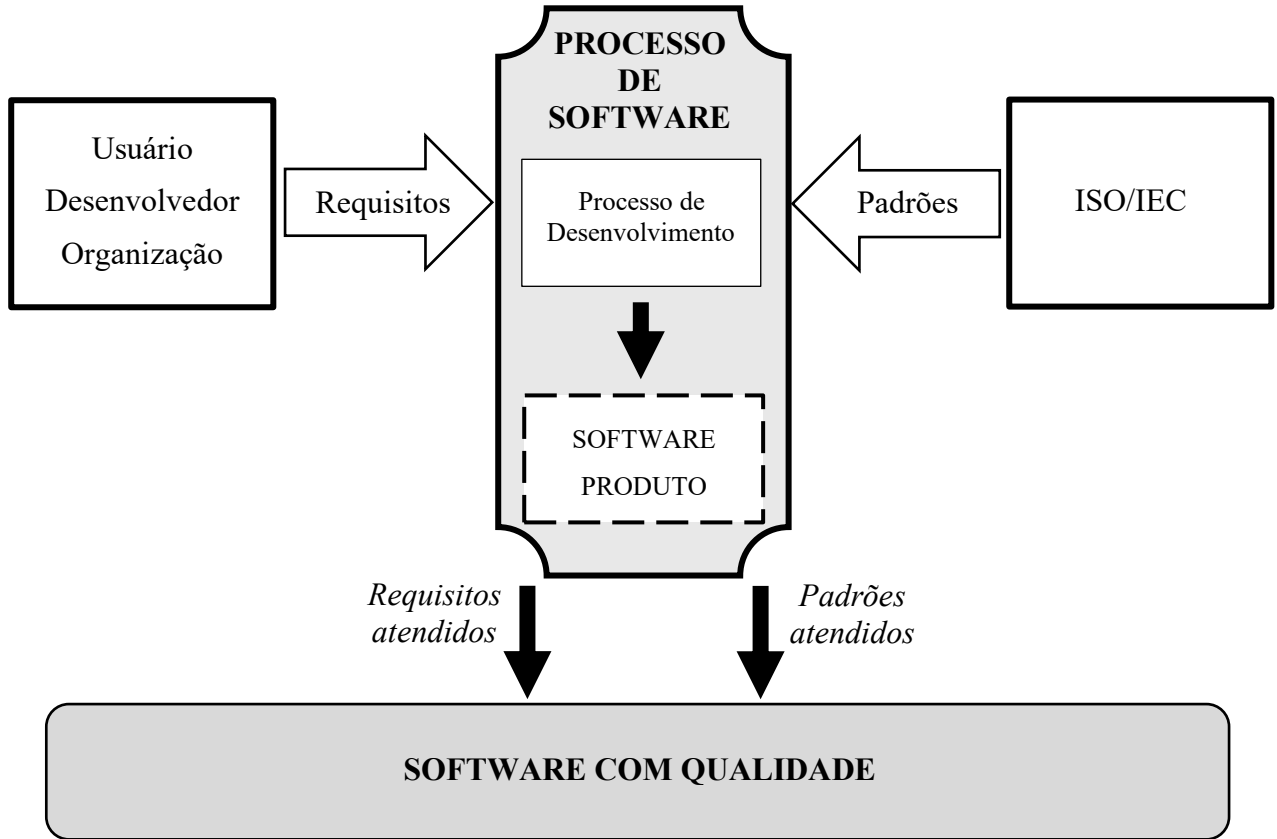


Figura 3.1 - Requisitos de software (Base da Qualidade)

Por meio da FIGURA 3.1, compreende-se que a qualidade de software é referente a atender plenamente os requisitos do cliente, ou seja, atingir e se possível superar as expectativas do cliente. Garvin [GARVIN, 2002] descreve as oito dimensões da qualidade de software na FIGURA 3.2:



Figura 3.2 - As oito dimensões da qualidade de Garvin

Por meio da figura acima, compreende-se que a qualidade de software é referente a atender plenamente os requisitos do cliente, ou seja, atingir e se possível superar as expectativas do cliente. Garvin [GARVIN, 2002] descreve as oito dimensões da qualidade de software na FIGURA 3.2:

- **Desempenho:** refere-se à capacidade do software de atender às expectativas dos usuários em termos de velocidade, precisão, eficiência e capacidade de resposta. Um software com bom desempenho é capaz de executar suas tarefas de forma rápida e eficiente, sem atrasos ou erros significativos.
- **Características:** são as funcionalidades e atributos essenciais do software que atendem às necessidades dos usuários. As características do software devem ser projetadas para proporcionar uma experiência satisfatória ao usuário, oferecendo as ferramentas e recursos necessários para realizar suas tarefas de forma eficaz.
- **Confiabilidade:** a probabilidade de o software funcionar sem falhas durante um período de tempo determinado concerne à confiabilidade. Um software confiável é aquele que pode ser utilizado de forma consistente e previsível, sem apresentar erros ou interrupções inesperadas que possam prejudicar a experiência do usuário.
- **Conformidade:** ao relacionar ao grau em que o software atende às normas, especificações e requisitos predefinidos, contempla a conformidade. Um software que está em conformidade com os padrões estabelecidos é aquele que cumpre todas as diretrizes e regulamentações relevantes, garantindo sua compatibilidade e interoperabilidade com outros sistemas e tecnologias.
- **Durabilidade:** é a capacidade do software de resistir ao desgaste e ao uso contínuo por um longo período de tempo. Um software durável é aquele que pode ser utilizado de forma consistente ao longo do tempo, sem perder sua eficácia ou desempenho devido ao uso prolongado ou condições adversas.
- **Facilidade:** relaciona-se à facilidade com que os usuários podem aprender, operar e entender o software. Um software fácil de usar é aquele que possui uma interface intuitiva e amigável, permitindo que os usuários realizem suas tarefas sem a necessidade de treinamento extensivo ou assistência técnica.
- **Estética:** são os aspectos visuais do software que influenciam a percepção do usuário em termos de atratividade e prazer de uso. Um software esteticamente atraente é aquele que possui um design visualmente agradável, com layouts, cores e elementos gráficos que tornam a experiência do usuário mais agradável e envolvente.
- **Qualidade percebida:** pertence a avaliação subjetiva da qualidade do software feita pelo usuário, considerando todos os outros fatores mencionados. A qualidade percebida de um software é influenciada pela experiência do usuário, suas expectativas e percepções sobre a eficácia, confiabilidade, facilidade de uso e outros aspectos do software.

A partir do conteúdo exposto na FIGURA 3.1, compreende-se que a qualidade de software pode ser definida como a relação entre os requisitos e os resultados alcançados, ou seja, pode ser considerada como a diferença entre o que se espera, e o que se consegue.

Por fim, conforme Roos, Sartori e Godoy [ROOS, SARTORI E GODOY, 2009], utilizando o modelo de Kano, define a qualidade de software em três níveis: qualidade por excelência, qualidade unidimensional e qualidade esperada. Onde:

- **Qualidade por excelência** é definida como qualquer característica do software que exceda o nível de desempenho esperado.
- **Qualidade unidimensional** é definida como qualquer característica do software desejada por usuários específicos, porém não inteiramente necessária para todos as categorias de usuários.
- **Qualidade esperada** é definida como qualquer característica necessária ao funcionamento normal do produto.

Diante de tudo isso, é possível verificar que é imprescindível o desenvolvimento de softwares de qualidade que sejam capazes de suprir e atender a todos os requisitos exigidos pelo cliente, e como meio de reger a qualidade faz-se necessário o uso de normas que estabelecem regras e critérios para assegurar que os produtos de software estejam adequados para o devido uso.

3.3.1 Conceitos Fundamentais em Avaliação de Software

Conceitos fundamentais na avaliação de software abrangem vários aspectos, como obter solicitações de fontes externas, analisar saídas de log, testar com diversos dados de entrada, identificar modificações na estrutura de dados e avaliar vulnerabilidades de softwares [MINOV, 2012] [YAMAKOSHI; NISHIYAMA, 2018]. A avaliação de desempenho envolve pontos de vista, métricas, caracterização da carga de trabalho e benchmarking [OBAIDAT; BOUDRIGA, 2010].

Além disso, os modelos de estimativa de software desempenham um papel crucial no gerenciamento de software, enfatizando o design e a avaliação de modelos de estimativa [ABRAN, 2015]. O teste, uma parte vital da avaliação de software, visa garantir que os sistemas atendam aos requisitos especificados por meio de execução sistemática e identificação de erros [YADAV, 2015]. Esses conceitos combinados formam a base da avaliação de software, abrangendo testes, métricas de desempenho, análise de entradas externas e avaliação de vulnerabilidades.

3.3.2 Métodos e Critérios de Avaliação

A avaliação de produto de software com base em normas de qualidade tem sido uma das formas empregadas por organizações que produzem ou adquirem software para aferirem a qualidade de seus produtos. Assim, para que a avaliação seja mais efetiva, é importante a utilização de modelos de qualidade que permitam estabelecer e avaliar requisitos de qualidade, e que o processo de avaliação seja bem definido e estruturado.

Assim, para que esses requisitos sejam atendidos e a qualidade atingida, existem as normas ISO/IEC 9126, ISO/IEC 14598 e ISO/IEC 25000 que trabalha com especificações técnicas, que estabelecem regras e critérios para assegurar que os produtos de software estejam adequados para o devido uso, desta forma essas normas tornam-se essenciais, pois elas orientam para o devido desenvolvimento e avaliação dos produtos de software com o objetivo de atender a expectativa e qualidade exigida pelo cliente [ISO 9126, 2001] [ISO 14598, 1998] [ISO 25000, 2014].

Algumas normas foram criadas para padronizar internacionalmente as características de implementação e avaliação da qualidade de software.

Para melhor entendimento, segue nos próximos itens especificações sobre as normas ISO/IEC 9126, ISO/IEC 14598 e ISO/IEC 25000 que trabalham com modelo de avaliação de software e métricas para aperfeiçoamento.

3.3.2.1 ISO/IEC 9126, define um modelo de qualidade para avaliação de produtos de software

A norma ISO/IEC 9126 recomenda que os produtos de software sejam avaliados usando um modelo de qualidade, que tem a finalidade de aprimorar o produto, para que dessa forma seja possível cumprir os requisitos para atingir a expectativa do cliente [ISO 9126, 2001]. A série de normas ISO/IEC 9126 descreve um modelo de qualidade para produtos de software categorizando a qualidade hierarquicamente em um conjunto de características e subcaracterísticas. Esta série também propõe métricas que pode ser utilizada durante a avaliação dos produtos de software (medição, pontuação e julgamento dos produtos de software).

A ISO/IEC 9126 é dividida em quatro partes (FIGURA 3.3): (9126-1) modelo de qualidade; (9126-2) métricas externas; (9126-3) métricas internas e, (9126-4) métricas de qualidade de uso.

- **ISO/IEC 9126-1** é definido um modelo de qualidade para produtos de software.
- **ISO/IEC 9126-2** é proposto métricas externa para medir os atributos das características de qualidade definidas na ISO/IEC 9126-1.
- **ISO/IEC 9126-3** apresenta métricas internas para medir os atributos das características de qualidade definidas na ISO/IEC 9126-1.
- **ISO/IEC 9126-4** descreve sobre métricas de qualidade em uso para medir os atributos das características de qualidade definidas na ISO/IEC 9126-1.

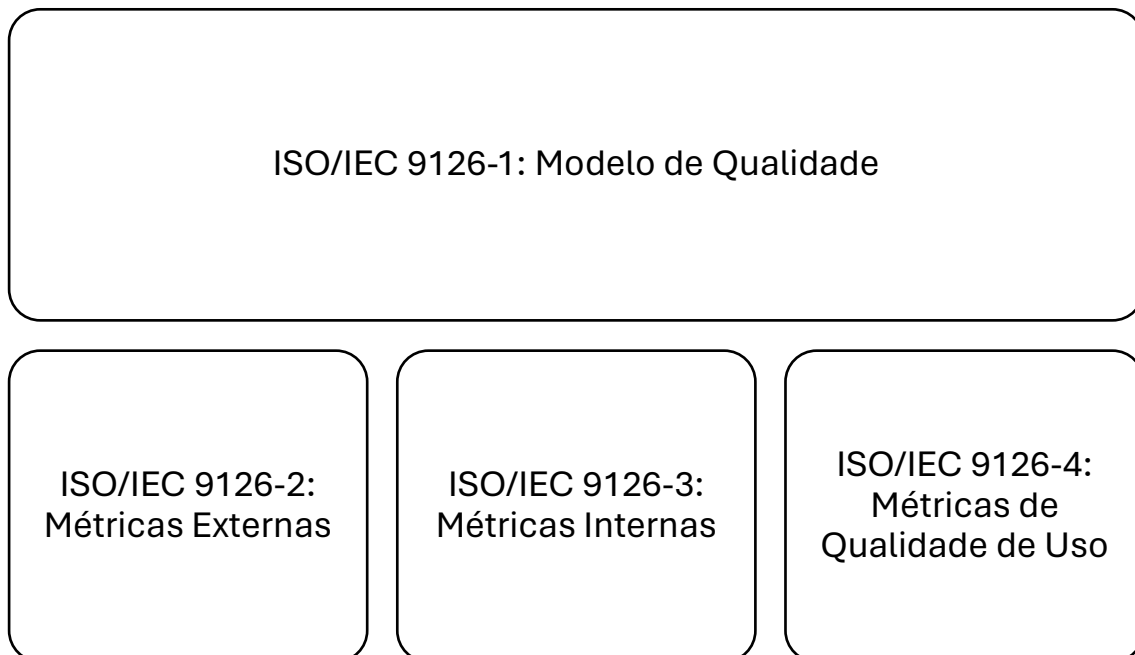


Figura 3.3 - Estrutura da série de normas ISO/IEC 9126

Nas ISO/IEC 9126-2 e 9126-3 apresenta um modelo de qualidade externa e interna que separa os atributos de qualidade de software em seis características: (i) funcionalidade; (ii) confiabilidade; (iii) usabilidade; (iv) eficiência; (v) manutenibilidade e (vi) portabilidade (FIGURA 3.4).

Essas características são subdivididas em subcaracterísticas, as quais podem ser mensuradas por métricas internas e externas, conforme exposto na FIGURA 3.4:

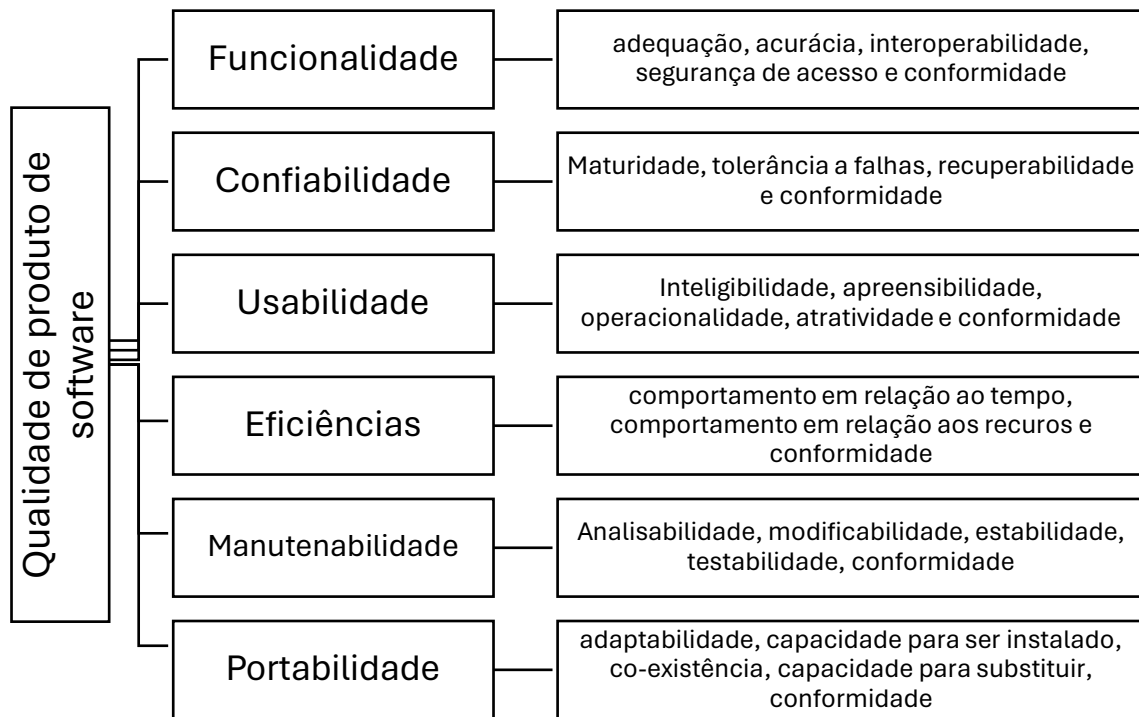


Figura 3.4 - ISO/IEC 9126-2 e 9126-3: Características e Subcaracterísticas

A norma ISO/IEC 9126 – 4 apresenta um modelo para qualidade. O modelo de qualidade em uso categoriza os atributos de qualidade de software em quatro características (Eficácia, Produtividade, Segurança e Satisfação). A qualidade em uso é a visão da qualidade sob a perspectiva do usuário. A obtenção de qualidade em uso é dependente da obtenção da necessária qualidade externa, a qual, por sua vez, é dependente da obtenção da necessária qualidade interna. Normalmente, são necessárias medidas em todos os três níveis, pois atender aos critérios para medidas internas em geral não é suficiente para garantir o atendimento aos critérios para medidas externas, e atender aos critérios para medidas externas de subcaracterísticas em geral não é suficiente para garantir o atendimento aos critérios para qualidade em uso [ISO 9126, 2001].

A ISO/IEC 9126 define as características do modelo da qualidade em uso da seguinte maneira:

- **Eficácia:** capacidade do produto de software de permitir que usuários atinjam metas especificadas com acurácia e completude, em um contexto de uso especificado.
- **Produtividade:** capacidade do produto de software de permitir que seus usuários empreguem quantidade apropriada de recursos em relação à eficácia obtida, em um contexto de uso especificado.

- **Segurança:** capacidade do produto de software de apresentar níveis aceitáveis de riscos de danos a pessoas, negócios, software, propriedades ou ao ambiente, em um contexto de uso especificado.
- **Satisfação:** capacidade do produto de software de satisfazer usuários, em um contexto de uso especificado

Em síntese, a ISO/IEC 9126 é uma norma composta por um conjunto de características que devem ser verificadas em um software para que ele seja considerado um “software de qualidade”, bem como, métricas usadas na sua avaliação (medição, pontuação e julgamento) dos softwares.

3.3.2.2 ISO/IEC 14598-1 - Tecnologia de informação – Avaliação de produto de software

O processo de avaliação da ISO/IEC 14598 é baseado na norma ISO/IEC 9126, a qual estabelece as métricas de qualidade de software e pode ser usado tanto para avaliar produtos prontos como produtos em desenvolvimento.

A avaliação de produtos de software deve ser objetiva, ou seja, baseada em observação e não em opinião. Também deve ser reprodutiva, de forma que avaliações do mesmo produto, para a mesma especificação de avaliação, executadas por diferentes avaliadores produzam resultados aceitos como idênticos e repetíveis. Para isso, um processo de avaliação deve ser definido. Este processo deve seguir, basicamente, cinco passos: análise dos requisitos de avaliação, especificação da avaliação, projeto e planejamento da avaliação, execução da avaliação e documentação dos resultados. A série de normas ISO/IEC 14598 descreve um processo para avaliação de produtos de software, que consiste em quatro passos: análise especificação, projeto e execução (TABELA 3.1). O padrão definido distingue três perspectivas de avaliação: desenvolvedor, adquirente e avaliador [ISO 14598, 1998].

Tabela 3.1 - Definição do processo de avaliação segundo a ISO/IEC 14598.

	Análise	Especificação	Projeto	Execução
Processo para desenvolvedores	Definição de requisitos de qualidade e análise de sua exequibilidade	Quantificação dos requisitos de qualidade	Planejamento da avaliação durante o desenvolvimento	Monitoramento da qualidade e controle durante o desenvolvimento
Processo para adquirentes	Estabelecimento do propósito e escopo da avaliação	Definição de métricas externas e medições correspondentes a serem realizadas	Planejar, programar e documentar a avaliação	A avaliação deveria ser realizada, documentada e analisada
Processo para avaliadores	Descrição dos objetivos da avaliação	Definição do escopo da avaliação e das medições	Documentação dos processos a serem usados pelo avaliador	Obtenção dos resultados a partir da realização de ações de medição e verificação do produto

O processo de avaliação da ISO/IEC 14598 é realizado em seis etapas, sendo a primeira uma abordagem geral de como é realizada a avaliação, seguida do planejamento

e intercursos, dos papéis dos desenvolvedores, dos adquirentes e dos avaliadores, e por último da documentação dos módulos de avaliação.

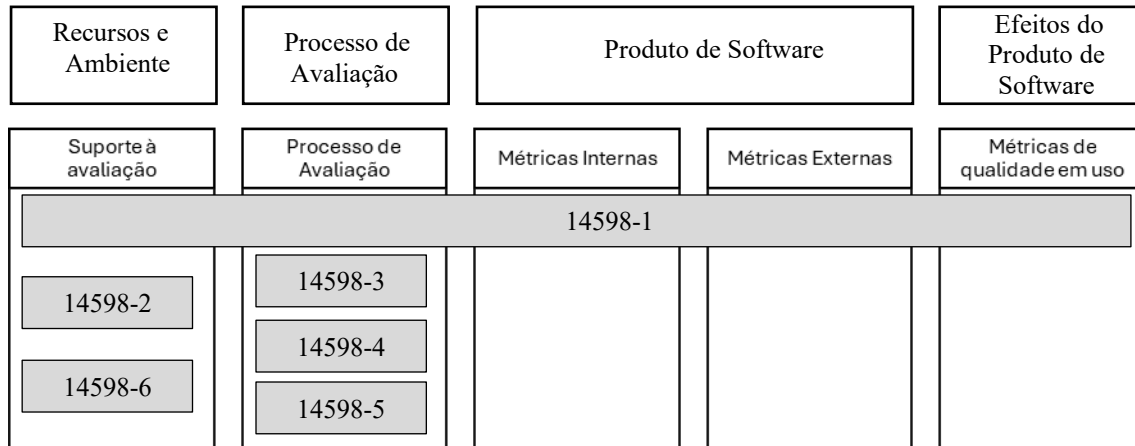


Figura 3.5 - Estrutura ISO/IEC 14598

A série de normas ISO/IEC 14598 é dividida em seis partes (FIGURA 3.5 e FIGURA 3.6):

- **ISO/IEC 14598-1 - Visão geral:** fornece uma visão geral do processo de avaliação da qualidade dos produtos de software e define toda a estrutura de funcionamento da série de normas ISO/IEC 14598.
- **ISO/IEC 14598-2 - Planejamento e gestão:** refere-se ao planejamento e gestão do processo de avaliação apresentando requisitos, recomendações e orientações para uma função de suporte ao processo
- **ISO/IEC 14598-3 - Processo para desenvolvedores:** define o processo para desenvolvedores. Destina-se ao uso durante o processo de desenvolvimento e manutenção de software
- **ISO/IEC 14598-4 - Processo para adquirentes:** define o processo para adquirentes, estabelecendo um processo sistemático para avaliação de: produtos de software tipo pacote (com equivalência a NBR ISO/IEC 12119), produtos de software sob encomenda, ou ainda modificações em produtos já existentes
- **ISO/IEC 14598-5 - Processo para avaliadores:** define o processo para avaliadores, fornecendo orientações para a implementação prática de avaliação de produtos de software (quando diversas partes necessitam entender, aceitar e confiar em resultados da avaliação)
- **ISO/IEC 14598-6 - Documentação de módulos para avaliação:** fornece orientação para documentação de módulos de avaliação. Estes módulos contêm a especificação do modelo de qualidade, as informações e dados relativos à aplicação prevista do modelo e informações sobre a real aplicação do modelo

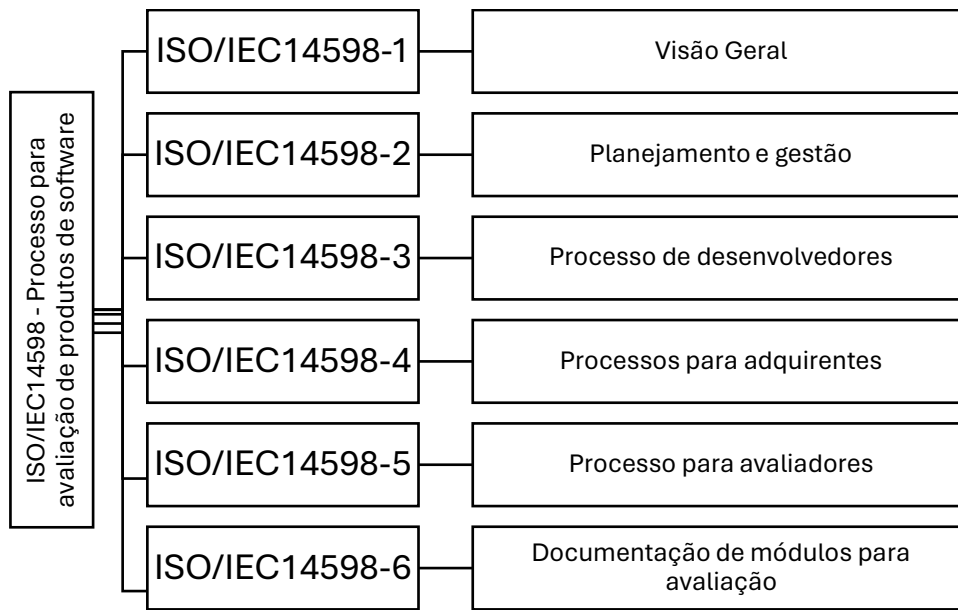


Figura 3.6 - Estrutura da série de normas ISO/IEC14598

Diante dessas séries da norma, compreende-se que para avaliar a qualidade do software, primeiramente devem ser estabelecidos os requisitos da avaliação, para então poder especificar, projetar e executar a avaliação, como exposto na FIGURA 3.7:

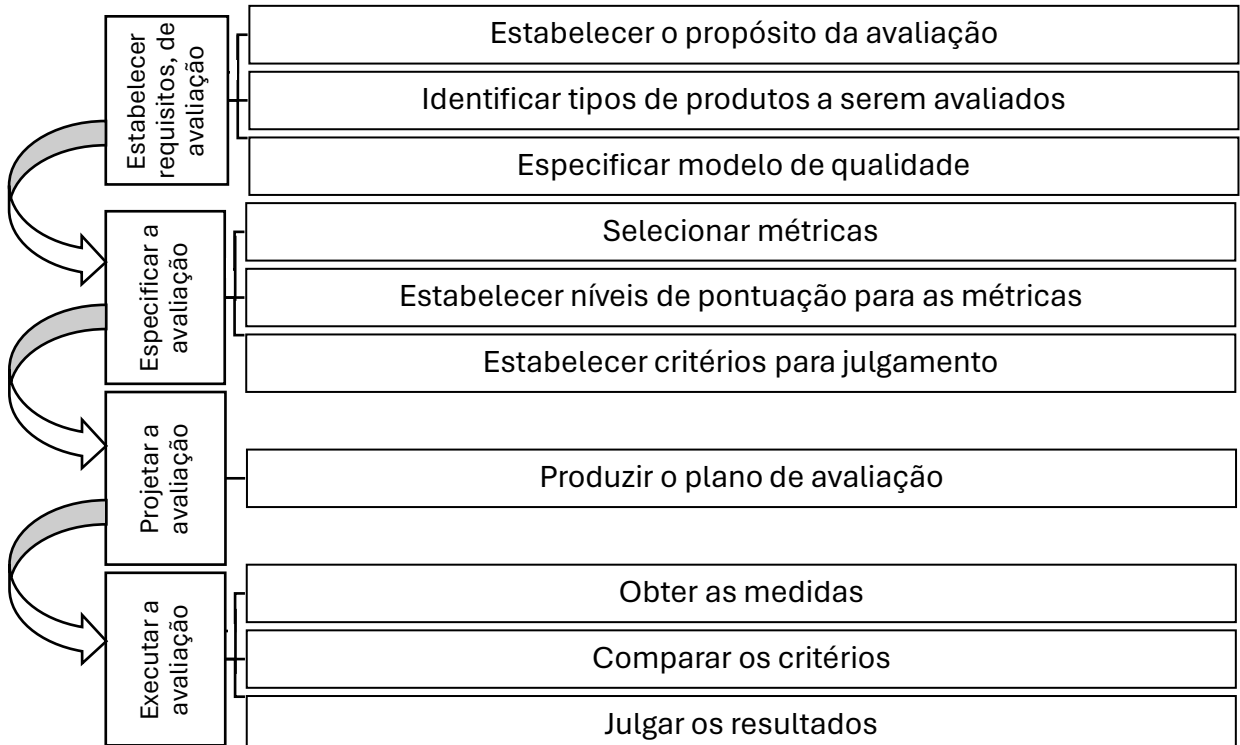


Figura 3.7 - Processo de avaliação segundo a ISO/IEC 14598-1

Compreende-se que a intenção da norma ISO 14598 é estabelecer uma padronização nos módulos de avaliação para que eles possam ser reusáveis. Criam-se assim, bibliotecas desses módulos, fornecendo também guias para documentação dos módulos e suporte ao seu desenvolvimento. O modelo com que as informações são manipuladas ou quais informações são necessárias na realização da avaliação, ilustram o processo de documentação dos módulos de avaliação.

3.3.2.3 ISO/IEC 25000 - SQuaRE (Requisitos e Avaliação de Qualidade de Sistemas e Software)

Seguindo a ISO 25000, que unifica as normas ISO/IEC 9126 e ISO/IEC 14598, desenvolver os processos principais: especificação de requisitos e a avaliação da qualidade.

O ISO 25000 é um padrão bem conhecido no campo da qualidade de software, que inclui o padrão ISO/IEC 25030 que ajuda a especificar os requisitos de qualidade de software. Os padrões ISO visam garantir que o desenvolvimento de software atenda a critérios de alta qualidade e produza aplicativos eficazes e versáteis. Esses padrões fornecem uma estrutura para entender, medir e prever a qualidade do software, facilitando para as organizações o desenvolvimento de produtos de software de qualidade. O trabalho da ISO na padronização da qualidade de software ajuda as organizações a reduzirem os custos de manutenção e melhorar a qualidade geral de seus produtos de software. Seguindo os padrões ISO, as organizações podem estabelecer uma base para a qualidade desejada e se beneficiar de um conceito contínuo e hierárquico de qualidade em todo o processo de desenvolvimento de software [KUI; ALI; SURYN, 2016] [ZHANG et al., 2018] [BOUGROUN et al., 2020] [BARZEGAR, 2021].

A série de normas ISO/IEC 25000:2014 (Software Quality Requirements and Evaluation, SQuaRE), dividida em cinco partes, foi desenvolvida para fornecer suporte à definição de requisitos e ao processo de avaliação da qualidade de softwares. O modelo de qualidade apresentado na ISO/IEC 25010:2014 aborda duas dimensões da qualidade de software.

A primeira delas é a qualidade em uso, que especifica características de qualidade relacionadas à interação humana com o sistema. A qualidade em uso se refere à capacidade de atender a requisitos para atingir metas específicas com produtividade, efetividade, segurança e satisfação do usuário, em um contexto de uso especificado (ISO 25000, 2014). A segunda dimensão proposta pela ISO/IEC 25010 é a qualidade do produto, que define um conjunto de oito características da qualidade relacionadas à atributos internos e externos do software. A adequação funcional é a capacidade do software de prover funções que atendam às necessidades implícitas e explícitas quando usado em condições especificadas. A eficiência de desempenho se refere ao desempenho em relação à quantidade de recursos utilizados. A confiança é a capacidade de manter um nível de desempenho especificado em um determinado período de tempo. A usabilidade é a capacidade de ser usado para atingir metas específicas com efetividade, produtividade e satisfação do usuário. A manutenibilidade é a capacidade de ser modificado visando à melhoria, correção ou adaptação a mudanças no ambiente ou nos requisitos. A portabilidade é a capacidade de ser transferido de um hardware, sistema operacional ou ambiente de uso para outro. A compatibilidade é a capacidade de trocar informações com outros softwares e desempenhar as funções que lhe forem requeridas enquanto

compartilha recursos de hardware e software. A segurança diz respeito à capacidade do software de proteger informações para que pessoas ou outros softwares tenham o nível de acesso apropriado aos seus níveis de permissão [ISO 25000, 2014]. Conforme descreve a Tabela 3.2, o modelo de qualidade apresentado na ISO/IEC 25010 desdobra as características da qualidade em um conjunto de subcaracterísticas relacionadas.

Tabela 3.2 - Características e subcaracterísticas da qualidade de software.

CARACTERÍSTICA	SUBCARACTERÍSTICA	DEFINIÇÃO DA SUBCARACTERÍSTICA
Adequação funcional	Integridade funcional	Capacidade de atender às tarefas e aos objetivos específicos do usuário a que foi destinado
	Correção funcional	Capacidade de fornecer resultados corretos e com precisão
	Adequação funcional	Capacidade de realizar tarefas e certos objetivos de maneira fácil
Eficiência de desempenho	Comportamento em relação ao tempo	Capacidade de fornecer tempos de resposta e processamento apropriados quando o software executa suas funções
	Utilização de recursos	Grau com que os tipos e as quantidades de recursos usados atendam aos requisitos
	Capacidade	Grau com que a capacidade máxima de parâmetros do software atende aos requisitos
Usabilidade	Reconhecibilidade	Grau com que o usuário é capaz de reconhecer se o produto é adequado às suas necessidades
	Apreensibilidade	Capacidade que o software possui de ser usado para alcançar objetivos específicos de aprendizagem, de forma eficiente, eficaz e sem riscos, garantindo que o usuário se sinta satisfeito no contexto em questão
	Operacionalidade	Capacidade do software de permitir ao usuário operá-lo e controlá-lo de forma fácil
	Proteção de erro	Capacidade de proteger os usuários de cometer falhas
	Estética da interface	Capacidade de possuir uma interface que seja satisfatória ao usuário
	Acessibilidade	Capacidade de ser usado por usuários com diferentes características e habilidades para alcançar objetivos especificados

CARACTERÍSTICA	SUBCARACTERÍSTICA	DEFINIÇÃO DA SUBCARACTERÍSTICA
Compatibilidade	Coexistência	Capacidade de compartilhar recursos com outro software sem causar impactos sobre qualquer outro produto
	Interoperabilidade	Capacidade de interagir com um ou mais sistemas especificados
Confiança	Maturidade	Capacidade de atender às necessidades de confiabilidade quando operado em condições normais
	Disponibilidade	Capacidade de ser operacional e acessível quando requerido para uso
	Tolerância a falhas	Capacidade de garantir um nível de desempenho especificado em caso de ocorrência de falhas de software ou hardware
	Capacidade de recuperação	Capacidade de restabelecer seu nível de desempenho especificado e recuperar os dados diretamente afetados no caso de uma falha
Segurança	Confidencialidade	Capacidade de permitir acesso de dados somente a usuários autorizados
	Integridade	Capacidade de bloquear acesso e modificações de usuários não autorizados
	Não repúdio	Capacidade de comprovar ações, eventos, alterações e envio de informações para que não possam ser repudiados futuramente
	Responsabilidade	Capacidade de rastrear as ações de entidades específicas
	Autenticidade	Capacidade de comprovar a identidade de um sujeito ou recurso caso seja requerido
Manutenibilidade	Modularidade	Capacidade de alterar elementos do software com impacto mínimo
	Reutilização	Capacidade que os componentes do software possuem de serem utilizados por outros sistemas existentes ou em construção
	Analisabilidade	Capacidade de avaliar o impacto de uma mudança em uma ou mais partes de um sistema, diagnosticar partes do sistema em que pode haver falhas e identificar componentes a serem modificados
	Modificabilidade	Capacidade de permitir que uma modificação seja implementada sem causar defeitos no produto existente

CARACTERÍSTICA	SUBCARACTERÍSTICA	DEFINIÇÃO DA SUBCARACTERÍSTICA
	Testabilidade	Capacidade de estabelecer critérios de teste para um sistema que foi modificado e de determinar se esses critérios foram cumpridos de forma eficaz e eficiente
Portabilidade	Adaptabilidade	Capacidade de ser adaptado para ambientes de operação especificados, sem a necessidade de aplicação de outras ações ou meios além daqueles fornecidos para essa finalidade pelo software considerado
	Instalabilidade	Capacidade de ser corretamente instalado e/ou desinstalado em um ambiente especificado
	Substituibilidade	Capacidade de substituir outro software no mesmo ambiente e para o mesmo fim

Embora a ISO/IEC 25010 sugira um conjunto de 31 subcaracterísticas da qualidade, a realização da avaliação de um ou mais softwares usando todas essas subcaracterísticas pode se tornar inviável devido ao tempo consumido. Dessa forma, essas características e subcaracterísticas devem servir de referência para a criação de modelos de avaliação adequados ao tempo, aos recursos humanos e às informações disponíveis [JUNG, 2007].

Na literatura podem ser encontrados diversos modelos de avaliação da qualidade para tipos de software específicos. Silva et al. [2015] desenvolveram um modelo para avaliar a qualidade de um sistema acadêmico com base nos critérios facilidade de uso, comportamento em relação a falhas, eficiência e satisfação das necessidades. Lima Junior e Martimiano [2010] propuseram um modelo de apoio à avaliação da qualidade de softwares PMIS com base na norma ISO/IEC 9126. Uma limitação desse modelo é a impossibilidade de avaliar os níveis de importância relativa (ou pesos) das características e subcaracterísticas da qualidade. Como os requisitos dos usuários variam de acordo com o contexto de aplicação do software, é desejável atribuir pesos diferentes para as características da qualidade [YANG, 2012]. Esse processo de ponderação pode ser apoiado por métodos de tomada de decisão multicritério [TRENDOWICZ; KOPCZYNSKA, 2014].

3.4 Avaliação de Tecnologia em Saúde Digital

A quantidade de medicamentos, dispositivos, procedimentos e outras tecnologias em saúde é vasta, e há um contínuo desenvolvimento de novas tecnologias para a atenção à saúde. Esse desenvolvimento de novas tecnologias contribui para a melhoria da qualidade de vida, e pressionam os sistemas de saúde a incorporá-las, que contam com recursos limitados. A inovação, regulação e garantia de acesso de tecnologias em saúde são fatores importantes para a sustentabilidade dos sistemas de saúde.

Neste cenário, torna-se necessário que haja procedimentos para nortear o desenvolvimento e adoção dessas novas tecnologias de forma adequada e segura nos sistemas de saúde, garantindo efetividade, segurança e eficácia, ao mesmo tempo que

aquelas tecnologias que não são tão eficazes como outras, possam ser substituídas por estas.

Segundo a Portaria Nº 2.510/GM de 19 de dezembro de 2005 [GABINETE DO MINISTRO, 2005], considera-se tecnologias em saúde: medicamentos, materiais, equipamentos e procedimentos, sistemas organizacionais, educacionais, de informações e de suporte, e programas e protocolos assistenciais, por meio dos quais a atenção e os cuidados com a saúde são prestados à população.

No tocante à inovação, o desenvolvimento de tecnologias da informação e comunicação ajudou a cunhar o termo saúde digital. A saúde digital se refere ao uso de tecnologia para diagnosticar, monitorar, tratar e prevenir doenças. Abrange vários avanços, como saúde móvel, registros eletrônicos de saúde, dispositivos vestíveis, telessaúde, inteligência artificial e robótica [BUCHNER, 2022].

O conceito de saúde digital está transformando os modelos tradicionais de prestação de cuidados de saúde, levando a uma transformação cultural no campo [MESKÓ et al., 2017]. No entanto, a implementação da saúde digital enfrenta desafios como regulamentações rígidas, resistência à mudança e a necessidade de considerações culturais e de fatores humanos [SHAFFER; PREZIOSI, 2019]. Além disso, há riscos associados à saúde digital, incluindo ameaças cibernéticas, perda de privacidade e disseminação de desinformação médica [SIMON LIN LINWOOD, 2022].

No geral, a saúde digital tem o potencial de revolucionar a saúde, mas uma consideração cuidadosa de seus riscos e benefícios é fundamental. A afirmação destaca a transformação potencialmente revolucionária proporcionada pela saúde digital no cenário da saúde. A expressão "saúde digital" refere-se à aplicação de tecnologias digitais, como dispositivos móveis, softwares e plataformas online, para melhorar a prestação de serviços de saúde. Embora essa revolução prometa avanços significativos, é importante abordá-la com uma consideração cuidadosa dos riscos e benefícios envolvidos. Isso implica avaliar minuciosamente as implicações éticas, privacidade dos dados, segurança cibernética e equidade no acesso, garantindo que o impacto positivo na saúde seja maximizado, ao mesmo tempo em que se minimizam possíveis desafios e ameaças. A busca por inovações na saúde digital deve, portanto, ser guiada por uma abordagem equilibrada e responsável para garantir seu pleno potencial em benefício da sociedade.

A saúde digital está relacionada à avaliação de várias maneiras. Em primeiro lugar, tecnologias digitais de saúde e dispositivos de biotelemetria estão sendo usados para prevenir, detectar, diagnosticar, monitorar e prever a evolução de várias doenças, sem exigir procedimentos invasivos ou interação face a face com o pessoal médico [BUCHNER, 2022].

Em segundo lugar, métodos e sistemas de avaliação de saúde digital discretos foram desenvolvidos, os quais integram características físicas, hábitos de vida e condições médicas prevalentes com atividades físicas monitoradas e medidas fisiológicas para avaliar a saúde de indivíduos de alto risco.

Em terceiro lugar, o impacto da digitalização dos serviços de saúde precisa ser avaliado com base no alcance das metas do sistema de saúde, incluindo qualidade, acessibilidade, eficiência e equidade [EUROPEAN COMMISSION. DIRECTORATE GENERAL FOR HEALTH AND FOOD SAFETY, 2019].

Além disso, a saúde digital tem beneficiado muito o sistema público de saúde, e a acessibilidade e a confiança dos pacientes nas informações da Internet influenciam suas decisões e avaliações dos profissionais de saúde [VUONG et al., 2022].

Finalmente, a Avaliação de Tecnologia de Saúde (ATS) de aplicativos móveis de saúde, incluindo aplicativos digitais de saúde, envolve uma avaliação rápida de sua eficácia e potencial de economia de custos para informar os tomadores de decisão no sistema de saúde [LESQUELEN et al., 2017].

Esta subseção apresenta os principais aspectos relacionados à avaliação de tecnologia em saúde digital.

3.4.1 Definição de Tecnologia em Saúde

O termo Tecnologias em Saúde abrange produtos, procedimentos e insumos utilizados nos serviços de saúde, e abrange medicamentos, vacinas, testes diagnósticos, equipamentos e dispositivos, modelos de organização de serviços de saúde, bem como os sistemas de apoio.

Segundo a Organização Mundial da Saúde, o termo Tecnologias de Saúde se aplica a próteses, diagnósticos in vitro, equipamentos de imagem e soluções de e-saúde utilizados para diagnosticar, monitorar, avaliar, prevenir e indicar o tratamento aos pacientes que sofrem de uma vasta gama de doenças. Reconhecendo que há muitas interpretações para os termos que tratam do setor, são apresentadas abaixo as definições estabelecidas pela Organização Mundial de Saúde (OMS) e pela Global Harmonization Task Force (GHTF) [SAÚDE 4.0, 2015].

- **Tecnologias médicas ou Dispositivos Médicos no sentido Amplo (DMAs):** envolvem a aplicação de conhecimento organizado e de práticas médicas sob a forma de aparelhos, procedimentos e sistemas desenvolvidos para identificar, prevenir e solucionar um problema de saúde e melhorar a qualidade de vida. Os DMAs incluem uma ampla variedade de produtos, indo desde os mais simples, como luvas, linhas de sutura, adesivos, camas hospitalares e lentes, até aplicativos para smartphones, implantes cardíacos, monitores de glicose e scanners de ressonância magnética. O prolongamento da vida mais ativa e independente, que vemos nos dias de hoje, é em larga medida o resultado das inovações na área da tecnologia médica. Abrangem qualquer instrumento, máquina, aparelho, implante, reagente para uso in vitro, software, material ou outro artigo similar ou relacionado, destinado pelo fabricante a ser utilizado, isoladamente ou em combinação pelos seres humanos, para um ou mais dos efeitos médicos específicos de:
 - diagnóstico, prevenção, controle, tratamento ou atenuação de uma doença;
 - diagnóstico, controle, tratamento, atenuação ou compensação de uma lesão;
 - estudo, substituição, modificação ou apoio da anatomia ou de um processo fisiológico;
 - apoiar ou sustentar a vida;
 - controle da concepção;
 - desinfecção de dispositivos médicos;
 - fornecimento de informações por meio de exame in vitro de amostras derivadas do corpo humano.

- **Dispositivos médicos no sentido restrito:** são artigos, instrumentos, aparatos ou máquinas utilizados na prevenção, diagnóstico ou tratamento de sintoma ou doença, com o objetivo de detectar, medir, restaurar, corrigir ou modificar a estrutura ou função do corpo para algum propósito de saúde. Tipicamente, o propósito de um dispositivo médico não é alcançado por meios farmacológicos, imunológicos ou metabólicos.
- **Equipamentos médicos:** são aparelhos médicos que requerem calibração, manutenção, reparo e treinamento de quem o utiliza. O equipamento médico é usado com o propósito específico de diagnosticar e tratar um problema de saúde. Pode ser usado sozinho ou em combinação com acessórios e produtos para consumo, e, dessa definição, são excluídos os produtos para a saúde implantáveis, descartáveis ou de uso único.
- **Diagnóstico in vitro:** que trata-se de qualquer dispositivo médico que consista num reagente, instrumento, aparelho ou sistema usado, isoladamente ou combinado, a ser utilizado in vitro no exame de amostras provenientes do corpo humano, de forma a obter informações sobre estados fisiológicos, de saúde, de doença, ou de anomalia congênita.
- **E-Saúde:** abrange as áreas e produtos para a saúde que se utilizam das tecnologias de informação e comunicação para diagnóstico, atendimento, automonitoramento, educação e prestação de serviços remotos em saúde. A informatização na saúde, tanto no controle de custos e na organização dos procedimentos, como na aferição da qualidade do atendimento ao paciente, configura-se como objeto de grande interesse em face dos significativos investimentos privados e públicos que requer e do seu potencial de bons resultados.

Na literatura são encontradas algumas formas de classificação de tecnologias de saúde. Em [LIAROPOULOS, 1997] é proposto uma hierarquia de tecnologias de saúde.

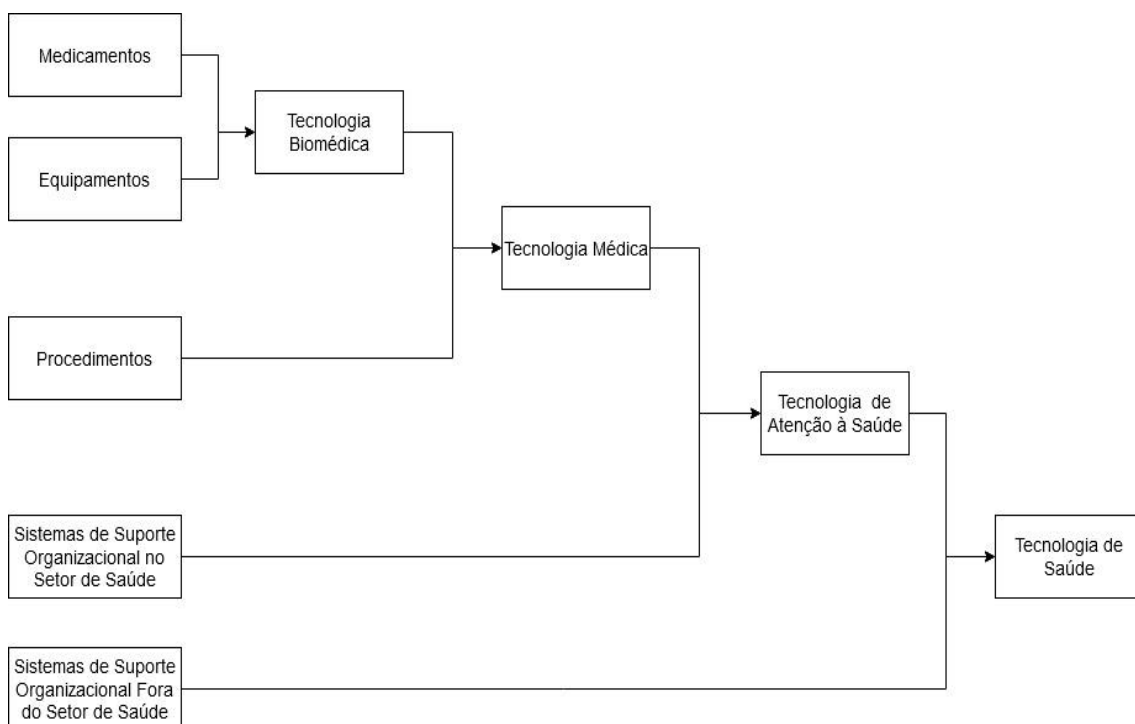


Figura 3.8 - Hierarquia de tecnologias de saúde

Medicamentos e equipamentos que interagem diretamente com pacientes são designados como tecnologias biomédicas. Os procedimentos, como anamnese, técnicas cirúrgicas, normas técnicas de uso de tecnologias biomédicas, são designadas, juntamente com as tecnologias biomédicas, tecnologias médicas. O uso das tecnologias médicas está inserido dentro de uma estrutura técnica e administrativa, contando também com sistemas de informação, denominados como Sistemas de Suporte Organizacional no Setor de Saúde. Os Sistemas de Suporte Organizacional no Setor de Saúde e as Tecnologias Médicas são designadas Tecnologias de Atenção à Saúde.

Por fim, há componentes organizacionais e de apoio que atuam fora do setor de saúde e que de alguma forma influenciam a saúde como, por exemplo, saneamento, controle ambiental e leis trabalhistas.

O uma outra classificação é proposta em [GOODMAN, 1998], de acordo com a natureza do material, o propósito, e o estágio de difusão:

- Medicamentos:
 - Equipamentos e suprimentos como ventiladores, marcapassos cardíacos, luvas cirúrgicas, kits de diagnóstico, etc;
 - Procedimentos médicos e cirúrgicos;
 - Sistemas de suporte: bancos de sangue, sistemas de prontuário eletrônico, etc.;
 - Sistemas gerenciais e organizacionais: sistema de informação, sistema de garantia de qualidade, etc.
- Quanto ao propósito:
 - Prevenção: visa proteger os indivíduos contra uma doença ou limitar a extensão de uma seqüela (exemplo: imunização, controle de infecção hospitalar, etc.);
 - Triagem: visa detectar a doença, anormalidade, ou fatores de risco em pessoas assintomáticas (mamografia, exame de Papanicolau);
 - Diagnóstico: visa identificar a causa e natureza ou extensão de uma doença em pessoas com sinais clínicos ou sintomas (eletrocardiograma, raios X para detectar fraturas ósseas);
 - Tratamento: visa melhorar ou manter o estado de saúde, evitar uma deterioração maior ou atuar como paliativo;
 - Reabilitação: visa restaurar, manter ou melhorar a função de uma pessoa com uma incapacidade física ou mental.
- Quanto ao estágio de difusão:
 - Futura: em estágio de concepção ou nos estágios iniciais de desenvolvimento;
 - Experimental: quando está submetida a testes em laboratório usando animais ou outros modelos;
 - Investigacional: quando está submetida a avaliações clínicas iniciais (em humanos);
 - Estabelecida: considerada pelos provedores como um enfoque padrão para uma condição particular e difundida para uso geral;
 - Obsoleta/abandonada/desatualizada: sobrepujada por outras tecnologias ou foi demonstrado que elas são inefetivas ou prejudiciais.

Embora não haja uma classificação padrão para tecnologias de saúde, nota-se que o termo de tecnologia de saúde é bem amplo. A próxima sub-seção é apresentada a diferença entre dispositivo médico e aplicativo de saúde.

3.4.2 Diferença entre dispositivo médico e aplicativo de saúde

Segundo a Resolução da Diretoria Colegiada (RDC) 751/2022 da Anvisa [MINISTÉRIO DA SAÚDE, 2022b] que dispõe sobre classificação de riscos, os regimes de notificação e de registro, e os requisitos de rotulagem e instruções de uso de dispositivos médicos, um Dispositivo Médico é qualquer instrumento, aparelho, equipamento, implante, dispositivo médico para diagnóstico in vitro, software, material ou outro artigo, destinado pelo fabricante a ser usado, isolado ou conjuntamente, em seres humanos, para algum dos seguintes propósitos médicos específicos, e cuja principal ação pretendida não seja alcançada por meios farmacológicos, imunológicos ou metabólicos no corpo humano, mas que podem ser auxiliados na sua ação pretendida por tais meios:

1. diagnóstico, prevenção, monitoramento, tratamento (ou alívio) de uma doença;
2. diagnóstico, monitoramento, tratamento ou reparação de uma lesão ou deficiência;
3. investigação, substituição, alteração da anatomia ou de um processo ou estado fisiológico ou patológico;
4. suporte ou manutenção da vida;
5. controle ou apoio à concepção; ou
6. fornecimento de informações por meio de exame in vitro de amostras provenientes do corpo humano, incluindo doações de órgãos e tecidos.

Já para a OMS, por meio do GHTF [SAÚDE 4.0, 2015], um dispositivo médico é definido como qualquer instrumento, aparelho, equipamento, software, material ou outro produto que possa ser utilizado, sozinho ou combinados, em seres humanos, para um ou mais dos seguintes propósitos médicos específicos:

1. diagnóstico, prevenção, monitoramento, tratamento ou alívio de doenças,
2. diagnóstico, monitoramento, tratamento, alívio ou compensação de uma lesão,
3. investigação, substituição, modificação ou suporte da anatomia ou de um processo fisiológico,
4. apoiando e sustentando a vida,
5. controle da concepção,
6. desinfecção de dispositivos médicos,
7. fornecimento de informações através de exames in vitro de amostras derivadas do corpo humano.

De acordo com a OMS, saúde móvel [MAASS et al., 2022] é um termo geral que abrange a saúde pública e a prática médica através do uso de smartphones, sensores, assistentes pessoais digitais, dispositivos de monitoramento sem fio ou outros dispositivos sem fio. Dentre as tecnologias de saúde móvel, os aplicativos móveis de saúde são softwares para smartphones e tablets que podem promover a saúde das pessoas. Esses aplicativos podem ser utilizados para suporte a pessoas com doenças crônicas, melhorar a adesão a tratamentos, aumentar a autonomia dos pacientes, além de oferecerem oportunidades de diminuição de custos sem comprometer a qualidade dos serviços de saúde.

Pelas definições tanto da OMS quanto da Anvisa, um aplicativo móvel de saúde pode ser considerado um dispositivo médico. Segundo a RDC 657/2022 da Anvisa [MINISTÉRIO DA SAÚDE, 2022a] que dispõe sobre a regularização de softwares como dispositivos médicos, um software é considerado um dispositivo médico (Software as a Medical Device - SaMD) caso ele atenda à definição de dispositivo médico, podendo ser de diagnóstico *in vitro* (IVD) ou não, sendo destinado a uma ou mais indicações médicas, e que realizam essas finalidades sem fazer parte de hardware de dispositivo médico. Inclui os aplicativos móveis e softwares com finalidades *in vitro*, se suas indicações estiverem incluídas na definição geral de dispositivos médicos. Incluem-se nesta definição, entre outros, o software licenciado por assinatura e hospedado centralmente (Software as a Service), que se enquadre na definição de dispositivos médicos.

A RDC 751/2022 da Anvisa [MINISTÉRIO DA SAÚDE, 2022b] acrescenta características de um SaMD: o SaMD pode ser executado em uma plataforma computacional de propósito geral (finalidade não médica); a "plataforma computacional" inclui recursos de hardware e software (sistema operacional, hardware de processamento, armazenamento, banco de dados, dispositivos de visualização, dispositivos de entrada, linguagem de programação, etc.); "sem fazer parte de" significa que o programa não precisa do hardware de um dispositivo médico para atingir sua finalidade de uso; um software não é considerado SaMD se seu objetivo é controlar o hardware de um dispositivo médico; um SaMD pode ser usado em combinação (por exemplo, como módulo) com outros produtos, incluindo outros dispositivos médicos; um SaMD pode interagir com outros dispositivos médicos, incluindo hardware de outros dispositivos médicos e outro SaMD, bem como software de uso geral; e os aplicativos móveis de saúde que atendem à definição são considerados SaMD.

Pela definição apresentada de SaMD, nem todos os aplicativos móveis de saúde podem ser considerados um SaMD. A RDC 657/2022 da Anvisa [MINISTÉRIO DA SAÚDE, 2022a], segundo o parágrafo 2 do seu artigo 1, não se aplica aos seguintes softwares: softwares de bem estar (softwares destinados a encorajar e manter o bem-estar, incluindo atividades saudáveis como exercícios físicos, ou a encorajar e manter o controle da saúde e um estilo de vida saudável que não são destinados à prevenção, diagnóstico, tratamento, reabilitação ou anticoncepção); relacionados em lista disponibilizada pela Anvisa de produtos não regulados; utilizado exclusivamente para gerenciamento administrativo e financeiro em serviço de saúde; que processa dados médicos demográficos e epidemiológicos, sem qualquer finalidade clínica diagnóstica ou terapêutica; e embarcado em dispositivo médico sob regime de vigilância sanitária.

A discussão sobre a distinção de aplicativos de saúde categorizados como SaMD daqueles que não são categorizados como SaMD é foco do trabalho apresentado em MAASS [2022], onde os aplicativos médicos correspondem ao SaMD dentro de um e os aplicativos de saúde correspondem aos não SaMD. Nele, são apresentadas as definições de aplicativos médicos e de aplicativos de saúde em diversos países, e evidencia que os SaMD são um grupo particular dos aplicativos de saúde, os quais requerem regulação legal para serem utilizados como dispositivos médicos.

Essa distinção de SaMD como um subgrupo de aplicativos de saúde é importante para a sua adequada regulação. A RDC 848/2024 da Anvisa que dispõe sobre os requisitos essenciais de segurança e desempenho aplicáveis aos dispositivos médicos e dispositivos médicos para diagnóstico *in vitro*, descreve em sua subseção VIII um conjunto de itens a

serem atendidos pelos desenvolvedores de SaMD. Uma vez que aplicativos de saúde podem ser categorizados como SaMD, é importante instrumentos para averiguar se tais aplicativos atendem às resoluções e recomendações estabelecidas. A próxima sub-seção apresenta brevemente os principais aspectos sobre Avaliação de Tecnologias de Saúde, enfatizando a Avaliação de Tecnologias Digitais em Saúde.

3.4.3 Importância e Impacto da Avaliação de Tecnologia em Saúde Digital

Para concretizar os benefícios e minimizar os riscos, é essencial um processo de avaliação robusto para informar os decisores sobre o financiamento e o planejamento de todas as novas tecnologias de saúde, incluindo as intervenções TDS. A avaliação das tecnologias de saúde (ATS) oferece um quadro para avaliar de forma abrangente as novas tecnologias de saúde. Muitos países utilizam habitualmente a ATS para determinar a relação custo-benefício e o impacto orçamental das novas tecnologias de saúde para informar as decisões de financiamento/investimento público.

Assim é preciso a adoção de mecanismos de regulação e de tecnologias que sejam embasadas por evidências científicas. A Lei 12.401 de 2011 [PRESIDÊNCIA DA REPÚBLICA - CASA CIVIL, 2011] estabelece que o uso de evidências de eficácia, segurança e custo-efetividade deve nortear a incorporação de tecnologias no SUS. A Avaliação de Tecnologias em Saúde (ATS) busca subsidiar as decisões quanto à adoção de tecnologias pelos sistemas de saúde. Em Goodman [1998], a ATS é descrita como um campo multidisciplinar de análise de políticas, que estuda as implicações clínicas, sociais, éticas e econômicas do desenvolvimento, difusão e uso da tecnologia em saúde.

Uma vez que tecnologia em saúde abrange uma grande variedade de componentes como medicamentos, procedimentos, instrumentos e equipamentos, é preciso compreender as dimensões a serem avaliadas [ELIAS, 2013]:

- **Eficácia:** benefício das tecnologias em condições ideais de utilização, e os resultados são avaliados pela interpretação de risco relativo, risco absoluto, redução de risco absoluto, e quantidade de vezes que a tecnologia deve ser aplicada para tratamento;
- **Segurança:** existência de eventos adversos provenientes da tecnologia;
- **Efetividade:** benefício das tecnologias em condições reais de utilização dos serviços;
- **Custo-benefício:** análise comparada entre os custos e os resultados em saúde gerados pela tecnologia. Busca estimar em que medida o resultado gerado compensa os custos decorrentes de sua utilização;
- **Impacto no orçamento:** estimativa do aumento ou redução de gastos ao se introduzir e difundir a tecnologia no serviço de saúde;
- **Equidade:** análise da proporção da população a ser beneficiada em razão de suas necessidades sócio-culturais, biológicas e de gênero;
- **Ética:** adequação ao código de princípios de moralidade definidos pela sociedade ou cultura, considerados ideais no caráter e na cultura;
- **Bases Legais:** compatibilidade e adequação à legislação vigente, necessidade de alteração de normas;
- **Logística:** adequação e necessidades de alteração de infraestrutura, de pessoal, de transporte, armazenamento, e todos os aspectos relacionados a cobertura e acesso de qualidade;

- **Macroeconomia:** análise de alocação de recursos no sistema de saúde e dos efeitos nas políticas de propriedade intelectual, na regulação, no investimento em inovação, na transferência de tecnologias e no aumento ou diminuição de empregos;
- **Meio ambiente:** geração de resíduos poluentes, condições e recursos necessários para mitigação de possíveis danos ao meio ambiente.

Essas dimensões normalmente são ponderadas sob a perspectiva de custo-oportunidade, investigando as consequências de curto e médio prazos, avaliando os efeitos desejados, indesejados e incertezas em torno da adoção da tecnologia.

As tecnologias digitais de saúde (TDS) que permitem a monitorização remota ou a autogestão dos pacientes oferecem muitos benefícios potenciais, especialmente para os pacientes que sofrem de doenças crônicas. Exemplos de potenciais benefícios incluem a redução dos tempos de viagem e de espera clínica, redução da exposição e propagação de infecções, ganhos de confiança nas condições de autogestão e melhor ligação do paciente com a sua equipa de saúde em momentos críticos.

No entanto, a utilização de TDSs como intervenção de saúde pode representar uma mudança significativa na prestação de serviços de saúde. Os benefícios só podem ser obtidos se o pessoal for apoiado por facilitadores apropriados, por exemplo, recursos adequados e apoio à gestão de mudanças, interoperabilidade de sistemas e gestão adequada da qualidade dos dados, financiamento apropriado e defensores da nova intervenção dentro da organização

Além disso, com os benefícios potenciais vêm riscos exclusivos dos TDSs, como ameaças à privacidade do paciente devido à segurança cibernética inadequada, uso de informações incorretas do paciente devido à má interoperabilidade/gestão da qualidade dos dados, falta de acesso em áreas de conectividade/infraestrutura deficiente e ameaças éticas. como uma falsa sensação de segurança por ser monitorado ou interpretação incorreta dos resultados dos testes sem supervisão médica [MAAß, FREYE, PAN... 2022]. Além disso, os TDSs geralmente coletam uma grande quantidade de dados pessoais, e os TDSs que promovem mudanças comportamentais costumam usar esses dados para desenvolver algoritmos de comportamento preditivo. As empresas de tecnologia também podem usar transmissões de dados com identificadores de usuário vinculáveis de TDSs para fins indesejáveis de marketing e análise.

As ameaças à privacidade resultantes da forma como os TDSs são concebidas são desafios únicos para avaliar os riscos versus benefícios destas tecnologias. Estes riscos únicos devem ser considerados juntamente com os riscos comuns a outras tecnologias, por exemplo, tratamentos menos eficazes, diagnósticos errados e classificação errada da gravidade da doença. Na próxima subseção são abordados brevemente os processos e protocolos de ATS.

3.4.4 Processos e Protocolos de Avaliação de Tecnologia em Saúde Digital

As primeiras iniciativas para estabelecer formas de avaliação para Tecnologias Digitais em Saúde (TDS) estavam inicialmente centradas na telemedicina. Foram desenvolvidas propostas de protocolos baseados em sistemas de pontuação para avaliar a qualidade dos estudos, com adaptações das avaliações ao contexto da telemedicina, englobando participantes de diversas disciplinas relevantes. Por meio do financiamento da Comissão Europeia, foi desenvolvido um quadro específico para a telemedicina, chamado Modelo

de Avaliação de Aplicações de Telemedicina (MAST). Esse modelo incorporou o EUnetHTA (European Network for Health Technology Assessment) Model Core, que é uma rede estabelecida para criar uma estrutura eficaz e sustentável para avaliações de tecnologias em saúde em toda a Europa [EUNETHTA, 2016]. O EUnetHTA Model Core possui nove domínios, com 51 temas e 145 questões. O conteúdo das estruturas incluídas foi mapeado para as 145 questões do Modelo Principal de ATS num processo de duas fases. Inicialmente, foram incluídos tópicos e questões específicas do TDS levantadas pelas estruturas, mas ainda não incluídas no modelo, para garantir uma compilação abrangente do conteúdo do TDS. Para novos tópicos específicos do TDS, novos nomes de tópicos foram propostos, e para novos problemas específicos do TDS, novos identificadores de problemas foram atribuídos usando um prefixo TDS. Posteriormente, todo o conteúdo recomendado por cada framework foi mapeado para o conjunto ampliado de questões.

Por sua vez, a Organização Mundial da Saúde (OMS) produziu o Kit de Ferramentas da Estratégia Nacional de eSaúde em 2012 [OMS, 2012], uma vez que vários países solicitaram orientações acerca do uso de tecnologias de eSaúde. Essas iniciativas resultaram em 2019 na Diretriz da OMS sobre Intervenções Digitais para o Fortalecimento do Sistema de Saúde. No entanto, tal diretriz não contempla as TDSs destinadas ao uso individual do paciente.

No âmbito da regulamentação, o Fórum Internacional de Reguladores de Dispositivos Médicos (IMDRF) estabeleceu um grupo de trabalho sobre "Software como Dispositivo Médico" (SaMD), publicando orientações sobre definições, categorização de riscos, gestão de qualidade e avaliação clínica [IMDRF, 2014]. Estas diretrizes foram adotadas pela Coordenação de Dispositivos Médicos da União Europeia em 2019, ampliando o escopo da regulamentação para todas as tecnologias digitais com finalidade médica.

Apesar desses esforços, havia poucas orientações sobre os critérios de avaliação para o financiamento público das TDSs, embora algumas agências europeias de Avaliação de Tecnologias em Saúde tenham publicado critérios de lista de verificação para avaliar aplicações de saúde em seus sites. O quadro Digi-HTA foi implementado na Finlândia para avaliações de TDSs, enquanto a Portaria sobre Aplicações de Saúde Digital agora regula o reembolso de aplicações de saúde digitais na Alemanha. No entanto, nenhum desses quadros recentes abrange todos os domínios recomendados pelo modelo principal da EUnetHTA e não são informados por estudos sobre as preferências das partes interessadas ao tomar decisões de financiamento/reembolso público sobre TDSs.

O trabalho apresentado em [VON HUBEN et al., 2021], apresenta uma lista de propostas de protocolos de avaliação de tecnologias digitais em saúde. Dada a complexidade do tema, reconhecida pela OMS e Comissão Europeia, o restante desta subseção dedica-se a instrumentos de avaliação de aplicativos de saúde. Especificamente para aplicativos móveis de saúde, existem protocolos específicos de avaliação que serão descritos a seguir.

A *Mobile App Rating Scale* (MARS) é um protocolo padronizado para avaliar a qualidade de aplicativos móveis, especialmente na área da saúde. A MARS contém 23 itens divididos em cinco subescalas que analisam diferentes aspectos dos aplicativos: Engajamento que avalia o quão interessante e atraente o aplicativo é para os usuários, considerando envolvimento, entretenimento, personalização, interatividade e adequação

ao público-alvo; Funcionalidade que avalia o desempenho do aplicativo, navegação, design gestual e facilidade de uso; Estética que mede o apelo visual do aplicativo, incluindo design e gráficos; Qualidade da informação que avalia a precisão e integridade das informações fornecidas; Qualidade subjetiva: Avalia a satisfação geral do usuário, incluindo a intenção de usá-lo novamente e recomendá-lo a outros [STOYANOV et al., 2015]. A MARS é amplamente utilizada para avaliar aplicativos em várias áreas da medicina, como cardiologia, reumatologia e obstetria. Também é empregada no design e desenvolvimento de novos aplicativos de saúde. Além disso, a MARS avalia a segurança dos aplicativos em termos de precisão das informações e privacidade do usuário.

A *Adapted MARS (A-MARS)* é uma versão adaptada da MARS, específica para aplicativos e ferramentas eletrônicas relacionadas à saúde. Inclui novas seções para avaliar: Se o aplicativo ou ferramenta eletrônica oferece recursos atuais e relevantes; Estratégias recomendadas relacionadas à área de preocupação; Soluções para sintomas identificados; Capacidade de uso em tempo real; Facilidade de acesso à ajuda ou suporte [ROBERTS et al., 2021]. O *User Version of MARS (uMARS)* é uma versão simplificada da MARS, focada no usuário final. Contém 20 itens divididos em seis subescalas: engajamento, funcionalidade, estética, qualidade da informação, qualidade subjetiva e impacto percebido pelos usuários. O uMARS é útil para obter a percepção dos usuários sobre a qualidade dos aplicativos durante o desenvolvimento e teste, contribuindo para melhorias gerais [STOYANOV et al., 2016].

A Escala de Usabilidade do Sistema (SUS) é um protocolo de avaliação de usabilidade para uma ampla gama de produtos, incluindo aplicativos móveis. Desenvolvida em 1986, a SUS avalia a efetividade, eficiência e satisfação da experiência do usuário, mas não identifica problemas específicos de usabilidade [ZHOU et al., 2019]. O PSSUQ, desenvolvido pela IBM, avalia as percepções dos usuários sobre a usabilidade de um sistema após o uso. Com 19 itens, mede utilidade do sistema, qualidade da informação e da interface de usuário [FRUHLING; LEE, 2005]. O *mHealth App Usability Questionnaire (MAUQ)* foi projetado para avaliar a usabilidade de aplicativos móveis de saúde. Com 30 questões, ele avalia: a Facilidade de através de questões sobre navegação, legibilidade e clareza das instruções; Utilidade por meio de questões sobre benefícios percebidos e utilidade para gerenciar e rastrear informações de saúde; Satisfação com questões sobre design, funcionalidade e experiência geral; Apelo visual com questões sobre esquema de cores e design geral; Qualidade da informação com questões sobre precisão, relevância e abrangência das informações [ZHOU et al., 2019].

O *Mobile Health Evidence Reporting and Assessment (mERA)* é uma escala com 16 itens focada em relatórios sobre intervenções de saúde móvel. Promove a clareza e completude na divulgação de pesquisas envolvendo ferramentas móveis na área da saúde, abordando especificações técnicas necessárias para um relatório completo [AGARWAL et al., 2016]. A ISO/TS 82304-2 é uma especificação técnica da ISO que orienta o desenvolvimento e avaliação de software de saúde, com foco na usabilidade. Avalia eficácia, eficiência e satisfação do usuário, distribuindo a avaliação entre 82 itens que cobrem aspectos como acessibilidade, segurança dos dados e robustez técnica [ISO 82304-2, 2021]. O ARIA (*Alberta Rating Index for App*) é um protocolo composto por 06 itens para avaliar a apresentação do aplicativo pelo hospedeiro para ajudar o usuário a decidir pela instalação ou não do aplicativo, e 12 itens para avaliar a qualidade do conteúdo e usabilidade do aplicativo após a instalação. Foram desenvolvidas duas versões

do ARIA, uma voltada para os usuários de aplicativos e outra para o provedor de serviços de saúde [AZAD-KHANEHGAH; RODUTA ROBERTS; LIU, 2022].

Essas ferramentas ajudam desenvolvedores, pesquisadores e profissionais de saúde a avaliar e melhorar a qualidade, segurança e usabilidade de aplicativos móveis de saúde.

3.4.5 Análise do Resultado da Avaliação de Tecnologia em Saúde Digital

A análise dos resultados de avaliação de uma tecnologia de saúde visa garantir a confiabilidade e a própria validade a que se propõe a medir, monitorar e gerar informações sobre variáveis ou condições do usuário.

A avaliação prévia por especialistas das áreas envolvidas em tecnologia e saúde é substancial. Assim, a busca por especialistas pode ser realizada através de consultas em bases de currículos, plataformas de redes de profissionais, entre outras, cabendo a ressalva da importância da descrição dos critérios de seleção dos especialistas bem como da descrição das qualificações dos mesmos na seção de resultados do protocolo [ALEXANDRE; COLUCI, 2011]. A norma ABNT/ISO/IEC 25062-2011 [2011], recomenda uma amostragem mínima de oito avaliadores para cada categoria de usuários, a fim de garantir representatividade em relação ao grupo de usuários pretendidos. Para a obtenção dos dados, neste rol de avaliadores, sugere-se como método, a Técnica Delphi, que consiste em obter consenso entre os especialistas – chamados de “juízes – através de rodadas entre as quais realiza-se adequações nas questões a partir dos feedbacks dos “juízes” [KEENEY; HASSON; MCKENNA, 2011]. Na Técnica Delphi, o número de rodadas é determinado visando um consenso final, através dos ajustes das divergências justificados pelas sugestões dos “juízes”. Assim, na primeira rodada, os especialistas respondem quantitativamente a cada questão, acompanhada de justificativas qualitativas. Após a obtenção dos resultados, os especialistas participam de novas rodadas, sucessivamente, repetindo o processo até que o consenso seja alcançado sobre o objeto de estudo. Para que o processo não se estenda por muitas rodadas, têm-se optado pela realização de uma versão modificada da Técnica Delphi, limitando o número de rodadas de duas a quatro, com o intuito de gerenciar o tempo disponível para a construção do consenso [ZARILI et al., 2021].

Para a avaliação quantitativa das apreciações dos “juízes” em busca do consenso final, um método indicado é o Índice de Validade de Conteúdo (IVC), muito utilizado na área de saúde na avaliação de constructos [ALEXANDRE; COLUCI, 2011]. O IVC mede a proporção ou porcentagem de concordância entre os especialistas, permitindo uma análise individual de cada item do instrumento e posterior análise global.

O método consiste em indicar para o especialista qual característica deverá ser analisada para cada item ou constructo e classifica-la numa escala de Likert com pontuação de 1 a 4, tomando como exemplo a acessibilidade, através da Questão X: “O aplicativo possui os recursos de acessibilidade”:

- 1 – NÃO POSSUI;
- 2 – POSSUI, MAS INADEQUADO;
- 3 – POSSUI, ADEQUADO;
- 4 – POSSUI, TOTALMENTE ADEQUADO.

O cálculo do índice é realizado pela soma do número dos itens que receberam “3” ou “4” dividida pelo número total de respostas [POLIT; BECK, 2006] [ALEXANDRE; COLUCI, 2011]. Juntamente com as justificativas dos especialistas, os itens que receberam notas “1” ou “2” devem ser ajustados ou retirados do instrumento para nova rodada [ALEXANDRE; COLUCI, 2011]. Os cálculos demonstrados até aqui são para a análise individual de cada item do instrumento. Para a análise global, embora não haja consenso na literatura, uma opção prática é calcular o IVC médio entre os itens que receberam avaliações “3” e “4” [YUSOFF, 2019].

A literatura indica que o IVC individual aceitável é de 0,78 e, no mínimo, 0,80 para o IVC global, porém, valores acima de 0,90 são consideravelmente excelentes para um instrumento [YUSOFF, 2019].

Uma outra opção para a análise da concordância entre especialistas é o coeficiente de Kappa de Cohen, que não somente avalia concordância entre os especialistas, mas também a confiabilidade e reprodutibilidade entre os observadores (no caso, os especialistas). O Kappa de Cohen permite avaliar as respostas de apenas dois observadores, o que na versão modificada, o Kappa de Fleiss, pode ser estendida para um número maior de avaliadores [MCHUGH, 2012]. O coeficiente do teste de Kappa (Cohen ou Fleiss) varia de 0 a 1, sendo valores próximos ao zero indicam nenhuma concordância ou confiabilidade o que no caso contrário indica alta confiabilidade [LANDIS; KOCH, 1977].

A confiabilidade e a concordância são quesitos importantes na classificação, no desenvolvimento de escalas e instrumentos visando garantir a qualidade e validação dos protocolos, fornecendo informações sobre o erro inerente a qualquer diagnóstico, pontuação ou medição, na qual a quantidade de erro de medição pode determinar a validade dos resultados ou pontuações obtidas [KOTTNER et al., 2011].

Para verificar a consistência interna dos resultados, ou seja, o quanto as questões estão correlacionadas para atender e medir determinado constructo, sugere-se o coeficiente Alfa de Cronbach. Seu cálculo considera a variância dos valores de cada questão respondida e a variância de cada indivíduo respondente. Naturalmente, o coeficiente Alfa de Cronbach varia entre 0 e 1, sendo o valor mínimo aceitável para o alfa é 0,70.

A consistência interna dos itens da escala ou instrumento abaixo desse valor, indica uma confiabilidade interna fraca, necessitando de ajustes consideráveis. Quanto mais os valores se aproximam de 1, maior a consistência interna. Contudo, valores muito próximos de 1 podem indicar a presença de redundância, ou seja, vários itens podem estar medindo o mesmo elemento [STREINER, 2003]. Uma vez passado pelas análises anteriores, pode-se ampliar para uma amostragem de usuários da tecnologia e, os dados coletados podem, então, ser analisados para fornecer insights sobre a usabilidade, satisfação e eficácia do aplicativo.

3.5 Desenvolvimento, validação e aplicação do protocolo

A fim de garantir a qualidade de um produto de software, foi adotado um protocolo de desenvolvimento fundamentado nos padrões estabelecidos pela ISO 25000, que integra as diretrizes previstas tanto na ISO/IEC 9126 quanto na ISO/IEC 14598 [ISO 14598, 1998] [ISO 9126, 2001] [ISO 25000, 2014].

3.5.1 Processo de desenvolvimento do Protocolo de Avaliação

A ISO 25000, cria um conjunto de regras, que envolve dois processos principais: especificação de requisitos e a avaliação da qualidade. A especificação de requisitos, envolve desde a análise comparativa de métodos existente na literatura, quanto o objetivo, configuração, medidas e níveis de pontuação e critérios de julgamento. A avaliação da qualidade abrange a validação da estrutura do protocolo quanto a eficácia do protocolo.

3.5.2 Especificação de requisitos

Para a especificação dos requisitos, é composto por cinco etapas:

- (i) análise e comparação de métodos de avaliação de aplicativos de saúde: identificação das características;
- (ii) definição dos objetos e contextualização da avaliação do protocolo;
- (iii) definição e configuração do modelo de qualidade do protocolo;
- (iv) identificação das medidas da avaliação do protocolo;
- (v) definição de níveis de pontuação e critérios de julgamento do protocolo.

3.5.2.1 Análise e comparação de métodos de avaliação de aplicativos de saúde

Para a etapa análise e comparação de métodos de avaliação de aplicativos de saúde, é feita uma revisão sistemática da literatura para a identificação das características de cada um dos métodos de avaliação.

Com o intuito de compreender e verificar quais as características estão presentes em cada método de avaliação, identificam-se todas as características de cada método e elaboram-se um quadro comparativo. Um quadro comparativo é uma ferramenta valiosa para a tomada de decisões, permitindo a análise de semelhanças e diferenças entre entidades [GIACOMETTI; MARKHOFF; SOULET, 2021]. Neste caso específico, é possível identificar as características presentes, assim como as ausentes, dentro de cada método pesquisado.

Para um método de avaliação atualizado é necessário abranger características gerais e específicas, conforme o objetivo. Além destes, a qualidade subjetiva também é levada em consideração por referir-se à percepção individual de qualidade, muitas vezes relacionada a aspectos de experiências individuais.

3.5.2.2 Definição dos objetos e contextualização da avaliação do protocolo

Na etapa de definição dos objetos e contextualização da avaliação do protocolo são inicialmente divididas em quatro grandes áreas. Cada uma dessas áreas tem subdivisões em categorias, as quais, posteriormente, constituem as categorias a serem avaliadas dentro do protocolo. É essencial também definir os objetivos específicos de extração dos dados de cada categoria antes da elaboração do protocolo.

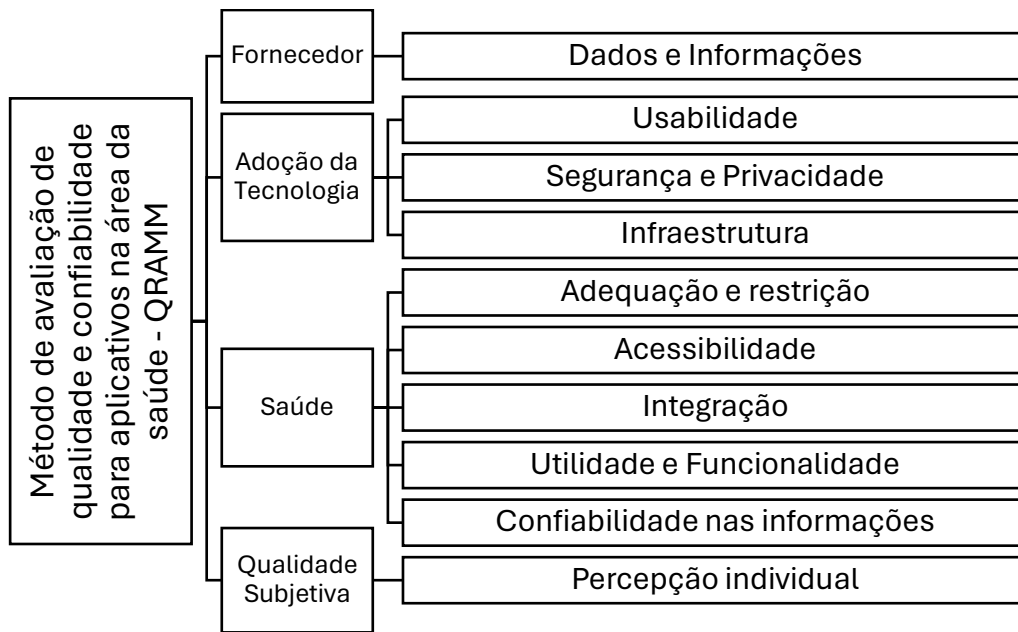


Figura 3.9 - Visão esquemática do Método de Avaliação para Aplicativos na área da Saúde

Considerando a importância de manter a continuidade do pensamento durante a avaliação de um aplicativo, é fundamental estabelecer uma ordem adequada para as categorias. A sequência das categorias deve ser organizada de forma a promover uma fluidez no pensamento e na compreensão das questões, a fim de obter respostas mais coerentes e precisas. Dessa forma, a ordenação das categorias deve ser cuidadosamente planejada, levando em consideração a interdependência dos aspectos avaliados e a progressão natural das áreas de avaliação [FERAUCHE; SILVA; ITO, 2024].

A ordem das categorias em um formulário de avaliação de aplicativos de saúde pode variar dependendo da abordagem e prioridades específicas de cada criador. No entanto, a sequência melhor analisada é a seguinte:

- **Informações iniciais:** Avaliar as informações sobre funcionalidade, uso, fornecedor do aplicativo, atualizações e fonte de download.
- **Acessibilidade:** Avaliar recursos de acessibilidade, suporte para pessoas com deficiências especiais e conformidade com as diretrizes de acessibilidade, como as WCAG.
- **Adequação e restrição:** Avaliar classificação indicativa, informações sobre riscos e contraindicações, atendimento às necessidades e restrições do usuário, orientações claras e capacidade de uso imediato.
- **Usabilidade:** Avaliar organização das informações, facilidade de recuperação de erros, consistência na navegação, acessibilidade às funções e intuitividade das interações.
- **Segurança e privacidade:** Avaliar medidas de segurança, proteção de dados, transparência no tratamento das informações e conformidade com regulamentações de privacidade.
- **Utilidade e funcionalidade:** Avaliar gerenciamento de saúde, acesso a serviços, desempenho do aplicativo e capacidade de personalização para atender às necessidades individuais dos usuários.

- **Integração com outros sistemas de saúde e/ou profissional de saúde:** Avaliar capacidade de compartilhamento de informações, integração com dispositivos e sistemas existentes e facilidade de comunicação e colaboração entre profissionais e usuários.
- **Confiabilidade das informações:** Avaliar precisão, clareza, base científica, linguagem utilizada e processo de revisão das informações fornecidas pelo aplicativo.
- **Infraestrutura:** Avaliar segurança das informações dos pacientes, sincronização, integração com sistemas de comunicação, proteção de dados e compatibilidade com dispositivos.
- **Qualidade Subjetiva:** Avaliar o grau de satisfação subjetiva do usuário com o aplicativo.

Seguir uma sequência lógica e coerente permite que os avaliadores conduzam a avaliação de maneira mais eficiente e facilita a compreensão das informações coletadas [FERAUCHE; SILVA; ITO, 2024].

3.5.2.3 Definição e configuração do modelo de qualidade do protocolo

Para a definição e configuração do modelo de qualidade do protocolo, busca atribuir um valor às características de qualidade e confiabilidade estabelecidas pelo protocolo. É essencial realizar uma medição adequada, portanto essa abordagem permite quantificar e interpretar a qualidade do critério a ser avaliado, fornecendo uma compreensão mais precisa do desempenho do aplicativo de saúde. As medidas são definidas mapeando as categorias em atributos mensuráveis, permitindo assim que elas sejam medidas e pontuadas.

Com o propósito de atingir esse objetivo, é escolhida comumente a abordagem da utilização de questionários como técnica de coleta de dados. Os questionários de coleta de dados são um método comum usado em pesquisas para coletar informações sobre variáveis-alvo. Eles podem ser administrados por vários meios, como métodos baseados em papel ou coleta eletrônica de dados (EDC) [MARSHALL, 2005].

As abordagens de EDC, que envolvem o uso de dispositivos eletrônicos para coletar e armazenar dados, oferecem vantagens sobre os métodos baseados em papel, incluindo pontuação e interpretação automatizadas, economia de tempo e custos e maior eficiência [MHIMDI; KARAA; GHEZALA, 2022]. Essa técnica envolve a elaboração de um conjunto de perguntas objetivas, que possibilitam a coleta de informações por meio de respostas claras e concisas.

Ainda nessa etapa, é estudado e definido a forma de coleta dos dados de cada uma das questões.

As escalas de medição em pesquisas são importantes para determinar a natureza das relações entre variáveis e orientar análises estatísticas apropriadas [DALATI, 2018] [STEVENS, 1946]. A classificação de escalas de Stevens [1946] enfatiza a importância de compreender o nível de medição, como nominal, ordinal, intervalo e razão, na seleção de métodos estatísticos. A distinção entre escalas intervalares e ordinais depende da crença na linearidade das relações entre medidas e quantidades reais de atributos, influenciando a escolha entre análises paramétricas e não paramétricas. Compreender as escalas de medição é fundamental para a interpretação precisa dos dados e a tomada de decisão estatística apropriada na pesquisa. Os tipos de escalas de medição incluem escalas

Likert, Guttman, diferencial semântico, classificação, Thurstone e Borgadus [AINI; ZULIANA; SANTOSO, 2018].

- **Escala Likert:** desenvolvida pelo psicólogo Rensis Likert, na década de 1930, e é amplamente utilizada em pesquisas e questionários. A escala Likert geralmente consiste em uma série de afirmações ou itens aos quais os respondentes devem indicar seu grau de concordância ou discordância. Essa escala é composta por uma série de níveis de resposta que variam de acordo com o número de opções oferecidas. A escala pode variar de 3 pontos (por exemplo, concordo, neutro, discordo) até 7 pontos (por exemplo, concordo plenamente, concordo moderadamente, neutro, discordo moderadamente, discordo plenamente) [BARUA, 2013] [PERES; BUENO; PIASSON, 2016].
- **Escala Guttman:** desenvolvida pelo sociólogo Louis Guttman, que era conhecido por suas contribuições significativas para a teoria da medição e a análise de dados na ciência social. Ele propôs essa técnica de medição na década de 1940 como uma maneira de medir e classificar a ordem de itens ou conceitos de acordo com sua dificuldade ou severidade. Muito parecida com a escala de Likert, porém a principal diferença entre a escala de Guttman e a de Likert é que na escala de Guttman as afirmações selecionadas incorporam a ideia da afirmação anterior, ou seja, é uma escala acumulativa. O respondente só responde questões posteriores se os escores das questões anteriores satisfizerem um determinado nível pré-estabelecido ou considerado limite. A Escala Guttman é frequentemente usada em pesquisas sociais, psicológicas e educacionais para avaliar atitudes, opiniões e comportamentos dos respondentes [PERES; BUENO; PIASSON, 2016].
- **Diferencial Semântico:** é uma técnica utilizada em pesquisa para medir a opinião ou atitude dos respondentes em relação a um determinado objeto, conceito ou ideia. Desenvolvida por Charles Osgood na década de 1950, essa escala busca capturar a valência, intensidade e direção das emoções ou avaliações associadas ao objeto em questão. A estrutura básica da escala consiste em uma série de pares de adjetivos antônimos, geralmente colocados em uma escala linear, em que os respondentes são solicitados a marcar um ponto que melhor represente sua opinião ou sentimentos. Por exemplo, os extremos da escala podem incluir adjetivos como "bom" e "ruim", "positivo" e "negativo", "satisfatório" e "insatisfatório". Os respondentes são instruídos a selecionar um ponto na escala que melhor descreva sua percepção em relação ao objeto em análise. A distância entre os pontos escolhidos pode refletir a intensidade da atitude ou sentimento do respondente. Além disso, a direção em que o ponto é marcado indica se a avaliação é favorável ou desfavorável. Essa escala é amplamente utilizada em pesquisas de marketing, psicologia, ciências sociais e outras áreas para avaliar atitudes em relação a produtos, marcas, serviços, personalidades e conceitos abstratos. A precisão e validade dos resultados dependem da seleção cuidadosa dos adjetivos antônimos, além da clareza das instruções fornecidas aos respondentes [PASSOS, 2015].
- **Classificação:** é uma técnica comum de mensuração usada em pesquisas e avaliações. Nesse método, os respondentes são solicitados a classificar os itens de acordo com um critério específico, geralmente em uma escala

ordinal. Essa escala permite uma rápida coleta de dados e é relativamente simples de administrar e entender. No entanto, ela pode não fornecer informações detalhadas sobre a intensidade ou a direção das opiniões dos participantes, já que geralmente é limitada a uma série de categorias pré-definidas. Além disso, a interpretação dos resultados pode ser influenciada pela forma como as categorias são definidas e pela ordem em que são apresentadas aos participantes. A diferença principal entre a escala de classificação e a escala de Likert está na forma como são estruturadas e como os respondentes são solicitados a responder. Na escala de classificação, os respondentes são solicitados a classificar os itens de acordo com um critério específico em uma escala ordinal. Na escala de Likert, os respondentes são apresentados a uma série de afirmações ou itens e são solicitados a indicar seu nível de concordância com cada item em uma escala de múltiplos pontos: como “discordo totalmente”, “discordo”, “neutro”, “concordo” e “concordo totalmente”, com diferentes pontos atribuídos a cada opção.

- **Thurstone:** desenvolvida por Louis Leon Thurstone na década de 1920, é uma técnica de medição de atitudes que visa quantificar o grau de concordância ou discordância dos participantes em relação a uma série de afirmações ou itens. Esta escala é construída com base no pressuposto de que as pessoas podem diferenciar entre diferentes graus de atitude ou opinião. A principal característica é a utilização de uma série de afirmações que representam uma gama de atitudes, que são classificadas pelos participantes em termos de concordância ou discordância. Os itens são geralmente selecionados e organizados em uma ordem gradual de acordo com o nível de aceitação social. Após a seleção dos itens, um grupo de juízes especialistas atribui uma pontuação a cada um deles com base em sua objetividade e representatividade em relação ao construto que está sendo medido. Em seguida, os participantes são apresentados a uma série dessas afirmações e solicitados a classificá-las de acordo com sua concordância ou discordância. Uma vez que todas as respostas são coletadas, as pontuações atribuídas pelos juízes são usadas para calcular uma média ponderada para cada item, resultando em uma pontuação média de concordância ou discordância para cada afirmação. Essas pontuações são então usadas para ordenar as afirmações em uma escala unidimensional, representando a intensidade da atitude dos participantes em relação ao construto em questão. A escala de Thurstone é amplamente utilizada em pesquisas de opinião, estudos de mercado e pesquisas sociais para medir atitudes em relação a questões específicas. É uma ferramenta valiosa para capturar nuances nas opiniões dos participantes e fornecer insights detalhados sobre as atitudes de um grupo em relação a determinados tópicos [PASSOS, 2015].
- **Bogardus:** também conhecida como Escala de Borg, é uma ferramenta de avaliação subjetiva da percepção de esforço durante o exercício físico. Ela foi desenvolvida por Gunnar Borg na década de 1980 e é amplamente utilizada em contextos de monitoramento da intensidade do exercício em estudos científicos e na prática clínica. Esta escala é uma técnica simples e eficaz para que os indivíduos avaliem a sensação subjetiva de esforço percebido durante o exercício. Ela consiste em uma série de números ou palavras que

representam uma graduação de esforço, que vai desde "nenhum esforço" até "máximo esforço". Geralmente, os participantes são instruídos a escolher o número ou palavra que melhor descreve a intensidade do esforço percebido no momento. A Escala de Borg é uma ferramenta subjetiva, pois depende da percepção individual do esforço pelo participante, levando em consideração fatores como a sensação de cansaço, dificuldade respiratória, desconforto muscular e ritmo cardíaco percebido. Por isso, ela pode variar de pessoa para pessoa e mesmo para uma mesma pessoa em diferentes situações. Uma das vantagens dessa escala é a sua simplicidade de uso e aplicação, sendo fácil para os participantes compreenderem e utilizarem durante o exercício. Além disso, ela pode ser adaptada para diferentes populações e tipos de exercício, tornando-se uma ferramenta versátil para monitorar a intensidade do esforço percebido em uma ampla gama de situações. No entanto, é importante ressaltar que a Escala de Borg é uma medida subjetiva e, portanto, pode não ser tão precisa quanto medidas objetivas de intensidade do exercício, como a frequência cardíaca ou o consumo de oxigênio. Assim, ela geralmente é utilizada em conjunto com outras medidas para fornecer uma avaliação mais abrangente da intensidade do exercício [SANTOS et al., 2014].

A escala de Borg e a escala de Thurstone são ferramentas de medição subjetiva que dependem da percepção individual dos participantes. Embora sejam úteis em contextos específicos, como avaliação de esforço físico ou opiniões pessoais sobre determinado tema, podem ser eliminadas na criação de um protocolo para a avaliação. Essas escalas podem não ser adequadas para avaliar aspectos específicos dos sistemas, como usabilidade, acessibilidade, eficácia ou confiabilidade. Elas são mais adequadas para avaliações subjetivas de sensações físicas ou opiniões pessoais. As escalas de Borg e Thurstone podem gerar dados difíceis de interpretar devido à natureza subjetiva das respostas. Isso pode dificultar a análise e a tomada de decisões com base nos resultados obtidos. Portanto, ao criar um protocolo para a avaliação de sistemas, é preferível utilizar métodos de avaliação mais objetivos e padronizados [SANTOS et al., 2014].

3.5.2.4 Identificação das medidas da avaliação do protocolo

Nesta etapa estabelecem-se critérios de avaliação, determinando o nível específico e objetivo da avaliação do protocolo. Essa definição possibilitará a comparação entre diferentes produtos, ao mesmo tempo em que identifica oportunidades de aprimoramento.

Para definir o nível específico e objetivo da avaliação do protocolo é necessário atribuir valores quantitativos. Os valores quantitativos podem ser obtidos com a combinação de perguntas e respostas. As pontuações são atribuídas a cada pergunta. Essas perguntas contemplam as categorias definidas na etapa de definição dos objetos e contextualização da avaliação do protocolo (3.6.2.2).

As respostas do questionário podem ser de escolha única ou múltipla, dependendo do objetivo da pergunta. Com a escolha definida é aplicado a escala escolhida em definição e configuração do modelo de qualidade do protocolo (3.6.2.3). Cada questão é elaborada cuidadosamente, incluindo diferentes opções de resposta, a fim de fornecer uma compreensão abrangente das avaliações. A variedade de respostas disponíveis permite capturar de forma mais precisa as opiniões e percepções dos avaliadores.

3.5.2.5 Definição de níveis de pontuação e critérios de julgamento do protocolo

Na última etapa, é definida a fórmula para resolução de cada categoria. A partir dos resultados obtidos para cada categoria, é possível calcular o nível de qualidade para cada uma das categorias definidas na etapa de definição dos objetos e contextualização da avaliação do protocolo.

3.5.3 Colocando em prática o protocolo de avaliação

Como prática de elaboração de protocolo de avaliação, segue como exemplo, o método QRAMHe de FERAUCHE; SILVA; ITO [2024]. Este foi proposto para avaliar aplicativos de saúde quanto a informações iniciais; adequação e restrição; usabilidade; segurança e privacidade; utilidade e funcionalidade; integração com outros sistemas de saúde e/ou profissional de saúde; confiabilidade das informações; infraestrutura; acessibilidade e qualidade subjetiva. Apesar de novo, este método de avaliação encontra-se disponível em <https://qramhe.org/>. Este método de avaliação, conforme FERAUCHE; SILVA; ITO [2024], segue sua estruturação segundo as normas ISO/IEC 25000 [2014], que unifica as normas ISO/IEC 9126 [2001] e ISO/IEC 14598 [1998].

As quatro etapas, segundo a ISO 25000 [2014], serão descritas separadamente a seguir.

3.5.3.1 Definição dos Objetos e Contextualização da Avaliação do Protocolo

Conforme proposto por FERAUCHE; SILVA; ITO [2024], a preparação do método de avaliação seguiu com a nomeação de protocolo QRAMHe; visto que este ainda não está pronto e habilitado para ser uma avaliação.

Inicialmente, para os objetos e contextualização da avaliação do protocolo QRAMHe, foi fundamentada com uma análise da revisão da literatura. O modelo do protocolo QRAMHe foi dividido em quatro grandes áreas: fornecedor, adoção da tecnologia, saúde e qualidade subjetiva. Além das grandes áreas, o método possui uma estrutura hierárquica composta por categorias, dentro de cada área [FERAUCHE; SILVA; ITO, 2024]:

- **Área Fornecedor:** foi composta pela categoria Dados e Informações; que foram consideradas as informações iniciais sobre o aplicativo, atualização e descrição do aplicativo e plataformas de distribuição;
- **Área Adoção da tecnologia:** foi dado maior ênfase a facilidade de uso, interface amigável e em ter funcionalidades claras. Além de garantir a segurança e a privacidade dos dados daqueles que utilizam o aplicativo;
- **Área Saúde:** com quatro domínios: adequação e restrição; integração; utilidade e funcionalidade e, por fim, confiabilidade nas informações. A integração tem o propósito das trocas de informações entre o aplicativo e outros relacionamentos, sejam estes outros sistemas ou o profissional de saúde. Com a finalidade de verificar se o aplicativo condiz com o proposto além de verificar se há restrições por meio de compensações monetárias, o domínio de utilidade e funcionalidade estão dispostos para esse tipo de avaliação. Para garantir a autenticidade das informações dispostas no aplicativo, o último domínio desta área, é verificado se há confiabilidade nas informações.

- **Área Qualidade Subjetiva:** refere-se à percepção individual de qualidade que pode variar de pessoa para pessoa. É uma medida subjetiva baseada em opiniões, preferências e experiências pessoais. É importante reconhecer que a qualidade subjetiva não é uma medida objetiva e pode variar de acordo com a perspectiva e o contexto de cada indivíduo.

Com a definição dos objetos e contextualização da avaliação do protocolo avança com a definição e configuração do modelo de qualidade do protocolo.

3.5.3.2 Definição e Configuração do Modelo de Qualidade do Protocolo

Para o protocolo QRAMHe, foi utilizado o questionário, que foram definidas as categorias em atributos mesuráveis, permitindo assim que elas fossem medidas e pontuadas [FERAUCHE; SILVA; ITO, 2024].

Ao final do esboço do protocolo, foi adotado o questionário composto por 50 questões, abordando as 10 categorias, utilizando a Escala Likert com variação de 3 a 5 pontos. A pontuação total pode ser calculada somando-se as respostas de cada item, permitindo uma análise quantitativa das atitudes ou opiniões dos participantes [FERAUCHE; SILVA; ITO, 2024].

3.5.3.3 Identificação das Medidas da Avaliação do Protocolo

As pontuações são atribuídas a cada pergunta do questionário. As respostas do questionário são de escolha única, variando de 1 a 5, na qual 1 representa um nível muito baixo naquele quesito e, no outro extremo, a resposta 5 representa um nível muito elevado [FERAUCHE; SILVA; ITO, 2024].

Com três etapas finalizadas, continua com a última etapa da elaboração do protocolo.

3.5.3.4 Definição de Níveis de Pontuação e Critérios de Julgamento do Protocolo

Para o protocolo QRAMHe, optou-se para o cálculo final pela média aritmética da pontuação de cada categoria. Definindo NQC como Nível de Qualidade por Categoria [FERAUCHE; SILVA; ITO, 2024].

A partir dos resultados obtidos para cada categoria, calcula-se o resultado que representa o nível de qualidade geral do aplicativo, pelo cálculo da média aritmética. Considerando NQCA como o Nível de Qualidade do Aplicativo [FERAUCHE; SILVA; ITO, 2024].

E, por fim, para definir o nível de qualidade, é dividido em três com seus respectivos valores padrões: (i) Superior, pontuação igual ou maior que 80%; (ii) Médio, pontuação igual ou maior que 50% e menor que 80%; e (iii) Baixo, pontuação menor que 50%. A Tabela 1 demonstra o nível de qualidade conforme o resultado obtido da pontuação por categoria.

Tabela 3.3 - Pontuações definidas para as respostas por média aritmética por categoria

Média Aritmética por Categoria	Significado da resposta	Nível de Qualidade
5	Eficiente	Superior
4	Aceitável	
3	Regular	Médio
2	Deficiente	Inferior
1	Ineficiente	

De acordo com os possíveis níveis de qualidade do aplicativo, considera-se que um aplicativo com nível superior possui um bom nível de qualidade, já que para obter essa pontuação o aplicativo conseguiu satisfazer os critérios de qualidade avaliados. Por outro lado, um nível de qualidade médio revela que o aplicativo necessita de melhorias. Por fim, um nível de qualidade baixo sugere que o aplicativo não conseguiu alcançar níveis satisfatórios de qualidade. Dessa maneira, o resultado da avaliação sugere que esse aplicativo deva passar por uma reestruturação, com base nos critérios de qualidade não satisfeitos a partir da avaliação conduzida [FERAUCHE; SILVA; ITO, 2024].

3.5.3.5 Validação do protocolo

Para avaliar o protocolo, os autores escolheram por dois processos de validação:

a. Primeira Avaliação

Para a primeira avaliação, estrutura do protocolo (a), opta-se pela aplicação de entrevista com especialistas. A entrevista é uma técnica usada em pesquisas qualitativas para obter credibilidade, precisão e a confiabilidade dos dados coletados [SIMPSON; MCDOWELL, 2019].

As etapas da entrevista de pesquisa envolvem planejamento, gerenciamento, análise e representação das entrevistas. Para o planejamento da entrevista, adaptando Mann (2016), são necessários os seguintes passos: (aa) objetivo; (ab) participantes; (ac) preparação; (ad) realização; e (ae) pós-entrevista.

Para o passo ‘aa’ do planejamento da entrevista, o objetivo da entrevista é validar o protocolo. No passo ‘ab’, para a seleção dos participantes é necessário profissionais qualificados que possam contribuir no processo de qualificação do protocolo [CHANG; MORREALE; BALON, 2017]. Para a preparação das entrevistas, passo ‘ac’, é necessário o prévio agendamento com previsão da duração e ambiente de condução. Na realização da entrevista, passo ‘ad’, as entrevistas devem ser conduzidas conforme a plataforma escolhida, respeitando o tempo agendado. No passo referente a pós-entrevista, passo ‘ae’, após o andamento da entrevista, ocorre a análise qualitativa e quantitativa. Para a análise qualitativa, é feita a transcrição, tradução e análise temática do discurso [BRITTEN, 1995]. A estatística Kappa de Fleiss é a medida estatística usada para avaliar a confiabilidade da concordância entre vários avaliadores [GWET, 2021]. Portanto, com o auxílio desta técnica é possível avaliar a reprodutibilidade do questionário.

Neste primeiro momento da análise, o resultado das entrevistas foi convertido em tabela classificando e resumindo os pareceres dos entrevistados, assim composta:

Tabela 3.4 - Conversão dos dados das entrevistas para análise quantitativa

Entrevistado	Questão 1	Questão 2	Questão 3	Questão 4	Questão 5	Questão 6	Questão 7	Questão 8	Questão 9	Questão 10
A	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Parc.	Não	Sim
B	Não	Sim	Sim	Sim	Sim	Sim	Sim	Não	Não	Sim
C	Sim	Sim	Sim	Sim	Sim	Parc.	Sim	Não	Sim	Sim

Para a análise, foi utilizado o software IBM SPSS Statistics para Windows (versão 28.0) com nível de significância adotado de $\alpha = 0,05$. O teste Kappa de Fleiss identificou uma confiabilidade geral razoável significativa entre os três avaliadores ($k = 0,298$ [IC 95%: 0,011-0,586]; $z = 2,032$; $p < 0,043$) [FERAUCHE; SILVA; ITO, 2024].

Ao analisar a categoria de respostas (SIM, NÃO, PARCIALMENTE), houve uma concordância significativa entre os entrevistados para a resposta SIM, representando uma confiabilidade/reprodutibilidade moderada ($k = 0,441$ [IC 95%: 0,083-0,799]; $z = 2,415$; $p < 0,017$), possibilitando a continuação do estudo. Diante destes resultados, seguiu-se para o segundo momento da análise que, após a seleção do aplicativo, o instrumento foi direcionado aos profissionais da área para aplicação do instrumento [FERAUCHE; SILVA; ITO, 2024].

Após a análise detalhada da estrutura do protocolo, foram realizados ajustes com base no feedback fornecido pelos entrevistadores [FERAUCHE; SILVA; ITO, 2024].

b. Segunda Avaliação

Para a segunda avaliação, para verificar a eficácia do protocolo, segue-se os seguintes procedimentos:

- **seleção do aplicativo:** para a seleção do aplicativo foi o Protocolo de Mapeamento Sistemático para Busca de Aplicativos de Saúde em Repositórios Não Acadêmicos (PBAS) escrito por Lima e Peres. Este protocolo faz uma adaptação do protocolo de mapeamento sistemático da literatura para o levantamento de aplicativos de software de saúde em repositórios não acadêmicos de aplicativos e código-fonte [LIMA; PERES, 2021].
- **aplicação da ferramenta QRAMHe no aplicativo selecionado:** foi utilizada a ferramenta QRAMHe em formato online (<https://qramhe.org/>), em um período determinado com uma seleção de participantes qualificados para contribuir no processo de avaliação de eficácia do protocolo [FERAUCHE; SILVA; ITO, 2024].
- **análise da aplicação:** fez-se a utilização das análise de coeficiente Alfa de Cronbach e o teste de Kappa de Fleiss [FERAUCHE; SILVA; ITO, 2024].

Foi utilizado o coeficiente Alfa de Cronbach, medida estatística que avalia a consistência e confiabilidade interna de uma escala, questionário ou instrumento de avaliação composto por múltiplas questões de um constructo. Para esse fim, o coeficiente Alfa de Cronbach varia de 0 a 1, indicando que os valores mais próximos de 1 evidenciam maior a consistência interna dos itens do questionário, ou seja, mede o quanto as questões referem-se ao mesmo assunto a que se propõe o construto e se são consistentes entre si [BLAND; ALTMAN, 1997] [GLIEM; GLIEM; OTHERS, 2003] [GASPAR; SHIMOYA, 2017] [FERAUCHE; SILVA; ITO, 2024]

O nível de consistência interna é considerado substancial para os valores acima de 0,60, e para valores acima de 0,80, quase perfeita [LANDIS; KOCH, 1977].

A análise desta etapa partiu das respostas de 10 profissionais da área de Saúde, especificamente, Educação Física, tendo como coeficiente Alfa de Cronbach significativo (Cronbach = 0,97; $p < 0,001$), evidenciando uma consistência interna e confiabilidade muito alta [FERAUCHE; SILVA; ITO, 2024].

3.5.4 Lições Aprendidas e Melhores Práticas

Diversas lições foram aprendidas no desenvolvimento do protocolo, alguns pontos podem ser destacados:

- Identificar quais os métodos de avaliações existentes para aplicativos na área da saúde é algo trabalhoso
- A colaboração entre profissionais de saúde, desenvolvedores de software e especialistas em avaliação foi essencial para garantir que o protocolo aborde todas as necessidades e especificidades dos usuários, pacientes e profissionais;
- A necessidade de adaptar o protocolo a diferentes tipos de aplicativos de saúde e contextos de uso foi uma lição fundamental. O protocolo deve ser suficientemente flexível para acomodar inovações tecnológicas e mudanças nas práticas de saúde;
- A validação contínua do protocolo permitiu ajustes e refinamentos antes da implementação final, garantindo maior precisão e relevância nas avaliações;
- fornecer treinamento adequado aos avaliadores sobre como aplicar o protocolo e interpretar os resultados identifica que o protocolo não foi suficiente independente para assegurar a consistência e a validade das avaliações;
- não se limitar à utilização de uma única escala para a avaliação, mas sim empregar diversas escalas conforme necessário, incluindo opções de seleção única e múltipla, dependendo da natureza de cada questionamento;
- cada um dos questionamentos pode ter valores diferenciados, permitindo que alguns aspectos sejam ponderados mais fortemente enquanto outros recebem menor peso, dependendo da importância relativa de cada critério na avaliação geral.

Apesar das lições compreendidas, algumas práticas foram identificadas:

- estabelecer critérios claros e específicos para cada dimensão de avaliação, como usabilidade, segurança e confiabilidade, facilita a aplicação consistente do protocolo e melhora a comparabilidade entre diferentes aplicativos;
- basear o desenvolvimento do protocolo em normas internacionais como a ISO 25000 proporcionou uma estrutura sólida e reconhecida globalmente, assegurando que o protocolo estivesse alinhado com as melhores práticas de qualidade de software;

- desenvolver questionários estruturado ajudou a sistematizar a coleta de dados e garantir que todos os aspectos relevantes fossem considerados durante a avaliação.

Incorporar essas lições aprendidas e melhores práticas pode não apenas fortalecer a aplicação do protocolo, mas também contribuir para o avanço da qualidade e segurança dos aplicativos de saúde em geral.

3.6 Desafios e Perspectivas da avaliação de Tecnologia em Saúde Digital

A avaliação de tecnologias digitais em saúde é um tema amplamente discutido em diversos estudos. O desenvolvimento dessas tecnologias ocorre em um ritmo bastante acelerado, o que torna crucial que os processos de avaliação sejam igualmente ágeis. Um dos desafios é como acelerar o desenvolvimento de processos e protocolos para a avaliação dessas tecnologias digitais diante de seu ritmo acelerado de desenvolvimento, de forma que a população seja beneficiada pela adoção dessas tecnologias.

Além disso, conforme destacado no estudo de MEZEI et al. [2023], há uma necessidade de padronização para a classificação, avaliação e financiamento dessas tecnologias. A falta de padrões dificulta a comparação entre tecnologias e a tomada de decisões informadas por parte dos stakeholders.

As tecnologias digitais aplicadas na saúde abrangem uma vasta gama, incluindo aplicativos móveis, algoritmos de inteligência artificial e aprendizado de máquina, robótica e Internet das Coisas médicas. Essa variedade dificulta o desenvolvimento de uma única ferramenta de avaliação que possa ser aplicada a todas essas tecnologias, como evidenciado pelas diversas propostas de frameworks existentes. A diversidade dessas tecnologias requer abordagens específicas e adaptadas para cada tipo de inovação.

Ainda conforme mencionado no estudo de MEZEI et al. [2023], vários países têm desenvolvido seus próprios frameworks para a avaliação de tecnologias digitais. Embora esses frameworks tenham dimensões similares, eles possuem diferenças que refletem as necessidades e contextos específicos de cada país. Isso sugere que, enquanto a padronização é desejável, a flexibilidade para acomodar variabilidades locais também é necessária.

Mesmo quando focados em tecnologias específicas, como aplicativos móveis de saúde, existem várias propostas de protocolos de avaliação. Esses protocolos, porém, não demonstram uniformidade nas dimensões avaliadas, evidenciando a necessidade de padronização e especificidade nos processos de avaliação.

A diversidade dos quadros de Avaliação de Tecnologias em Saúde (ATS) para Tecnologias de Saúde Digital (TDS) decorre da complexidade da avaliação dessas tecnologias em rápida evolução, que utilizam métodos diversos e transformam significativamente os sistemas e comportamentos sociais e de saúde. A rápida evolução das TDSs exige métodos de avaliação dinâmicos e adaptáveis que possam acompanhar essas mudanças.

Nas propostas de frameworks de avaliação de tecnologias em saúde, é comum encontrar uma dimensão dedicada à cibersegurança dos dados de saúde. No entanto, não é explicitamente observada uma dimensão dedicada à segurança do paciente ao avaliar essas tecnologias digitais. A segurança do paciente é uma preocupação crítica que deve ser integrada nos processos de avaliação para garantir que as tecnologias digitais não

apenas protejam os dados, mas também melhorem a segurança e a qualidade do cuidado ao paciente.

References

ABIODUN, T.; RAMPERSAD, G.; BRINKWORTH, R. Driving Industrial Digital Transformation. **Journal of Computer Information Systems**, v. 63, n. 6, p. 1345–1361, 2 nov. 2023.

ABNT NBR ISO/IEC 25062. **ABNT NBR ISO/IEC 25062 NBRISO/IEC25062 Engenharia**. Disponível em: <<https://www.normas.com.br/visualizar/abnt-nbr-nm/30823/abnt-nbriso-iec25062-engenharia-de-software-requisitos-e-avaliacao-da-qualidade-de-produto-de-software-square-formato-comum-da-industria-fci-para-relatorios-de-teste-de-usabilidade>>. Acesso em: 28 maio. 2024.

ABRAN, A. Software Project Estimation: The Fundamentals for Providing High Quality Information to Decision Makers. 6 abr. 2015.

AGARWAL, S. et al. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. **BMJ**, v. 352, p. i1174, 17 mar. 2016.

AINI, Q.; ZULIANA, S. R.; SANTOSO, N. P. L. Management Measurement Scale As A Reference To Determine Interval In A Variable. **Aptisi Transactions on Management (ATM)**, v. 2, n. 1, p. 45–54, 2 jan. 2018.

AKBAR, S.; COIERA, E.; MAGRABI, F. Safety concerns with consumer-facing mobile health applications and their consequences: a scoping review. **Journal of the American Medical Informatics Association**, v. 27, n. 2, p. 330–340, 1 fev. 2020.

ALEXANDRE, N. M. C.; COLUCI, M. Z. O. Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. **Ciência & Saúde Coletiva**, v. 16, n. 7, p. 3061–3068, jul. 2011.

AZAD-KHANEGHAH, P.; RODUTA ROBERTS, M.; LIU, L. Alberta Rating Index for Apps: Study of Reliability and Validity. **Canadian Journal of Occupational Therapy**, v. 89, n. 3, p. 326–338, 1 set. 2022.

BARUA, A. Methods for Decision-Making in Survey Questionnaires Based on Likert Scale. **Journal of Asian Scientific Research**, v. 3, n. 1, p. 35–38, 1 jan. 2013.

BARZEGAR, A. **Measuring Software Quality Product Based on Fuzzy Inference System Techniques in ISO Standard**. Preprints, , 27 abr. 2021. Disponível em: <<https://www.preprints.org/manuscript/202104.0721/v1>>. Acesso em: 9 maio. 2024

BOUGROUN, Z. et al. **A Survey for Validation Concepts to Measure Quality as Well Their Application on the Maintainability of ISO.** (A. El Moussati et al., Eds.) *Advances in Smart Technologies Applications and Case Studies. Anais...* Cham: Springer International Publishing, 2020.

BRUCE, C. et al. Evaluating Patient-Centered Mobile Health Technologies: Definitions, Methodologies, and Outcomes. **JMIR mHealth and uHealth**, v. 8, n. 11, p. e17577, 11 nov. 2020.

BUCHNER, B. Datenschutz und Datensicherheit - DuD. **DIGITAL HEALTH**, v. 46, n. 12, p. 729–729, 1 dez. 2022.

DALATI, S. Measurement and Measurement Scales. Em: MARX GÓMEZ, J.; MOUSELLI, S. (Eds.). **Modernizing the Academic Teaching and Research Environment: Methodologies and Cases in Business Research.** Cham: Springer International Publishing, 2018. p. 79–96.

ELIAS, F. T. S. A importância da avaliação de tecnologias para o Sistema Único de Saúde. 2013.

ENTRINGER, T. C.; FERREIRA, A. DA S. Software quality evaluation: a bibliometric analysis and future perspectives. **Independent Journal of Management & Production**, v. 10, n. 5, p. 1499–1515, 1 out. 2019.

EUNETHTA. **EUnethTA Joint Action 2 Work Package 8.** , 2016. Disponível em: <<https://www.eunethta.eu/wp-content/uploads/2018/03/HTACoreModel3.0-1.pdf>>. Acesso em: 14 maio. 2024

EUROPEAN COMMISSION. DIRECTORATE GENERAL FOR HEALTH AND FOOD SAFETY. Opinion on assessing the impact of digital transformation of health services. **LU: Publications Office**, 2019.

FERAUCHE, T. M. Y.; SILVA, J. J.; ITO, M. Proposal for the Protocol for Assessment of mHealth: QRAMHe. **CONTECSI - International Conference on Information Systems and Technology Management**, n. 20, 2024.

FRUHLING, A.; LEE, S. Assessing the Reliability, Validity and Adaptability of PSSUQ. p. 10, 2005.

FUKUYAMA, M. Society 5.0: Aiming for a new human-centered society. **Japan Spotlight**, v. 27, n. 5, p. 47–50, 2018.

GABINETE DO MINISTRO. **Portaria n. 2.510.** Disponível em: <https://bvsms.saude.gov.br/bvs/saudelegis/gm/2005/prt2510_19_12_2005.html>.

GARVIN, D. A. **Gerenciando A Qualidade**. 3ª edição ed. [s.l.] Qualitymark, 2002.

GIACOMETTI, A.; MARKHOFF, B.; SOULET, A. Comparison Table Generation from Knowledge Bases. Em: VERBORGH, R. et al. (Eds.). **The Semantic Web**. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021. v. 12731p. 179–194.

GOODMAN, C. S. **Introduction to health care technology assessment: ten basic steps**. , 1998. Disponível em: <<http://www.nlm.nih.gov/nichsr/ta101>>

IMDRF, S. AS A M. D. (SAM) W. G. “Software as a Medical Device”: Possible Framework for Risk Categorization and Corresponding Considerations. **International Medical Device Regulators Forum (IMDRF)**, 2014.

ISO 9126. **ISO/IEC 9126-1:2001**. Disponível em: <<https://www.iso.org/standard/22749.html>>. Acesso em: 15 out. 2023.

ISO 14598. **ISO/IEC 14598-5:1998**. Disponível em: <<https://www.iso.org/standard/24906.html>>. Acesso em: 15 out. 2023.

ISO 25000. **ISO/IEC 25000:2014**. Disponível em: <<https://www.iso.org/standard/64764.html>>. Acesso em: 9 maio. 2024.

ISO 82304-2. **ISO/TS 82304-2:2021 Health software — Part 2: Health and wellness apps — Quality and reliability**. Disponível em: <<https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/07/81/78182.html>>. Acesso em: 16 set. 2022.

ITO, M. Patient-centered care. Em: **Fundamentals of Telemedicine and Telehealth**. [s.l.] Elsevier, 2020. p. 115–126.

ITO, M.; PEREIRA, D. **Proposta de um Ambiente Colaborativo para a Coordenação de Cuidado de Pacientes Crônicos**. Anais Principais do Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2015). **Anais...** Em: NCIPAI DO SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO APLICADA À SAÚDE. Brasil: Sociedade Brasileira de Computação - SBC, 20 jul. 2015. Disponível em: <<https://sol.sbc.org.br/index.php/sbcas/article/view/10382>>. Acesso em: 26 maio. 2024

JUNG, H.-W. Validating the external quality subcharacteristics of software products according to ISO/IEC 9126. **Computer Standards & Interfaces**, v. 29, n. 6, p. 653–661, 1 set. 2007.

KAPUR, R. et al. **Training Software Engineers for Qualitative Evaluation of Software Architecture**. Disponível em: <<https://arxiv.org/abs/2105.09595v1>>. Acesso em: 9 maio. 2024.

KARUNA, P.; DIVYA, M.; MANGALA, N. Statistical Analysis of Metrics for Software Quality Improvement. **International Journal of Software Engineering & Applications**, v. 9, n. 1, p. 77–88, 30 jan. 2018.

KEENEY, S.; HASSON, F.; MCKENNA, H. The Delphi Technique in Nursing and Health Research. Em: **The Delphi Technique in Nursing and Health Research**. [s.l.] John Wiley & Sons, Ltd, 2011. p. 1–17.

KOTTNER, J. et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. **International Journal of Nursing Studies**, v. 48, n. 6, p. 661–671, jun. 2011.

KUI, K. M.; ALI, K. B.; SURYN, W. The Analysis and Proposed Modifications to ISO/IEC 25030—Software Engineering—Software Quality Requirements and Evaluation—Quality Requirements. **Journal of Software Engineering and Applications**, v. 9, n. 4, p. 112–127, 18 abr. 2016.

LANDIS, J. R.; KOCH, G. G. The Measurement of Observer Agreement for Categorical Data. **Biometrics**, v. 33, n. 1, p. 159–174, 1977.

LESQUELEN, A. et al. OP124 Can Registry Failures Be Compensated By Medico-Administrative Database. **International Journal of Technology Assessment in Health Care**, v. 33, n. S1, p. 56–57, jan. 2017.

LIAROPOULOS, L. Do we need ‘care’ in technology assessment in health care, letter to the editor. **International Journal of Technology Assessment in Health Care**, v. 13, n. 1, p. 125–127, 1997.

LIMA JUNIOR, F.; MARTIMIANO, L. Avaliação da qualidade de softwares voltados à gestão de projetos. **XXX Encontro Nacional de Engenharia de Produção, São Carlos. Anais do XXX Encontro Nacional de Engenharia de Produção**, 2010.

LIMA, L. F. D.; PERES, L. M. **Protocolo de Mapeamento Sistemático para Busca de Aplicativos de Saúde em Repositórios Não-acadêmicos**. Anais do I Workshop de Práticas de Ciência Aberta para Engenharia de Software (OpenScienSE 2021). **Anais...** Em: WORKSHOP DE PRÁTICAS DE CIÊNCIA ABERTA PARA ENGENHARIA DE SOFTWARE. Brasil: Sociedade Brasileira de Computação, 28 set. 2021. Disponível em: <<https://sol.sbc.org.br/index.php/opensciense/article/view/17138>>. Acesso em: 10 nov. 2023

LU, Y. et al. **A Quality Evaluation Method for Software Testing about Safety-Critical Software**. 2022 IEEE 13th International Conference on Software Engineering and Service Science (ICSESS). **Anais...** Em: 2022 IEEE 13TH INTERNATIONAL CONFERENCE ON SOFTWARE ENGINEERING AND SERVICE SCIENCE (ICSESS). out. 2022. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/9930258>>. Acesso em: 9 maio. 2024

MAASS, L. et al. The Definitions of Health Apps and Medical Apps From the Perspective of Public Health and Law: Qualitative Analysis of an Interdisciplinary Literature Overview. **JMIR mHealth and uHealth**, v. 10, n. 10, p. e37980, 31 out. 2022.

MARSHALL, G. The purpose, design and administration of a questionnaire for data collection. **Radiography**, v. 11, n. 2, p. 131–136, maio 2005.

MAXIM, B.; PRESSMAN, R. S. **Engenharia de software: uma abordagem profissional**. Porto Alegre:[sn], , 2021.

MCHUGH, M. L. Interrater reliability: the kappa statistic. **Biochemia Medica**, v. 22, n. 3, p. 276–282, 15 out. 2012.

MESKÓ, B. et al. Digital health is a cultural transformation of traditional healthcare - PMC. 2017.

MEZEI, F. et al. International practices in health technology assessment and public financing of digital health technologies: recommendations for Hungary. **Frontiers in Public Health**, v. 11, p. 1197949, 31 ago. 2023.

MHIMDI, N.; KARAA, W. B. A.; GHEZALA, H. B. Data Collection in Hospital Emergencies: A Questionnaire Survey. **International Journal of Computer and Information Engineering**, v. 16, n. 3, p. 48–53, 7 mar. 2022.

MINGLIANG, L. et al. **Research on the Evaluation Method of Software Static Quality based on AHP**. Proceedings of the 2021 1st International Conference on Control and Intelligent Robotics. **Anais...**: ICCIR '21. New York, NY, USA: Association for Computing Machinery, 13 ago. 2021. Disponível em: <<https://doi.org/10.1145/3473714.3473836>>. Acesso em: 9 maio. 2024

MINISTÉRIO DA SAÚDE. Resolução da Diretoria Colegiada (RDC) 657/2022 da Anvisa. . 24 mar. 2022 a.

MINISTÉRIO DA SAÚDE. Resolução da Diretoria Colegiada (RDC) 751/2022 da Anvisa. . 15 set. 2022 b.

MINOV, J. Evaluation of software applications. 14 dez. 2012.

OBAIDAT, M. S.; BOUDRIGA, N. Fundamentals of Performance Evaluation of Computer and Telecommunication Systems. 26 jan. 2010.

OMS. World Health Organisation. . 2012.

PASSOS, M. F. D. Elaboração e validação de escala de diferencial semântico para avaliação de personalidade. 27 abr. 2015.

PERES, L. DA S.; BUENO, A. F.; PIASSON, D. **MATURIDADE LOGÍSTICA DE PLS: AVALIAÇÃO DA ADERÊNCIA DAS ESCALAS DE LIKERT E GUTTMAN EM UM MODELO ESTAGIADO XXXVI ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO.** [s.l: s.n.].

POLIT, D. F.; BECK, C. T. The content validity index: Are you sure you know what's being reported? critique and recommendations. **Research in Nursing & Health**, v. 29, n. 5, p. 489–497, out. 2006.

PRESIDÊNCIA DA REPÚBLICA - CASA CIVIL. Lei 12.401 de 2011. **Diário Oficial da União**, 28 abr. 2011.

ROBERTS, A. E. et al. Evaluating the quality and safety of health-related apps and e-tools: Adapting the Mobile App Rating Scale and developing a quality assurance protocol. **Internet Interventions**, v. 24, p. 100379, 1 abr. 2021.

ROOS, C.; SARTORI, S.; GODOY, L. P. Portaria n. 2.510. **Revista Produção Online**, v. 9, n. 3, 17 ago. 2009.

SANTOS, C. M. DOS et al. A utilização da escala de Borg em atividade física com gestantes. **ConScientiae Saúde**, v. 13, n. 2, p. 241–245, 14 jul. 2014.

SAÚDE 4.0. **PROPOSTAS PARA IMPULSIONAR O CICLO DAS INOVAÇÕES EM DISPOSITIVOS MÉDICOS (DMAs) NO BRASIL.** Saúde 4.0, , 2015.

SHAFFER, F. A.; PREZIOSI, P. Digital Health: Join the Movement. **Nursing Administration Quarterly**, v. 43, n. 4, p. 351–353, out. 2019.

SHEPHERD, J. What is the Digital Era? **Social and Economic Transformation in the Digital Era**, 2004.

SILVA, S. V. et al. **Avaliação da qualidade de um sistema acadêmico: estudo de caso no Q-Acadêmico.** Congresso de Tecnologia da Informação. **Anais...**2015.

SIMON LIN LINWOOD, M. D. Preface. Exon Publications. p. ix–x, 29 abr. 2022.

STEVENS, S. S. On the Theory of Scales of Measurement. **Science**, v. 103, n. 2684, p. 677–680, 7 jun. 1946.

STOUMPOS, A. I.; KITSIOS, F.; TALIAS, M. A. Digital Transformation in Healthcare: Technology Acceptance and Its Applications. **International Journal of Environmental Research and Public Health**, v. 20, n. 4, p. 3407, 15 fev. 2023.

STOYANOV, S. R. et al. Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. **JMIR mHealth and uHealth**, v. 3, n. 1, p. e3422, 11 mar. 2015.

STOYANOV, S. R. et al. Development and Validation of the User Version of the Mobile Application Rating Scale (uMARS). **JMIR mHealth and uHealth**, v. 4, n. 2, p. e5849, 10 jun. 2016.

TRENDOWICZ, A.; KOPCZYŃSKA, S. Chapter 4 - Adapting Multi-Criteria Decision Analysis for Assessing the Quality of Software Products. Current Approaches and Future Perspectives. Em: MEMON, A. (Ed.). **Advances in Computers**. Advances in Computers. [s.l.] Elsevier, 2014. v. 93p. 153–226.

VON HUBEN, A. et al. Health technology assessment for digital technologies that manage chronic disease: a systematic review. **International Journal of Technology Assessment in Health Care**, v. 37, n. 1, p. e66, 2021.

VUONG, Q.-H. et al. The psychological mechanism of internet information processing for post-treatment evaluation. **Heliyon**, v. 8, n. 5, p. e09351, maio 2022.

WETTER, T. **Consumer health informatics**. [s.l.] Springer, 2012.

WHO, W. H. O. Global strategy on digital health 2020–2025. 2021. **License: CC BY-NC-SA**, v. 3, 2021.

YADAV, K. K. Software Testing and Evaluation. 1 jan. 2015.

YAMAKOSHI, K.; NISHIYAMA, M. Software evaluation method and software evaluation device. 27 set. 2018.

YANG, H. Measuring Software Product Quality with ISO Standards Base on Fuzzy Logic Technique. Em: LUO, J. (Ed.). **Affective Computing and Intelligent Interaction**. Berlin, Heidelberg: Springer, 2012. p. 59–67.

YUSOFF, M. S. B. ABC of Content Validation and Content Validity Index Calculation. **Education in Medicine Journal**, v. 11, p. 49–54, 28 jun. 2019.

ZARILI, T. F. T. et al. Técnica Delphi no processo de validação do Questionário de Avaliação da Atenção Básica (QualiAB) para aplicação nacional. **Saúde e Sociedade**, v. 30, p. e190505, 18 jun. 2021.

ZHANG, Y. et al. **Development and Reconstitution of Software Quality Measurement and Evaluation Standards**. 2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD). **Anais...** Em: 2018 19TH IEEE/ACIS INTERNATIONAL CONFERENCE ON SOFTWARE ENGINEERING, ARTIFICIAL INTELLIGENCE, NETWORKING AND PARALLEL/DISTRIBUTED COMPUTING (SNPD). jun. 2018. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/8441040>>. Acesso em: 9 maio. 2024

ZHOU, L. et al. The mHealth App Usability Questionnaire (MAUQ): Development and Validation Study. **JMIR mHealth and uHealth**, v. 7, n. 4, p. e11500, 11 abr. 2019.

Chapter

4

Learned Index: Perspectivas e Desafios na Gestão de Dados de Saúde

Guilherme Waldschmidt (IC-UFBA), Alberto Sironi (IC-UFBA), Laís Sacramento (IC-UFBA/Cidacs-Fiocruz), Mirlei Moura (IC-UFBA), Maíra Lima Souza (Cidacs-Fiocruz), Nívea Bispo (IME-UFBA), Robespierre Pita (IC-UFBA/Cidacs-Fiocruz)

Abstract

Healthcare-focused computational solutions, such as record linkage, image pattern recognition, data mining for diagnostic support, or natural language information extraction, often utilize indexing methods to optimize the retrieval and access of database records. This course introduces the concept of the Learned Index, which leverages machine learning models to create index structures that outperform traditional algorithms. The engagement of the Applied Health Computing community in this new research area has the potential to drive the development of various solutions and applications with a direct impact on clinical practice, health surveillance, management, and public policy formulation. This mini-course is designed for computing and healthcare students and researchers while offering valuable insights for database professionals and managers.

Resumo

Soluções computacionais voltadas à saúde, tais como record linkage, reconhecimento de padrões em imagens, mineração de dados para suporte a diagnósticos ou extração de informação em linguagem natural frequentemente recorrem a métodos de indexação para otimizar a recuperação e acesso aos registros de uma base de dados. Este curso apresenta o conceito de Learned Index, que emprega modelos de machine learning para produzir estruturas de índices capazes de superar algoritmos tradicionais. Acreditamos que a participação da comunidade de Computação Aplicada à Saúde neste novo tópico de pesquisa tem o potencial de levar ao desenvolvimento de diversas soluções e

aplicações com impacto direto na prática clínica, da vigilância em saúde, gestão e formulação de políticas públicas. Tendo um caráter teórico apoiado por demonstrações práticas de criação, uso e validação de índices, este minicurso se dirige a estudantes e pesquisadores da computação e saúde, além de trazer insights para profissionais e gestores de bancos de dados.

4.1. Introdução

Sistemas Gerenciadores de Bancos de Dados (SGBD) são coleções de programas que compartilham o objetivo de apoiar o armazenamento, acesso e recuperação de dados [Silberschatz et al., 2020]. Estes sistemas são responsáveis por suportar diversas tarefas de gestão e administração de dados, implementando todas as funcionalidades e restrições impostas pelos diferentes modelos. Um modelo de dados estabelece um conjunto de regras e abstrações para representação, descrição, relacionamento, semântica, consistência e coerência dos fenômenos registrados num mini-mundo. O modelo relacional define tabelas — ou relações — para orientar o armazenamento físico dos dados. Neste sentido, cada registro numa tabela descreve os fenômenos ou entidades observados usando um número fixo de atributos. Administradores de bancos de dados (DBA, do inglês *database administrator*) e Engenheiros de Dados são responsáveis por projetar, manter e manipular dados nestes sistemas.

Os SGBD tradicionais são comumente utilizados pelos Sistemas de Informação projetados para administração do Sistema Único de Saúde (SUS), seja na atenção primária, vigilância em saúde ou na administração de políticas públicas diversas^{1,2}. Dentre eles figuram o Sistema de Informações sobre Nascidos Vivos (Sinasc), Registro de Eventos em Saúde Pública (Resp-Microcefalia), Sistema de Informação de Agravos de Notificação (Sinan), e-SUS Notifica, Sistema de Informações sobre Mortalidade (SIM), etc. Como descrito no sítio eletrônico do principal mantenedor das bases de dados de saúde do país, o Departamento de Informática do Sistema Único de Saúde (DATASUS), as diretrizes de gestão dos bancos de dados que suportam seus sistemas de informação são estabelecidas pela Metodologia de Administração de Dados (MAD), que normatiza regras gerais de uso e otimização. As especificações encontradas no MAD conformam os principais aspectos do modelo de dados, incluindo quais os SGBDs, definições de padrões, regras técnicas, normas de modelagem e dicionários de dados [DATASUS, 2024]. Fora deste ambiente de produção, versões pseudonimizadas destas bases relacionais são disponibilizadas pelo DATASUS para fins de transparência e suporte à pesquisa em saúde pública³.

O Módulo de ‘Notificação Covid-19’ do SI e-SUS Notifica, por exemplo, concentra a coleta de dados sobre todos os casos leves de Síndrome Gripal (SG), execução e resultados dos testes laboratoriais, a classificação final em relação à

1

<https://www.gov.br/saude/pt-br/composicao/svsa/vigilancia-de-doencas-chronicas-nao-transmissiveis/sistemas-de-informacao-em-saude>

2

<https://www.gov.br/saude/pt-br/assuntos/noticias/2023/julho/entenda-as-principais-caracteristicas-dos-sistemas-de-informacao-do-ministerio-da-saude>

³ <https://datasus.saude.gov.br/informacoes-de-saude-tabnet/>

Covid-19 (suspeito, confirmado ou descartado) e o desfecho (cura ou óbito). Esta adaptação suportou a vigilância epidemiológica durante a pandemia, adaptando rapidamente um SI para cumprir com a notificação compulsória imediata da Covid-19, definida em maio de 2022 [Ministério da Saúde, 2022]. Na Figura 4.1 é possível verificar o formulário físico utilizado para coletar os dados primários. Em seu preâmbulo, a ficha inclui informações para a sua própria identificação, incluindo seu número único, unidade federativa e município de notificação. Os três segmentos subsequentes concentram dados de identificação do indivíduo, suas informações clínicas e epidemiológicas e desfecho do caso.

Para fins de demonstração, o modelo lógico exposto na Figura 4.2., denominado `covid_bd`, mimetiza o banco de dados do Módulo de 'Notificação Covid-19' do e-SUS Notifica, orientado a partir dos segmentos da ficha na Figura 4.1. O `covid_bd` é composto por quatro tabelas. A tabela `tbl_identificacao` concentra os dados derivados do primeiro segmento da ficha na Figura 4.1, incluindo nome, cadastro de pessoa física (CPF), a raça/cor autodeclarada, sexo e data de nascimento. Na `tbl_info_clinicas` estão as variáveis de aferição de temperatura, saturação de oxigênio no sangue, data da coleta, data do resultado, tipo e resultado de cada teste. Parte das informações que subsidiam a investigação epidemiológica estão na tabela `tbl_info_epidemiologicas`, que concentra variáveis binárias indicando se existem sintomas gripais, distúrbios olfativos ou gustativos e dispneia. Por fim, as variáveis de data de admissão, evolução do caso, diagnóstico/classificação final e data de encerramento estão na tabela `tbl_adm_encerramento`.

É importante ressaltar que a variável `id_caso`, original da tabela `tbl_identificacao` é exportada para as demais tabelas para garantir o relacionamento entre elas. Estes relacionamentos consistem na associação dos fenômenos e entidades fisicamente armazenadas sob orientação do SGBD. Por exemplo, a partir do `id_caso = 17781`, cujo paciente é Alfredo Casanova, é possível varrer toda a tabela `tbl_info_clinicas` para verificar se existem coletas e resultados de testes ou procurar pelo seu diagnóstico e desfecho desta infecção na `tbl_adm_encerramento`. Outro aspecto da notação estabelece entre colchetes as restrições de mínimo e máximo de associações nestas relações. No caso da relação entre as tabelas `tbl_identificacao` e `tbl_info_clinicas`, um mesmo caso pode ter nenhuma ou muitos testes associados, denotado pelo $[0, N]$. Contudo, cada teste nesta relação só pode ser registrado se já existe um caso cadastrado e estar associado a apenas um caso, representado pela notação $[1, 1]$. Esta restrição permite que haja dois testes para o `id_caso = 29819`.

Os SGBD que implementam o modelo relacional oferecem uma linguagem de consulta estruturada chamada (SQL, do inglês *structured query language*) usada para recuperar a informação armazenada pelo sistema [Silberschatz et al., 2020]. Em seu livro, Elmasri et al., 2016 argumenta que a maioria das consultas aos dados recupera apenas uma pequena porção de todos os seus registros, tornando ineficiente a estratégia de varrer todo o dado disponível para responder a cada busca. Os *datasets* resultantes de consultas SQL frequentemente suportam tarefas de processamento analítico online (OLAP, do inglês *online analytical processing*). Por exemplo, após a saída do SGBD para a consulta “*quais os casos confirmados e descartados mediante testes coletados*”

entre 01/01/2022 e 31/12/2022?”, diversas perguntas sobre os casos suspeitos de Covid-19 ou modelos analíticos podem ser elaborados. Analistas de dados podem

MINISTÉRIO DA SAÚDE SECRETARIA DE VIGILÂNCIA EM SAÚDE		Nº	
e-SUS Notifica - 05/10/2020			
FICHA DE INVESTIGAÇÃO DE SG SUSPEITO DE DOENÇA PELO CORONAVÍRUS 2019 – COVID-19 (B34.2)			
<small>Definição de caso: indivíduo com quadro respiratório agudo, caracterizado por pelo menos dois (2) dos seguintes sinais e sintomas: febre (mesmo que referida), calafrios, dor de garganta, dor de cabeça, tosse, coriza, distúrbios olfativos ou distúrbios gustativos. Em crianças: além dos itens anteriores considera-se também obstrução nasal, na ausência de outro diagnóstico específico. Em idosos: deve-se considerar também critérios específicos de agravamento como síncope, confusão mental, sonolência excessiva, irritabilidade e inapetência. Observação: Na suspeita de COVID-19, a febre pode estar ausente e sintomas gastrointestinais (diarreia) podem estar presentes.</small>			
UF de notificação: <input type="text"/>		Município de Notificação: <input type="text"/>	
Tem CPF? (Marcar X) <input type="checkbox"/> Sim <input type="checkbox"/> Não		Estrangeiro: (Marcar X) <input type="checkbox"/> Sim <input type="checkbox"/> Não	
Profissional de saúde (Marcar X) <input type="checkbox"/> Sim <input type="checkbox"/> Não		Profissional de segurança (Marcar X) <input type="checkbox"/> Sim <input type="checkbox"/> Não	
CBO: <input type="text"/>		CPF: <input type="text"/>	
CNS: <input type="text"/>			
Nome Completo: <input type="text"/>			
Nome Completo da Mãe: <input type="text"/>			
Data de nascimento: <input type="text"/>		País de origem: <input type="text"/>	
Sexo: (Marcar X) <input type="checkbox"/> Masculino <input type="checkbox"/> Feminino		Raça/COR: (Marcar X) <input type="checkbox"/> Branca <input type="checkbox"/> Preta <input type="checkbox"/> Amarela <input type="checkbox"/> Parda <input type="checkbox"/> Indígena - Etnia: <input type="text"/> <input type="checkbox"/> Ignorado	
Passaporte: <input type="text"/>			
CEP: <input type="text"/>			
Estado de residência: <input type="text"/>		Município de Residência: <input type="text"/>	
Logradouro: <input type="text"/>		Número: <input type="text"/> Bairro: <input type="text"/>	
Complemento: <input type="text"/>			
Telefone Celular: <input type="text"/>		Telefone de contato: <input type="text"/>	
Data da Notificação: <input type="text"/>		Data do início dos sintomas: <input type="text"/>	
Sintomas: (Marcar X) <input type="checkbox"/> Assintomático <input type="checkbox"/> Febre <input type="checkbox"/> Dor de Garganta <input type="checkbox"/> Dispneia <input type="checkbox"/> Tosse <input type="checkbox"/> Coriza <input type="checkbox"/> Dor de Cabeça <input type="checkbox"/> Distúrbios gustatórios <input type="checkbox"/> Distúrbios olfativos <input type="checkbox"/> Outros: <input type="text"/>			
Condições: (Marcar X) <input type="checkbox"/> Doenças respiratórias crônicas descompensadas <input type="checkbox"/> Diabetes <input type="checkbox"/> Obesidade <input type="checkbox"/> Doenças renais crônicas em estágio avançado (graus 3, 4 e 5) <input type="checkbox"/> Imunossupressão <input type="checkbox"/> Portador de doenças cromossômicas ou estado de fragilidade imunológica <input type="checkbox"/> Gestante <input type="checkbox"/> Doenças cardíacas crônicas <input type="checkbox"/> Puérpera (até 45 dias do parto)			
Estado do Teste: (Marcar X) <input type="checkbox"/> Solicitado <input type="checkbox"/> Coletado <input type="checkbox"/> Concluído <input type="checkbox"/> Exame Não Solicitado		Tipo de Teste: (Marcar X) <input type="checkbox"/> RT – PCR <input type="checkbox"/> Teste rápido – anticorpo <input type="checkbox"/> Teste rápido – antígeno <input type="checkbox"/> Testes sorológico	
Resultado (PCR/Rápidos): (Marcar X) <input type="checkbox"/> Negativo <input type="checkbox"/> Positivo <input type="checkbox"/> Inconclusivo ou Indeterminado		Data do Teste (PCR/Rápidos): <input type="text"/>	
Resultado (IgG): (Marcar X) <input type="checkbox"/> Reagente <input type="checkbox"/> Não Reagente <input type="checkbox"/> Inconclusivo ou Indeterminado		Data do Teste (Sorológico): <input type="text"/>	
Resultado (IgM): (Marcar X) <input type="checkbox"/> Reagente <input type="checkbox"/> Não Reagente <input type="checkbox"/> Inconclusivo ou Indeterminado		Resultado (IgA): (Marcar X) <input type="checkbox"/> Reagente <input type="checkbox"/> Não Reagente <input type="checkbox"/> Inconclusivo ou Indeterminado	
Resultado (Anticorpos Totais): (Marcar X) <input type="checkbox"/> Reagente <input type="checkbox"/> Não Reagente <input type="checkbox"/> Inconclusivo ou Indeterminado		Resultado (Anticorpos Totais): (Marcar X) <input type="checkbox"/> Reagente <input type="checkbox"/> Não Reagente <input type="checkbox"/> Inconclusivo ou Indeterminado	
Evolução do caso: (Marcar X) <input type="checkbox"/> Cancelado <input type="checkbox"/> Ignorado <input type="checkbox"/> Em tratamento domiciliar <input type="checkbox"/> Internado em UTI <input type="checkbox"/> Internado <input type="checkbox"/> Óbito <input type="checkbox"/> Cura		Classificação final: (Marcar X) <input type="checkbox"/> Descartado <input type="checkbox"/> Confirmado Clínico Imagem <input type="checkbox"/> Confirmado Clínico-Epidemiológico <input type="checkbox"/> Confirmado Por Critério Clínico <input type="checkbox"/> Confirmado Laboratorial <input type="checkbox"/> Síndrome Gripal Não Especificada	
Data de encerramento: <input type="text"/>			
Informações complementares e observações			
<input type="text"/>			
<input type="text"/>			

Figura 4.1. e-SUS Notifica: Ficha de investigação utilizada para coletar dados de vigilância sobre casos de Síndrome Gripal suspeitos de doença pelo Covid-19.

estimar a média móvel de casos de síndrome gripal no município, calcular o volume de testes feitos, verificar a duração média dos casos, computar o risco de óbito com diferentes sintomas, características sociais ou regionais. Cientistas de dados podem conceber modelos preditivos capazes de auxiliar no diagnóstico precoce, estimar a quantidade de leitos ou insumos para atender a elasticidade da demanda ou encontrar

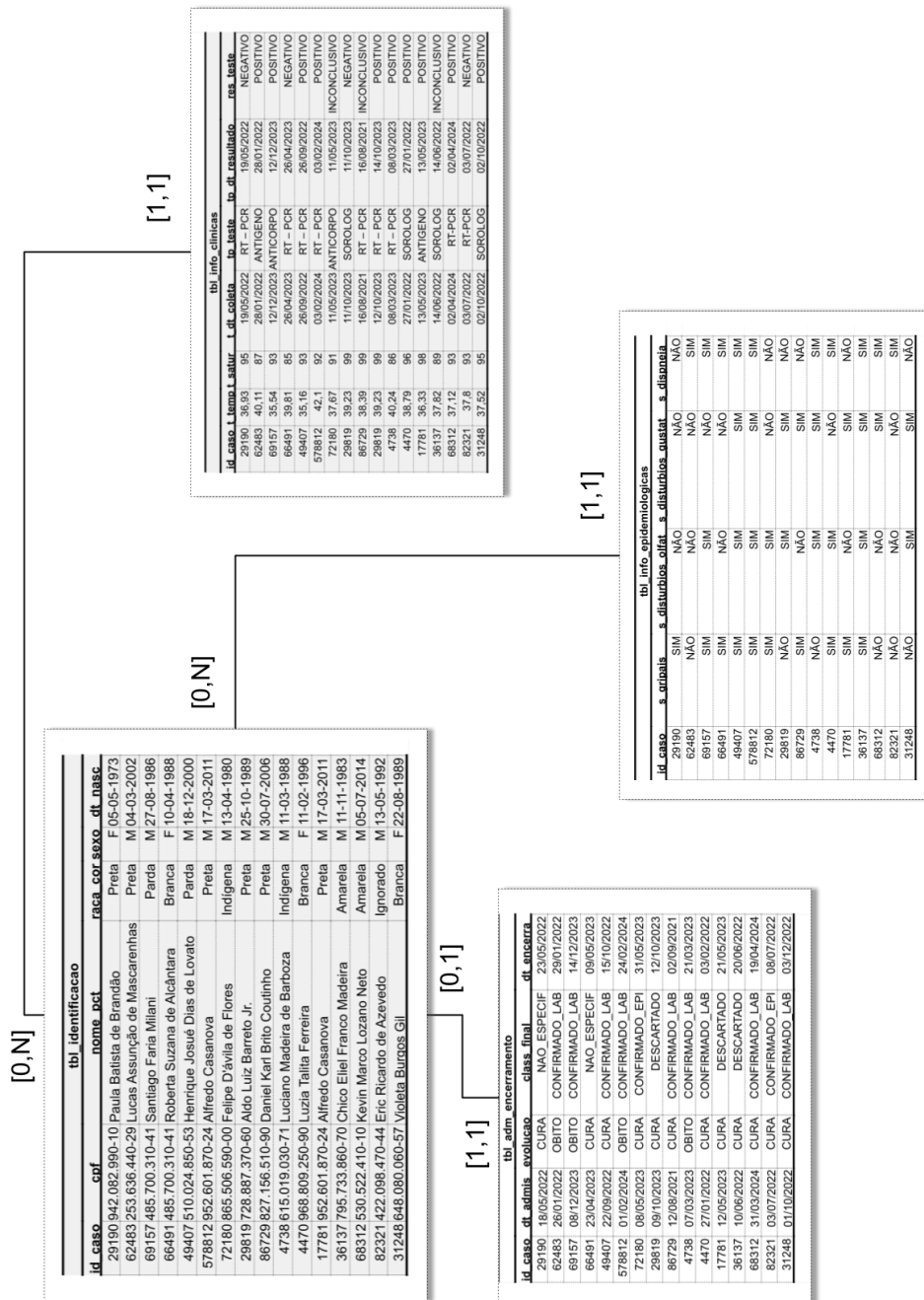


Figura 4.2. Modelo lógico de uma base de dados inspirada na Ficha de investigação de casos da Síndrome Gripal para doença pelo Coronavírus 2019.

padrões de casos através do tempo para definir estratégias de intervenção medicamentosa ou de assistência básica em saúde. Este mesmo *dataset* pode ser utilizado por pesquisadores interessados em estudar iniquidades em saúde, auxiliando na obtenção de conhecimento sobre o acesso desproporcional a testes por pessoas com diferentes marcadores demográficos, a exemplo da raça ou sexo. Alternativamente, o *dataset* pode ser enriquecido a partir da integração com outros sistemas de informação ou bancos de dados, possibilitando análises ainda mais complexas e valiosas. Estas

poderosas ferramentas de suporte à decisão, contudo, possuem alta complexidade computacional e frequentemente classificadas como rotinas intensivas de dados.

Aplicações intensivas de CPU (*central processing unit*) demandam arranjos computacionais de alta capacidade, frequentemente medida em FLOPS (*float operations per second*), como os supercomputadores da lista TOP500⁴. Por outro lado, aplicações *data-intensive* são desenvolvidas para endereçar a complexidade inerente ao dado, seja pelo seu volume, sua estrutura ou a velocidade com que as observações mudam [Kleppmann, 2017]. Em geral, sistemas computacionais capazes de suportar aplicações *data-intensive* são confiáveis, escaláveis e sustentáveis. A confiabilidade é invariavelmente alcançada por algoritmos e componentes de hardware tolerantes a falhas. A propriedade de escalabilidade está relacionada a capacidade de manter boas métricas de latência e vazão em cenários com volume variável de requisições. Por fim, a sustentabilidade é garantida pela incorporação adequada de boas práticas da Engenharia de Software, a exemplo da modularidade, reuso e interoperabilidade. SGBDs adequados para atender os requisitos destas aplicações devem oferecer métodos eficientes de recuperação e acesso aos registros de uma base de dados [Kaur, 2018]. Modelos analíticos e preditivos, como algoritmos de agrupamento de dados, sistema de recomendação, *record linkage* e extração de informação são exemplos de aplicações intensivas que demandam acesso eficiente aos registros de um banco de dados.

Segundo Silberschatz et al., (2020), as tarefas de acesso a registros em SGBD frequentemente extraem pequenas frações de tuplas armazenadas em discos ou alocadas em memória. Especialmente quando estes sistemas de dados são usados para suportar aplicações *data-intensive*. O processamento destas consultas pode ser ineficiente se considerarmos varrer todas as observações, selecionando apenas aquelas que correspondem com o valor buscado. Neste caso, métodos de indexação são aplicados para auxiliar organizar os dados, orientando o armazenamento e otimizando operações de recuperação ou acesso, tais como consultas, filtros e varreduras.

Neste capítulo, exploraremos as propriedades das principais estruturas de índices, sua evolução e uso em aplicações intensivas de dados, especialmente na gestão de dados da saúde. No curso desta exposição, focaremos em soluções habilitadas por algoritmos de aprendizado de máquina, chamadas de Índices Aprendidos (LI, do inglês *Learned Index*) [Kraska et al., 2018]. Este novo conceito, apresentado em 2018 à comunidade de gestão de dados, é precursor de uma nova classe de sistemas de dados autoajustáveis, denominada *instance optimized data systems* (IODS). No cerne destes sistemas e estruturas de índices estão três premissas: i) algoritmos de gestão de dados são modelos; ii) é possível superar os algoritmos tradicionais a partir da obtenção de conhecimento, ou estimativa, da distribuição de uma coleção de dados; e iii) os inúmeros parâmetros para otimização da gestão de dados representa um volume de trabalho cada vez mais desafiador para os DBA engenheiros de dados. Sendo assim, o objetivo geral deste texto é apresentar o estado-da-arte de LI e suas aplicações na gestão de dados de saúde. Acreditamos que a participação da comunidade de Computação Aplicada à Saúde neste novo tópico de pesquisa tem o potencial de levar ao

⁴ <https://top500.org/lists/top500/>

desenvolvimento de diversas soluções e aplicações com impacto direto na prática clínica, da vigilância em saúde, gestão e formulação de políticas públicas.

Este texto está dividido em três principais seções. Ainda na Introdução (Seção 4.1), abordaremos o papel da indexação nas atividades de governança e gestão de dados em saúde, focando em experiências documentadas de entidades governamentais e centros de dados nacionais. Na Seção 4.2, aprofundaremos a exposição do conceito de LI, seus avanços e limitações. Por fim, na Seção 4.4, apresentaremos uma discussão sobre tendências, desafios de pesquisa e aplicações para a saúde que podem ser endereçadas por métodos inovadores de indexação.

4.1.1 Indexação na gestão de dados em saúde

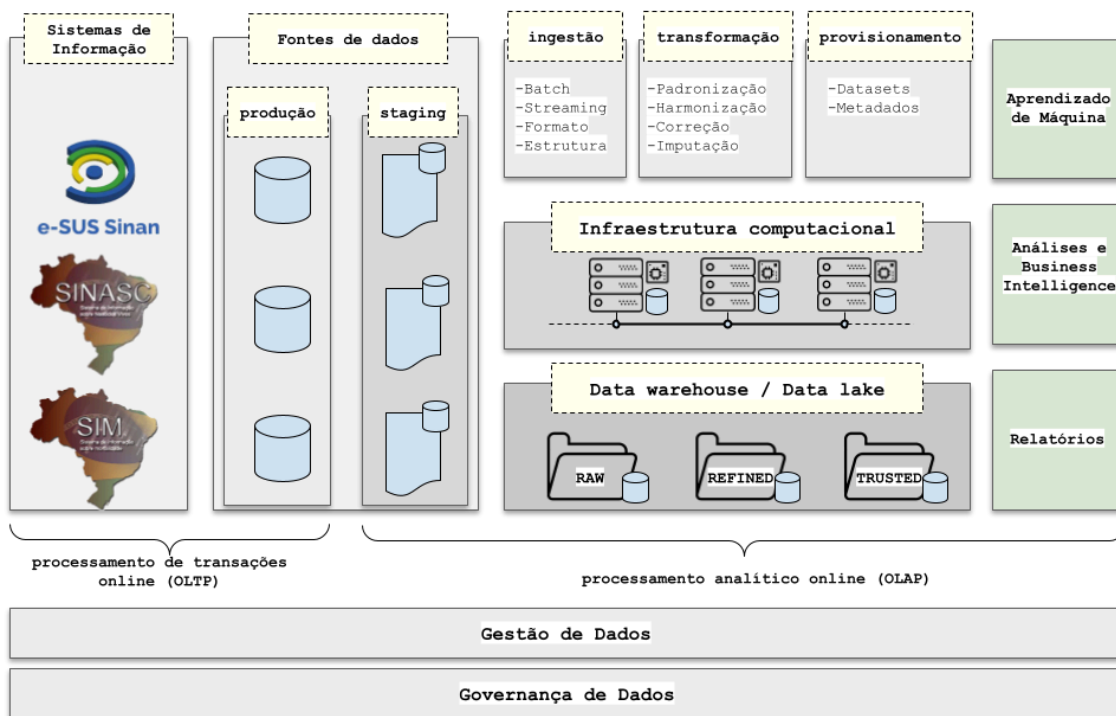


Figura 4.3. Modelo de arquitetura de dados e seus principais componentes, adaptado de [Reis et al, 2022;].

O termo governança de dados agrega diversas diretrizes e técnicas que devem ser adotadas por uma organização para garantir a qualidade, integridade, usabilidade e segurança dos dados que são produzidos ou coletados [Eryurek et al., 2021]. Estas diretrizes, quando bem aplicadas e suportadas pela cultura organizacional, tem o potencial de estabelecer processos dirigidos a dados que maximizam o valor das decisões ao passo em que estabelece uma plataforma de segurança adequada para a preservação da privacidade. Derivada da governança, a disciplina de gestão de dados consiste em práticas e conceitos aplicados no desenvolvimento, execução e supervisão do manejo da informação através do seu ciclo de vida [Dama, 2017]. Administradores de bancos de dados e engenheiros de dados são os principais responsáveis por desenhar,

planejar e materializar em práticas cotidianas todas as determinações de um modelo de gestão.

A arquitetura demonstrada da Figura 4.3 ilustra, conceitualmente, como se dá a gestão dos dados dos sistemas de informação da saúde, dirigida pelo DATASUS. As aplicações em produção demandam tecnologias e SGBD capazes de suportar *workloads* dinâmicos, onde as observações podem ser produzidas, coletadas, integradas, alteradas e incluídas. Este processamento de transações online (OLTP, do inglês *online transaction processing*) obedece a restrições muito bem definidas pela modelagem do sistema de informação e deve suportar uma alta taxa de transações com características de atomicidade, consistência, isolamento e durabilidade. Por outro lado, fora da produção, o DATASUS fornece acesso ao *backup* dos bancos de dados destes sistemas de informação em seus canais online, ou área de *staging*. Como uma demanda das diretrizes de governança que determinam a transparência e preservação da privacidade, estabelecidas pelas leis de acesso à informação e de proteção dos dados pessoais, estes dados disponibilizados são pseudonimizados, processo que extrai todas as variáveis com potencial de identificação direta dos indivíduos. A partir deste ponto, o DATASUS, gestores, pesquisadores, técnicos e organizações podem estabelecer ambientes auxiliares para efetuar suas análises. Neste ambiente são executadas as tarefas de processamento analítico online (OLAP, do inglês *online analytical processing*).

O modelo de gestão OLAP ilustrado na Figura 4.3 compreende as principais tarefas da engenharia de dados. Tarefas analíticas mais complexas, *data-intensive*, se beneficiam de arranjos computacionais dispostos em cluster ou computação em nuvem para garantir o armazenamento e processamento massivo de grandes volumes de dados. Contudo, rotinas extração, transformação e carga (ETL) antecedem estas análises, produzindo *datasets* e metadados que servirão de entrada para modelos e *dashboards*. Cabe ao modelo de gestão de dados estabelecer quais são as melhores práticas para o manejo de dados neste ambiente. Por exemplo, um engenheiro de dados deverá estabelecer políticas de *backup* ou de nomenclatura para documentos, pastas, bases e zonas no *data lake*. Estratégias de indexação são elementos centrais para garantir a eficiência, escalabilidade e confiabilidade de todas as etapas presentes na Figura 4.3. Engenheiros de dados e DBAs são responsáveis por aplicar os métodos adequados para cada tipo de dado, volume de operações, tecnologia ou recurso computacional disponível.

Como descrito no sítio eletrônico do principal mantenedor das bases de dados de saúde do país, o Departamento de Informática do Sistema Único de Saúde (DATASUS), as diretrizes de gestão dos bancos de dados que suportam seus sistemas de informação são estabelecidas pela Metodologia de Administração de Dados (MAD), que normatiza regras gerais de uso e otimização. As especificações encontradas no MAD conformam os principais aspectos do modelo de dados, incluindo quais os SGBDs, definições de padrões, regras técnicas, normas de modelagem e dicionários de dados [DATASUS, 2024]. Fora deste ambiente de produção, as tarefas de Extração, Transformação e Carga (ETL) e a construção de painéis analíticos são suportados pelas ferramentas PowerCenter, PowerBI e Qlik Sense.

Na saúde, a estratégia de gestão de dados, juntamente com a governança, desempenha papel crucial ante a diversidade dos sistemas de informação e seus modelos de negócio. Sistemas que quando combinados permitem o mapeamento da linha da vida, levantamento dos aspectos clínicos e farmacológicos de uma população, estudo sobre a estrutura genética e o percurso social e hospitalar de um indivíduo. Cada tema e finalidade de construção do Sistema de Informação para a saúde repercute na forma como a informação será estruturada. Há um aspecto *ex ante* que antecede a construção do sistema, e que repercute nele; e que deve ser considerado na estratégia ampliada de gestão, e na proposta metodológica de administração de dados. Essas considerações possibilitam montar uma organização informacional a partir de estruturação de domínios e combinações capazes de relacionar os dados tanto fisicamente, quanto logicamente.

A estratégia e métodos de análise e organização de dados para saúde, devem considerar um modelo de negócio que enfatiza tanto o ciclo de vida dos dados, quanto a noção de cauda longa do dado. A primeira estabelece as considerações sobre a evolução de acesso e compartilhamento dos dados, enquanto a segunda enfatiza a frequência de uso e consumo dos dados. Neste sentido, a arquitetura de governança e, por conseguinte, a de gestão deve considerar (a) um pipeline de dados que se retroalimenta e interage; (b) os requisitos de "*informatics*", segundo os termos de Gray, ou da ecologia computacional e informacional necessárias para o contexto e o campo intelectual do negócio [Gray et al., 2007; Hey, Tansley, Tolle, 2009] e (c) uma arquitetura de dados orientada a serviços que deve ser delineada com vistas às operações repetitivas, objetivando a troca e o reaproveitamento de aplicativos, recursos e fluxos de trabalho. O ecossistema debatido deve projetar uma administração de banco de dados que considere a gestão da informação, o modelo computacional, a orquestração do fluxo de dados e a estratégia de documentação, catalogação e publicação de metadados.

Considerando que a gestão de dados aborda a funcionalidade necessária para administração eficiente e eficaz dos dados e, além da coleta e manipulação de dados, inclui ações de controle e avaliação de qualidade, gestão de metadados, gestão de catálogo de dados; a indexação na gestão de dados em saúde atrela-se à organização e à categorização desses modelos informacionais e dos metadados. Uma indexação eficaz apoia a recuperação de dados, suporta iniciativas de pesquisa, facilita o fluxo de trabalho e a eficiência operacional, apoia a integração de dados, e contribui para manutenção da qualidade e integridade de dados.

As considerações sobre o impacto da indexação na gestão de dados em saúde ampliam-se quando consideramos as ações para a Saúde Digital. Iniciativa que estimula o debate em torno de confederações e formação de redes segundo uma perspectiva de compartilhamento de dados de saúde que envolve atores distintos.

Iniciativas na pesquisa, como o Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS/Fiocruz), utilizam estas diversas bases de dados de saúde para compreender o impacto das iniquidades em saúde e avaliar políticas de proteção social na saúde populacional [Barreto et al., 2019]. As investigações conduzidas no CIDACS são suportadas por diversas etapas de computação intensiva de dados e conta com tecnologia de ponta de *big data*, computação de alto desempenho e segurança da

informação para apoiar a resposta das diversas questões levantadas pelos projetos do Centro.

A relevância das técnicas de indexação para garantir a gestão e manejo dos *datasets* em seu *data lake* podem ser ilustradas em três exemplos. Em primeiro lugar, sua principal ferramenta de integração de dados, o CIDACS-RL [Barbosa et al., 2020], tem suas propriedades de escalabilidade e desempenho baseadas numa solução eficiente de busca em bases indexadas para definir quais registros correspondem a mesma pessoa em bases de dados distintas. No segundo exemplo, gestão de grandes coleções de dados integrados, incluindo a Coorte de 100 Milhões de Brasileiros, a Coorte de Nascimentos [Paixão et al., 2021] e a Plataforma de Dados da Covid-19 [Junior et al., 2021] é apoiada pelo Dataverse⁵, que utiliza indexação para garantir encontrabilidade e visibilidade dos datasets. Um terceiro exemplo consiste no aplicativo de visualização da Coorte de 100 Milhões de Brasileiros⁶ que é capaz de sintetizar em gráficos uma quantidade considerável de dados sob demanda.

4.1.2. Um overview sobre indexação, sua evolução e uso prático em tarefas intensivas de dados

Índices são estruturas capazes de organizar os dados, orientando o armazenamento e otimizando operações de recuperação ou acesso, tais como consultas, filtros e varreduras [Kaur, 2018]. Um dos modelos mais populares, a *Árvore-B*, pode permitir consultas de intervalo, baseadas em ordenação e multidimensionais [Graefe et al., 2010]. Este modelo consiste em extrair um conjunto de amostras para uma série de valores temporalmente decrescentes geometricamente. Por mais que a *Árvore-B* tenha a capacidade de responder a consultas top-k(t) em tempo linear e com um custo de E/S esperado, ao lidar com fluxos de dados online que apresentam comportamentos desconhecidos a utilização da *Árvore-B* se torna inviável. Seguindo com índices baseados em árvore, o modelo KR+ introduziu um conceito de divisão e mesclagem de nós fazendo com que o acesso a múltiplos atributos fossem melhorados e com uma capacidade de recuperação de dados espaciais deturpados. Nesse modelo já é considerado tratativas para dados em nuvem e fornece suporte às em buscas de intervalo e de vizinhos próximos [Gani et al., 2014]. Já os índices de bitmap são capazes de acelerar operações em data warehousing, processamentos de análises online e diversas tarefas de gerenciamento de dados, mas não conseguem responder às consultas em um tempo viável [Wu et al., 2010]. Na indexação de hash, a busca de similaridade aproximada para dados de alta dimensão utilizando codificação de binários compactados para obtenção de resultados mais rápidos é o principal foco [Wu et al., 2010].

Os índices estão disponíveis nativamente em diversos sistemas gerenciadores de bases de dados (SGBDs). Uma pessoa responsável por administrar SGBDs consegue incluir facilmente os algoritmos tradicionais no seu projeto de otimização de processamento de consultas, definindo uma ou mais colunas de cada relação. Estruturas baseadas em intervalos, como *árvores-B*, são capazes de produzir uma organização de dados auto-balanceada que ordena o dado em páginas dispostas hierarquicamente

⁵ <https://dataverse.cidacs.org/>

⁶ <https://appcoorte.cidacs.org/home>

[Taranpreet et al., 2018]. Desta forma, ao buscar uma chave, o índice retornará em tempo logarítmico a página onde aquele valor se encontra, permitindo que aplicações desenvolvam suas próprias estratégias para buscar o valor de interesse naquele intervalo. Soluções posicionais, como os índices baseados em hash, deterministicamente mapeiam pontos de dados num *array* finito. A eficiência desta alternativa está associada à sua capacidade de evitar colisões que promovam o desbalanceamento, incluindo muitos registros num *array* com espaço limitado. Estruturas baseadas em filtros de Bloom são capazes de informar se uma determinada chave de busca está presente num conjunto de dados, esta característica permite evitar buscas irrelevantes e com resultados livres de falsos negativos [Kraska et. al, 2018].

Consultas do mundo real envolvem frequentemente a busca por múltiplos atributos usando uma chave de busca composta em um SGBD. Esse contexto pode exigir múltiplos índices ou um índice construído a partir de uma chave de busca multiatributo, um índice multivariado ou multidimensional (MI) [Silberschatz et al., 2020]. Geralmente, é possível produzir um MI a partir de uma chave de busca composta recorrendo à ordenação lexicográfica, *hashing* particionado ou arquivos de grade. Índices ordenados podem depender da ordenação lexicográfica e oferecer suporte a consultas de intervalo em atributos de dados. O *hashing* particionado pode retornar prontamente a interseção de *buckets* de cada valor de chave de busca codificado envolvido na consulta. No entanto, essas estratégias anteriores podem falhar ou levar a operações de E/S desnecessárias em consultas de condição de comparação. Os arquivos de grade fornecem particionamento uniforme em todas as dimensões de dados, abordando as deficiências das soluções ordenadas e baseadas em hash. Devido às operações intensivas de OLTP de dados dinâmicos, a principal limitação dos arquivos de grade relaciona-se ao *overhead* de manter a distribuição uniforme de instâncias em células e *buckets* [Silberschatz et al., 2020]. Casos especiais de MI se esforçam para oferecer suporte a consultas em dados espaciais ou temporais. A R-tree estende a B+-Tree para indexar recursivamente caixas delimitadoras ou tuplas bidimensionais que representam polígonos, pontos ou séries temporais. Kraska et al. (2021) argumenta que a adoção de algoritmos em SGBD frequentemente requer eficiência, maturidade e capacidade de lidar com diversas situações. Ainda assim, os desafios do gerenciamento de big data impõem uma revisão do modelo relacional e de suas soluções de indexação tradicionais embutidas.

Avanços em alternativas de SGBD para endereçar os requisitos do big data produziram modelos de dados de documentos e vetores. Sistemas de gerenciamento de banco de dados vetoriais (VDBMS) [Kukreja et al., 2023] e sistemas de gerenciamento de banco de dados orientados a documentos (DODBMS) [Banker et al., 2016] são responsáveis por fornecer armazenamento eficiente, processamento de consultas, transações, privacidade e escalabilidade para um grande volume de registros e atributos que não possuem um esquema fixo. Com relação à indexação, DODBMS se beneficia das estruturas de *B+-Tree* para permitir a recuperação rápida de informações [Banker et al., 2016], processamento de consultas de similaridade e intervalo. VDBMS, por outro lado, frequentemente suporta consultas em um conjunto mais amplo de modalidades de dados, armazenados após codificação em vetores de características n-dimensionais por meio de modelos de *embedding*. As estruturas de índice em

VDBMS são categorizadas como baseadas em tabelas [Datar et al., 2004], baseadas em árvores [Dasgupta et al., 2008] ou baseadas em gráficos [Harwood et al., 2016]. Essas soluções inovadoras enfrentam o custo da comparação de distâncias e a ausência de uma estrutura explícita para particionar vetores, permitindo buscas de similaridade e de vizinho mais próximo. Fora desses sistemas, os pesquisadores no campo de gerenciamento de dados têm discutido diversos índices adequados ao *big data*.

A mudança de técnicas de indexação para *big data* é dupla: extensões para algoritmos tradicionais e técnicas aprimoradas de aprendizado de máquina [Gani et al., 2014]. As extensões de ponta são projetadas para produzir estruturas de MI sofisticadas para aplicações específicas, frequentemente focando em grandes volumes de dados não-estruturados e multimodais. Por exemplo, o índice T-PARINET combina particionamento de gráficos e a árvore *B+ -Tree* para permitir busca e recuperação rápida em consultas espaço-temporais [Sandu et al. 2014]. *KR+ -Tree* é implementado no Cassandra, um sistema de banco de dados distribuído de código aberto, para permitir acesso a dados espaciais através de consultas de faixa e vizinhança mais próxima em um ambiente de computação em nuvem [Wei et al. 2014]. Por outro lado, estruturas construídas a partir de computação suave são principalmente derivadas de métodos baseados em gráficos, usando decisões binárias aprendidas a partir de dados rotulados para navegar e escanear o índice [Gani et al., 2014]. Em [LAZARIDIS et al., 2013] um mecanismo de recuperação de informações, consultas de similaridade em bancos de dados multimídia, incluindo objetos 3D, imagens, vídeo e áudio, são permitidas. Sua metodologia abrange o uso de técnicas de aprendizado de variedades para extrair descritores de itens de mídia, construir uma estrutura de índice que vincula cada entrada de dados transformada ao objeto original respectivo e usar anotação de propagação para atribuir automaticamente conteúdo amigável ao humano aos registros. Apesar dos resultados de desempenho promissores, principalmente apoiados pela extensão do trabalho de [Amato et al., 2013], os autores apontam para a falta de suítes de *benchmark* e *baselines* para fundamentar a vantagem em relação às alternativas tradicionais e escaláveis.

Para fins de exemplo, considere um cenário hipotético em que um conjunto de gestores da saúde desejam construir um modelo preditivo capaz de antecipar o desfecho de um caso suspeito. Inicia-se então um *pipeline* OLAP, ilustrado na Figura 4.3. Neste contexto, uma pessoa DBA ou Engenheira de Dados deve: i) alinhar com especialistas quais são as bases de dados, atributos e filtros necessários para suportar este modelo; ii) executar tarefas de aquisição dos dados nos portais e serviços do DATASUS; iii) armazenar estas bases no *data lake* de seu domínio; escrever e executar rotinas de pré-processamento dos dados (padronização, harmonização e correção); e iv) tornar disponível em uma zona do *data lake* o *dataset* pronto para treinar e avaliar o modelo. Após algumas rodadas de discussão, decidiu-se então que as variáveis independentes devem incluir a temperatura e a saturação sanguínea obtidos na triagem, a quantidade de sintomas gripais e aspectos de temporalidade do caso. Os intervalos de tempo entre a admissão e os eventos de coleta, resultado do exame e encerramento do caso devem ser considerados. Os códigos SQL a seguir foram usados para produzir a nova variável que calcula a quantidade de sintomas gripais e o conjunto de dados — amostra ilustrada na Tabela 4.1 — que será usado para treinar o modelo, doravante denominado *Dataset 1*:

```

-- código SQL para adicionar uma nova coluna "qt_sintomas" à
-- tabela tbl_info_epidemiologicas contendo a contagem de
valores "SIM"
-- nas colunas "s_dispneia", "s_disturbios_gustat",
"s_disturbios_olfat"
-- e "s_gripais".

ALTER TABLE tbl_info_epidemiologicas
ADD COLUMN qt_sint INT;

UPDATE tbl_info_epidemiologicas
SET qt_sintomas =
CASE WHEN s_dispneia = 'SIM' THEN 1 ELSE 0 END +
CASE WHEN s_disturbios_gustat = 'SIM' THEN 1 ELSE 0 END +
CASE WHEN s_disturbios_olfat = 'SIM' THEN 1 ELSE 0 END +
CASE WHEN s_gripais = 'SIM' THEN 1 ELSE 0 END;

-- código de consulta SQL que recupera, para todos os casos em
2022,
-- 2023 e 2024:
-- - temperatura no momento da triagem (t_temp);
-- - saturação de oxigênio no sangue no momento da triagem
(t_satur);
-- - quantidade de sintomas gripais (qt_sint);
-- - a diferença entre as datas dt_admis e t_dt_coleta
(tmp_coleta);
-- - a diferença entre dt_admis e tp_dt_resultado
(tmp_result1);
-- - a diferença entre t_dt_coleta e tp_dt_resultado
(tmp_result1);
-- - a diferença entre dt_admis e desencerra (tmp_total)
-- - a evolução final do caso (evolucao)

CREATE TABLE dataset1 AS
SELECT
i.id_caso,
i.t_temp,
i.t_satur,
e.qt_sint,
DATEDIFF(a.dt_admis, c.t_dt_coleta) AS tmp_coleta,
DATEDIFF(a.dt_admis, c.tp_dt_resultado) AS tmp_result1,
DATEDIFF(c.t_dt_coleta, c.tp_dt_resultado) AS tmp_result2,
DATEDIFF(a.dt_admis, a.dt_encerra) AS tmp_total,
a.evolucao
FROM
tbl_identificacao i
JOIN
tbl_info_epidemiologicas e ON i.id_caso = e.id_caso
JOIN
tbl_adm_encerramento a ON i.id_caso = a.id_caso
JOIN

```

```
tbl_info_clinicas c ON i.id_caso = c.id_caso
WHERE
YEAR(i.dt_nasc) IN (2022, 2023, 2024) AND
c.res_teste = 'POSITIVO';
```

Uma vez disponível na zona “REFINED” do *data lake*, o *Dataset 1* ilustrado na Tabela 4.1 pode ser submetido para uma análise de grupos por analistas ou cientistas. Esta estratégia objetiva segmentar os casos num número arbitrário de grupos, revelando relações escondidas entre eles [Kaufman et al., 2009]. O AGNES (*Agglomerative Nesting*) é um dos modelos de aprendizado não-supervisionado tradicionalmente usados para agrupamento de dados. Em resumo, o modelo estabelece numa primeira iteração que todos os objetos de dados são *singletons*, ou seja, grupos formados por apenas um registro. As iterações seguintes efetuam fusões sucessivas dos grupos de registros mais similares até que haja um único grupo. Cada iteração fornece uma quantidade decrescente de grupos, dispostos hierarquicamente, e cientista de dados é responsável por fazer inspeções que subsidiem a decisão do melhor número de grupos para um determinado conjunto de dados [Kaufman et al., 2009; Witten et. al, 2002].

Tabela 4.1. Dataset 1 resultado da consulta SQL para recuperar o perfil dos casos com resultado positivo do teste laboratorial no banco de dados da Figura 4.2.

id_caso	t_temp	t_satur	qt_sint	tmp_coleta	tmp_result 1	tmp_result2	tmp_tota 1	evolucao
62483	40,11	87	3	0	0	0	5	OBITO
69157	35,54	93	3	1	2	2	0	CURA
49407	35,16	93	4	1	0	0	2	CURA
578812	42,1	92	4	1	0	0	2	CURA
86729	38,39	99	2	0	0	0	5	CURA
4738	40,24	86	4	0	0	0	2	OBITO
4470	38,79	96	2	0	1	0	3	CURA
17781	39,33	98	4	0	0	0	2	OBITO
68312	37,12	93	2	0	0	0	3	OBITO
31248	37,52	95	2	0	0	0	2	CURA

Uma implementação do AGNES pode incluir uma estrutura de dados que auxilia a decisão sobre qual par de grupos deve ser mesclado a cada iteração. Esta estrutura mimetiza uma matriz de distâncias que consiste na combinação de todos os pares únicos

de registros, sua respectiva distancia euclideana. Uma coluna adicional *membership* do *id_caso_a*. A primeira estrutura é a matriz de distancias, de cada par de registro. Considerando uma combinação $C(n, 2)$, sendo o número de registros no Dataset 1 $n = 10$, então a quantidade de pares desta estrutura de dados, expressa na Tabela 4.2, será de $\frac{n \times (n-1)}{2} = \frac{10 \times (10-1)}{2} = 45$ registros.

```
-- Calcular a matriz de distâncias

CREATE TABLE matrix_distancias AS
SELECT
  a.id_caso AS id_caso_a,
  b.id_caso AS id_caso_b,
  a.id_caso AS membership,
  SQRT(
    POW(a.t_temp - b.t_temp, 2) +
    POW(a.t_satur - b.t_satur, 2) +
    POW(a.qt_sint - b.qt_sint, 2) +
    POW(a.tmp_coleta - b.tmp_coleta, 2) +
    POW(a.tmp_result1 - b.tmp_result1, 2) +
    POW(a.tmp_result2 - b.tmp_result2, 2) +
    POW(a.tmp_total - b.tmp_total, 2)
  ) AS distancia_euclideana,
  MOD(ROW_NUMBER() OVER (ORDER BY a.id_caso), 10) AS
membership
FROM
  dataset1 a
JOIN
  dataset1 b
ON
  a.id_caso < b.id_caso
ORDER BY
  a.id_caso, b.id_caso;
```

A Tabela 4.2 mostra o resultado do código SQL acima. Esta estrutura já traz o resultado da primeira iteração do AGNES, em que se atribui para cada registro, identificado pelo *id_caso_a*, um número de grupo contento apenas ele.

Tabela 4.2. Resultado da consulta SQL para produzir uma estrutura de dados com as distâncias de todos os pares únicos do Dataset 1 (Figura 4.1), sua distancia e cluster.

id_caso_a	id_caso_b	dist_euclideana	membership
62483	69157	9,40	0

62483	49407	9,16	0
62483	578812	6,98	0
62483	86729	12,08	0
62483	4738	3,18	0
62483	4470	9,62	0
62483	17781	11,61	0
62483	68312	7,47	0
62483	31248	9,11	0
69157	49407	2,83	1
69157	578812	7,56	1
69157	86729	8,54	1
69157	4738	10,91	1
69157	4470	7,21	1
69157	17781	8,77	1
69157	68312	4,03	1
69157	31248	3,74	1
49407	578812	7,55	2
49407	86729	10,87	2
49407	4738	5,72	2
49407	4470	7,72	2
49407	17781	10,63	2
49407	68312	5,20	2
49407	31248	7,17	2
578812	86729	13,61	3
578812	4738	6,27	3
578812	4470	9,68	3
578812	17781	12,40	3
578812	68312	8,11	3
578812	31248	10,30	3
86729	4738	16,08	4
4470	86729	3,00	5

17781	86729	2,65	6
68312	86729	7,21	7
31248	86729	7,34	8
4738	4470	9,11	9

Nas

iterações subsequentes, utilizando o parâmetro de *single linkage*, os grupos mesclados serão aqueles que possuem registros de menor distância. O seguinte código SQL será, portanto, executado exaustivamente até que só reste um cluster.

```
-- encontrar os grupos a ser mesclados: registros com
-- menor distancia -- entre si que se encontram em grupos
-- diferentes (single linkage).

SELECT MIN(dm1.dist_euclideana)

FROM
  distance_matrix dm1
INNER JOIN
  distance_matrix dm2
ON
  dm1.id_caso_b = dm2.id_caso_a
AND
  dm2.membership <> dm1.membership
```

A consulta SQL acima é central para execução completa do algoritmo AGNES. Portanto, cabe entender quais as possibilidades de otimização do seu processamento. Diferentes SGBDs oferecem ferramentas que permitem DBAs e Engenheiros de Dados analisar o plano de execução de uma *query* complexa. Suponha que o grafo de processamento da Figura 4.4 expressa o plano de execução real de um SGBD.

As etapas realçadas em vermelho na Figura 4.4 concentram os trechos mais complexos do grafo de processamento. Em casos com um número maior de registros, varrer todas as tuplas da tabela *distance_matrix* para efetivar o filtro que suporta a função `INNER JOIN` é proibitivo, principalmente num contexto em que esta mesma *query* será executada inúmeras vezes.

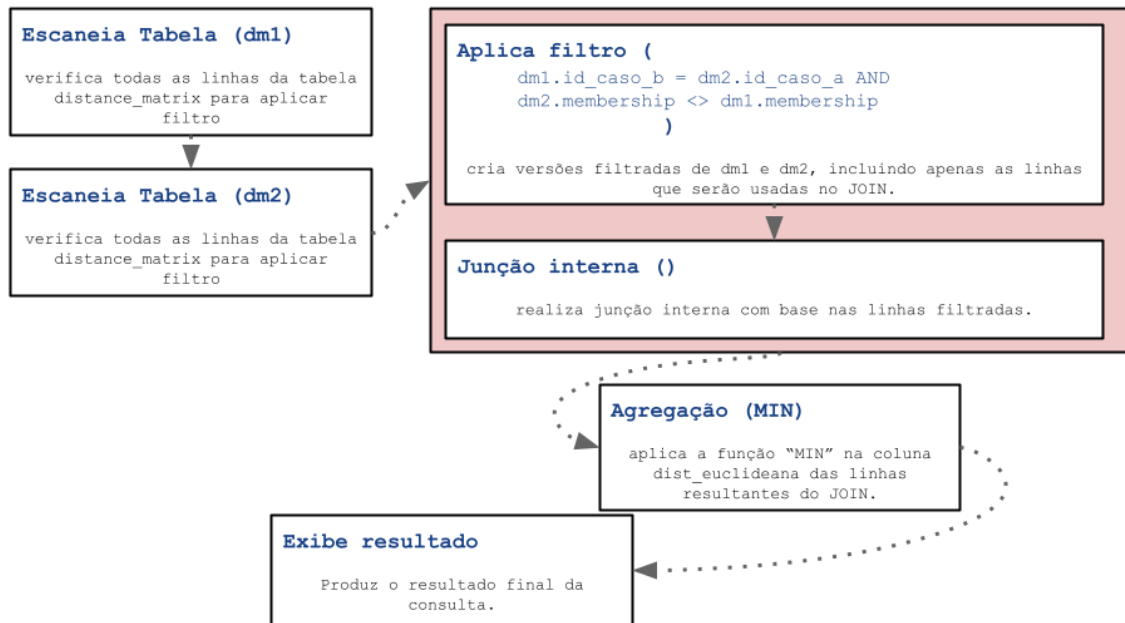


Figura 4.4. Grafo de processamento de query SQL que busca na tabela `distance_matrix` pelo par de registros com menor distância entre si até então atribuídos a grupos diferentes.

-- Definindo índice uma estrutura de índice a partir do atributo "id_caso_b" para auxiliar o processamento da consulta SQL expressa na Figura 4.4.

```
CREATE INDEX idx_dm1_id_caso_b ON distance_matrix
(id_caso_b) USING BTREE;
```

-- Definindo índice uma estrutura de índice composta a partir dos atributos "id_caso_a" e "membership" para auxiliar o processamento da consulta SQL expressa na Figura 4.4.

```
CREATE INDEX idx_dm2_id_caso_a_membership ON
distance_matrix (id_caso_a, membership) USING BTREE;
```

Após a definição do índice B+ -Tree — padrão para índices B-Tree no PostgreSQL e MySQL — nos atributos de interesse, o processamento ilustrado na Figura 4.4 pode ser sensivelmente otimizado [Silberschatz et al., 2020]. O principal ganho é a possibilidade de escanear toda a tabela `distance_matrix` duas vezes para efetivar o filtro. A estrutura de índice B+ -Tree dispõe os valores de uma ou mais colunas numa árvore balanceada, garantindo que a distância, expressa em saltos, entre o nó raiz e o qualquer nó folha seja a mesma. Cada nó deste índice multinível possui três componentes: i) um ponteiro que indica chaves, nó, ou nível com valores que antecedem

a chave corrente; ii) o valor da chave corrente; e iii) um ponteiro que indica chaves, nó, ou nível com valores que sucedem à chave corrente. O Algoritmo 1 ilustra um pseudo-algoritmo adaptado de [Silberschatz et al., 2020] para construção de uma estrutura B+ -Tree.

Algoritmo 1: pseudo-código para construção de estrutura B+ - Tree.

```

funcao insere(valor K, ponteiro p):
  se arvore == vazio:
    cria um nó folha L que também é o raiz
  senão:
    encontra nó folha L que deverá conter o valor K

  se L possui menos de n-1 chaves:
    insere_na_folha(L, K, P)

  ou se L já possui n-1 valores:
    cria-se o nó L'
    Copia L.P1 ... L.Kn-1 para um bloco T que pode segurar
    n pares de chave-valor

    insere_na_folha(T,K,P)
    L'.Pn ← L.Pn
    deleta L.P1 até L.Kn-1 de L

    Copia T.P1 até T.K(n/2) de T em L iniciando em L.P1
    Copia T.P(n/2)+1 até T.Kn de T em L' iniciando em L'.P1
    O menor valor de K' fica em L'
    insere_no_pai(L,K',L')

funcao insere_no_folha(nó L, valor K, ponteiro P):
  se(K < L.K1):
    insere P, K em L após L.P1
  senão:
    Ki se torna o maior valor em L que é menor ou igual a
    K
    Insere P,K em L após L.Ki

funcao insere_no_pai(nó N, Valor K', nó N'):
  se(N é o nó raiz):
    cria-se um novo nó R contendo N, K', N' #N e N' são
    ponteiros
    R se torna a raiz da árvore
  return

```

```

P <- pai(N)

se (P possui menos de n ponteiros):
    insere(K', N') em P após N

senão:
    Copia P para um bloco de memória T que pode receber P
    e (K', N')
    Insere (K', N') em T após N
    Apaga todas as entradas de P
    Cria um nó P'
    Copia T.P' ... T.P'_{⌈(n+1)/2⌉} em P
    K'' <- T.K'_{⌈(n+1)/2⌉}
    Copia T.P'_{⌈(n+1)/2⌉+1} ... T.P'_{n+1} em P'
    insere_no_pai(P, K'', P')

```

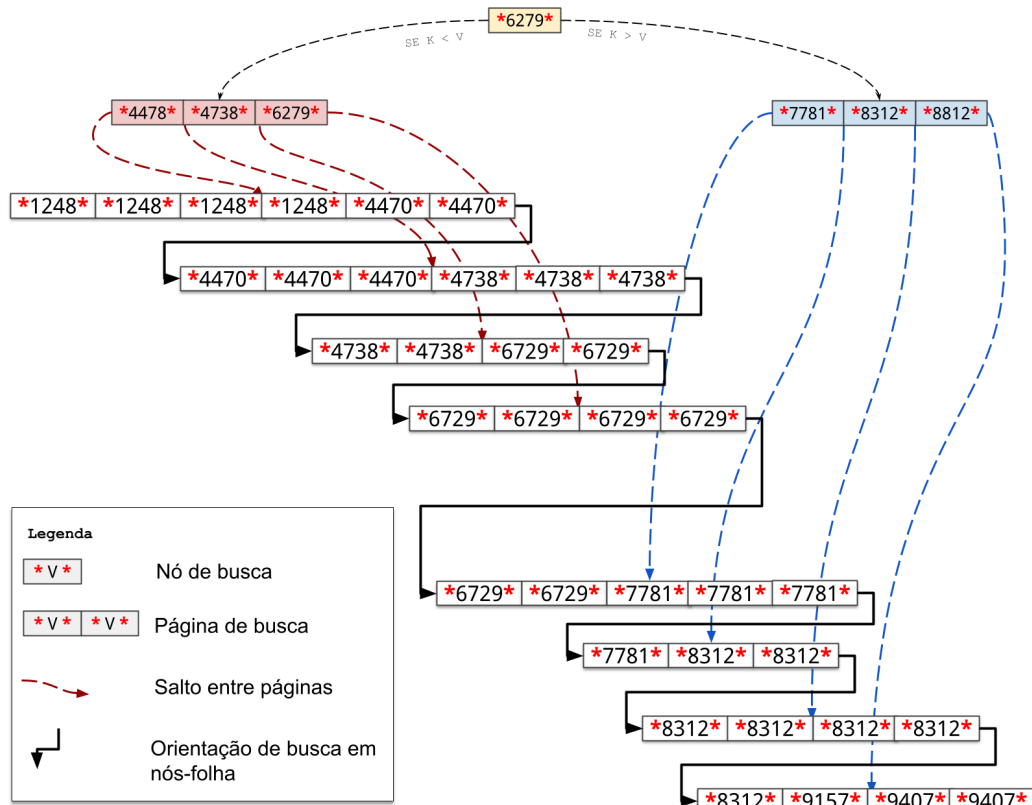


Figura 4.5. Estrutura de índice B+-Tree produzida a partir do atributo `id_caso_b`.

Um DBA ou Engenheiro de Dados é responsável por efetuar tarefas de *tuning* que inclui, entre outros parâmetros, avaliar quais atributos de cada tabela podem ser indexados para otimizar a maioria das consultas (*workload*) de uma determinada equação. Ao utilizar o pseudo-código do Algoritmo 1 na coluna `id_caso_b`, como sugere o código SQL acima, indicando um valor de $K = 7$, referente ao grau da

árvore a ser construída, obtemos uma estrutura de índice expressa na Figura 4.5. Note que, considerando o grafo de processamento da Figura 4.4, a estrutura de índice construída limitará o escaneamento de valores em `id_caso_b` e `id_caso_a` a um espaço de busca muito menor, tornando esta tarefa intensiva de dados ainda mais eficiente.

Estruturas de índices produzidas a partir de algoritmos tradicionais como árvores-B, hash e filtros de Bloom são capazes de organizar os dados de maneira a facilitar sua recuperação posterior, auxiliando o processamento de queries aos registros de uma determinada base de dados, provendo a eficiência necessária para diversas tarefas em casos de alta-dimensionalidade. Os índices estão disponíveis nativamente em diversos sistemas gerenciadores de bases de dados (SGBDs). Uma pessoa responsável por administrar SGBDs consegue facilmente incluir os algoritmos tradicionais no seu projeto de otimização de processamento de consultas, definindo uma ou mais colunas de cada relação. Estruturas baseadas em intervalos, como árvores-B, são capazes de produzir uma organização de dados auto-balanceada que ordena o dado em páginas dispostas hierarquicamente [Taranpreet et. al., 2018]. Desta forma, ao buscar uma chave, o índice retornará em tempo logarítmico a página onde aquele valor se encontra, permitindo que aplicações desenvolvam suas próprias estratégias para buscar o valor de interesse naquele intervalo. Soluções posicionais, como os índices baseados em hash, deterministicamente mapeiam pontos de dados num array finito. A eficiência desta alternativa está associada à sua capacidade de evitar colisões que promovam o desbalanceamento, incluindo muitos registros num *array* com espaço limitado. Estruturas baseadas em filtros de Bloom são capazes de informar se uma determinada chave de busca está presente num conjunto de dados, esta característica permite evitar uma buscas irrelevantes e com resultados livres de falsos negativos [Kraska et. al, 2018].

4.2. Learned index: usando machine learning para superar modelos tradicionais de indexação

Sistemas de dados de propósito geral demandam o ajuste de uma miríade de hiper-parâmetros, ou *knobs*, no esforço de otimizar seus componentes para atender conjuntos de dados ou *workloads* específicos. A estrutura de índice ilustrada na Figura 4.5 é um exemplo de como pequenas alterações na entrada do algoritmo podem alterar o resultado. Na figura, o DBA decidiu construir uma árvore de grau 7. Isso implica que cada nó da árvore deve conter, no máximo, 6 valores de chaves ordenadas. Uma rápida inspeção nos valores do atributo `id_caso_b` da Tabela 4.2 permite verificar que vários valores se repetem mais que 4 vezes. Esta duplicação implica numa fragmentação dos valores em mais de um nó de segundo nível, aumentando o espaço de busca e tornando o método menos eficiente. Por outro lado, outro algoritmo, como o *Hash* poderia ser aplicado neste caso e alcançar melhores resultados de latência, uso de memória e eficiência de armazenamento. Por décadas, SGBDs de propósito geral proviam versatilidade a partir destes domínios de intervenção, os algoritmos e os parâmetros de ajuste disponíveis para a adequação em certos cenários.

O uso de inteligência artificial em sistemas de dados não é uma prática recente. Diversos SGBDs ou soluções *ad hoc* já incorporam componentes inteligentes [Li et al. 2021]. As soluções inteligentes abordam, principalmente, aspectos de configuração e otimização. Consultores automáticos de indexação, *tuning* e *views* são exemplos de

componentes habilitados por aprendizado de máquina para configuração. Por outro lado, otimizadores inteligentes focam em obter um maior desempenho no processamento de *queries*, oferecendo estimativas de cardinalidade e ordenação para suportar operações de junção [Li et al. 2021].

Uma terceira vertente do uso de inteligência artificial em sistemas de dados foca no redesenho completo de seus componentes, tornando-os autoajustáveis para qualquer distribuição de dados ou padrão de carga de trabalho [Kraska, 2021; Ding et al., 2022]. Nas seções seguintes, apresentaremos com detalhes o Learned Index e seus principais avanços desde sua concepção em 2018. Contudo, Por mais que existiram alguns esforços para implementar os componentes otimizados por instância, como o Bourbon que substituiu índices de blocos em uma árvore LSM por Learned Index e o Google que integrou índices aprendidos no *BigTable*, grande parte dos componentes otimizados foram criados e avaliados isoladamente [Ding et al., 2022].

O conceito de utilizar modelos de aprendizado de máquina para substituir algoritmos também foi explorado em diversos outros componentes de sistemas de dados. Soluções como otimizadores de consultas [Marcus et al., 2019; Mikhaylov et al., 2022], uso de aprendizado por reforço para escalonadores de tarefas [Mao et al., 2016], layout de dados [Nathan et al., 2020; Campero Durand et al., 2019] e ajuste adaptativo de parâmetros para tuning do sistema de dados [Zhao et al., 2023; Li et al., 2021]. Alguns desses desenvolvimentos alcançaram maturidade suficiente para serem integrados em serviços relevantes de computação em nuvem. Hussam et al., (2020) relata em seu trabalho a solução de LI capaz de armazenar sua estrutura em disco, em oposição à solução baseada em memória na proposta original. A Amazon incorporou componentes de sistemas de dados habilitados por ML em seus serviços da AWS e Redshift [Fanggohans et al., 2023; Armenatzoglou et al., 2022]. A Azure também incluiu estratégias de ML para efetuar tuning automático no SQL Server da Azure [Fritchey et al., 2022]. Além disso, a exploração desta lacuna na área de gestão de dados serviu de base para uma nova classe de soluções, chamadas de sistemas de dados otimizados para instância (IODS, do inglês instance-optimized data systems), capazes de se ajustar automaticamente a qualquer distribuição de dados ou workload. Vários IODS empregam ML em seus componentes principais para aliviar o trabalho quase proibitivo de administradores e engenheiros de dados, oferecendo sistemas de gerenciamento de dados autônomos [Ding et al., 2022; Li et al. 2019; Ma et al., 2022; Li et al., 2021; Zhou et al., 2021; Ma et al., 2021].

4.2.1. O conceito de Learned Index

Em Kraska et al. (2018) é proposto o conceito de Índice Aprendido (LI, do inglês *Learned Index*), uma estrutura de índices produzida a partir da estimação da distribuição dos dados para otimizar as consultas em banco de dados. O principal argumento para o uso de um método de indexação habilitado por ML é que a estimação adequada de uma distribuição de dados utilizando aprendizado de máquina (ML, do inglês *machine learning*) tem o potencial de superar a eficiência dos métodos tradicionais. Para fins de exemplo, considere um vetor $v_j = \{0, 1, 2, 3, \dots, N\}$ contendo N chaves inteiras ordenadas do $j^{\text{ésimo}}$ atributo. Como mencionado na Seção 4.1.2, o uso de um algoritmo

de indexação, como o B+ -Tree, pode reduzir a complexidade da busca de uma determinada chave de $O(n)$ para $O(\log n)$. O conhecimento sobre esta distribuição, contudo, pode permitir que o valor da própria chave de busca seja utilizado como offset, reduzindo a complexidade para $O(1)$. Kraska et al. (2018) argumenta que este tipo de otimização pode ser alcançado em outras distribuições mais complexas através do uso do LI.

O LI é trazido visando a substituição dos modelos tradicionais conhecidos (árvores, hash, bitmap e etc) para um novo que constroi índices a partir de funções que descrevem como os dados estão distribuídos. Por mais que seja difícil imaginar que modelos de aprendizagem de máquina sejam capazes de garantir a mesma eficiência e assertividade de modelos tradicionais, Kraska et al. (2018) conceitua uma árvore B como um “modelo”, o qual tem a capacidade de “prever” a localização de um valor dentro de um conjunto com chaves ordenadas. Em uma base de dados analítica em memória, onde a chave primária está ordenada, a indexação mapeia uma posição dentro do vetor ordenado que contém um valor igual ou menor à da chave de consulta. Com isso, o modelo da Árvore B contém um erro máximo do tamanho da folha. Sendo assim, a ideia do LI é substituir esse modelo por outros, como redes neurais, desde que sejam capazes de fornecer garantias semelhantes ao erro máximo e mínimo.

Uma comparação realizada entre o modelo LI e a Árvore-B é feita em Kraska et al. (2018), onde é exemplificado que o modelo tradicional necessita re-equilibrar a disposição dos dados a cada inserção e, de forma semelhante, um modelo que utiliza de aprendizado de máquina para indexação necessitaria realizar um re-treinamento dos dados para continuar apresentando resultados satisfatórios de busca. Outra semelhança entre os modelos é que ambos realizam um mapeamento que retorna um intervalo onde o possível dado está (estará caso exista), porém com a abordagem de LI o espaço amostral do erro é consideravelmente menor, visto que a disposição dos dados é conhecida e a previsão do índice acaba apresentando um apontamento mais eficaz.

Outro ponto abordado por Kraska et al. (2018) é a busca pontual. Tradicionalmente, mapas de hash são utilizados para esse tipo de busca e usam uma função de hash para mapear as chaves para posições dentro do vetor. Porém esses modelos podem acabar gerando conflitos ao mapear várias chaves para a mesma posição dentro do mapa de hash, visando evitar esse tipo de problema, vários estudos foram desenvolvidos, inclusive soluções onde é utilizado aprendizado de máquina. Porém nenhuma dessas soluções traz consigo a utilização do aprendizado da disposição de dados subjacentes. Com isso em vista foi proposto a utilização da CDF como uma maneira de aprender uma função de hash, nesses casos os dados não precisam ser guardados de forma ordenada. O modelo utilizado para nessa nova versão de busca pontual continua sendo o recursivo e depende de dois fatores importantes para funcionar com um bom desempenho: a representação da CDF deve ser precisa e a arquitetura do mapa de hash deve ser bem pensada, onde características de resolução de conflitos e a quantidade de memória que pode ser alocada.

No que diz respeito às estruturas LI, estudos focados em investigar diferentes abordagens estão disponíveis para aproximar a CDF dos dados subjacentes [Kipf et al. 2020; Ding et al., 2020; Ferragina et al. 2020], a indexação de dados multidimensionais

[Nathan et al., 2020] e de strings [Wang et al., 2020], a escalabilidade fornecida por arquiteturas *multicore* [Tang et al., 2020] e de GPU [Zhong et al., 2022; Liu et al., 2024], e permitir a persistência da CDF aproximada no disco [Hussam Abu-Libdeh et al., 2020]. Nas próximas seções, traremos um overview sobre o conceito de função de distribuição cumulativa, aplicando para os dados sintéticos que apresentamos na Figura 4.1. Em seguida, focaremos nossa exposição nos LI que compõem o estado-da-arte, estendidos a partir da sua primeira publicação.

4.2.2. Estimando a distribuição dos dados

Para introduzir o conceito de função de distribuição acumulada (CDF, do inglês *Cumulative Distribution Function*), denotaremos X como uma variável aleatória, que pode ser contínua, ou seja, assume valores reais em um intervalo, como, por exemplo, a altura, em centímetros, de uma pessoa, ou discreta, onde assume valores finitos ou contáveis, como por exemplo, a idade de uma pessoa em anos completos [Ross, 2014]. Para cada variável aleatória X , seja ela discreta ou contínua, existirá uma função $p(x)$ ou $F(x)$, respectivamente, que irá associar a cada possível valor $x \in X$ uma probabilidade. Tal função será chamada de função de probabilidade ou função densidade de probabilidade, se certas condições forem satisfeitas.

Outro conceito importante associado à variável aleatória X é o da função de distribuição acumulada, comumente denotada por $F(x)$, que representa a probabilidade de que X assumira um valor menor ou igual a x . Formalmente, descrevemos

$$F(X) = \int_{-\infty}^x f(x)dx, \text{ no caso contínuo e } F(X) = \sum_{i=1}^x p(x), \text{ para o caso discreto}$$

[Wong et al., 2001]. Como exemplo, consideremos que um dado não viciado é lançado, denotaremos como a variável X , o valor da face do dado, na tabela abaixo são mostrados os possíveis resultados, a probabilidade e a probabilidade acumulada:

Tabela 4.3: Tabela de resultados e probabilidades referente ao lançamento de um dado não viciado.

Resultados possíveis	$p(x)$	$F(x)$
1	1/6	1/6
2	1/6	2/6
3	1/6	3/6
4	1/6	4/6
5	1/6	5/6
6	1/6	6/6 = 1

Dados esse exemplo consideremos a seguintes questões:

1) Qual a probabilidade de sair a face 6?

$$R: p(x = 6) = 1/6$$

2) Qual a probabilidade do resultado ser no máximo 3?

$$R: p(x \leq 3) = 3/6 = F(x = 3)$$

3) Qual a probabilidade do resultado ser maior do que 3 ?

$$R: p(x > 3) = 3/6 = 1 - F(x = 3)$$

Agora, consideremos uma variável aleatória X que tem função densidade de probabilidade dada pela função:

$$f(x) = \begin{cases} 0, & x < 1 \\ \frac{2}{x^3}, & x \geq 1 \end{cases}$$

Para podermos encontrar a função distribuição cumulativa precisamos dividir em duas partes:

$$\text{Se } x < 1, \text{ temos } \int_{-\infty}^x f(t) dt = 0$$

$$\begin{aligned} \text{Se } x \geq 1, \text{ temos } & \int_{-\infty}^x f(t) dt \\ &= \int_1^x f(t) dt \\ &= \int_1^x \frac{2}{t^3} dt \\ &= \left[-\frac{1}{t^2} \right]_1^x \\ &= -\frac{1}{x^2} + 1 \end{aligned}$$

Então:

$$F(x) = \begin{cases} 1 - \frac{1}{x^2}, & x \geq 1 \\ 0, & \text{c.c} \end{cases}$$

A função de distribuição acumulada possui algumas propriedades específicas, como:

i) é uma função não decrescente;

ii) é uma função contínua à direita;

iii) $\lim_{X \rightarrow -\infty} F(X) = 0$ e $\lim_{X \rightarrow +\infty} F(X) = 1$ [Magalhães, 2004].

Considerando as propriedades demonstradas anteriormente, é possível, desde que essa distribuição seja definida para as variáveis do banco de dados, treinar um determinado LI para que este, ao aprender com a distribuição desses dados, seja capaz

de executar as mesmas tarefas que já são realizadas pelos índices tradicionais de maneira mais eficiente e assertiva [Kraska et al., 2018].

Peguemos do exemplo inicial a variável referente a temperatura do paciente na hora da admissão na unidade hospitalar. Usando essas informações queremos estimar a função de distribuição de probabilidade dessa variável, para posteriormente utilizarmos essa informação para estimar a função de distribuição acumulado que servirá como base para o nosso índice aprendido. Essa implementação será feita no R, como mostrado abaixo:

```
library(ggplot2)

dados = data.frame(tem = c(36.93, 40.11, 35.54, 39.81, 35.16,
42.1, 37.67, 39.23, 38.23, 40.24, 38.79, 39.23, 40.24, 38.79,
36.33, 37.82, 37.12, 37.38, 37.52))

ggplot(dados) +
  aes(tem) +
  geom_histogram(aes(y= ..density..), fill="lightblue", colour
="black") +
  stat_function(fun = dnorm, args = list(mean =
mean(dados$tem), sd = sd(dados$tem)))

fn = ecdf(dados$tem)

plot(fn, main = " Distribuição Acumulada Temperatura" , xlab
= "Temperatura", ylab="F(x) ")
```

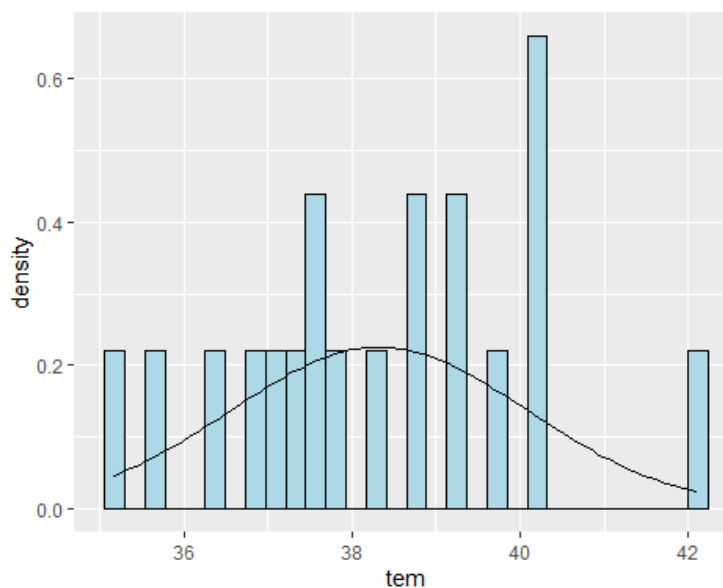


Figura 4.7. Distribuição de frequência da temperatura de 19 pacientes

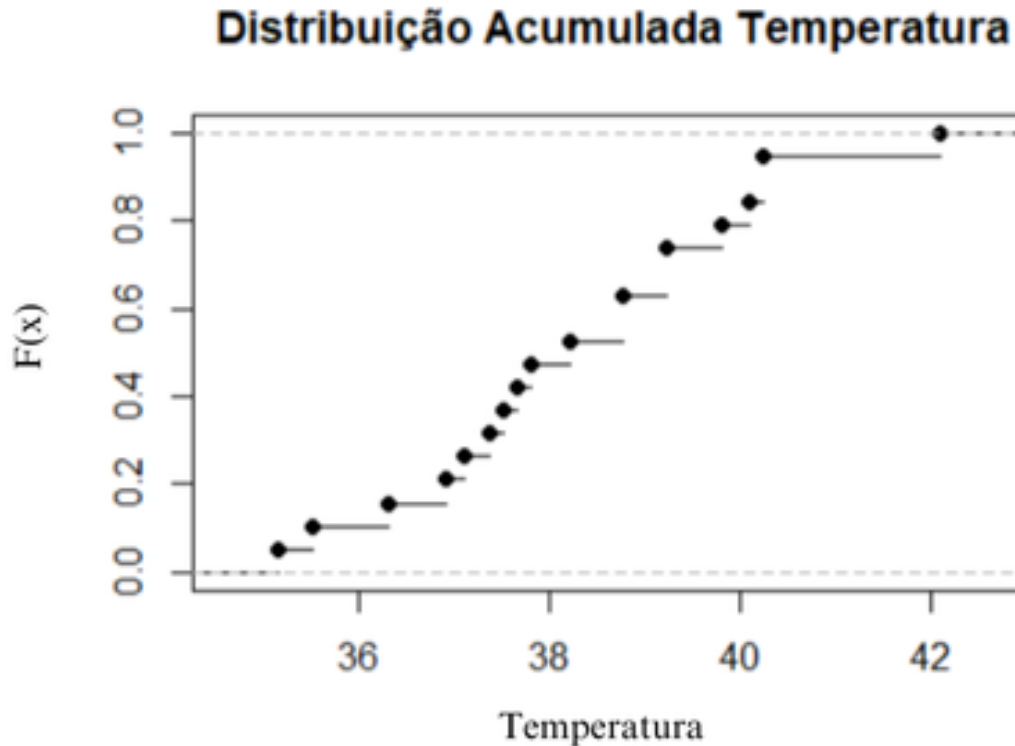


Figura 4.8. Distribuição Acumulada da temperatura de 19 pacientes

O formato no gráfico é diferente daquilo que conhecemos como formato de uma distribuição acumulada, isso porque nesse exemplo temos poucas observações o que não permite suavizar a curva da distribuição acumulada. Para tentarmos visualizar melhor a cara dessa distribuição, simularemos uma distribuição normal com média 38,33 e desvio padrão 1,77 (mesma distribuição dos dados de temperatura), mas aumentamos o tamanho amostral para 10.000.

```
amostra = rnorm(10000, 38.33, 1.77)

fn_amostra = ecdf(amostra)

plot(fn_amostra, main = "Distribuição Acumulada
Temperatura", xlab = "Temperatura", ylab="F(x)")
```

4.2.3. Recursive Model Index: o primeiro índice proposto

O framework proposto em Kraska et al. (2018) inclui o RMI, uma hierarquia de regressões estruturada para permitir que o modelo de cada nível receba a chave como entrada e a transfira para o modelo da camada seguinte mais apto a prever qual a posição de saída referente à chave buscada. O funcionamento do RMI é baseado na observação da localização de uma chave em um vetor ordenado e a capacidade de conseguir calculá-la usando uma função de distribuição acumulada dos dados. Este método, portanto, é responsável por estimar a CDF de um conjunto de dados, mapeando uma chave de busca para a sua posição num vetor ordenado [Maltry and Dittrich 2021]. Com isso, cada modelo é capaz de fazer uma previsão com um certo erro para a posição da chave e com isso é escolhido o próximo modelo que é responsável por outra área do

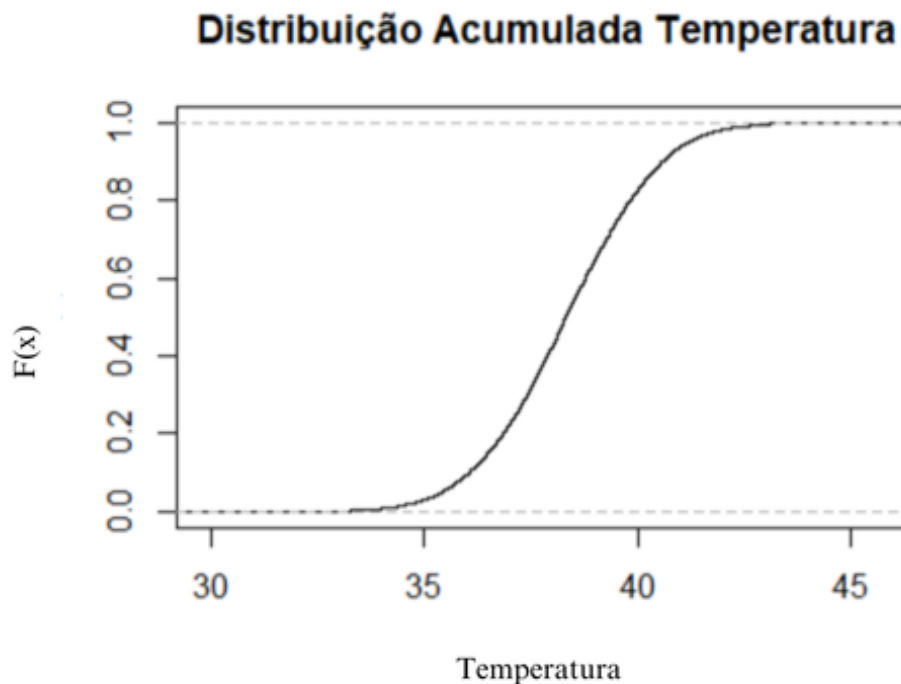


Figura 4.9. Distribuição Acumulada de dados simulados da temperatura de 10.000 pacientes

espaço de chaves e esse consegue fazer uma previsão mais assertiva com um erro menor que o modelo anterior. O RMI pode ser considerado como um grafo acíclico dirigido, diferentemente de um árvore, um nó (ou uma folha) do modelo recursivo pode ter múltiplos predecessores diretos e cada camada abaixo pode consistir em um número arbitrário de modelos [Maltry and Dittrich 2021].

A utilização do modelo RMI acarreta em alguns benefícios para o framework, Kraska et al. (2018) traz 4 deles, sendo: o tamanho e a complexidade do modelo são separadas para o custo de execução; o aprendizado da distribuição de dados é algo considerado fácil; o espaço de dados é dividido de maneira eficaz em subfaixas menores

(assim como na B-Tree) e não há uma busca entre os estágios dos modelos, fazendo com que a saída de um seja diretamente utilizada para selecionar o modelo da próxima etapa, fazendo com que o número de instruções gerenciais sejam reduzidas. Outra vantagem que a utilização de LI traz com o RMI é a capacidade de modelos híbridos. Por exemplo: em camadas superiores a utilização de uma RN ReLU pode se demonstrar como o mais eficaz, porém nas camadas intermediárias modelos de regressão linear podem se encaixar melhor por serem mais econômicos e, além desses modelos, a própria árvore B-Tree pode ser utilizada nos nós folhas, caso a distribuição dos dados apresente uma dificuldade de aprendizagem.

4.2.4. Índices para suportar OLTP

Por mais que os índices aprendidos apresentem uma boa resposta para atividades de leitura, as atividades de escrita eram um desafio para o primeiro modelo. Visando contornar esse problema, Ding et al. (2020) traz o ALEX, um modelo de learned index que oferece suporte a buscas pontuais, consultas de espaço, inserções, atualizações, exclusões e workloads, operações comuns em processamento de transações online (OLTP). Resumidamente, esse LI reserva espaços em memória para que novos valores possam ser inseridos sem a necessidade de deslocamento de chaves para cada inserção. O ALEX mantém os modelos precisos utilizando de mecanismos de expansão adaptativa e divisão dos nós atrelados com retreinamento de modelos, fazendo com que a estrutura seja possível de reagir a cargas de trabalho reais com mudanças dinâmicas.

Ding et al. (2020) traz algumas diferenças entre o LI e o ALEX. A primeira é a estrutura, onde o ALEX se assemelha com a árvore B+ para armazenar os dados nas folhas, utilizando apenas um nó por folha e com isso é possível realizar divisões e expansões mais facilmente para uma inserção. O ALEX usa de espaços livres nas folhas organizadas de forma estratégica para que a inserção e busca sejam mais eficientes. A próxima diferença é que o LI usa busca binária dentro dos limites fornecidos pelo modelo, enquanto o ALEX utiliza de busca exponencial para encontrar os valores dentro das folhas. Essa alteração permite uma busca mais rápida que a binária e também elimina a necessidade do armazenamento de limites de erro nos modelos do RMI. A terceira diferença é a inserção do ALEX ser feita na localização de onde a chave deve estar, essa é prevista pelo modelo e foi chamada de “Inserção Baseada em Modelo” (do inglês *model-based insertion*). A quarta diferença é que o ALEX consegue ajustar a forma e a altura do RMI de forma dinâmica. Por fim, a quinta diferença é que o ALEX não precisa ter seus parâmetros reajustados por carga de trabalho, diferentemente do LI que precisa ajustar o número de modelos.

O *layout* do ALEX apresenta dois tipos diferentes de nós: o folha e o interno. No nó folha (nós de dados), são armazenados os registros de dados. Esse nó possui um modelo de regressão linear que mapeia a chave para uma posição e possui dois vetores com lacunas, onde um mapeia a chave para uma posição e o outro é responsável pelas cargas de dados. Para não perder eficiência na busca, os vetores com lacunas ocupam os espaços vazios com a chave mais próxima à direita, fazendo com que a busca não seja afetada por esses espaços. Para passar pelas lacunas na busca, cada nó possui um bitmap que verifica se o próximo espaço está ocupado por uma chave ou se é um espaço vazio.

Já os nós internos são todos os que estão na estrutura do RMI, nesses são armazenados um modelo de regressão linear juntamente com um array contendo os ponteiros dos nós filhos. Os nós internos calculam a localização no array de ponteiros do próximo filho.

Já para a etapa de busca, o ALEX inicia no nó raiz do RMI e o modelo é utilizado de maneira iterativa para calcular uma localização nos ponteiros para o próximo filho, essa atividade é repetida até chegar no nó folha. Ao chegar no nó de dados é realizado uma previsão da chave de busca e — caso necessário — uma busca exponencial que retorna a posição real da chave. Caso essa seja encontrada, o valor correspondente é retornado e caso não o encontre, um registro nulo é retornado. Já para a busca de intervalo é realizada uma busca para a primeira chave onde o valor não é menor que o valor inicial do intervalo e depois é percorrido até chegar ao valor final.

Como explicado acima, o ALEX é um modelo capaz de realizar inserções no banco de dados e para essa etapa dois casos possíveis: em nós cheios e em nós não cheios. Na primeira, o algoritmo de busca é executado e ao achar o nó no qual o dado deveria estar, é feita a inserção. Nesse caso, caso o espaço seja uma lacuna, o dado é inserido e caso o espaço possua valor, um espaço vazio é criado e os elementos são empurrados em direção à lacuna mais próxima.

Na inserção de dados em nós cheios são executados alguns passos: primeiramente o nó é expandido criando um novo vetor de lacunas maior e o modelo é re-treinado ou escalado e os dados são inseridos nesse espaço. Ao realizar a divisão do nó, duas vertentes podem ser aplicadas: a lateral que se assemelha a B+ -Tree onde caso o nó pai ainda não tenha atingido o tamanho máximo, o nó de dados é dividido e o ponteiro do pai para o nó dividido é substituído para ponteiros para os novos nós de dados. E caso o nó do pai tenha atingido o tamanho máximo do nó, esse pode ser dividido e novos ponteiros são gerados para os nós filhos. Esse modelo de divisão se assemelha à B+ -Tree e a divisão pode chegar até o nó raiz. A segunda opção de divisão é a descendente, onde um nó de dados é convertido em um interno com dois filhos e esses são treinados com suas chaves.

Por fim, as exclusões e atualizações no ALEX são mais simples que as inserções, visto que para deletar um valor a busca é realizada para encontrar a chave e após isso faz-se a remoção. Essa atividade não realiza o deslocamento de chaves, tornando a operação mais simples que a inserção, caso o nó que possuía o valor acabe chegando ao limite inferior de densidade, o nó é retraído. Já as atualizações mesclam uma inserção com uma remoção.

4.2.5. Indexando dados multidimensionais

Índices multidimensionais são alternativas para os índices secundários no armazenamento de dados. Eles podem ser baseados em árvores (como a árvore k-d ou a árvore R) ou podem ser baseados em múltiplos atributos, como o Z-ordering. Sistemas como o Redshift, Spark SQL, Vertica e IBM Informix são alguns dos exemplos de banco de dados que usam técnicas de indexação de dados multidimensionais. Porém, esse modelo de indexação ainda apresentam desvantagens, como a dificuldade de ajuste (nesse caso os administradores de banco de dados precisam ajustar manualmente qual a

ordem de classificação e precisa saber quais colunas são acessadas juntas); outra desvantagem é que o índice multidimensional não possui uma universalidade de padrão, visto que esses variam conforme a distribuição dos dados e com a carga de *workload* [Nathan et al., 2020].

Com as limitações mostradas anteriormente, um framework de LI multidimensional em memória foi proposto em Natan et al. (2020). Nomeado de Flood, esse modelo objetiva buscar registros que correspondam a uma consulta de maneira

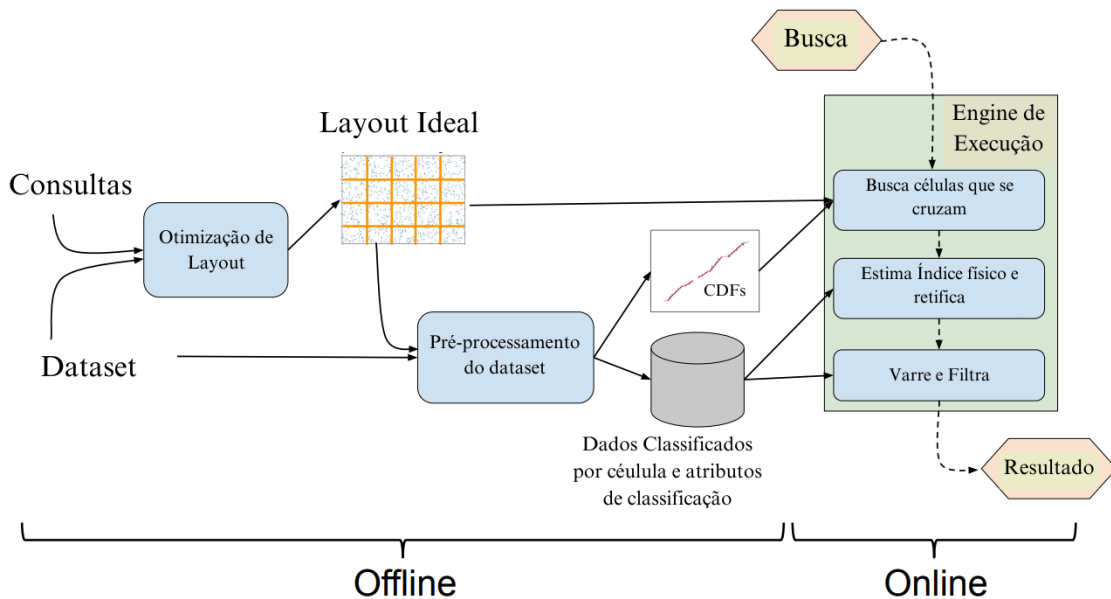


Figura 4.10. Arquitetura do Flood, é possível ver a distinção entre as partes offline e on-line. Fonte: Traduzido de [Natan et al, 2020]

mais rápida que os demais modelos de indexação existentes, para isso a estrutura do índice e o layout dos dados conseguirão ser ajustados automaticamente para uma distribuição de dados e consultas específicas. De maneira resumida, o Flood analisa uma carga de trabalho amostral para aprender a frequência que as dimensões são usadas, quais dessas são usadas juntas e quais são mais seletivas, com isso o layout do Flood é personalizado para otimizar a consulta dada. Outra característica do Flood é usar modelos empíricos de CDF para “projetar a distribuição de dados multidimensional e potencialmente enviesada em um espaço mais uniforme” Natan et al. (2020), com isso uma limitação do número de pontos pesquisados é aplicada e um melhor desempenho pode ser encontrado.

O Flood consegue suportar desde transações em memória OLTP à data warehouses em disco, porém o objetivo central do framework proposto é melhorar o desempenho de índices multidimensionais para um armazenamento em coluna em memória (os quais estão se tornando cada vez mais comum pela queda dos preços da memória RAM e aumento da memória principal alocada em máquinas). Outra característica do Flood é a otimização da leitura à custa da escrita, com isso a performance é maior em ambientes analíticos relativamente estáticos.

Por assumir que as consultas OR podem ser feitas através de diversas consultas de intervalo e que consultas de igualdade (ex: $R.Z == f$) pode ser reescrita como “ $f \leq R.Z \leq f$ ”, o Flood foca em cláusulas AND [Natan et al, 2020]. Sendo assim, o framework funciona em duas partes: primeiramente há um pré-processamento offline em que um layout ótimo é escolhido e o índice será baseado neste. A segunda etapa é online, onde um componente é responsável por executar as consultas na medida que chegam [Natan et al, 2020]. A arquitetura do Flood pode ser vista na Figura 4.10.

Para a estruturação dos dados, levando em consideração que ao contrário de dados unidimensionais, os pontos em múltiplas dimensões não possuem uma ordem de classificação natural, o Flood inicia com tratativas para impor uma ordenação sobre os dados. Para isso, os d atributos do dado são classificados. Com isso, as $d-1$ dimensões são usadas na ordenação para sobrepor a grade $(d-1)$ -dimensional dos dados, onde i -ésima dimensão na ordenação é dividida em c_i colunas igualmente espaçadas entre seus valores mínimos e máximos. Caso M_i e m_i sejam os valores máximo e mínimos dos dados na i -ésima dimensão, o intervalo da dimensão é $r_i = M_i - m_i + 1$ [Natan et al, 2020]. Com isso, a célula para o ponto $p = (p_1, \dots, p_d)$ é:

$$\text{célula}(p) = \left(\left\lfloor \frac{p_1 - m_1}{r_1} \cdot c_1 \right\rfloor, \dots, \left\lfloor \frac{p_{d-1} - m_{d-1}}{r_{d-1}} \cdot c_{d-1} \right\rfloor \right)$$

Na fórmula apenas as $d - 1$ dimensões aparecem, pois a dimensão d é utilizada para ordenação dos pontos dentro da célula. Durante essa etapa, uma travessia em profundidade é utilizada nas células ao longo da ordenação das dimensões, sendo assim, as células são classificadas por cardinalidade dentro da tupla (1º valor, 2º valor, ..., n -ésimo valor). Já dentro da célula os pontos são classificados pelo seu valor na

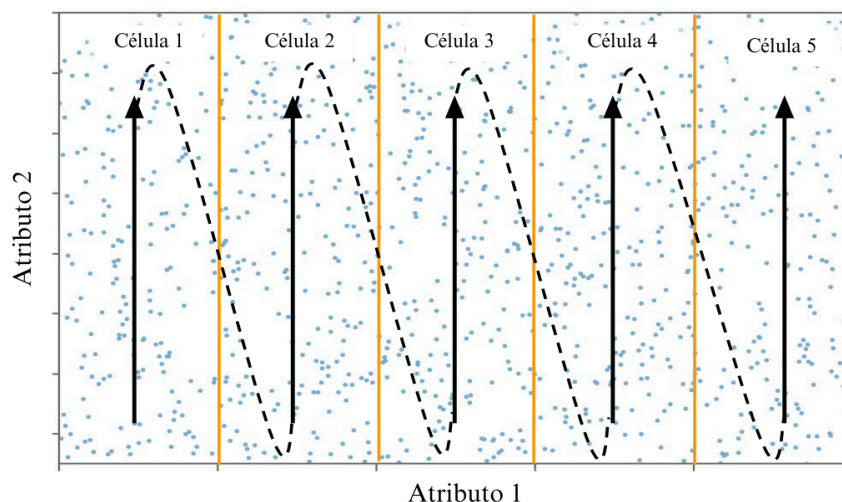


Figura 4.11. Estrutura de ordenação de 5 células em um cenário com dados de duas dimensões. Fonte: Traduzido de [Natan et al, 2020]

dimensão d . Na Figura 4.11 é possível ver essa movimentação da classificação [Natan et al, 2020].

O Flood opera em um fluxo de três etapas: projeção, refinamento e escaneamento. Nesse processo, uma entrada com um predicado de filtro de intervalos sobre um ou mais atributos (unidos por ANDs) é recebida. O Flood encontra e processa os pontos dentro do hiper-retângulo formado pela interseção dos intervalos de entrada [Natan et al, 2020]. A Figura 4.10 exemplifica esse fluxo.

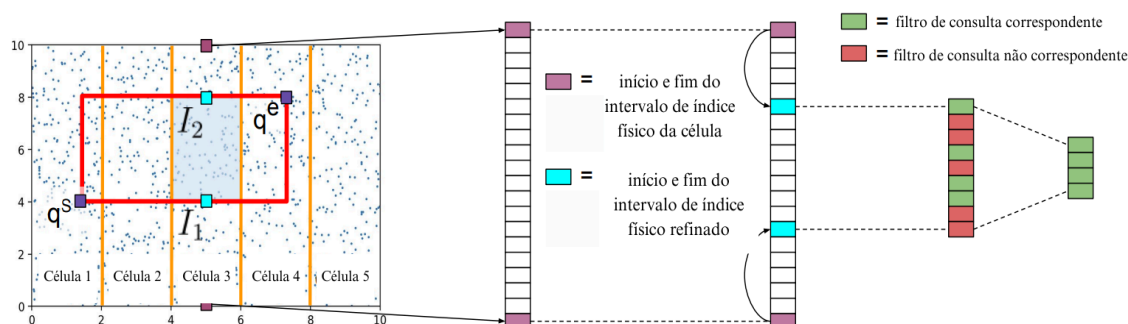


Figura 4.12. Passo-a-passo da operação do Flood

A etapa de projeção consiste em determinar quais pontos correspondem a um filtro. Para isso, o Flood inicia determinando quais as células que possuem os pontos correspondentes. Caso uma dimensão indexada não esteja na consulta, os pontos de início e fim são tomados como $-\infty$ e $+\infty$. Porém, caso a consulta inclua uma dimensão que não está indexada, o filtro é ignorado nessa etapa do processamento da consulta.

A Figura 4.12 também mostra a formação do hiper-retângulo (etapa de projeção), onde o canto inferior esquerdo é $q^s = (q^s_0, \dots, q^s_{d-1})$ e o mesmo para o canto superior direito q^e . Sendo assim, as células que estão no intervalo do hiper-retângulo são representadas pelo conjunto:

$$\{C_i \mid \text{célula}(q^s)_i \leq C_i \leq \text{célula}(q^e)_i\}$$

E o Flood usa uma tabela de células que registra o índice físico do primeiro ponto em cada célula. O conhecimento das células que estão no intervalo delimitado se traduz em um conjunto de intervalos físicos de índices para escanear [Natan et al, 2020].

Quando uma consulta traz um filtro que usa da dimensão de ordenação (a dimensão d , relatada anteriormente), o Flood utiliza do fator de ordenação dos pontos da célula e refina ainda mais os intervalos de índice físico que precisam ser varridos. Por exemplo: em uma consulta que inclua um filtro sobre a dimensão de ordenação $R.S$ onde $a \leq R.S \leq b$. O Flood encontrará os índices físicos do ponto I_1 onde $R.S \geq a$ e encontrará os índices físicos do ponto I_2 onde $R.S \leq b$. Essa etapa ajuda com que o intervalo de busca de índices físicos para $[I_1, I_2]$ seja reduzido. Como os pontos em C são armazenados de forma classificada pela dimensão d , uma busca binária em C pode ser utilizada para encontrar o $[I_1, I_2]$.

Por mais que o Flood consiga trazer soluções concretas para a utilização de learned index em dados multidimensionais, algumas limitações foram encontradas na sua operação, como: existe uma restrição para se adaptar de maneira eficiente em cargas de trabalho com consultas enviesadas, onde a frequência de consulta e a seletividade dos filtros variam no espaço dos dados. Outra dificuldade é com dimensões correlacionadas,

o Flood não consegue manter células de grade de tamanho uniforme, o que acaba piorando o desempenho e o uso de memória do framework [Diang et al. 2020]. Visando superar essas limitações, Diang et al. (2020) propõe o Tsunami, um learned index otimizado para leitura na memória para dados multidimensionais.

O Tsunami consegue um ótimo desempenho para workloads de consultas enviesadas com a utilização de uma árvore de decisão leve que é chamada de Grid Tree (Árvore de Grade), a qual particiona o espaço em regiões não sobrepostas de forma que o enviesamento das consultas seja reduzido. Outro diferencial apresentado pelo Tsunami é a capacidade de atuar com correlações, através da indexação por regiões, alcança um alto desempenho em conjunto de dados correlacionados utilizando uma grade aumentada (essa usa mapeamentos funcionais e CDFs condicionais) [Diang et al. 2020].

Como relatado anteriormente, o Tsunami é uma composição de duas estruturas de dados independentes, árvore de grade e a grade aumentada. A árvore de grade é uma árvore de decisão de partição espacial e divide o espaço de dados com d dimensões em regiões não sobrepostas. Em cada das dimensões existe uma grade aumentada e essa indexa os pontos de sua região. Porém, caso nenhuma consulta faça interseção com sua região, essa não terá uma grade aumentada. Baseado no Flood, a grade aumentada é uma generalização do seu sistema de indexação, porém elas utilizam $F: Y \rightarrow X$ e $CDF(Y|X)$ enquanto o Flood utiliza $CDF(Y)$ [Diang et al. 2020].

De forma mais detalhada, a Grid Tree é uma árvore de decisão de particionamento de espaço, a qual se assemelha com a árvore k-d. Cada nó interno da árvore de grade faz a divisão espacial com base nos valores de uma dimensão específica. Essa nova estrutura se difere da árvore k-d na granularidade dos nós, enquanto a k-d é uma árvore binária, os nós internos da árvore de grade podem se dividir em mais de um valor. No processamento de consultas, a árvore de grade é percorrida para encontrar todas as regiões que fazem interseção com os predicados da consulta. Caso haja um índice sobre os pontos nessa região, uma consulta é enviada para esse índice e os resultados são agregados e retornados. Caso não haja o índice nessa região, todos os pontos serão percorridos [Diang et al. 2020].

Com isso, Diang et al. (2020) apresentaram um índice aprendido multidimensional que utiliza de duas novas estruturas de dados (Árvore de Grade e Grade Aumentada) e é capaz de superar os demais learned index para dados multidimensionais em até 6x na taxa de transferência de consulta e 8x no espaço.

4.2.6. Indexando strings

Outra limitação que o primeiro framework de LI apresenta é a indexação de dados textuais. O suporte a workloads com dados desse tipo é custoso na inferência do modelo e no acesso dos dados que aumentam com o comprimento da chave. Outra questão é que a precisão do modelo diminui e a quantidade de acesso aos dados aumentam, isso pode ser resolvido utilizando modelos mais sofisticados como deep learning (aprendizagem profunda), porém esse benefício acaba por acarretar um tempo maior de treinamento e inferência, ocasionando um desempenho mais baixo que o esperado [Wang et al., 2020].

Wang et al. (2020) compara a indexação do XIndex com o Masstree para dados de string e o desempenho de escrita e de leitura caem significativamente quando o tamanho da chave é aumentado. Por exemplo: quando o Xindex possui uma chave de 128 bytes em uma carga de leitura, o desempenho cai em 70% quando comparado com uma chave de 8 bytes. Já em uma carga de leitura e escrita, o desempenho do XIndex cai em mais de 66%, enquanto o Masstree apresenta uma queda de apenas 34%.

Essas diferenças expressivas de desempenho em chaves de texto são causadas por alguns motivos, sendo eles: o custo de processamento do modelo aumenta significativamente com o comprimento da chave, até mesmo para um modelo simples como o linear, onde uma chave de 128 bytes pode apresentar um aumento de 23 vezes que a computação de uma chave de 8 bytes. Outro fator que influencia nessa queda de desempenho é o aumento do erro do modelo, onde aumenta de 24 para 68 quando a chave cresce de 8 para 128 bytes, essa característica faz com que os mais registros sejam acessados durante as buscas. Mesmo que modelos complexos como as redes neurais reduzem esse erro, o tempo de computação deles não apresenta melhora no desempenho. Outro fato é que a busca binária também acaba sendo custosa em buscas de strings e são afetadas pelo aumento dessas [Wang et al., 2020].

Visando superar essas dificuldades e buscando, o framework de índice aprendido escalável Sindex é proposto por Wang et al. (2020) com o objetivo de indexar eficientemente chaves de string. Em resumo esse modelo agrupa de forma gananciosa as chaves com prefixos comuns e em seguida cada grupo usa a parte única das chaves e a chave parcial para treinar o modelo e indexar os dados.

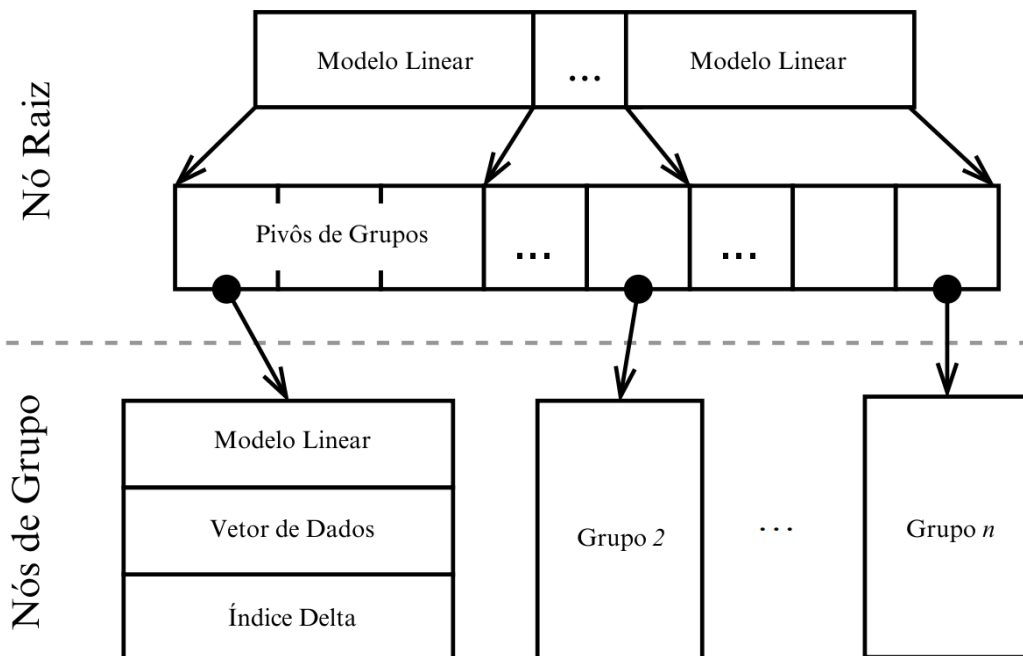


Figura 4.13. Arquitetura do Sindex. Traduzida de Wang et al. (2020)

O Xindex possui uma arquitetura separada em duas camadas, um nó raiz e vários grupos de nós na camada abaixo. Nesse layout cada grupo armazena registros de um

intervalo específico de chaves, onde ficam armazenados em um vetor ordenado e indexados por um modelo linear, ou ficam em delta buffers caso tenham sido oriundos de inserção. Já o nó raiz usa de vários modelos lineares para indexar os grupos e conta com o armazenamento das chaves pivô de cada grupo, que limitam o intervalo inferior inferior da divisão. Rodando em background, uma verificação contínua de todos os grupos é uma das atividades do SIndex, onde é analisado se o tamanho do índice delta de um determinado grupo está maior que o limite configurado pelo usuário. Caso seja encontrado valores que correspondam à essa busca, uma compactação é realizada e o índice delta é mesclado com o antigo vetor ordenado em um novo e os modelos passam por novo treinamento para se adequar à nova disposição dos dados [Wang et al., 2020].

A primeira etapa do XIndex é a extração das chaves parciais, as quais são substrings que possuem comprimento fixo extraídas das chaves originais de cada grupo. Com essa etapa é possível remover os prefixos comuns das chaves, ou seja, é possível eliminar a parte inicial das chaves que é igual entre todas, fazendo com que o comprimento da chave seja reduzido. Também é possível manter a distinção entre as substrings, fazendo com que essas possam ser suficientemente diferentes a fim de garantir a unicidade e ordenação correta em relação às demais. Para as chaves parciais pode-se utilizar a notação: para qualquer chave k em um grupo, a chave parcial é $k[pl:pl+el]$ onde pl é o comprimento do maior prefixo comum e el é o menor comprimento de prefixo entre as chaves após a remoção que garanta que todas as chaves ainda sejam iguais [Wang et al., 2020]. Tomando o grupo [“cardiologista”, “cardiopatia”, “cardiograma”] como exemplo, o pl será “cardio” (6 letras). Já o menor comprimento necessário para distinguir as chaves restantes será $el = 1$, pois as substrings “l”, “p” e “g” são suficientes para diferenciá-las. As parciais melhoram a eficiência da computação do modelo e do seu treinamento, além de reduzir a complexidade da busca binária, fazendo-a apenas nas chaves parciais devido a preservação da ordem.

A próxima etapa é o particionamento dos grupos de dados, visando que o comprimento da chave parcial e os erros dos modelos estejam dentro dos limites especificados pelo usuário e para isso uma estratégia gananciosa é utilizada. Os parâmetros são: et - limite de erro do modelo; pt - limite de comprimento da chave parcial; fs - quantidade de registros inseridos por vez; bs - quantidade de registros removidos por vez. Essa etapa tem início na interação de todos os registros do conjunto de dados e em cada grupo é determinado se novos registros devem ser inseridos ou se valores já existentes devem ser removidos, com base nos limites de erro. Ao adicionar os registros, um modelo linear é treinado no grupo para calcular o erro médio. Caso a inserção fs cause uma violação no limite do comprimento da chave parcial ou no limite do erro do modelo, são removidos os bs registros mais recentes, essa etapa é repetida até o comprimento da chave parcial e o erro do modelo estejam dentro dos limites definidos. Com esses passos o SIndex ajusta dinamicamente os grupos de dados e otimiza a eficiência do modelo de indexação, garantindo que os grupos estejam formados de maneira a manter o erro do modelo e o comprimento da chave parcial dentro dos limites aceitáveis. Com isso é alcançado uma indexação mais eficiente e precisa, enquanto melhora o desempenho do sistema como um todo [Wang et al., 2020].

No nó raiz é utilizado um modelo linear segmentado que tem objetivo dividir o espaço de chaves em segmentos menores e mais gerenciáveis, onde cada segmento é

tratado por um modelo linear separado, fazendo com que a busca e a precisão da previsão sejam melhoradas para as chaves de string. O treinamento desses modelos é feito usando o pivô de cada grupo (esse é definido usando o mesmo algoritmo ganancioso do agrupamento, assim também é determinado o intervalo de chaves de cada modelo). Já a inferência é feita através de busca binária para encontrar qual o modelo linear que é responsável pelo intervalo de chaves da solicitação. O SIndex substitui o RMI pelo modelo linear segmentado visando solucionar duas dificuldades: no RMI, dois pivôs distantes podem ser classificados pelo mesmo modelo folha, resultando em um grande intervalo de chaves nesse nó, resultando em um intervalo grande e gerando dificuldade na remoção dos prefixos comuns que reduz o benefício das chaves parciais e acarreta em uma perda de eficiência.

4.2.7. Índices de alto desempenho

Visando superar desafios de atividades concorrentes, o XIndex foi desenvolvido para ser uma estrutura de learned index concorrente, acelerando a busca com modelos aprendidos e lidando com escritas concorrentes. Uma das características de sua funcionalidade é particionar os dados em diversos grupos e cada um está associado a um modelo aprendido para busca e a um índice delta para realizar inserções. Outra vantagem do XIndex é atualizar sua estrutura de acordo com a carga de trabalho recebida, para isso são realizadas fusões e divisões de grupos parametrizadas por valores (limite de erro, limite de tamanho do delta e etc) ajustados pelo usuário [Tang et al., 2020].

Assim como foi falado anteriormente, o primeiro modelo de LI não apresentava maneiras eficientes de escrita de dados, principalmente quando trata-se de operações concorrentes. Para isso o XIndex traz a utilização do índice delta para armazenar chaves inseridas, neste os dados recém inseridos são gravados e no momento da busca o delta é consultado apenas caso a chave não seja encontrada no learned index. Para compactar o índice delta, foi implementada uma Compactação de Duas Fases e essa é realizada em segundo plano, visando não bloquear operações concorrentes [Tang et al., 2020].

A estrutura do XIndex é semelhante à apresentada pelo SIndex, onde ambos contam com uma arquitetura de duas camadas. Na camada do nó raiz os nós de cada grupo da camada inferior são indexados através de um modelo de RMI. Já nos nós folha, modelos lineares são utilizados para indexar seus dados e um índice delta armazena temporariamente as inserções do seu referente grupo [Tang et al., 2020].

Com isso, é notável que o layout do XIndex consiste de três estruturas básicas: record_t que é a representação dos dados, consiste das chaves, valores e de metadados (is_ptr indica se o valor é real ou é uma referência de memória; removed que indica se o registro foi removido logicamente; lock e version que são usados para controle de concorrência, garantindo que as operações concorrentes serão executadas de maneira exclusiva e segura. A segunda estrutura é o root_t (nó raiz), esse guarda as informações sobre os grupos de dados e utiliza de um RMI para prever qual o grupo que possui uma chave específica. Ele guarda três informações sobre os grupos: endereço dos grupos (groups), menor chave de cada grupo (pivots) e o número total de grupos (group_n). Para realizar o treinamento do RMI, o XIndex faz uso dos pivots dos grupos e seus

índices. A última estrutura é o nó de grupo (`group_t`), o qual é responsável por gerenciar os dados reais, os modelos de indexação e os índices delta encarregados de receber as inserções. Os dados (`data_array`) armazenam todos os registros do grupo de maneira contínua. Os modelos (`models`) contém pelo menos um modelo linear para indexar os registros do `data_array` e o `model_t` inclui os parâmetros do modelo linear e a menor chave do intervalo de dados que o modelo cobre. Já o índice delta (`buf`) armazena todas as inserções temporárias, o `buf_frozen` é o indicador de que o `buf` está congelado durante a compactação e o `tmp_buf` é um índice delta temporário utilizado durante a compactação. Todo esse layout auxilia o XIndex gerenciar operações de leitura e de escrita, enquanto mantém a integridade dos dados e alcança um desempenho otimizado [Tang et al., 2020].

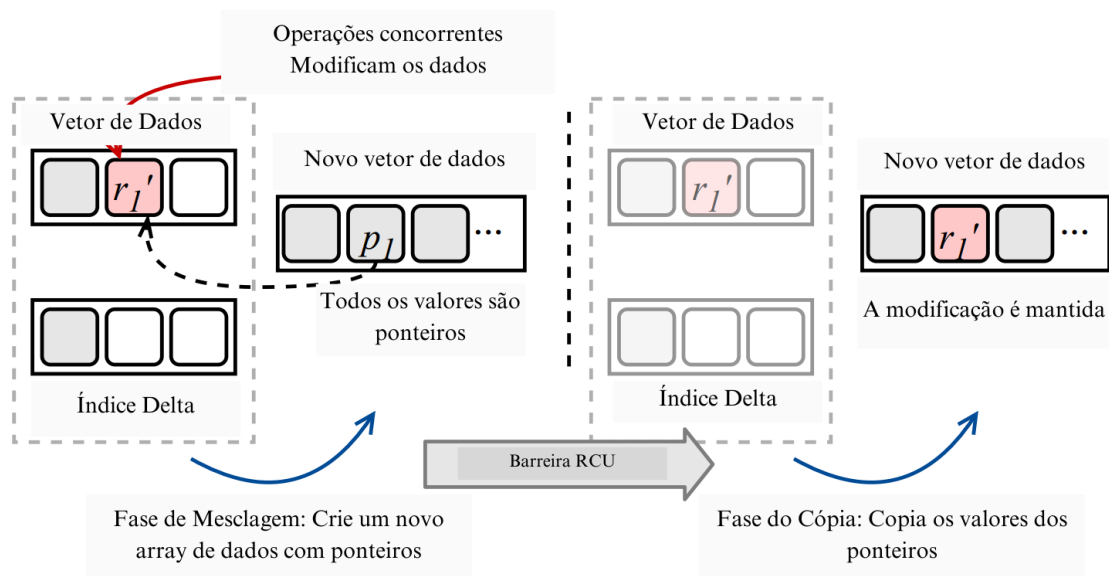


Figura 4.14. Compactação de duas fases do XIndex. Fonte: modificada de Tang et al (2020).

O XIndex também apresenta a capacidade de ajustar sua estrutura em tempo de execução, dessa forma um grupo que apresenta um alto erro de previsão recebe mais modelos lineares para uma melhora de performance. Caso um grupo esteja com muitos modelos ou o índice delta esteja sobrecarregado, o XIndex consegue dividi-lo em dois novos grupos. A mesclagem de modelos e grupos podem ser feitas desde que não apresentem queda da precisão de previsão e o re-treinamento do modelo de RMI pode ser treinado para se readequar às novas configurações dos grupos [Tang et al., 2020].

O XIndex apresenta as operações básicas de índices (`get`, `put`, `remove` e `scan`) que permitem a manipulação dos dados. No geral, todas as operações são realizadas tendo início com a identificação do grupo apropriado que contém o registro buscado, para isso o nó raiz prevê um grupo usando o RMI e em seguida ocorre uma busca binária nos pivôs dos grupos dentro de um intervalo limitado por erro. Caso o grupo encontrado pela operação `get_group` possua o ponteiro `next`, o XIndex seguirá em busca do grupo correto, pois os grupos podem ter sido recém criados. Quando o grupo é

achado, para a operação de busca no `data_array`, o `XIndex` usa o modelo linear apropriado para prever a posição do dado dentro do vetor e novamente uma busca binária é realizada com base no erro do modelo. Caso um registro seja encontrado e a operação seja `get`, o `XIndex` lê o valor, se o registro estiver logicamente removido, o `XIndex` continua a busca no `buf` do nó e talvez no `temp_buf`, um valor é retornado caso a busca encontre um resultado não vazio. Para as operações de `put` e `remove`, o `XIndex`

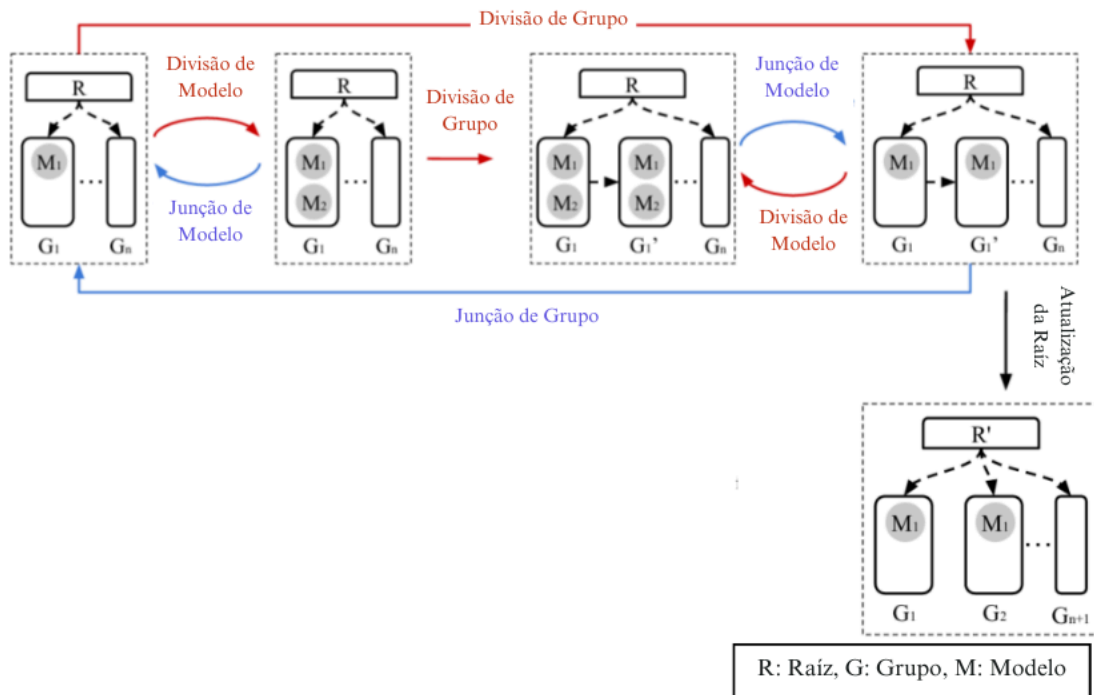


Figura 4.15. Exemplos das demonstrações de partição de grupo, nó e modelo do XIndex. Fonte: modificada de Tang et al. (2020).

tenta atualizar ou remover o valor diretamente no `data_array` e, caso a operação não possa ser feita diretamente, essa será realizada no `buf` ou no `tmp_buf`, caso a flag `frozen_buf` esteja ativa. Já a operação `scan` localiza o menor registro que é maior ou igual à chave solicitada e lê consistentemente um número definido de registros consecutivos [Tang et al., 2020].

Como dito anteriormente, o `XIndex` realiza uma compactação em duas etapas com o objetivo de suportar operações concorrentes. A primeira etapa é a mesclagem, onde o `data_array` de um grupo e o `buf` serão juntados em um novo array ordenado que possui ponteiros para os registros existentes. Para isso, o marcador (flag) `buf_frozen` é ativado para evitar novas operações de `put` no `buf` que está sendo mesclado e em seguida o `tmp_buf` é inicializado para armazenar inserções durante a compactação. Feito isso, um novo grupo é criado para armazenar o resultado da junção do `data_array` e do `buf` do grupo antigo em um novo `data_array`. No novo vetor de valores, cada valor de cada registro é uma referência para o registro correspondente do grupo antigo e a flag `is_ptr` de cada registro é atualizada para verdadeiro. Nesse novo grupo, a posição de `buf` será ocupada pelo antigo `tmp_buf` e após treinar os modelos lineares e finalizar de configurar os metadados do novo grupo, o `XIndex` substitui o grupo antigo pelo novo fazendo com

que a referência do grupo em `groups` seja atualizada. A segunda etapa é a cópia, onde cada referência do `data_array` é substituída pelo valor mais recente do registro, em paralelo, o XIndex usa da `rcu_barrier` para esperar que cada worker processe uma solicitação, fazendo com que o grupo antigo não seja acessado após a barreira. Finalizadas todas as etapas, a memória usada pelo grupo antigo pode ser liberada [Tang et al., 2020]. A Figura 4.14 demonstra a funcionalidade da compactação de duas etapas.

Outra das propriedades do XIndex que foi comentada anteriormente é a capacidade de atualizar sua estrutura em divisão ou fusão de modelos; divisão ou fusão de grupos e atualização no nó raiz. Na divisão de modelos o XIndex clona o nó responsável do grupo e ambos os nós estão referenciando os mesmos dados. Dando sequência, o número de modelos no novo nó é incrementado e os dados dos grupos são redistribuídos de maneira uniforme para cada modelo, feito isso todos são re-treinados e a referência no group do nó raiz é atualizada para apontar para o novo grupo. Já a fusão de modelos é a operação inversa da separação. A divisão de grupos consiste em criar dois grupos lógicos que compartilham os mesmos dados, porém possuem `buf` e `tmp_buf` diferentes para armazenar as novas inserções durante a divisão. Depois disso a atualização dos metadados é realizada e as referências de pivô e organização dos ponteiros entre os grupos mudam. A segunda etapa da divisão de grupos inicia com a mesclagem dos `data_array` e `buf` em um novo `data_array`, dividindo-o em dois novos grupos com base no pivot. Finalizado isso, as referências do `data_array` são substituídas pelos valores reais e os ponteiros de referência são atualizados. A fase de fusão tem um processo semelhante à divisão de grupos onde os `data_arrays` e os `bufs` são mesclados e as referências são substituídas pelos valores reais, finalizado essas etapas, o nó raiz é atualizado com as novas informações da fusão. Por fim, a etapa de atualização do nó raiz passa pelo achatamento da estrutura dos grupos para reduzir a latência de acesso dos ponteiros, o modelo RMI é re-treinado para garantir precisão nas previsões e o ponteiro global do nó raiz é substituído para garantir a consistência e integridade dos dados. Essas modificações estruturais estão representadas graficamente na Figura 4.15.

Visando tratar operações concorrentes de escrita com a mesma chave, o XIndex adiciona um bloqueio por registro ao `data_array` e a operação concorrente é realizada no `buf`. Ou seja, primeiramente as escritas são endereçadas ao vetor, porém, se a atualização nesse não estiver disponível, o escritor recorre ao índice delta, que é protegido por um bloqueio de leitura-escrita. Já para a leitura e escrita, o XIndex garante que a leitura sempre alcança um resultado consistente quando executado concorrente à uma escrita utilizando bloqueios e versionamento no `data_array` e no `buf`. Com isso, o `get` tentará ler um valor do vetor e realizar uma captura do número da versão antes da leitura. Quando o valor é obtido, é verificado se o bloqueio ainda existe e se o número da versão atual corresponde ao número da captura. Caso haja um número errado de versão, o `get` repete o procedimento até uma validação bem sucedida, fazendo com que resultados inconsistentes ou obsoletos sejam descartados. Caso a leitura do `data_array` seja impossibilitada, o `buf` será lido [Tang et al., 2020].

Sendo assim, esse é um modelo de Learned Index que consegue um bom desempenho em uma abordagem multicore trazendo a Compactação em Duas Fases junto com algumas abordagens clássicas de concorrência.

Saindo de abordagens de software e buscando uma diretriz de hardware, Zhong et al (2022) traz o modelo de learned index apresentado anteriormente (PGM) em um enfoque de execução em GPU, diferindo da tradicional que realiza suas tratativas em CPU. Uma das desvantagens amplamente citadas em diversos trabalhos sobre Learned Index e reforçada nos tópicos anteriores é a limitação do poder de processamento computacional da CPU. Porém, as Unidades de Processamento Gráfico (GPU) e as Unidades de Processamento Tensorial (TPU) se mostraram indispensáveis na área de pesquisa em Machine Learning e Deep Learning, porém, mesmo diante desse padrão, os LI acabam por rodar apenas em CPU e aderem modelos simples como regressão linear ou redes neurais de duas camadas.

Enquanto a arquitetura da CPU é composta por alguns núcleos com cache para lidar com uma vasta série de tarefas, a GPU é composta por centenas de núcleos que operam apenas com tarefas simples. Por outro lado, o Learned Index substitui a busca tradicional por uma formada por computação de modelos de aprendizado de máquina, o que se encaixa na arquitetura da GPU. Por mais que exista esse nexo entre processamento em GPU e LI, existem algumas dificuldades encontradas na junção de ambos: a latência da transferência de dados entre memória principal e recuperação dos resultados da memória da GPU é alta; a perda de desempenho ocasionada por sincronização de operações entre CPU e GPU pode ser considerável; a arquitetura do LI precisa se adaptar ao espaço de memória hierárquico da GPU (ZHONG et al., 2022).

Portanto, a proposta do PGM-GPU é ser armazenado totalmente na memória da GPU, evitando problemas de desempenho causados por sincronização entre CPU e GPU. Antes de iniciar as consultas, um lote de chaves de busca é coletado na CPU e, caso as solicitações não sejam concorrentes, as chaves são coletadas dentro de um intervalo de tempo caracterizado por “lote”. Como as consultas de entrada são realizadas na memória principal, o processo do PGM-GPU é iniciado com a transferência destas para a memória da GPU [Zhong et al., 2022].

O GPU-PGM é beneficiado com a capacidade de execuções paralelas da GPU, onde é atribuído consultas aos kernels para execuções e cada um é responsável por iniciar threads que realizam as atividades de consultas, permitindo que várias operações sejam executadas simultaneamente. Quando *threads* do mesmo kernel operam de forma coordenada e sincronizada, eles acessam o PGM armazenado na memória global da GPU de maneira ordenada e executam as mesmas instruções de consulta. Ao finalizar todas as operações, os *threads* escrevem os resultados nas localizações de memória designadas simultaneamente. Ao finalizar as operações de busca, as posições exatas dos resultados das consultas são transferidas da GPU para a memória do sistema e a CPU utiliza as posições recebidas para acessar diretamente os registros específicos na base de dados. Ou seja, toda a tarefa intensiva de busca é realizada pela GPU e a CPU finaliza a operação acessando os dados específicos [Zhong et al., 2022]. Na Figura 4.16 há um esquema visual do funcionamento descrito acima.

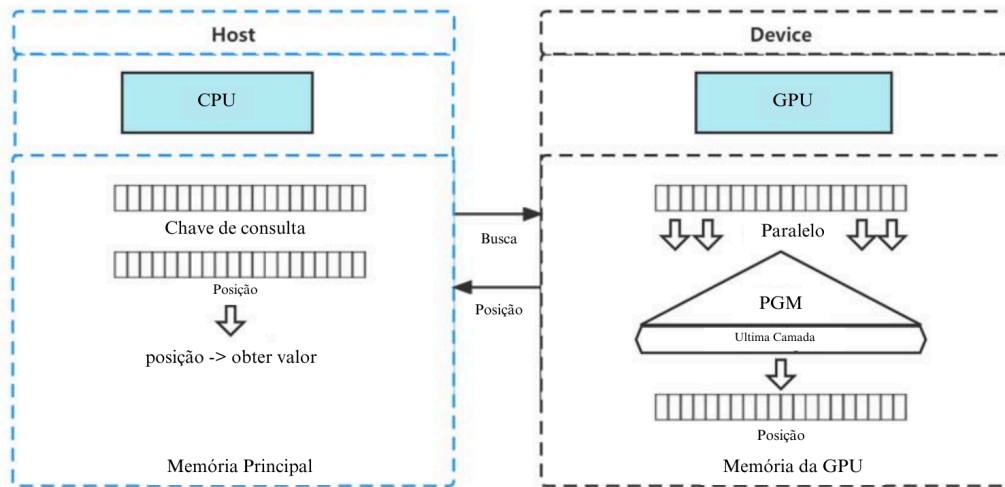


Figura 4.16. Funcionamento entre CPU (esquerda) e GPU (direita). Fonte: Traduzida de Zhong et al., 2022

Para montar o índice, a primeira etapa é a criação de um modelo linear por partes, onde o erro não exceda um limite pré-estabelecido. Com isso, o algoritmo garante que o modelo linear por partes seja ótimo. A principal ideia do algoritmo é reduzir o problema de construir o fecho convexo do conjunto de pontos (conjunto $\{key_i, pos(key_i)\}$ crescido para $i = 0, \dots, n-1$). Enquanto o fecho convexo puder ser fechado em um retângulo de altura não maior que 2, o índice i é incrementado e o conjunto estendido. Quando o retângulo que envolve o casco é maior que 2, a construção é interrompida e um segmento é determinado abraçando a linha que divide o retângulo em duas metades de tamanho iguais. O algoritmo é iterado a partir dos pontos de entrada restantes. Ao finalizar essa etapa, o próximo passo é construir a estrutura de índice hierárquica, o qual inicia com o modelo construído sobre todo o vetor de entrada e a primeira chave do vetor coberta por cada segmento é extraída. Agora um novo modelo linear é construído por partes sobre esse conjunto reduzido de chaves. Todo esse processo continua de forma recursiva até que o modelo final consista em apenas um segmento [Zhong et al., 2022]. A estrutura do índice PGM pode ser vista na Figura 4.17.

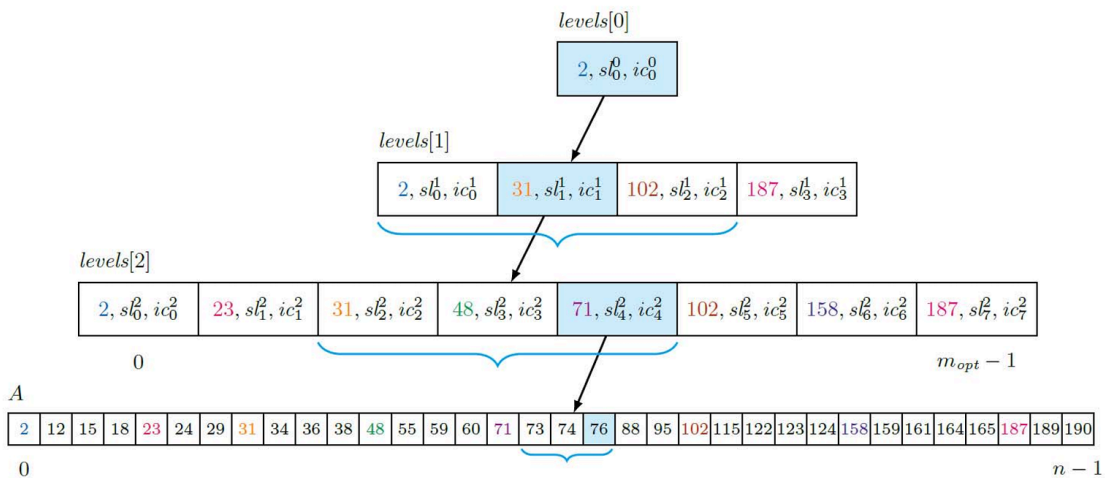


Figura 4.17. Estrutura do índice PGM. Fonte: [Zhong et al., 2022]

Já que o algoritmo de construção do PGM-Index necessita percorrer todo o conjunto de dados e não é possível realizar otimização paralela, o índice é construído na CPU e em seguida a estrutura de índice hierárquica é organizada como um array compacto e o deslocamento por nível é registrado. Finalizado esse processo, o índice é transferido para a GPU e o GPU-PGM é obtido.

Na etapa de consulta, o lote de chaves de entrada é enviado para a GPU e distribuído para um número designado de threads. Durante a operação é necessário verificar o número de threads no multiprocessador de streaming na GPU quando as threads são distribuídas, uma vez que a execução sincronizada é a maneira mais eficiente de execução na GPU, ou seja, threads do mesmo bloco executam as mesmas instruções de consulta e acessam a memória global de maneira síncrona e alinhada.

Quando cada thread recebe a chave de consulta correspondente, a busca funciona estimando a posição da chave e em cada nível é usado o segmento referente ao nó visitado para fazer essa estimativa entre as chaves do nível inferior. A posição real é encontrada por uma busca binária de tamanho dois que tem como centro a posição estimada. Portanto, como cada chave do próximo nível é a primeira chave coberta por um segmento, o próximo segmento que será consultado é encontrado e esse processo é repetido até que o último nível seja alcançado. Tendo os resultados, esses são escritos na localização especificada e são transferidos de volta para a memória principal.

4.3. Tendências, desafios de pesquisa e aplicações para a saúde

Desde sua concepção, o LI apresenta suas limitações como principais lacunas para investigação, desenvolvimento e avanço de suas soluções. Apesar de ser tema assíduo nas principais conferências sobre gestão de dados, big data e banco de dados, são inúmeras as oportunidades de evolução do conceito e suas aplicações. Neste minicurso, apresentaremos estes desafios e oportunidades em duas categorias: i) a de pesquisa básica, que inclui esforços para avanços teóricos, de assimilação de estrutura de dados inovadoras e escalabilidade; e ii) a de aplicação, que explora as otimizações já demonstradas empiricamente das diversas estruturas de LI já propostas nas outras áreas do conhecimento. A seguir, focaremos nossa apresentação em pesquisa aplicada, principalmente sob a ótica da Computação Aplicada à Saúde. É relevante mencionar, contudo, que diversos problemas de pesquisa relacionados com a gestão de dados desta comunidade ainda demandam avanços teóricos e um acúmulo mais robusto de evidências dos trabalhos de LI presentes na literatura.

4.3.1. Gestão de Big Data

Apesar dos vários artigos acrescentando evidências empíricas da superioridade dos modelos habilitados por ML para suportar grandes volumes de dados. Ainda existe uma lacuna evidente para a concepção de modelos de LI capazes de suportar a gestão de *big data*. Seja pela maturidade das estruturas de índice ou pela disponibilidade para o uso pelo público amplo, os modelos de LI escaláveis através de arquiteturas *multicore*, GPU

ou TPU ainda não são capazes de suprir a demanda da comunidade de gestão de big data. Por outro lado, estratégias de construção, uso e atualização de índices quando estes já não cabem em memória estão restritas a grandes provedores de computação em nuvem.

4.3.2. Dados multimodais

Do ponto de vista de pesquisa básica, especificamente no que diz respeito ao domínio de dados, ainda há uma carência muito grande de modelos existentes de LI capazes de lidar com o volume e a variedade das coleções de dados dos sistemas de informação em saúde. Por exemplo, a variedade pode ser definida através de uma determinada tipologia de dados, que estabelece escalas de mensuração e as operações matemáticas adequadas para cada uma delas. No contexto do Big Data, contudo, esta definição se mostra limitada. A variedade das variáveis numa coleção de dados pode ser avaliada também pelas modalidades presentes ou pela sua orientação de dependência (ex.: séries temporais).

Os dados que estão relacionadas à saúde do paciente podem apresentar diversas modalidades, a exemplificação na Figura \diamond mostra como um paciente pode estar associado aos diferentes tipos de dados, como os dados de imagens, série temporal ou sinal, dados genéticos, notas clínicas e informações laboratoriais e demográfica [Jensen et. al., 2012].

Ao agregar informações de diversas fontes e modalidades temos as bases de dados multimodais. Uma abordagem integrada das informações oferece uma visão mais abrangente dos pacientes, permitindo uma compreensão mais profunda dos casos em análise, bem como, melhores prognósticos e intervenções mais personalizadas [Acosta et. al., 2022]. Por exemplo, considere um sistema de diagnóstico, que recebe dados de registros médicos enriquecido de dados de imagem. Essa diversidade de informações permite que o sistema ofereça diagnósticos mais precisos com *insights* mais profundos.

A literatura apresenta trabalhos que utilizam conjunto de dados multimodais com objetivo de identificar e diagnosticar doenças. Nie et. al., 2017 propôs o uso de estruturas de deep learning para extrair características de imagens cerebrais multimodais (ou seja, ressonância magnética T1, fMRI e DTI) de pacientes com tumor cerebral. Esses dados foram utilizados juntamente com as principais características clínicas dos pacientes com o objetivo de prever se o paciente terá um tempo de sobrevivência global longo ou curto, os experimentos apresentaram uma precisão de até 89,9%.

Pillai et. al., 2017, combinam características extraídas de imagens cerebrais com dados do líquido cefalorraquidiano para distinguir pacientes com doença de Alzheimer de indivíduos saudáveis. El-Sappagh et. al., 2020 buscam detectar a progressão da doença Alzheimer utilizando dados de diferentes modalidades, para isso, é proposta uma arquitetura de deep learning multimodal multitarefa.

A utilização de dados multimodais pode aprimorar a detecção precoce da mortalidade e elevar os índices de sobrevivência em pacientes com sepse na unidade de terapia intensiva (UTI), como evidenciado por Shin et al. (2021). Esse estudo integra informações demográficas, medidas fisiológicas e anotações clínicas para aplicar

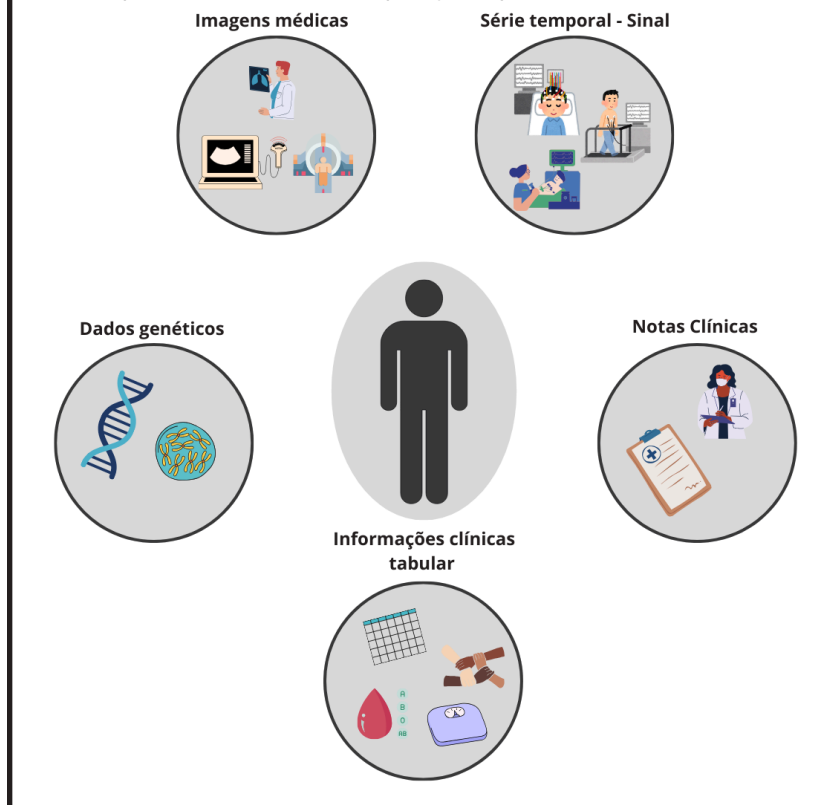


Figura 4.18. Exemplificação de dados multimodais em saúde

modelos de aprendizado de máquina visando prever o risco de mortalidade hospitalar nesses pacientes. Por fim, o estudo sugere que a integração de características clínicas estruturadas e não estruturadas pode ser efetivamente aplicada para auxiliar os médicos na identificação do risco de mortalidade em pacientes com sepse na UTI.

A fim de coletar e converter dados em informações e apoiar os profissionais na gestão de saúde, seja na esfera pública ou privada, existem os Sistemas de Informação em Saúde (SIS), os quais alimentam as bases de dados de saúde. No Brasil, os principais sistemas são Sistema do Cadastro Nacional de Estabelecimentos de Saúde (SCNES), Sistema de Informações Ambulatoriais (SAI/SUS), Sistema de Informações Hospitalares (SIH/SUS), Sistema Nacional de Agravos de Notificação (SINAN), Sistema de Informações sobre Nascidos Vivos (SINASC) e Sistema de Informação sobre Mortalidade (SIM) [Saraiva et al., 2021]. Esses dados se constituem como dados administrativos governamentais, recolhidas pelo serviço público como parte de seus processos contínuos e cotidianos.

4.3.2. Índices compostos ou multivariados

Índices multidimensionais consiste em estruturas de LI capazes de aprender uma mistura de CDFs ou um particionamento uniforme de um conjunto de dados com mais de uma dimensão, permitindo mimetizar a indexação multicoluna implementada em SGBDs tradicionais [Nathan et al., 2023]. Apesar dos diversos avanços nesta área, várias lacunas ainda devem ser exploradas nos próximos anos, principalmente para endereçar a ordenação monotônica que é desafiada quando se usa modelos de codificação das dimensões num único espaço dimensional, a escolha adequada de modelos de ML, concorrência e *benchmarking* [Al-Mamun et al., 2024]

4.3.1. Aplicações e oportunidades

Diversas iniciativas de gestão de dados em saúde incorporam modelos de indexação para garantir eficiência e escalabilidade das suas aplicações, como se pode observar no Modelo de Administração de Dados (MAD) do DATASUS e do Plano Diretor de Tecnologia de Informação (PDTI) do Ministério da Saúde. Várias destas soluções são proprietárias e oferecem pouca transparência sobre seus resultados ou confiabilidade de suas implementações. Neste sentido, o domínio de novas técnicas de indexação, especialmente pela comunidade de Computação Aplicada à Saúde, tem o potencial de promover uma adoção por soluções mais adequadas para as distribuições de dados existentes, ao passo que favorece a soberania de dados do país.

Aplicações de estruturas de índices habilitados por aprendizado de máquina na catalogação e curadoria de dados pode permitir uma maior maturidade na adoção dos princípios FAIR, perseguidos por centros de dados e pesquisa. Métodos de indexação estão diretamente ligados à melhoria de propriedades dos dados localizáveis (*Findable*), acessíveis (*Accessible*), interoperáveis (*Interoperable*) e reutilizáveis (*Reusable*). Por exemplo, serviços que suportam o modelos FAIR, como o SolR e o Dataverse, utilizam diversas técnicas avançadas de indexação para catalogar e encontrar *datasets*.

A integração de grandes volumes de dados é possível a partir da aplicação de técnicas poderosas de indexação. Por exemplo, em seu trabalho, Barbosa et al. (2020) argumenta que a escalabilidade e acurácia do CIDACS-RL — ferramenta de *record linkage* nacional responsável pela criação da Coorte de 100 Milhões de brasileiros — foi possível graças ao uso do Apache Lucene ou do Elasticsearch, duas soluções de indexação adequadas para sistemas de dados orientados a documentos (DODBMS). Contudo, novas técnicas de indexação aplicadas a ferramentas como estas podem garantir melhores resultados de desempenho, acurácia e mitigação de vieses.

Referências

- Acosta, J. N., Falcone, G. J., Rajpurkar, P., and Topol, E. J. (2022) “Multimodal biomedical AI” *Nature Medicine*, 28(9), 1773-1784.
- Al-Mamun, Abdullah, et al. "A Survey of Learned Indexes for the Multi-dimensional Space." *arXiv preprint arXiv:2403.06456* (2024).
- Amato, Giuseppe, and Pasquale Savino. "Approximate similarity search in metric spaces using inverted files." *3rd International ICST Conference on Scalable Information Systems*. 2010.
- Armenatzoglou, Nikos, et al. "Amazon Redshift re-invented." *Proceedings of the 2022 International Conference on Management of Data*. 2022.
- Banker, Kyle, et al. *MongoDB in action: covers MongoDB version 3.0*. Simon and Schuster, 2016.

- Barbosa, G. C., Ali, M. S., Araujo, B., Reis, S., Sena, S., Ichihara, M. Y., ... & Barreto, M. L. (2020). CIDACS-RL: a novel indexing search and scoring-based record linkage system for huge datasets with high accuracy and scalability. *BMC medical informatics and decision making*, 20, 1-13.
- Barreto, M. L., Ichihara, M. Y., Almeida, B. D. A., Barreto, M. E., Cabral, L., Fiaccone, R. L., ... and Smeeth, L. (2019). "The centre for data and knowledge integration for health (CIDACS): linking health and social data in Brazil" *International journal of population data science*, 4(2).
- Campero Durand, Gabriel, et al. "Automated vertical partitioning with deep reinforcement learning." *New Trends in Databases and Information Systems: ADBIS 2019 Short Papers, Workshops BBIGAP, QAUCA, SemBDM, SIMPDA, M2P, MADEISD, and Doctoral Consortium*, Bled, Slovenia, September 8–11, 2019, Proceedings 23. Springer International Publishing, 2019.
- Dama, International. *DAMA-DMBOK: data management body of knowledge*. Technics Publications, LLC, 2017.
- Dasgupta, Sanjoy, and Yoav Freund. "Random projection trees and low dimensional manifolds." *Proceedings of the fortieth annual ACM symposium on Theory of computing*. 2008.
- Datar, Mayur, et al. "Locality-sensitive hashing scheme based on p-stable distributions." *Proceedings of the twentieth annual Symposium on Computational Geometry*. 2004.
- DATASUS. (2024) "Orientações para Construção de Modelo de Dados com Qualidade". Coordenação de Gestão de Banco de Dados. Disponível em: <https://ced.saude.gov.br/arquivo/download?path=%2Fuploads%2F&file=daaed-cartilha-ad-20180815140452.pdf>. Acesso em: 08 de março de 2024.
- Ding, J., Marcus, R., Kipf, A., Nathan, V., Nrusimha, A., Vaidya, K., van Renen, A. and Kraska, T. (2022). "Sagedb: An instance-optimized data analytics system" *Proceedings of the VLDB Endowment*, v. 15, n. 13, p. 4062-4078.
- Ding, J., Nathan, V., Alizadeh, M., and Kraska, T. (2020). "Tsunami: a learned multi-dimensional index for correlated data and skewed workloads" *Proceedings of the VLDB Endowment*, 14(2), 74-86.
- Ding, Jialin, et al. "ALEX: an updatable adaptive learned index." *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*. 2020.
- Elmasri, R., & Navathe, S. B. (2016). "Fundamentals of Database Systems" 7th ed.

- El-Sappagh, S., Abuhmed, T., Islam, S. R., and Kwak, K. S. (2020). “Multimodal multitask deep learning model for Alzheimer’s disease progression detection based on time series data”. *Neurocomputing*, 412, 197-215.
- Eryurek, Evren, et al. *Data governance: The definitive guide*. " O'Reilly Media, Inc.", 2021.
- Fanggohans, Dean. *Defio: Instance-Optimized Fusion of AWS Database Services*. Diss. Massachusetts Institute of Technology, 2023.
- Ferragina, Paolo, and Giorgio Vinciguerra. "The PGM-index: a fully-dynamic compressed learned index with provable worst-case bounds." *Proceedings of the VLDB Endowment* 13.8 (2020): 1162-1175.
- Ferreira, J. E. D. S. M., Oliveira, L. R. D., Marques, W. S., Lima, T. S. D., Barbosa, E. D. S., Castro, R. R. D., and Guimarães, J. M. X. (2020). “Sistemas de Informação em Saúde no apoio à gestão da Atenção Primária à Saúde: revisão integrativa” *RECIIS - Revista Eletrônica de Comunicação, Informação e Inovação em Saúde*, Rio de Janeiro, v. 14, n. 4, p. 970-982.
- Fritchey, Grant. "Automated Tuning in Azure and SQL Server." *SQL Server 2022 Query Performance Tuning: Troubleshoot and Optimize Query Performance*. Berkeley, CA: Apress, 2022. 657-675.
- Gani, A. et al. A survey on indexing techniques for big data: Taxonomy and performance evaluation. *Knowledge and Information Systems*, v. 46, 04 2015.
- Graefe, G. (2010). A survey of B-tree locking techniques. *ACM Transactions on Database Systems*, 35(3), Article 16. DOI: 10.1145/1806907.1806908.
- Gray, J. (2007) “eScience -- A Transformed Scientific Method” eScience Talk at NRC-CSTB meeting, 11 January 2007. Apresentação do Power Point. Disponível em: https://jimgray.azurewebsites.net/talks/NRC-CSTB_eScience.ppt. Acesso em: 21 de dez. 2023.
- Hadian, A., Ghaffari, B., Wang, T., and Heinis, T. (2023). “COAX: Correlation-Aware Indexing” In 2023 IEEE 39th International Conference on Data Engineering Workshops (ICDEW) (pp. 55-59). IEEE.
- Harwood, Ben, and Tom Drummond. "Fanng: Fast approximate nearest neighbor graphs." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
- Hussam Abu-Libdeh, Deniz Altinbükten, Alex Beutel, Ed H. Chi, Lyric Doshi, Tim Kraska, Xiaozhou, Li, Andy Ly, and Christopher Olston. 2020. *Learned Indexes for a Google-scale Disk-based Database*. arxiv: 2012.12501
- Hey, T., Tansley, S., and Tolle, K. M. (2009). “Jim Gray on eScience: a transformed scientific method” *The Fourth Paradigm: Data-Intensive Scientific*

- Discovery. Redmond, Washington: Microsoft Research. 284pp. ISBN978-0982544204.
- Jensen, P. B., Jensen, L. J., and Brunak, S. (2012). Mining electronic health records: towards better research applications and clinical care. *Nature Reviews Genetics*, 13(6), 395-405.
- Junior, E. P. P., Normando, P., Flores-Ortiz, R., Afzal, M. U., Jamil, M. A., Bertolin, S. F., ... & Khalid, S. (2023). Integrating real-world data from Brazil and Pakistan into the OMOP common data model and standardized health analytics framework to characterize COVID-19 in the Global South. *Journal of the American Medical Informatics Association*, 30(4), 643-655.
- Kaufman, Leonard, and Peter J. Rousseeuw. *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons, 2009.
- Kaur, T. (2018). *Learned Index Structures: Practical Implementations and Future Directions*.
- Kleppmann, M. (2017). *Designing Data-Intensive Applications*. Beijing: O'Reilly. ISBN: 978-1-4493-7332-0.
- Kraska, T., Beutel, A., Chi, E. H., Dean, J., & Polyzotis, N. (2018, May). The case for learned index structures. In *Proceedings of the 2018 international conference on management of data* (pp. 489-504).
- Kraska, Tim. "Towards instance-optimized data systems." *Proceedings of the VLDB Endowment* 14.12 (2021).
- Kipf, Andreas, et al. "RadixSpline: a single-pass learned index." *Proceedings of the third international workshop on exploiting artificial intelligence techniques for data management*. 2020.
- Kukreja, Sanjay, et al. "Vector Databases and Vector Embeddings-Review." *2023 International Workshop on Artificial Intelligence and Image Processing (IWAIP)*. IEEE, 2023.
- Li, Guoliang, Xuanhe Zhou, and Lei Cao. "Machine learning for databases." *Proceedings of the First International Conference on AI-ML Systems*. 2021.
- Li, G., Zhou, X., & Cao, L. (2021). *AI Meets Database: AI4DB and DB4AI*. *Proceedings of the 2021 International Conference on Management of Data*. doi:10.1145/3448016.3457542
- Ma, Chaohong, et al. "FILM: a fully learned index for larger-than-memory databases." *Proceedings of the VLDB Endowment* 16.3 (2022): 561-573.
- Ma, Lin, et al. "MB2: decomposed behavior modeling for self-driving database management systems." *Proceedings of the 2021 International Conference on Management of Data*. 2021.

- Mao, Hongzi, et al. "Resource management with deep reinforcement learning." Proceedings of the 15th ACM workshop on hot topics in networks. 2016.
- Marcus, Ryan, et al. "Neo: A learned query optimizer." Proceedings of the VLDB Endowment 12.11 (2019).
- Magalhães, M. N. "Probabilidade e variáveis aleatórias" São Paulo: IME-USP, 2004.
- Maltry, M., and Dittrich, J. (2021). "A critical analysis of recursive model indexes" arXiv preprint arXiv:2106.16166.
- Mikhaylov, Artem, et al. "Learned Query Optimizers: Evaluation and Improvement." IEEE Access 10 (2022): 75205-75218.
- Ministério da Saúde, PORTARIA N° 4.279, DE 30 DE DEZEMBRO DE 2010. Estabelece diretrizes para a organização da Rede de Atenção à Saúde no âmbito do Sistema Único de Saúde (SUS). Gabinete do Ministro. Brasília, DF.
- Ministério da Saúde. PORTARIA GM/MS N° 1.102, DE 13 DE MAIO DE 2022. Dispõe sobre a inclusão da covid-19 na Lista Nacional de Notificação Compulsória de doenças, agravos e eventos de saúde pública e Altera o Anexo 1 do Anexo V à Portaria de Consolidação GM/MS n° 4, de 28 de setembro de 2017. Diário Oficial da União, Brasília, DF, 16 maio 2022.
- Nathan, Vikram et al. Learning Multi-Dimensional Indexes. In: Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data. New York, NY, USA: Association for Computing Machinery, 2020. p. 985–1000.
- Nie, D., Zhang, H., Adeli, E., Liu, L., and Shen, D. (2016). "3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients" Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19 (pp. 212-220). Springer International Publishing.
- Liu, Jiesong, et al. "G-Learned Index: Enabling Efficient Learned Index on GPU." IEEE Transactions on Parallel and Distributed Systems (2024)
- LAZARIDIS, Michalis et al. Multimedia search and retrieval using multimodal annotation propagation and indexing techniques. Signal Processing: Image Communication, v. 28, n. 4, p. 351-367, 2013.
- Paixao, E. S., Cardim, L. L., Falcao, I. R., Ortelan, N., Silva, N. D. J., Rocha, A. D. S., ... & Teixeira, M. G. (2021). "Cohort profile: Centro de Integração de Dados E Conhecimentos para Saúde (CIDACS) birth cohort" International journal of epidemiology, 50(1), 37-38.

- Pillai, P. S., Leong, T. Y., and Alzheimer's Disease Neuroimaging Initiative. (2015). "Fusing heterogeneous data for Alzheimer's disease classification" MEDINFO 2015: eHealth-enabled Health (pp. 731-735). IOS Press.
- Reis, Joe, and Matt Housley. Fundamentals of Data Engineering. " O'Reilly Media, Inc.", 2022.
- Ross, S. (2013), Simulation (Fifth Edition) - Chapter 2 - Elements of Probability, Pages 5-38.
- Ross, S. (2014), Introduction to probability models. Academic press.
- Sandu Popa, Iulian, et al. "Indexing in-network trajectory flows." The VLDB Journal 20 (2011): 643-669.
- Saraiva, L. I. M., Ramos, F. A. S., dos Santos, G. F., and Vetorazo, J. V. P. (2021). "Sistemas de informação em saúde, o instrumento de apoio à gestão do SUS: aplicabilidade e desafios" Revista Eletrônica Acervo Enfermagem, 9, e6418-e6418.
- Shin, J., Li, Y., and Luo, Y. (2021). "Early prediction of mortality in critical care setting in sepsis patients using structured features and unstructured clinical notes" In 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 2885-2890). IEEE.
- Silberschatz, Abraham, Henry F. Korth, and S. Sudarshan (2020). "Database System Concepts" 7th ed.
- TANG, Chuzhe; WANG, Youyun; DONG, Zhiyuan; HU, Gansen; WANG, Zhaoguo; WANG, Minjie; CHEN, Haibo. XIndex: a scalable learned index for multicore data storage. In: ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, 25., 2020, San Diego, California. Proceedings [...]. New York, NY, USA: Association for Computing Machinery, 2020. p. 308–320. Disponível em: <https://doi.org/10.1145/3332466.3374547>
- Wang, Y., Tang, C., Wang, Z., & Chen, H. (2020). "SIndex: a scalable learned index for string keys" In Proceedings of the 11th ACM SIGOPS Asia-Pacific Workshop on Systems (pp. 17-24).
- Wei, Ling-Yin, et al. "Indexing spatial data in cloud data management." Pervasive and Mobile Computing 15 (2014): 48-61.
- Witten, Ian H., and Eibe Frank. "Data mining: practical machine learning tools and techniques with Java implementations." Acm Sigmod Record 31.1 (2002): 76-77.
- Wong, D. W. (2001). Location-Specific Cumulative Distribution Function (LSCDF): An Alternative to Spatial Correlation Analysis. Geographical Analysis, 33(1), 76-93.

- Wu, K., Shoshani, A., & Stockinger, K. (2010). Analyses of Multi-Level and Multi-Component Compressed Bitmap Indexes. *ACM Transactions on Database Systems*, 35(1), 1-52. DOI: 10.1145/1670243.1670245.
- Zhao, Xinyang, Xuanhe Zhou, and Guoliang Li. "Automatic database knob tuning: a survey." *IEEE Transactions on Knowledge and Data Engineering* (2023).
- Zhou, Xuanhe, et al. "Dbmind: A self-driving platform in opengauss." *Proceedings of the VLDB Endowment* 14.12 (2021): 2743-2746.
- Zhong, X., Zhang, Y., Chen, Y., Li, C., and Xing, C. (2022). "Learned index on GPU" In *2022 IEEE 38th International Conference on Data Engineering Workshops (ICDEW)* (pp. 117-122). IEEE.

Capítulo

5

Harmonização Global de Dados de Saúde: O Papel dos Vocabulários Padronizados OHDSI

Maria Tereza Fernandes Abrahão (HIAE, OHDSI Latam), Pablo Jorge Madril (OHDSI Latam), Mateus de Lima Freitas (HIAE, OHDSI Latam)

Abstract

The Observational Health Data Sciences and Informatics (OHDSI) has emerged as a leading force in large-scale health data analysis, and its use of the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) plays a central role in this success. The standardization provided by the OMOP CDM is essential to enable large-scale observational analysis, allowing data from different sources to be harmonized and analyzed in a consistent and reliable manner. A vital component of this process is the integration of the OMOP CDM with a comprehensive and effective vocabulary system. This vocabulary plays a crucial role in structuring and organizing data, providing a solid foundation for integrating diverse data sources. By ensuring semantic consistency and interoperability between data, the ontology facilitates comprehensive and reliable analysis, even when dealing with complex and heterogeneous data sets.

Resumo

A Observational Health Data Sciences and Informatics (OHDSI) emergiu como uma força líder na análise de dados de saúde em larga escala, e sua utilização do Common Data Model (CDM) do Observational Medical Outcomes Partnership (OMOP) desempenha um papel central nesse sucesso. A padronização proporcionada pelo CDM OMOP é fundamental para viabilizar a análise observacional em grande escala, permitindo que dados de diferentes fontes sejam harmonizados e analisados de maneira consistente e confiável. Um componente vital desse processo é a integração do CDM OMOP com um sistema de vocabulário abrangente e eficaz. Esse vocabulário desempenha um papel crucial na estruturação e organização dos dados, fornecendo uma base sólida para a integração de diversas fontes de dados. Ao garantir a consistência semântica e a interoperabilidade entre os dados, o vocabulário facilita uma análise abrangente e confiável, mesmo quando lidando com conjuntos de dados complexos e heterogêneos.

5.1 OHDSI – Introdução

A OHDSI (Observational Health Data Sciences and Informatics)¹, é uma colaboração global que redefiniu a área de pesquisa observacional em dados de saúde trazendo a possibilidade de realizar análises sistemáticas em grandes massas de dados provindas de diversas fontes. Fundada em 2014, sua missão é capacitar uma comunidade de ciência aberta para gerar evidências médicas, promovendo decisões de saúde baseadas em dados sólidos. A iniciativa desenvolveu um Modelo Comum de Dados (Common Data Model CDM) padronizado, chamado CDM OMOP (Observational Medical Outcomes Partnership), que facilita a integração de dados de saúde de diferentes fontes. Além disso, a OHDSI fornece vocabulários padronizados e ferramentas de software livre que possibilitam a geração sistemática de evidências em larga escala para análise avançada de dados de saúde.

O Vocabulário Padronizado OHDSI é um dos componentes fundamentais desta iniciativa. A OHDSI desenvolve e mantém um vocabulário de referência centralizada em grande escala para harmonização internacional de dados de saúde. Esse vocabulário inclui uma ampla gama de termos e conceitos relacionados à saúde, abrangendo diferentes áreas, como doenças, procedimentos médicos, medicamentos, resultados clínicos e características demográficas dos pacientes. Ao padronizar os termos usados para descrever esses elementos, o Vocabulário Padronizado OHDSI permite que pesquisadores e profissionais de saúde compartilhem e comparem os resultados de análises estatísticas realizadas nos dados de diferentes fontes e sistemas de saúde de forma consistente e interoperável.

A padronização dos vocabulários de dados de saúde facilita a realização de estudos multicêntricos e meta-análises, além de promover uma melhor compreensão e interpretação dos resultados. Isso também é crucial para o desenvolvimento e aprimoramento de métodos analíticos em epidemiologia e pesquisa clínica.

O capítulo está estruturado como se segue. Primeiro, a seção 5.1 oferece uma breve fundamentação sobre a iniciativa OHDSI, desafios e soluções. A seção 5.2 apresenta o modelo comum de dados (CDM) e a seção 5.3 conceitos gerais dos vocabulários. A seção 5.4 os principais vocabulários, a seção 5.5 apresenta a estrutura dos vocabulários OHDSI. A seção 5.6, o Athena (consulta, seleção e download de vocabulários), a seção 5.7 a ferramenta Usagi (mapeamento dos vocabulários). A seção 5.8, o Atlas, a seção 5.9, apresenta as considerações finais e conclusões e a seção 5.10 as referências consultadas.

Na elaboração deste capítulo, seguiu-se o livro da OHDSI, The book of OHDSI², de domínio público, sob a licença Creative Commons Zero v1.0 Universal, (16/04/2020). O livro é um documento vivo, mantido pela comunidade por meio de ferramentas de desenvolvimento de código aberto e evolui continuamente. A versão

¹ <http://www.ohdsi.org>

² The book OHDSI <http://book.ohdsi.org>

online, disponível gratuitamente, sempre representa a versão mais recente. O texto do livro foi traduzido e complementado na escrita deste capítulo. As figuras e tabelas que ilustram os estudos, foram adaptadas do livro e de apresentações dos tutoriais da OHDSI. Os exemplos foram elaborados a partir dos conhecimentos adquiridos pelos autores na participação em eventos OHDSI.

5.1.1 Uma breve história da OHDSI, desafios e soluções

A iniciativa OHDSI (Observational Health Data Sciences and Informatics), fundada em 2014, tem sido fundamental na transformação da pesquisa observacional em dados de saúde. Seu modelo comum de dados (CDM OMOP) e suas ferramentas de software livre abriram caminho para análises sistemáticas em grandes volumes de dados de saúde, provenientes de diversas fontes. Isso não apenas facilitou a manipulação e análise desses dados, mas também promoveu a interoperabilidade e reprodutibilidade na geração de evidências médicas.

Ao estabelecer uma comunidade global de pesquisadores e bancos de dados observacionais de saúde, com um centro de coordenação na Universidade de Columbia, a OHDSI facilitou a colaboração em escala internacional. Com centenas de pesquisadores em mais de 30 países e registros de saúde de cerca de 600 milhões de pacientes únicos em todo o mundo, a OHDSI está capacitando a comunidade a gerar evidências que promovam melhores decisões e cuidados de saúde. Essa abordagem colaborativa e sistemática está mudando a forma como a pesquisa médica é conduzida, com o objetivo final de melhorar a saúde globalmente [Abrahão 2019].

A adesão dos pesquisadores às práticas de ciência aberta, como as promovidas pela iniciativa OHDSI, desencadeia uma série de benefícios significativos para a pesquisa e a sociedade em geral. Ao abraçar os valores de transparência, colaboração e acessibilidade, a ciência aberta amplia o alcance e o impacto do conhecimento científico de várias maneiras:

1. **Eficiência na pesquisa:** Ao compartilhar dados, materiais e resultados de pesquisa de forma aberta, os pesquisadores podem evitar a duplicação de esforços e acelerar a progressão do conhecimento, aumentando a eficiência da pesquisa.
2. **Confiabilidade dos resultados:** A transparência e a replicabilidade são fundamentais para a confiabilidade dos resultados científicos. Ao permitir que outros pesquisadores examinem e reproduzam os estudos, a ciência aberta promove a confiabilidade e a validade dos achados.
3. **Criatividade e inovação:** O acesso aberto ao conhecimento científico inspira novas ideias e abordagens, incentivando a criatividade e a inovação. Ao permitir que uma ampla gama de indivíduos acesse e contribua para a pesquisa, a ciência aberta pode levar a descobertas inesperadas e avanços significativos.

4. **Colaboração global:** A ciência aberta transcende fronteiras geográficas e disciplinares, facilitando a colaboração entre pesquisadores de diferentes países e áreas de especialização. Isso pode levar a parcerias produtivas e abordagens interdisciplinares para resolver desafios complexos.
5. **Benefícios para a sociedade:** Ao tornar o conhecimento científico amplamente acessível, a ciência aberta beneficia não apenas a comunidade acadêmica, mas também a sociedade em geral. Isso pode levar a avanços em saúde, tecnologia, meio ambiente e outras áreas que impactam diretamente a qualidade de vida das pessoas.

A adoção generalizada das práticas de ciência aberta é fundamental para maximizar o potencial da pesquisa científica e garantir que seus benefícios sejam compartilhados de forma ampla e equitativa.

5.1.2 Componentes da OHDSI

A interoperabilidade entre sistemas depende de 2 fatores [Mucheroni 2011]:

- Interoperabilidade sintática (forma)
- Interoperabilidade semântica (conteúdo)

Para ser possível efetuar comparações e aplicar métodos estatísticos em conjuntos de dados de fontes diversas, a OHDSI se fundamenta em 3 componentes principais:

- Modelo comum de dados: CDM OMOP
- Vocabulários Padronizados
- Ferramentas para preparação do CDM, bibliotecas para análises estatísticas e definição de estudos

As ferramentas disponibilizadas pela OHDSI operam de maneira coesa e padronizada para apoiar uma variedade de análises de dados observacionais no nível do paciente. Os principais pontos em destaque são:

1. **Interoperabilidade com o modelo CDM OMOP:** Todas as ferramentas da OHDSI são projetadas para interagir com bancos de dados estruturados no modelo CDM OMOP. Isso permite que diferentes ferramentas acessem e analisem os dados de forma consistente, independentemente do banco de dados subjacente.
2. **Padronização das análises:** As ferramentas da OHDSI padronizam as análises para vários casos de uso, garantindo consistência e uniformidade nos resultados. Por exemplo, ao calcular uma taxa de incidência, as ferramentas oferecem opções padronizadas para especificar os parâmetros necessários, garantindo que os cálculos sejam realizados de maneira consistente em diferentes projetos e por diferentes usuários.

3. **Facilidade de execução:** Ao padronizar as análises e fornecer uma interface amigável, as ferramentas da OHDSI tornam mais fácil a execução de análises complexas de dados observacionais.
4. **Melhoria da reprodutibilidade e transparência:** Ao garantir que as análises sejam padronizadas e bem documentadas, as ferramentas da OHDSI promovem a reprodutibilidade dos resultados. Isso significa que outros pesquisadores podem reproduzir os mesmos resultados usando as mesmas ferramentas e configurações. Além disso, a transparência é aprimorada, pois as escolhas metodológicas são claramente definidas e documentadas.

No geral, as ferramentas da OHDSI desempenham um papel crucial na promoção da colaboração, reprodutibilidade e transparência na análise de dados observacionais de saúde, permitindo avanços significativos na pesquisa e na prática clínica. O conjunto de ferramentas disponibilizadas pela OHDSI auxiliam a preparação da base no modelo CDM OMOP, na validação do processo de mapeamento, na verificação da qualidade dos dados que compõem o CDM, na elaboração e análise dos diferentes tipos de estudos, facilitando a exploração dos dados e a geração de evidências [Abrahão 2022]. Podemos citar:

- Ferramentas para geração do banco CDM OMOP
 - o ETL: White Rabbit, Rabbit-In-A-Hat³
 - o Vocabulários: Athena⁴/Usagi⁵
 - o Qualidade: Achilles⁶ e *Data Quality Dashboard (DQD)*⁷
- Ferramentas de análise
 - o Geração de Coortes/estudos - Atlas⁸
 - o Análises estatísticas: HADES⁹

5.1.3 Onde achar: principais referências

As informações a respeito dos componentes da OHDSI podem ser classificadas em:

- Informações gerais: <http://www.ohdsi.org> - Site principal;
- Código e instalações: <https://github.com/OHDSI/> Código fonte de todas as ferramentas. Em particular destacamos: *Common Data Model* (<https://github.com/OHDSI/CommonDataModel>), com a definição completa do modelo e as implementações para os diversos bancos suportados;

³ White Rabbit, Rabbit-In-A-Hat <http://ohdsi.github.io/WhiteRabbit/index.html>

⁴ ATHENA <http://athena.ohdsi.org>

⁵ Usagi <http://ohdsi.github.io/Usagi/>

⁶ ACHILLES <http://www.ohdsi.org/web/achilles>

⁷ DQD <https://ohdsi.github.io/DataQualityDashboard/articles/DataQualityDashboard.html>

⁸ Atlas <https://github.com/OHDSI/Atlas/wiki/Atlas-Setup-Guide>

⁹ HADES <https://github.com/OHDSI/Hades>

- Tutoriais e vídeos: (<https://www.google.com/search?q=youtube+ohdsi>), documentação e tutoriais em vídeo dos eventos anuais do grupo;
- Fórum: <http://forums.ohdsi.org/>
- Livro OHDSI: [The book of OHDSI](#)

5.2 Modelo Comum de Dados (CDM OMOP): Estrutura e Domínios

Um modelo de dados comum desempenha um papel fundamental em várias áreas cruciais da saúde, como:

1. **Colaboração:** Ao adotar um modelo de dados comum, diferentes instituições de saúde, pesquisadores e profissionais médicos podem compartilhar e entender os dados de forma consistente. Isso facilita a colaboração entre diferentes partes interessadas, permitindo a troca de informações e o trabalho conjunto em projetos e iniciativas de saúde.
2. **Pesquisa:** Um modelo de dados comum fornece uma estrutura padronizada para organizar e analisar dados de saúde. Isso simplifica o processo de pesquisa, permitindo que os pesquisadores combinem e analisem grandes conjuntos de dados de forma eficiente. Como resultado, os estudos de pesquisa podem ser realizados de maneira mais eficaz, levando a descobertas significativas e avanços na área da saúde.
3. **Qualidade do cuidado:** Com acesso a dados mais abrangentes e consistentes, os provedores de saúde podem tomar decisões mais informadas sobre o tratamento e a gestão da saúde de seus pacientes. Um modelo de dados comum ajuda a garantir que as informações relevantes estejam disponíveis quando necessário, o que pode levar a uma melhoria na qualidade do cuidado e melhores resultados para os pacientes.
4. **Abordagem integrada:** Ao padronizar a maneira como os dados são coletados, armazenados e compartilhados, um modelo de dados comum promove uma abordagem mais integrada para o gerenciamento da saúde. Isso significa que os dados podem fluir mais facilmente entre diferentes sistemas e partes interessadas, permitindo uma visão mais abrangente e holística da saúde de um indivíduo ou população.

O CDM OMOP é uma estrutura de dados padronizada que facilita o armazenamento, a padronização, integração e análise de informações de saúde. Compreender a estrutura do modelo OMOP é o primeiro passo para qualquer organização que deseje adaptar suas bases de dados para este padrão. Essa compreensão

envolve estudar a organização das tabelas do modelo, os tipos de dados que cada uma armazena e como esses dados são inter-relacionados. Alguns pontos-chave sobre o CDM OMOP são:

1. **Estrutura padronizada:** estabelece uma estrutura comum para representar uma variedade de dados clínicos, com informações centradas no paciente, procedimentos médicos, condições médicas, medicamentos prescritos e resultados de exames. Isso ajuda a organizar os dados de forma consistente e compreensível.
2. **Interoperabilidade:** ao adotar o CDM OMOP, diferentes sistemas de saúde e pesquisadores podem compartilhar e comparar dados de saúde de maneira mais eficiente. Isso facilita a integração e análise de grandes conjuntos de dados de diversas fontes, promovendo a colaboração e a troca de informações.
3. **Flexibilidade:** projetado para ser flexível e abrangente, permitindo a adaptação para atender às necessidades específicas de diferentes estudos e projetos de pesquisa. Ele pode acomodar uma ampla gama de dados clínicos e de pesquisa, proporcionando uma estrutura versátil para análise e investigação.
4. **Suporte à pesquisa clínica:** amplamente utilizado em estudos de pesquisa clínica e epidemiológica para investigar uma variedade de questões relacionadas à saúde, como segurança de medicamentos, eficácia de tratamentos e padrões de doenças. Ele fornece uma base sólida para análise de dados e geração de insights significativos.
5. **Comunidade ativa:** OMOP conta com uma comunidade ativa de pesquisadores, desenvolvedores de software e organizações de saúde que trabalham em conjunto para melhorar e expandir o modelo. Isso garante que o CDM do OMOP esteja sempre atualizado e alinhado com as melhores práticas da indústria, promovendo sua relevância e eficácia contínuas.

Esses pontos destacam a importância do CDM do OMOP como uma ferramenta valiosa para padronizar e harmonizar dados observacionais de saúde, facilitando a pesquisa e promovendo avanços na área da saúde.

O CDM OMOP define uma estrutura comum para representar informações clínicas, que facilita a análise comparativa e colaborativa. A Figura 5.1 apresenta o esquema do modelo CDM OMOP.

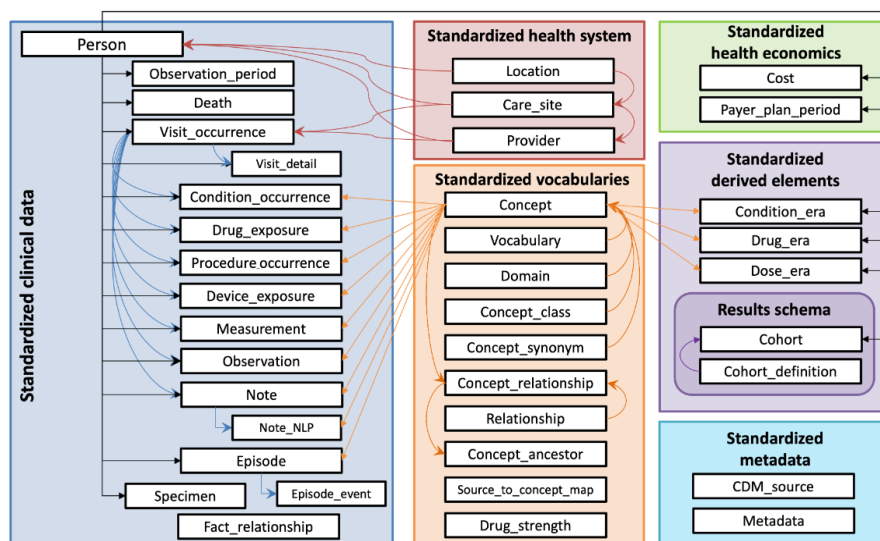


Figura 5.1 Visão geral da estrutura do CDM OMOP V5.4 ¹⁰

Os componentes do CDM OMOP são:

- Esquema – (tabelas onde são colocados os dados) é formado por:
 - 39 tabelas com 433 colunas
 - 17 tabelas com dados clínicos (centrados na pessoa)
 - 10 tabelas de vocabulários
 - 3 tabelas com informações de saúde, 2 com dados econômicos, 3 com dados derivados, 2 de resultados e 2 com metadados
- Vocabulário:
 - 153 vocabulários distribuídos em 41 domínios
 - Padrões: SNOMED, RxNorm, LOINC
 - 9 milhões de conceitos
 - >3,3 milhões de conceitos padrão
 - >5,1 milhões de códigos-fonte
 - 629.000 conceitos de classificação
 - >55 milhões de relacionamentos conceituais
 - >84 milhões de relacionamentos ancestrais
- Convenções – (regras de como os dados devem ser armazenados).

O CDM é otimizado para fins típicos de pesquisa observacional para identificar populações de pacientes com determinadas intervenções de saúde e resultados; na caracterização dessas populações de pacientes para vários parâmetros, como informações demográficas, história natural da doença, prestação de cuidados de saúde,

¹⁰ Fonte: <https://ohdsi.github.io/CommonDataModel/>

morbidades, tratamentos; prever a ocorrência destes resultados em pacientes individuais; estimar o efeito que estas intervenções têm na população.

Para atingir este objetivo, o desenvolvimento do CDM segue os seguintes elementos de design:

- **Adequação à finalidade:** O CDM visa fornecer dados organizados de uma forma ideal para análise;
- **Proteção de dados:** Todos os dados que possam comprometer a identidade e a proteção dos pacientes, como nomes, datas de nascimento são limitados.
- **Design de domínios:** Os domínios são modelados em um modelo de dados relacionais centrado na pessoa, onde para cada registro a identidade da pessoa e uma data são capturadas no mínimo. Aqui, um modelo de dados relacional é aquele em que os dados são representados como uma coleção de tabelas vinculadas por chaves primárias e estrangeiras.
- **Justificativa para domínios:** Os domínios são identificados e definidos separadamente em um modelo de relacionamento entre entidades se tiverem um caso de uso de análise (condições, por exemplo) e o domínio tiver atributos específicos que não são aplicáveis de outra forma. Todos os outros dados podem ser preservados como uma observação na tabela de observação em uma estrutura entidade-atributo-valor.
- **Vocabulários Padronizados:** Para padronizar o conteúdo desses registros, o CDM conta com os Vocabulários Padronizados contendo todos os conceitos de saúde padrão correspondentes, necessários e apropriados.
- **Reutilização de vocabulários existentes:** Se possível, esses conceitos são aproveitados de organizações ou iniciativas nacionais ou industriais de padronização ou definição de vocabulário, como a Biblioteca Nacional de Medicina, o Departamento de Assuntos de Veteranos, o Centro de Controle e Prevenção de Doenças, etc.
- **Manutenção de códigos-fonte:** embora todos os códigos sejam mapeados para os vocabulários padronizados, o modelo também armazena o código-fonte original para garantir que nenhuma informação seja perdida.
- **Neutralidade tecnológica:** O CDM não exige uma tecnologia específica. Pode ser realizado em qualquer banco de dados relacional, como Oracle, SQL Server etc., ou como conjuntos de dados analíticos SAS.
- **Escalabilidade:** O CDM é otimizado para processamento de dados e análise computacional para acomodar fontes de dados que variam em tamanho, incluindo bancos de dados com até centenas de milhões de pessoas e bilhões de observações clínicas.
- **Compatibilidade com versões anteriores:** todas as alterações dos CDMs anteriores estão claramente delineadas no repositório github¹¹. Versões mais

¹¹ <https://github.com/OHDSI/CommonDataModel>

antigas do CDM podem ser facilmente criadas a partir da versão atual e nenhuma informação que estava presente anteriormente será perdida.

A Figura 5.2 apresenta uma análise do esquema¹² do CDM OMOP com a documentação das tabelas, colunas, tipos de dados, tamanho, relacionamentos, um guia para o usuário, índices e convenções para o ETL dos dados. A Figura 5.3.1 e 5.3.2 ilustram o detalhamento da tabela DEATH com sua documentação.

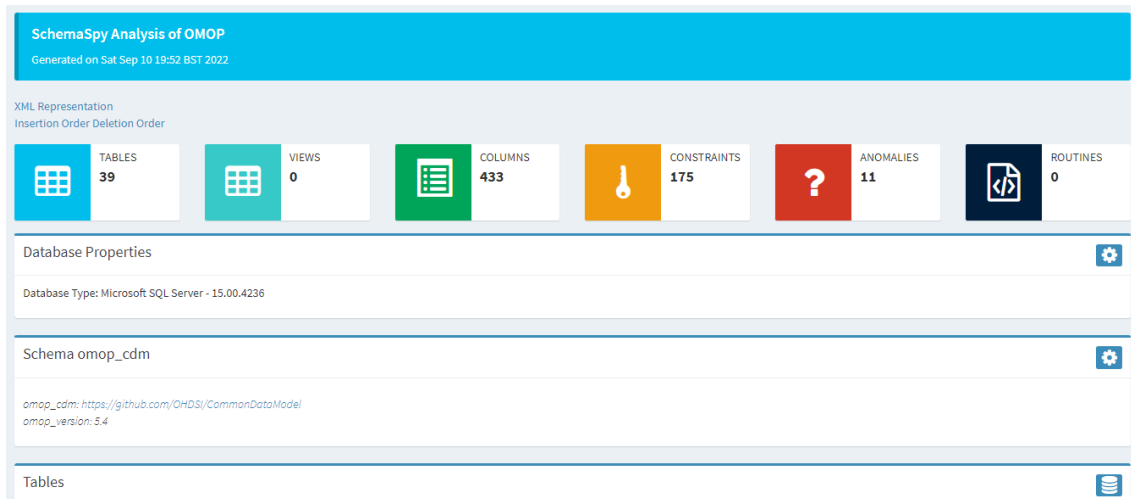


Figura 5.2 Análise do esquema do CDM OMOP

DEATH - OMOP Database

Column	Type	Size	Nulls	Auto	Default	Children	Parents	Comments
person_id	bigint	8			null		PERSON.person_id fpk_DEATH_person_idR	
death_date	date	6			null			User Guidance: The date the person was deceased. ETL Conventions: If the precise date include day or month is not known or not allowed, December is used as the default month, and the last day of the month the default day.
death_datetime	datetime	16,3	✓		null			ETL Conventions: If not available set time to midnight (00:00:00)
death_type_concept_id	int	4	✓		null		CONCEPT.concept_id fpk_DEATH_death_type_concept_idR	User Guidance: This is the provenance of the death record, i.e., where it came from. It is possible that an administrative claims database would source death information from a government file so do not assume the Death Type is the same as the Visit Type, etc. ETL Conventions: Use the type concept that reflects the source of the death record. Accepted Concepts. A more detailed explanation of each Type Concept can be found on the vocabulary wiki.
cause_concept_id	int	4	✓		null		CONCEPT.concept_id fpk_DEATH_cause_concept_idR	User Guidance: This is the Standard Concept representing the Person's cause of death, if available. ETL Conventions: There is no specified domain for this concept, just choose the Standard Concept Id that best represents the person's cause of death.
cause_source_value	varchar	50	✓		null			ETL Conventions: If available, put the source code representing the cause of death here.
cause_source_concept_id	int	4	✓		null		CONCEPT.concept_id fpk_DEATH_cause_source_concept_idR	ETL Conventions: If the cause of death was coded using a Vocabulary present in the OMOP Vocabularies put the CONCEPT_ID representing the cause of death here.

Figura 5.3.1 Detalhes da tabela DEATH

¹² Esquema OMOP Análise <https://omop-erd.surge.sh/index.html>

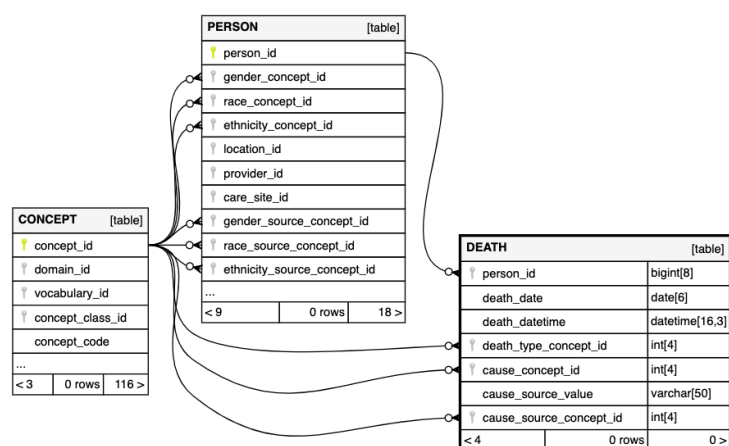


Figura 5.3.2 Fluxo de relacionamento da tabela DEATH

Essa padronização em um conjunto definido de tabelas e relacionamentos fornece um contexto comum para os elementos de dados clínicos, o que é necessário para criar métodos e algoritmos analíticos e de garantia de qualidade unificados que possam ser executados em toda a rede.

A harmonização do conteúdo de dados médicos é alcançada por meio do uso de vocabulários médicos ou esquemas de codificação, que são desenvolvidos e mantidos por diversas organizações e sociedades profissionais. Esses vocabulários garantem uma comunicação precisa e consistente sobre cuidados e tratamentos de pacientes. Eles variam desde conjuntos simples de códigos ou termos até hierarquias complexas ou vocabulários, frequentemente com cobertura cruzada de diversos domínios de saúde.

Para qualquer domínio específico, os membros de redes de dados distribuídas podem empregar diferentes vocabulários, versões distintas do mesmo vocabulário, vocabulários não públicos ou, em alguns casos, não utilizar vocabulário padronizado em seus dados. Essa diversidade pode complicar a integração e a análise dos dados, tornando a harmonização uma tarefa essencial para assegurar a interoperabilidade e a eficácia das trocas de informações na área da saúde.

5.3 Conceitos gerais dos vocabulários

Neste item vamos revisar, através de uma revisão bibliográfica, os principais conceitos relacionados com vocabulários como auxílio na compreensão da estrutura dos vocabulários da OHDSI.

Para isso reproduzimos partes das mais importantes fontes da literatura:

- O relatório do standard Guidelines for the construction, format, and management of monolingual controlled vocabularies (ANSI/NISO Z39.19-2005) [National 2005], fonte com as definições mais referenciadas da área
- Health Concept and Knowledge Management: Twenty-five Years of Evolution, uma atualização do trabalho seminal de Cornet e Chute [Cornet 2016]

- Desiderata for Controlled Medical Vocabularies in the Twenty-First Century, de Cimino [Cimino 2016]
- The MMI Guides: Navigating the World of Marine Metadata [Stocks 2010], de diversos autores, cada autor é indicado no correspondente parágrafo. Esta obra é um compêndio definitivo dos termos relacionados com a classificação de conceitos. É de destacar que este grupo apresenta um dos maiores trabalhos de criação de ferramentas gratuitas para criação e manutenção de vocabulários
- Outros autores são referenciados nos respectivos parágrafos.

A linguagem é uma das principais características que definem o ser humano. Única entre todas as formas de comunicação animal, se distingue por ser composicional, ela expressa pensamentos em sentenças abrangendo sujeito, verbo e objeto e reconhece passado, presente e futuro. Esta característica permite que a linguagem humana tenha uma capacidade ilimitada para gerar novas frases combinando conjuntos limitados de palavras, por exemplo, com 25 palavras de cada tipo podemos formular 15.000 frases diferentes. Outra característica é ser referencial, o que permite aos agentes falantes trocar informações específicas entre eles a respeito de outras pessoas ou objetos e as suas localizações ou ações [Pagel 2017].

À medida que cresce a quantidade de informação acessível globalmente, aumenta também a inerente dificuldade em encontrar itens de informação desejados. Dentro do contexto de arquivos documentais, uma das técnicas mais aplicadas para facilitar a descoberta de itens, tanto em sistemas manuais tradicionais quanto em sistemas mais recentes informatizados, tem sido a indexação.

A indexação consiste na atribuição de valores a atributos predefinidos para servir como base para pesquisa e descoberta de recursos. A combinação desses atributos e valores devem constituir informações suficientes para caracterizar com sucesso o conteúdo de um documento e permitir a futura recuperação deste documento apenas observando essas informações. Exemplos dos atributos comumente encontrados são: autor, título, assunto, resumo, etc. Estes são geralmente chamados de metadados [Ferreira 2005].

Mas esta flexibilidade também cria dificuldades resumidas na seguinte lista:

- Padronização: termos diferentes que representam o mesmo conceito (sinônimos)
- Ambiguidade: termos iguais para ideias diferentes. Exemplo:
 - Mercúrio: planeta
 - Mercúrio: metal
 - Mercúrio: Deus grego
- Pesquisa: dificuldade para achar o termo certo numa lista grande de termos
- Classificação: dificuldade para separar e agrupar conceitos em classes que facilitem a pesquisa

- Representação de conhecimento: Necessidade de transmitir informações a respeito da informação.

Assim aparecem as primeiras iniciativas de organização, no começo, a partir da definição de listas de termos para auxiliar na identificação unívoca de objetos ou conceitos.

Um termo é definido como uma ou mais palavras usadas para representar um conceito.

Estas listas de termos começam a crescer e agora precisamos achar os termos, distinguir sinônimos, evitar erros ortográficos; começamos assim a descrever o que hoje são chamados de vocabulários controlados.

Um vocabulário controlado é uma forma de inserir uma camada interpretativa de semântica entre o termo inserido pelo usuário e o banco de dados subjacente para melhor representar a intenção original do usuário [Karl 2002].

Isto permite que na hora da busca, possa ser apresentada uma lista de opções que capturem a intenção do usuário e facilitem a ele achar o que procura.

Um vocabulário controlado é uma lista de termos que foram enumerados explicitamente. Esta lista é controlada por uma autoridade de registro que se constitui na fonte autoritativa da verdade. Todos os termos de um vocabulário controlado devem ter uma definição inequívoca e não redundante, o que depende de quão rigorosa é a autoridade controladora em relação ao registro de novos termos [National 2005].

Um Authority File ou arquivo de referência ou autorizativo, é um tipo de vocabulário controlado que consiste em uma lista de rótulos e valores que estabelecem os valores aceitáveis que podem ser inseridos em um parâmetro específico. Nenhuma explicação ou informação adicional é fornecida sobre os valores aceitáveis.

Duas regras devem ser aplicadas:

- Se o mesmo termo é comumente usado para significar conceitos diferentes, então o seu nome é explicitamente qualificado para resolver esta ambiguidade. NOTA: Esta regra não se aplica a anéis de sinônimos.
- Se vários termos foram usados para significar a mesma coisa, um dos termos é identificado como o termo preferido no vocabulário controlado e os outros termos são listados como sinônimos ou pseudônimos.

A Figura 5.4 mostra os diversos tipos de vocabulários controlados em função da capacidade de representar conhecimento e dos problemas abordados [Heather 2022].

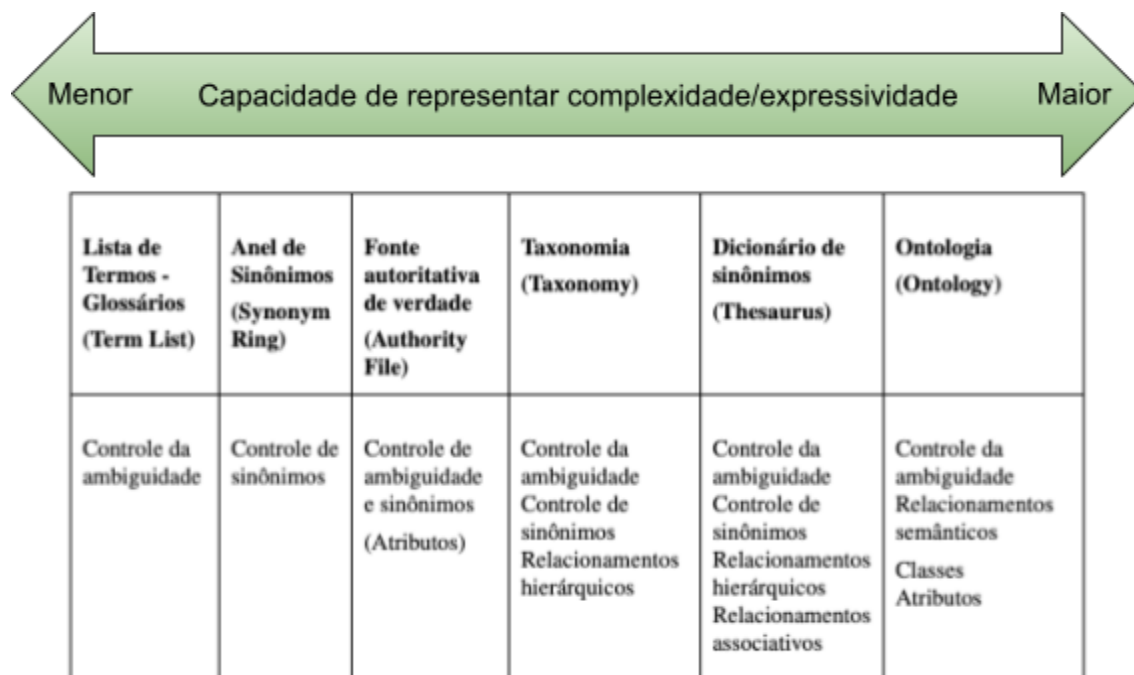


Figura 5.4 Classificação dos vocabulários

5.3.1 Classificação dos vocabulários controlados pela forma

Para permitir a gestão formal, um vocabulário controlado pode ser organizado estruturalmente de modo que se encaixe em uma dessas amplas categorias [Stocks 2010]:

- **Plana:** fornece um conjunto de termos obrigatórios que podem ser usados. Alguns vocabulários planos controlados fornecerão informações adicionais sobre cada termo.
- **Multinível:** baseia-se em um vocabulário plano e controlado, atribuindo cada termo a uma categoria.
- **Relacional:** fornece um conjunto de termos e captura como eles estão associados entre si.

A Tabela 5.1 apresenta esta organização.

Tabela 5.1 Classificação dos vocabulários pela forma

Categoria de Vocabulário Controlado	Tipo de vocabulário controlado	Descrição
Plano (Flat)	Arquivo de autoridade (Authority File)	Lista de Termos preestabelecidos
	Glossário	Lista de termos e definições dentro de um domínio específico
	Dicionário	Lista de termos, definições, e informações adicionais
	Lista de códigos	Lista de códigos (ex. abreviaturas) e definições
Multinível	Taxonomia	Termos classificados em categorias
	Subject Heading (Cabeçalho do Assunto)	Termos classificados em categorias, que podem ser classes abrangentes
Relacionais	Tesauros	Conjunto de termos e os relacionamentos entre valores individuais
	Redes Semânticas	Conjunto de termos/conceitos e relacionamentos direcionados
	Ontologias	Conjunto de termos e relações entre termos, aprimorado por informações adicionais fornecidas por regras e axiomas

5.3.1.1 Vocabulários planos: Arquivo de autoridade (Authority Files), Glossários, Dicionários, Diários, Lista de códigos

Todos os vocabulários simples contém um rótulo e um valor. Alguns vocabulários planos baseiam-se nesta base, adicionando uma definição ou informações adicionais sobre cada valor. Nenhum relacionamento é estabelecido, nenhuma hierarquia é estabelecida e nenhuma matriz complicada é criada.

5.3.1.2 Vocabulários multiníveis: Taxonomias, Título do assunto (Subject Headings)

Um vocabulário multinível é essencialmente uma forma de agrupar termos em classes com hierarquia. Uma classificação diz mais sobre os "termos" do que um vocabulário plano, colocando-os em subcategorias bem pensadas.

Em um vocabulário multinível, você pode examinar a qual subcategoria um termo pertence e também examinar as relações entre as subcategorias. Em alguns vocabulários multiníveis (taxonomias), a única conexão entre as subcategorias é uma

comparação "mais abrangente"/"mais específica" (BT-Broader Term/NT-Narrower Term).

Em outros, você pode comparar categorias semelhantes em categorias mais amplas (cabeçalho de assunto- Subject Headings).

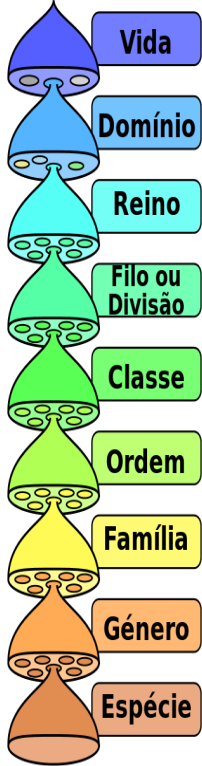
5.3.1.2.1 Taxonomia

É um vocabulário controlado multinível no qual os valores de metadados são agrupados de acordo com classes específicas de assunto, geralmente hierárquicas.

Táxon é derivado do grego taxis, ordenamento, por tanto taxonomia é a ciência que lida com a descrição, identificação e classificação de conceitos.

Talvez a taxonomia mais conhecida seja a taxonomia de Lineu, que classifica de forma única os seres vivos, extensamente usada nas ciências biológicas. Ela foi desenvolvida por Carolus Linnaeus (Conhecido normalmente como Carl von Linné ou em português como Carlos Lineu, botânico e médico sueco) no Século XVIII durante a grande expansão da história natural. A Tabela 5.2 mostra a classificação do ser humano.

Tabela 5.2 Classificação do ser humano na taxonomia de Linnean

 <p>(figura: Peter Halasz)</p>	<p>Domínio: Eukarya Característica: seres vivos compostos por células com um núcleo celular organizado</p> <p>Reino: Animalia Característica: células eucarióticas com membrana celular, mas sem parede celular, multicelulares, heterotróficas</p> <p>Filo: Chordata Característica: notocorda, cordão nervoso dorsal e fendas branquiais faríngeas</p> <p>Classe: Mamíferos Característica: glândulas endotérmicas, capilares e mamárias (que servem para nutrir os filhotes)</p> <p>Ordem: Primatas Característica: clavícula, olhos voltados para frente, agarrar as mãos com os dedos, dois tipos de dentes</p> <p>Família: Hominidae Característica: postura ereta, cérebro grande, visão estereoscópica, rosto achatado, mãos e pés com diferentes funções</p> <p>Gênero: Homo Característica: coluna curvada em S</p> <p>Espécie: sapiens Característica: testa alta, queixo bem desenvolvido, ossos do crânio finos</p>
---	--

Observe que, como seres humanos, recebemos um identificador único (Homo sapiens) e exibimos todas as características listadas (em outras palavras, como nossa

classificação está no final de uma lista aninhada, podemos herdar todas as características das “superclasses”). Para classificar completamente humanos nesta taxonomia, precisamos usar o termo "Homo sapiens", mas também pode ser usado "primatas". Esta não seria a classificação mais restrita, mas é uma classificação precisa.

5.3.1.3 Vocabulários Relacionais: Tesouros, Rede Semânticas, Ontologias

Os Vocabulários Relacionais, também chamados de listas de relacionamentos, contém um mecanismo para conectar termos. As relações são descritas por vários padrões e protocolos, como para tesouros no padrão ANSI [233]/NISO [109] Z39.19 - 2005, incluindo "mais abrangente"/"mais específico", "use" (*USE*, no sentido: "use no lugar", apontando para o termo principal), "usado para" (*UF-USE FOR*: apontado para os termos semelhantes) e "relacionado" (RT- Related Term).

5.3.1.3.1 Tesouro

No contexto de metadados, um tesouro é um tipo de vocabulário relacional controlado que fornece uma lista de termos de metadados com relações específicas entre os termos. De acordo com o padrão ANSI/NISO Z39.19 - 2003 [National 2005]:

Um tesouro é um vocabulário controlado organizado em uma ordem conhecida e estruturada de modo que relações de equivalência, homografia, hierarquia e associação entre valores são exibidas claramente e identificadas por indicadores de relacionamento padronizados que são empregados reciprocamente.

Os objetivos principais de um tesouro são:

- (a) facilitar a recuperação de documentos
- (b) alcançar consistência na indexação de documentos escritos ou registrados de outra forma e outros itens, principalmente para sistemas pós-coordenados de armazenamento e recuperação de informações.

Os tesouros basicamente pegam as taxonomias descritas acima e as estendem para torná-las mais capazes de descrever o mundo [Garshol 2004], não apenas permitindo que os assuntos sejam organizados em uma hierarquia, mas também permitindo que outras declarações sejam feitas sobre os assuntos. O padrão ISO 2788 [Garshol 1986] fornece as seguintes propriedades para descrever assuntos:

BT (Broader Term)

Abreviação de “termo mais amplo”, refere-se ao termo acima deste na hierarquia; esse termo deve ter um significado mais amplo ou menos específico. Na prática, alguns sistemas permitem múltiplos BTs para um período, enquanto outros não. (Existe uma propriedade inversa conhecida como NT (Narrower Term), para "termo mais restrito", que está implícita no BT.) Pode-se dizer que as taxonomias conforme descritas acima são tesouros que usam apenas as propriedades BT/NT para construir uma hierarquia, e

sem fazermos uso de nenhuma das propriedades descritas abaixo, então pode-se dizer que todo tesauro contém uma taxonomia.

SN (Scope Note)

Esta é uma string anexada ao termo explicando seu significado no dicionário de sinônimos. Isto pode ser útil nos casos em que o significado preciso do termo não é óbvio no contexto. "SN" significa "nota de escopo".

USE (e o complemento: UF Use For)

Refere-se a outro termo que deve ser preferido em vez deste termo; implica que os termos são sinônimos. (Existe uma propriedade inversa conhecida como UF: Use For, "use para"). Por exemplo, na figura Figura 5.5 do Tesauro da UNESCO, o termo "Dreams" aponta para "Sleep", representando o "USE" graficamente com uma seta. Na definição do termo "Sleep" vemos explicitamente a referência .

RT (Related Term / Related Concept)

Abreviação de "termo relacionado", refere-se a um termo que está relacionado a este termo, sem ser um sinônimo dele ou um termo mais amplo/restrito. No nosso exemplo: "Suggestopaedia"

Em suma, os tesauros fornecem um vocabulário muito mais rico para descrever os termos do que as taxonomias e, portanto, são ferramentas muito mais poderosas. Como pode ser visto, usar um tesauro em vez de uma taxonomia resolveria vários problemas práticos na classificação de objetos e também na busca por eles.

Na Figura 5.5 vemos o Tesauro da UNESCO¹³, apresentando os relacionamentos de "Dreams" para o termo principal "Sleep" (use preferred term).

The screenshot shows the UNESCO Thesaurus interface. The main content area displays the entry for "Sleep" as the preferred term. It lists related concepts, including "Suggestopaedia", and entry terms, including "Dreams". The interface also shows the term in other languages: Arabic (نوم), French (Sommeil), Russian (Сон (состояние)), and Spanish (Sueño). The URI for the concept is provided as <http://vocabularies.unesco.org/thesaurus/concept16462>.

Figura 5.5 Tesauro da Unesco

¹³ Unesco <https://vocabularies.unesco.org>

Por sua vez, a definição de "*Sleep*" aponta para "*Dreams*" (*use for*) mostrado como "*Entry Terms*".

Entry terms (pontos de entrada), no exemplo "*Dreams*", são os vocábulos que permitem que uma busca consiga recuperar mais facilmente o termo principal (preferred term) da definição de um termo (*recall*) [DeMars 2022].

O "MMI Guides: Navigating the World of Marine Metadata"[Stocks 2010] e o padrão (ANSI/NISO Z39.19-2005) [National 2005] contém definições para todos os elementos destas classificações.

5.3.1.3.2 Ontologias

Segundo Tom Gruber (1993):

"An ontology is a specification of a conceptualization"

(Uma ontologia é a especificação de uma conceitualização)

Ele descreve uma ontologia da seguinte maneira:

Uma ontologia especifica um vocabulário com o qual fazer afirmações, que podem ser entradas ou saídas de agentes de conhecimento (como um programa de software). Uma ontologia deve ser formulada em alguma linguagem de representação.

A demanda por uma linguagem de representação restringe o conceito, ou seja, a definição de Gruber (1995) descreve o conceito de ontologia formal [Madsen 2009].

No seu artigo de revisão, "Health Concept and Knowledge Management: Twenty-five Years of Evolution" [Cornet, R, and C G Chute 2016], os autores discursam sobre os conceitos já aqui descritos e definem a principal característica das classificações estatísticas:

Terminologia: Um sistema de conceitos com identificadores atribuídos e termos de linguagem humana, normalmente envolvendo algum tipo de hierarquia semântica. Alguns sistemas podem suportar a atribuição de múltiplos termos, ou sinónimos, a um determinado conceito; estes podem incluir termos em vários idiomas naturais, como inglês ou holandês.

Ontologia: Uma terminologia que invoca relações semânticas formais entre conceitos, normalmente manifestadas como um tipo de Lógica de Descrição.

Classificação: Um sistema terminológico destinado a descrever exhaustivamente um domínio ou tópico, normalmente invocando a colocação criteriosa de categorias residuais, como Não especificado ou Não classificado em outro lugar, para alcançar a abrangência.

Classificações Estatísticas:

Uma classificação onde todos os conceitos são mutuamente exclusivos para evitar contar as coisas duas vezes. Isso normalmente é conseguido usando uma mono-hierarquia, onde cada conceito tem um e apenas um pai.

É importante distinguir as classificações estatísticas das terminologias, uma vez que servem propósitos diferentes. A classificação estatística mais conhecida é Classificação Internacional de Doenças¹⁴ (CID).

O CID é importante porque fornece uma linguagem comum para registrar, notificar e monitorar doenças. Isto permite ao mundo comparar e partilhar dados de uma forma consistente e padronizada – entre hospitais, regiões e países e ao longo de períodos de tempo. Facilita a recolha e armazenamento de dados para análise e tomada de decisões baseadas em evidências.

O objetivo mais importante do CID é manter estatísticas mundiais de morbidade e mortalidade através do tempo. Por isso nada pode ser contado duas vezes.

Isso obriga que a classificação tenha categorias do tipo:

- NEC: Not Elsewhere Classified (Sem classificação)
- NOS: Not Otherwise Specified (Sem especificação)

Estas são chamadas de categorias residuais. Elas, junto com a mono-hierarquia, são imprescindíveis para a construção de uma classificação estatística.

5.3.2 Exemplificando as diferenças entre os tipos de vocabulários

As seguintes figuras exemplificam os diferentes tipos [Stocks 2010].

5.3.2.1 Dicionário



Figura 5.6 Dicionário

Cada termo é articulado com uma sigla. (1ª entrada, azul)

As siglas estão explicadas na descrição. (2ª entrada, amarelo)

¹⁴ World Health Organization. (2004). ICD-10 : international statistical classification of diseases and related health problems : tenth revision, 2nd ed. World Health Organization. <https://iris.who.int/handle/10665/42980>

Informações adicionais sobre como cada termo surgiu estão incluídas na etimologia. (3ª entrada, verde).

Isto forma a definição clássica da relação objeto/metainformação onde a metainformação tem 3 campos: sigla, descrição, etimologia.

5.3.2.2 Taxonomia

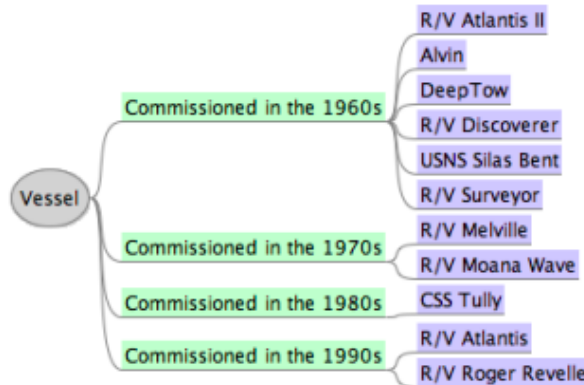


Figura 5.7 Taxonomia

Os termos reais (2ª entrada, azul) são colocados numa estrutura, de acordo com a década em que foram comissionados (1ª entrada, verde).

5.3.2.3 Ontologia

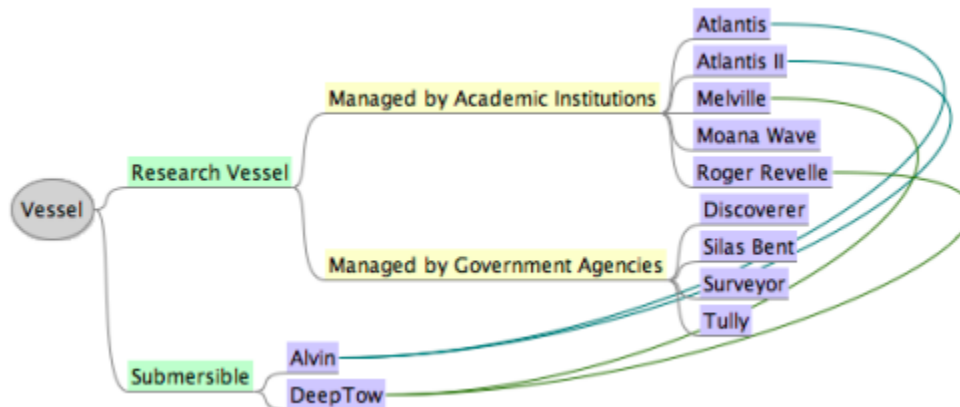


Figura 5.8 Ontologia

Os termos reais (3ª entrada, azul) são classificados em duas classes principais (1ª entrada, verde) e uma subclasse (2ª entrada, amarelo).

Observe que as embarcações estão conectadas a submersíveis, de acordo com a instituição operadora. Esta é uma inter-relação complexa que aumenta a hierarquia de classes.

Cada um desses vocabulários controlados representa a mesma lista de objetos do mundo real (ou seja, embarcações ou submarinos). Eles são apresentados com diferentes tipos de vocabulários controlados, usando termos diferentes para representar os mesmos objetos do mundo real e com informações ligeiramente distintas.

5.3.3 Importância dos vocabulários controlados

Temos então agora, um conjunto limitado de conceitos associados a termos específicos e com variabilidade controlada por uma autoridade que faz manutenção contínua deste vocabulário controlado.

A construção de vocabulários controlados é um processo demorado e trabalhoso, especialmente se o domínio a ser coberto é amplo e a terminologia utilizada é rica e complexa. O trabalho envolvido é justificado porque o uso de vocabulários controlados ajuda a garantir consistência na indexação e promove uma recuperação mais satisfatória.

Vocabulários controlados são importantes para pesquisadores por vários motivos [Stocks 2010]:

- Consistência
- Precisão
- Automação
- Simplificação de entrada
- Interoperabilidade
- Aprimoramento de pesquisas e descobertas
- Completude
- Gestão de longo e curto prazo
- Uso eficiente do tempo

Em muitos casos, termos de vocabulário controlado definem completamente o conteúdo permitido para um determinado elemento de metadados.

Além disso, um vocabulário controlado pode ser facilmente incorporado em procedimentos automatizados. Em um sistema de dados, um vocabulário controlado pode simplificar a entrada do sistema e contribuir para o controle de qualidade, fornecendo aos usuários ou outros sistemas uma lista de entradas permitidas e podem ser usados para verificar descrições de metadados existentes ou importadas quanto à consistência e correção, incluindo ortográfica e hifenização.

A necessidade de interoperabilidade surgiu logo após o desenvolvimento dos primeiros vocabulários controlados. Um trabalho considerável, tanto prático quanto acadêmico, foi feito para desenvolver métodos que permitam que vocabulários controlados sejam usados em vários bancos de dados e sistemas e compartilhados entre indexadores e pesquisadores.

Dentre os mecanismos que facilitam a interoperabilidade podemos citar os descritores e as palavras-chaves¹⁵.

Descritores (subject headings): São conjuntos padronizados e formalmente atribuídos de termos (também chamados de descritores) em um banco de dados para identificar os tópicos principais de um livro ou artigo. Os dois exemplos mais conhecidos de descritores são:

- Library of Congress Subject Headings (LCSH) [Library of Congress 2004]
- MeSH (Medical Subject Headings)¹⁶

Palavras-chave são palavras ou frases que podem aparecer em qualquer lugar do item (dados de citação ou texto completo de um livro ou artigo). A pesquisa por palavra-chave é a forma normal de pesquisa nos mecanismos de pesquisa na web e é um bom começo para encontrar títulos de assuntos relevantes em bancos de dados.

¹⁵ CEU Library, The Central European University Library, Vienna, Austria
<https://ceu.libguides.com/databasesearchtips/subjects>

¹⁶ MeSH (Medical Subject Headings) (2004) Bethesda (MD): National Library of Medicine. Available from: <http://www.nlm.nih.gov/mesh>

5.3.4 Knowledge organization system (KOS)

É um termo genérico usado para se referir a uma ampla gama de itens (por exemplo, títulos de assuntos, tesouros, esquemas de classificação e ontologias), que foram concebidos com relação a diferentes propósitos, em momentos históricos distintos. São caracterizados por diferentes estruturas e funções específicas, formas variadas de se relacionar com a tecnologia e são utilizados numa pluralidade de contextos por diversas comunidades [Mazzocchi 2018]. No entanto, o que todos têm em comum é que foram concebidos para apoiar a organização do conhecimento e da informação, de forma a facilitar a sua gestão e recuperação.

Uma das tipologias mais abrangentes tenha sido fornecida por Souza et al. [Souza 2012], que ainda identificam a estrutura como o principal critério de divisão, embora também esteja incluída uma divisão secundária, que leva em conta numerosos domínios de aplicação e casos de uso. A Figura a seguir apresenta este detalhamento.

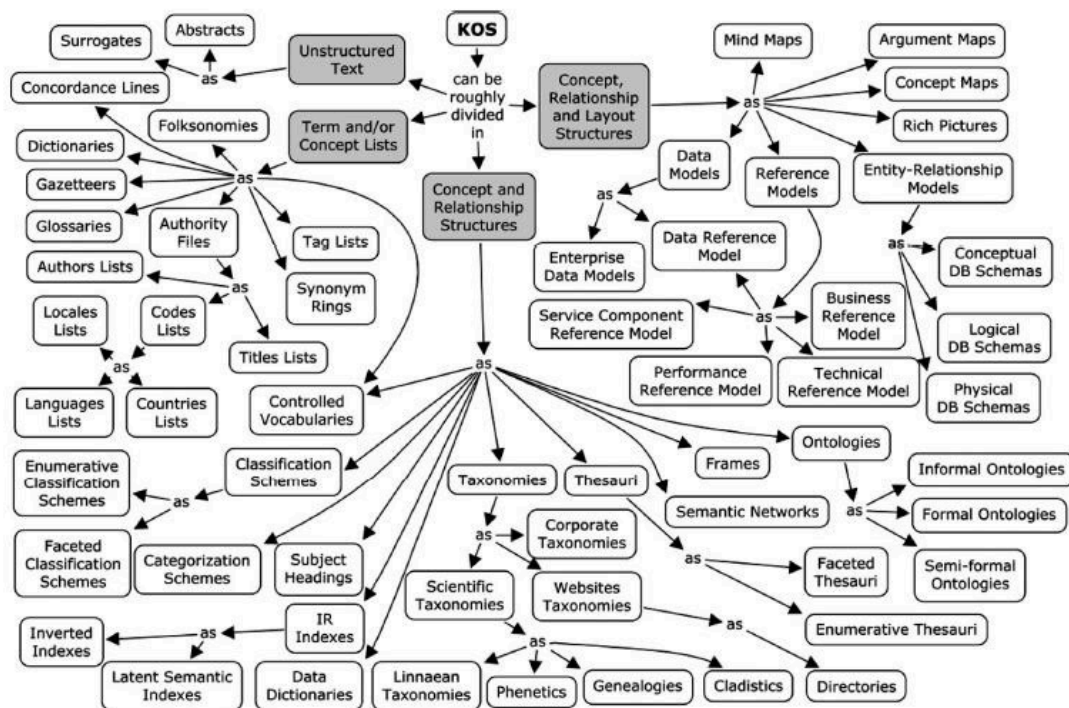


Figura 5.9 Souza et al.'s (2012) classification of KOSs

5.3.4.1 SKOS

O Simple Knowledge Organization System¹⁷ é uma recomendação do W3C [Mazzocchi 2018] projetada para representação de tesouros, esquemas de classificação, taxonomias,

¹⁷ Introduction to SKOS <https://www.w3.org/2004/02/skos/intro>

sistemas de controle de autoridade ou qualquer outro tipo de vocabulário controlado estruturado.

É uma área de trabalho que desenvolve especificações e padrões para apoiar o uso de sistemas de organização do conhecimento (KOS), como tesauros, esquemas de classificação, sistemas de cabeçalhos de assuntos e taxonomias no âmbito da Web Semântica.

A seguinte Figura descreve o avanço das terminologias no tempo [Lomax 2021]:

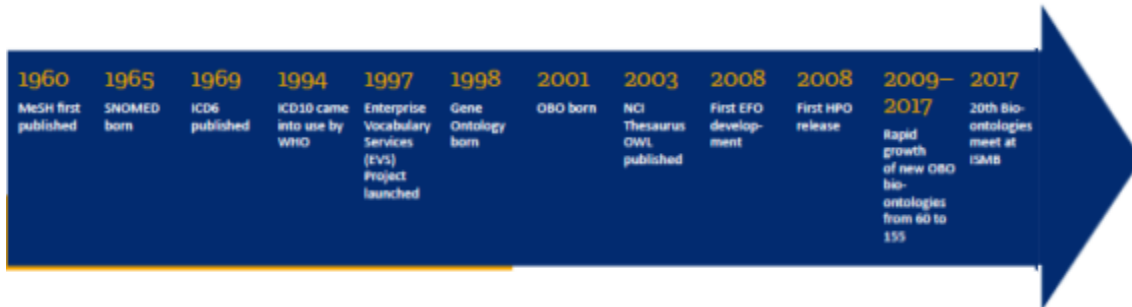


Figura 5.10 Terminologias no tempo (1960 a 2017)

5.4 Principais Vocabulários Padrão OHDSI

SNOMED CT¹⁸ é uma abreviação para "Systematized Nomenclature of Medicine Clinical Terms", que em português significa "Nomenclatura Sistematizada de Termos Clínicos em Medicina". É um sistema de terminologia clínica abrangente usado globalmente na área da saúde e pesquisa clínica. SNOMED CT é organizado em torno de conceitos em vez de termos individuais, possui uma estrutura hierárquica e é amplamente utilizado para codificação clínica em registros eletrônicos de saúde, faturamento, suporte a decisões e gerenciamento de saúde da população. É mantido pela IHTSDO e continuamente atualizado para refletir novos conhecimentos médicos e mudanças na prática clínica. Em resumo, SNOMED CT desempenha um papel fundamental na padronização e interoperabilidade das informações clínicas, melhorando a qualidade e eficiência dos cuidados de saúde.

LOINC¹⁹, ou Logical Observation Identifiers Names and Codes, é um sistema de codificação universalmente reconhecido para identificar testes laboratoriais e observações clínicas. Ele fornece códigos numéricos e alfanuméricos para uma ampla variedade de observações clínicas, como medidas de laboratório, observações clínicas, questionários e escalas de avaliação. LOINC é usado para padronizar a comunicação e intercâmbio de dados entre sistemas de saúde, facilitando a interoperabilidade e o compartilhamento de informações clínicas. Ele é mantido pelo Regenstrief Institute, Inc. e continua a evoluir para abranger novas áreas de observação clínica e laboratorial. Em resumo, LOINC desempenha um papel crucial na integração de dados clínicos e na promoção da qualidade e eficiência dos cuidados de saúde.

¹⁸ <https://www.snomed.org/> e <https://www.nlm.nih.gov/healthit/snomedct/index.html>

¹⁹ <https://loinc.org/>

O **RxNorm**²⁰ é um sistema de terminologia (tesauro) desenvolvido pelo National Library of Medicine (NLM) dos Estados Unidos. Ele fornece nomes padronizados para medicamentos e produtos relacionados à saúde, bem como seus ingredientes ativos, formas farmacêuticas, dosagens e rotas de administração. RxNorm é amplamente utilizado em sistemas de informação de saúde para facilitar a interoperabilidade e a troca de dados sobre medicamentos entre diferentes instituições e sistemas de informação. Ele ajuda a evitar ambiguidades na comunicação sobre medicamentos, promove a segurança do paciente e auxilia na pesquisa clínica e farmacêutica.

O RxNorm vincula nomes normalizados para medicamentos clínicos a muitos vocabulários de medicamentos comumente usados em software de gerenciamento de farmácia e interação medicamentosa, incluindo os do First Databank, Micromedex, Multum e Gold Standard Drug Database. Ao fornecer links entre esses vocabulários, o RxNorm pode mediar mensagens entre sistemas que não usam o mesmo software e vocabulário.

No entanto, o principal desafio envolve as limitações do RxNorm, um modelo originalmente concebido para representar medicamentos no mercado dos EUA. Seu foco está nos medicamentos atuais, capturando detalhes como conteúdo e dosagem de forma eficaz. No entanto, é insuficiente em dois aspectos críticos: não pode representar conceitos internacionais sobre medicamentos e não inclui dados sobre medicamentos que já não estão disponíveis. Isso coloca obstáculos adicionais ao trabalho dos usuários, como:

- Necessidade de acesso a um amplo espectro de dados sobre medicamentos, incluindo informações internacionais e anteriores sobre medicamentos para estudos abrangentes.
- Exigir informações precisas e completas sobre medicamentos de vários mercados internacionais para melhor atendimento ao paciente e planejamento de tratamento.
- Companhias farmacêuticas que buscam dados extensos sobre medicamentos para análise e pesquisa do mercado global.

Para resolver essas lacunas, o sistema RxNorm Extension²¹ foi desenvolvido.

Esta solução inovadora foi concebida para expandir o âmbito do RxNorm, incorporando diferentes dados sobre medicamentos de diferentes países, ampliando significativamente a sua aplicabilidade para além do mercado dos EUA. Além disso, incorpora dados sobre medicamentos descontinuados, cruciais para estudos longitudinais e para a compreensão da evolução dos medicamentos e dos perfis de

²⁰ <https://www.nlm.nih.gov/research/umls/rxnorm/index.html>

²¹ https://www.ohdsi.org/web/wiki/doku.php?id=documentation:international_drugs

<https://medium.com/sciforce/rxnorm-extension-tool-to-standardize-source-drug-data-using-omop-cdm-71fd87eddaa2>

segurança, tornando o RxNorm Extension uma ferramenta abrangente para análise global de medicamentos.

5.5 Vocabulários OMOP/OHDSI

Os Vocabulários Padronizados OHDSI são uma coleção de vocabulários públicos consolidados na estrutura da tabela CDM. Este processo envolve a atribuição de identificadores estáveis a códigos individuais, tornando-os únicos em todo o sistema, adicionando atributos e estabelecendo relações para integrar os vocabulários em uma estrutura ontológica global. Além disso, a OHDSI cria seus próprios vocabulários e relacionamentos para referência interna e padronização semântica. O conteúdo de um banco de dados CDM OMOP é um conjunto de eventos clínicos representados por um código que identifica o quê aconteceu ao paciente e em que data. Essa representação é definida num sistema de codificação.

Com o tempo, os sistemas desenvolvidos para descrever os eventos clínicos expandiram-se enormemente em tamanho e complexidade e espalharam-se por outros aspectos dos cuidados de saúde, tais como procedimentos e serviços, medicamentos, dispositivos médicos, etc. Os princípios fundamentais permaneceram os mesmos: são vocabulários controlados, terminologias, hierarquias ou ontologias com as quais algumas comunidades de saúde concordam com a finalidade de capturar, classificar e analisar dados de pacientes. Muitos destes vocabulários são mantidos por agências públicas e governamentais com um mandato de longo prazo para o fazer. Por exemplo, a Organização Mundial da Saúde (OMS) produz a Classificação Internacional de Doenças (CID).

Como resultado, cada país, região, sistema de saúde e instituição tende a ter as suas próprias classificações que muito provavelmente só seriam relevantes onde fossem utilizadas. Esta miríade de vocabulários impede a interoperabilidade dos sistemas em que são utilizados. A padronização é a chave que permite a troca de dados dos pacientes, desbloqueia a análise de dados de saúde a nível global e permite a investigação sistemática e padronizada, incluindo a caracterização do desempenho e a avaliação da qualidade. Para resolver esse problema, surgiram organizações multinacionais que começaram a criar padrões amplos, como a OMS mencionada acima, a Nomenclatura Padrão de Medicina (SNOMED) e Nomes e Códigos de Identificadores de Observação Lógica (LOINC). Nos EUA, o Comitê de Padrões de TI em Saúde (HITAC) recomenda o uso de SNOMED, LOINC e o vocabulário de medicamentos RxNorm como padrões ao Coordenador Nacional de TI em Saúde (ONC), para uso em uma plataforma comum para troca nacional de informações de saúde entre diversas entidades .

Normalmente, encontrar e interpretar o conteúdo dos dados observacionais de saúde, sejam dados estruturados usando esquemas de codificação ou dispostos em texto

livre, é passado até o pesquisador, que se depara com uma infinidade de maneiras diferentes de descrever eventos clínicos.

A OHDSI exige harmonização não apenas com um formato padronizado, mas também com um conteúdo que segue um padrão rigoroso.

A OHDSI desenvolveu o OMOP CDM, um padrão global para pesquisa observacional. Como parte do CDM, os Vocabulários Padronizados OMOP estão disponíveis para dois propósitos principais:

- Repositório comum de todos os vocabulários usados na comunidade
- Padronização e mapeamento para uso em pesquisa

Os Vocabulários Padronizados estão disponíveis gratuitamente à comunidade e devem ser utilizados para a instância do OMOP CDM como sua tabela de referência obrigatória.

Todos os vocabulários dos Vocabulários Padronizados são consolidados no mesmo formato comum. Isso evita que os pesquisadores tenham que compreender e lidar com vários formatos e convenções de ciclo de vida diferentes dos vocabulários originários. Ele é construído e administrado pela Equipe de Vocabulário OHDSI, que faz parte do Grupo de Trabalho CDM geral da OMOP.

A descrição do que aconteceu num evento clínico dentro de um banco CDM é representada por um CONCEPT. As tabelas dos vocabulários expressam todos os possíveis CONCEPT que podem ser utilizados para descrever um evento clínico. Este vocabulário representa um arquivo de autoridade (Authority File) no sentido que todos os eventos clínicos presentes num banco CDM-OMOP precisam ser identificados com esta codificação.

5.5.1 Representação de Conteúdo através de Conceitos

Nas tabelas de dados do CDM o conteúdo de cada registro é totalmente normalizado e representado por meio de Conceitos (CONCEPT). Os conceitos são armazenados em tabelas de eventos com seus valores CONCEPT_ID, que são chaves estrangeiras para a tabela CONCEPT, que serve como tabela de referência geral. Todas as instâncias do CDM utilizam a mesma tabela CONCEPT como referência dos Conceitos, que juntamente com o Modelo Comum de Dados é um mecanismo chave de interoperabilidade e a base da rede de pesquisa OHDSI. Se um Conceito Padrão não existir ou não puder ser identificado, o valor do CONCEPT_ID é definido como 0, representando um conceito inexistente, um valor desconhecido ou não mapeável.

Os registros da tabela CONCEPT contêm informações detalhadas sobre cada conceito (nome, domínio, classe etc.). Conceitos, Relacionamentos de Conceitos, Ancestrais de Conceitos e outras informações relativas aos Conceitos estão contidas nas tabelas dos Vocabulários Padronizados. Existem apenas três tipos de conceitos, que são:

STANDARD CONCEPT - Conceito Padrão:

De todas as possíveis codificações disponíveis para codificar um conceito, apenas uma é escolhida como STANDARD (padrão) para representar o significado de cada evento clínico.

Por exemplo, o código MESH D001281, o código CIEL 148203, o código SNOMED 49436004, o código ICD9CM 427.31 e o READ CODE G573000 definem “fibrilação atrial” no domínio da condição, mas apenas o conceito SNOMED é padrão e representa a fibrilação atrial nas tabelas de eventos do paciente. Os demais são designados ou como conceitos NON-STANDARD ou como SOURCE-CONCEPT e mapeados para os Padrões. Podemos pensar neles como "Entry Points" ou "Entradas" dos vocabulários controlados, usadas para ajudar na procura de um termo. Os Conceitos Padrão são indicados através de um “S” no campo STANDARD_CONCEPT. E apenas esses STANDARD_CONCEPT são usados para registrar dados nos campos do CDM que terminam em "_CONCEPT_ID".

NON-STANDARD CONCEPT - Conceitos Fora do Padrão:

Conceitos não padronizados não são utilizados para representar os eventos clínicos, mas ainda fazem parte dos Vocabulários Padronizados e são frequentemente encontrados nos dados de origem. Por esse motivo, também são chamados de “conceitos fonte” (SOURCE CONCEPT). A conversão de conceitos fonte em Conceitos Padrão é um processo denominado “mapeamento”. Conceitos não padronizados não possuem valor no campo STANDARD_CONCEPT (são deixados como NULL).

CLASSIFICATION CONCEPT - Conceitos de Classificação:

Esses conceitos não são padrão e, portanto, não podem ser usados para representar os dados. Mas eles participam da hierarquia com os Conceitos Padrão e podem, portanto, ser usados para realizar consultas hierárquicas. Por exemplo, a consulta de todos os descendentes do código MedDRA 10037908 irá recuperar o conceito SNOMED Padrão para Fibrilhação Auricular.

A seguinte figura exemplifica os vocabulários envolvidos nos 3 tipos de conceitos:

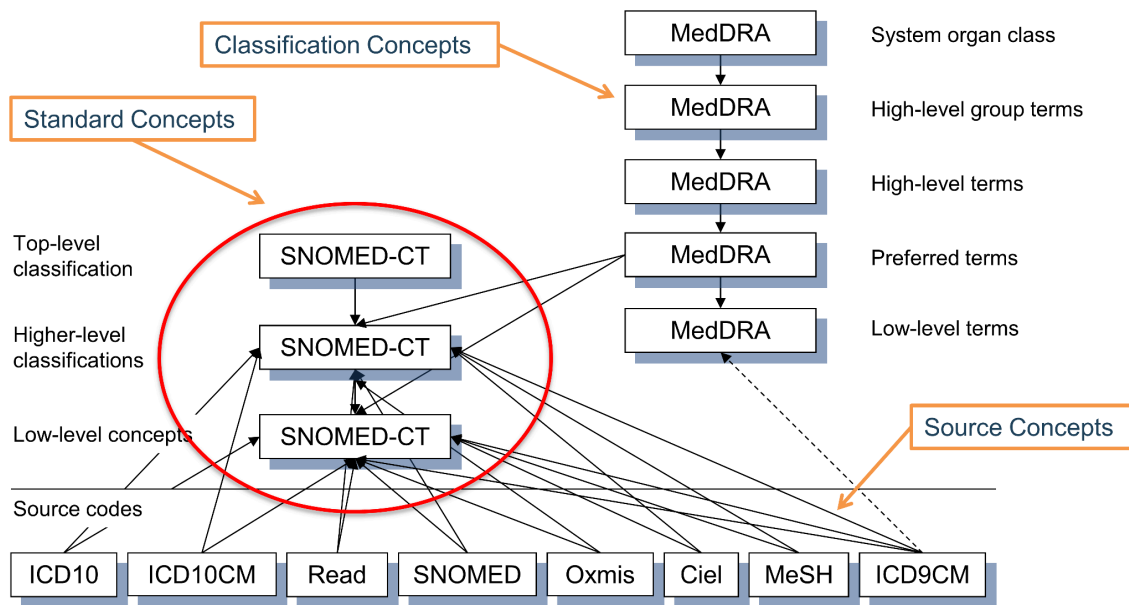


Figura 5.11 Exemplo dos tipos de conceitos

5.5.2 Estrutura

Nos Vocabulários Padronizados não há distinção por que uma informação não está disponível; pode ser devido a uma retirada ativa de informações pelo paciente, a um valor ausente, a um valor que não está definido ou padronizado de alguma forma ou à ausência de um registro de mapeamento em `CONCEPT_RELATIONSHIP`. Qualquer conceito deste tipo não está mapeado, o que corresponde por default, a uma mapeamento para o `STANDARD_CONCEPT` com o `ID_CONCEPT = 0`.

O conjunto dos códigos padrão (`STANDARD CODES`) forma um vocabulário controlado plano, em particular uma classificação estatística, dado que o `ID_CONCEPT = 0` representa a categoria residual (NEC ou NOS). Isso se deve a que para poder montar análises estatísticas precisamos garantir que nada é contado duas vezes.

A função dos Conceitos classificatórios (`CLASSIFICATION CONCEPTS`) é subsanar os defeitos de uma estrutura plana, auxiliando a definir hierarquias classificatórias dos termos. Porém, devemos ressaltar que nas tabelas de eventos de um banco CDM-OMOP apenas existem `STANDARD CODES`.

As hierarquias classificatórias são herdadas dos vocabulários que deram origem aos `STANDARD CODES`, como SNOMED e RxNorm.

5.5.3 Hierarquia

Dentro de um domínio, os conceitos padrão e de classificação são organizados em uma estrutura hierárquica e armazenados na tabela `CONCEPT_ANCESTOR`. Isso permite consultar e recuperar conceitos e todos os seus descendentes hierárquicos. Esses

descendentes possuem os mesmos atributos de seu ancestral, mas também atributos adicionais ou mais definidos.

A tabela `CONCEPT_ANCESTOR` é construída automaticamente a partir da tabela `CONCEPT_RELATIONSHIP` percorrendo todos os conceitos possíveis conectados através de relacionamentos hierárquicos. Estes são os pares “É um” - “Inclui” (ver Figura 5.6) e outros relacionamentos que conectam hierarquias entre vocabulários. A escolha se um relacionamento participa do construtor de hierarquia é definida para cada ID de relacionamento pelo sinalizador `DEFINES_ANCESTRY` na tabela de referência `RELATIONSHIP`. A Figura 5.12 apresenta estes relacionamentos.

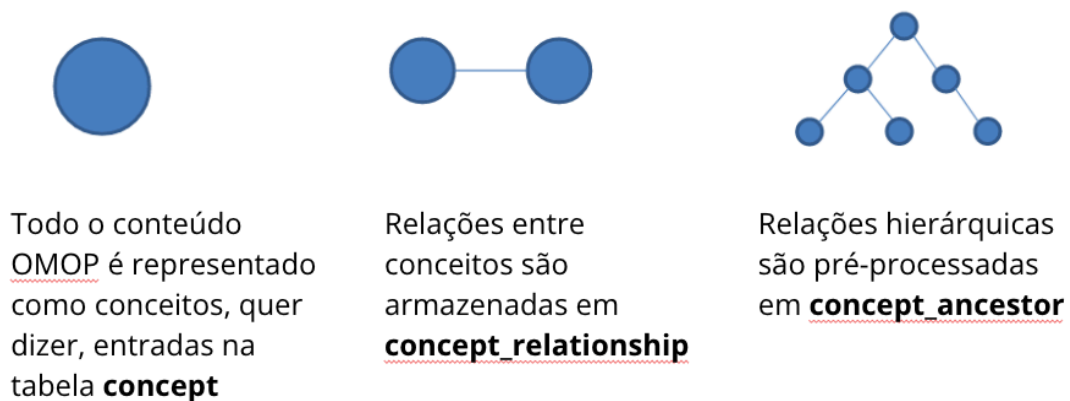


Figura 5.12 Representação dos conceitos e seus relacionamentos

A Figura 5.13 exemplifica os diversos relacionamentos que podem acontecer para o caso de "Fibrilação Atrial".

A ancestralidade de primeiro grau é definida através dos relacionamentos “É um” e “Inclui” (no sentido Paciente "é uma" Pessoa, Pessoa "inclui" Paciente) enquanto todas as relações de grau superior são inferidas e armazenadas na tabela `CONCEPT_ANCESTOR`. Cada conceito também é seu próprio descendente com ambos os níveis de separação iguais a 0.

O grau ancestral, ou número de passos entre ancestral e descendente, é capturado nos campos `MIN_LEVELS_OF_SEPARATION` e `MAX_LEVELS_OF_SEPARATION`, definindo a conexão mais curta ou mais longa possível. Nem todas as relações hierárquicas contribuem igualmente para o cálculo dos níveis de separação. Uma etapa contada para o grau é determinada pelo sinalizador `IS_HIERARCHICAL` na tabela de referência `RELACIONAMENTO` para cada ID de relacionamento.

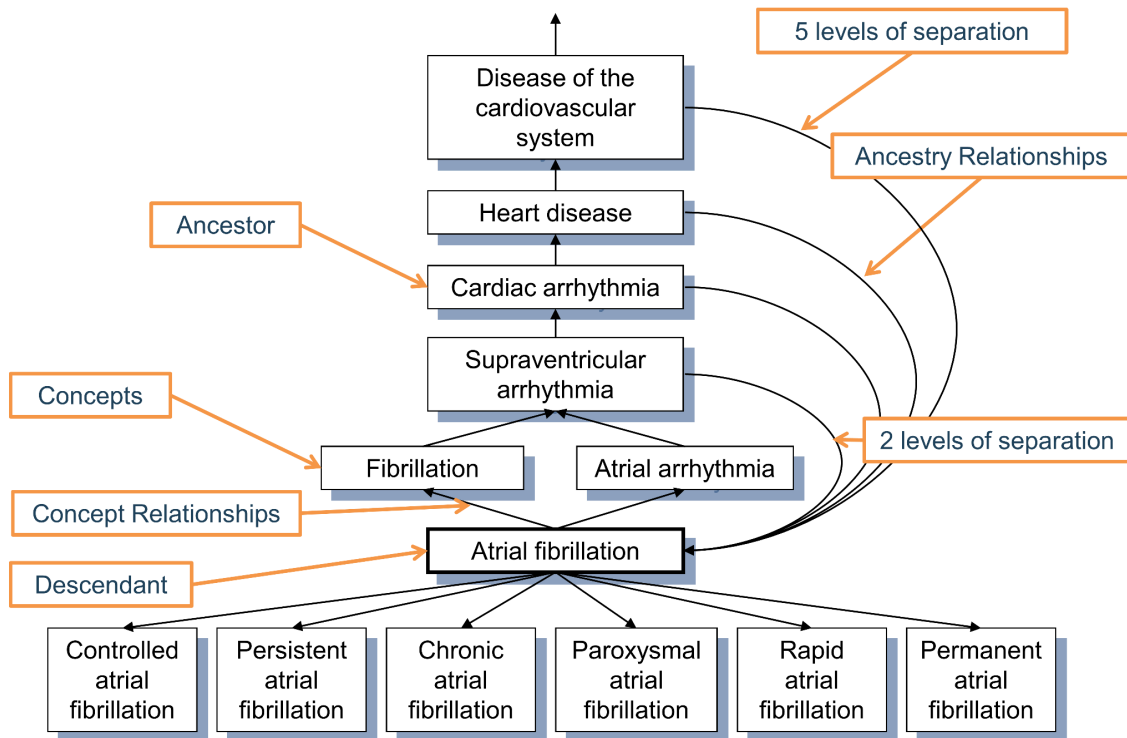


Figura 5.13 Exemplo de hierarquia da condição “Fibrilação atrial”

Neste momento, existe uma hierarquia abrangente de alta qualidade apenas para dois domínios: medicamentos (RxNorm) e doenças (SNOMED). Os domínios de procedimento, medição e observação estão apenas parcialmente cobertos e em processo de construção. A ancestralidade é particularmente útil para o domínio de medicamentos, pois permite navegar por todos os medicamentos com um determinado ingrediente ou membros de classes de medicamentos, independentemente do país de origem, marca ou outros atributos.

5.5.4 Domínios e conceitos

Eventos de diversas naturezas são organizados em domínios. Esses Eventos são armazenados em tabelas e campos específicos em cada domínio e são representados por conceitos padrão que também são específicos do domínio, conforme definido nos Vocabulários Padronizados. Cada Conceito Padrão tem uma atribuição de Domínio única, que define em qual tabela eles são registrados. Embora a atribuição correta de domínio seja objeto de debate na comunidade, esta regra estrita de correspondência domínio-tabela-campo garante que sempre haja um local inequívoco para qualquer código ou conceito. Por exemplo, sinais, sintomas e conceitos de diagnóstico são do Domínio de Diagnóstico (CONDITION), e são registrados no CONDITION_CONCEPT_ID da tabela CONDITION_OCCURRENCE.

Os Vocabulários Padronizados OHDSI são uma coleção de vocabulários públicos consolidados na estrutura da tabela CDM. Este processo envolve a atribuição de identificadores estáveis a códigos individuais, tornando-os únicos em todo o sistema, adicionando atributos e estabelecendo relações para integrar os vocabulários em uma estrutura ontológica global. Além disso, a OHDSI cria seus próprios vocabulários e relacionamentos para referência interna e padronização semântica.

Após essa preparação, os elementos individuais dos vocabulários são chamados de conceitos. Cada conceito possui um nome (descrição) e vários sinônimos, embora não haja uma tentativa de cobertura lexical abrangente para apoiar o processamento de linguagem natural ou a recuperação de informações. Todos os nomes dos conceitos estão em inglês, mas os sinônimos podem estar em qualquer idioma. Esses nomes e relacionamentos formam a estrutura da ontologia.

Os vocabulários padronizados são uma ontologia de referência comum obrigatória para todos os sites de dados na rede OHDSI. Compreende ontologias importadas e geradas de novo contendo conceitos e relacionamentos entre eles, e a prática de converter os dados de origem para o CDM OMOP com base neles. Permite a harmonização através de domínios atribuídos de acordo com categorias clínicas, cobertura abrangente de entidades dentro de cada domínio, suporte para esquemas de codificação internacionais comumente usados e padronização de conceitos semanticamente equivalentes.

Os Vocabulários Padronizados OMOP, são uma parte fundamental da rede de pesquisa OHDSI e parte integrante do Common Data Model (CDM). Eles permitem a padronização de métodos, definições e resultados definindo o conteúdo dos dados, para pesquisa e análise de rede. Normalmente, encontrar e interpretar o conteúdo dos dados observacionais de saúde, sejam dados estruturados usando esquemas de codificação ou dispostos em texto livre, é repassado ao pesquisador, que se depara com uma infinidade de maneiras diferentes de descrever eventos clínicos. A OHDSI exige a harmonização não apenas de um formato padronizado, mas também de um conteúdo padrão rigoroso.

Os Vocabulários estão em constante evolução e compreendem hoje mais de 11 milhões de conceitos de 142 vocabulários, distribuídos em 44 domínios. É atualizado regularmente para incluir novos termos, refletir mudanças na terminologia médica e atender às necessidades emergentes dos pesquisadores e profissionais de saúde em todo o mundo. Com cerca de 8.600 usuários, já foram realizados mais de 50.000 downloads do sistema. Essa abordagem dinâmica e colaborativa garante que a ontologia OHDSI continue sendo uma ferramenta eficaz para promover a harmonização internacional de dados de saúde²².

²² <https://www.ohdsi.org/wp-content/uploads/2023/11/OHDSI-Book2023.pdf>

5.5.5 Convenções das tabelas do vocabulário OMOP

As tabelas que compõem o vocabulários padrão, bem como as convenções para a criação dessas tabelas estão detalhadas em um repositório github²³.

Dez tabelas compõem os vocabulários, são elas:

Concepts (Conceitos): Os conceitos do Common Data Model são derivados de uma série de terminologias públicas ou proprietárias, como SNOMED-CT, LOINC e RxNorm, ou gerados de forma personalizada para padronizar aspectos dos dados observacionais. Ambos os tipos de Conceitos são integrados com base nas seguintes regras:

- Todos os conceitos são mantidos centralmente pelo Grupo de Trabalho CDM e Vocabulários. Conceitos adicionais podem ser adicionados, conforme necessário, mediante solicitação;
- Todos os Conceitos é atribuído um identificador numérico exclusivo `concept_id` que é usado como chave para vincular todos os dados observacionais aos dados de referência do Conceito correspondente;
- O `concept_id` de um Conceito é persistente, ou seja, permanece o mesmo para o mesmo Conceito entre versões dos Vocabulários Padronizados;
- Um nome descritivo é armazenado para cada Conceito. Nomes e descrições adicionais para o Conceito são armazenados como Sinônimos na tabela [CONCEPT_SYNONYM](#).
- Cada Conceito é atribuído a um Domínio e uma classe;
- O campo `concept_code` é usado para referenciar o vocabulário fonte;
- Os Conceitos Padrão (designados como 'S' no campo `standard_concept`) podem aparecer nas tabelas do CDM em todos os campos `*_concept_id`, enquanto os Conceitos de Classificação ('C') não devem aparecer nos dados do CDM, mas participar da construção da tabela `CONCEPT_ANCESTOR` e pode ser usado para identificar descendentes que podem aparecer nos dados.
- A vida útil de um Conceito é registrada através de seus campos `valid_start_date`, `valid_end_date` e `invalid_reason`;
- Os valores para `concept_ids` gerados como parte dos Vocabulários Padronizados serão reservados de 0 a 2.000.000.000. Acima desta faixa, `concept_ids` estão disponíveis para uso local e garantem que não entrarão em conflito com versões futuras dos Vocabulários Padronizados;

Vocabularies (vocabulários):

- Existe um registro para cada vocabulário. O campo `vocabulário_id` contém um identificador alfanumérico;
- O registro com `vocabulary_id = 'None'` é reservado para conter informações sobre a versão atual de todos os vocabulários padronizados;

²³ https://ohdsi.github.io/CommonDataModel/dataModelConventions.html#Data_Model_Conventions

- O campo `vocabulary_name` contém o nome oficial completo do Vocabulário, bem como a fonte ou fornecedor entre parênteses;
- Cada Vocabulário possui uma entrada na tabela `CONCEPT`, que é registrada no campo `vocabulário_concept_id`;

Domains (Domínios):

- Existe um registro para cada Domínio. Os domínios são definidos pelas tabelas e campos do CDM OMOP que podem conter conceitos que descrevem os aspectos de saúde de um paciente;
- O campo `domain_id` contém um identificador alfanumérico, que também pode ser utilizado como abreviatura do Domínio;
- O campo `domain_name` contém os nomes não abreviados do Domínio;
- Cada Domínio também possui uma entrada na tabela `Conceito`, que é registrada no campo `domain_concept_id`;

Concept Classes (Classes de Conceitos):

- Há um registro para cada classe conceitual. Classes de Conceito são usadas para criar estrutura adicional para os Conceitos dentro de cada vocabulário;
- O campo `concept_class_id` contém um identificador alfanumérico, que também pode ser utilizado como abreviatura da Classe Conceito;
- O campo `concept_class_name` contém os nomes não abreviados da classe conceitual.
- Cada Classe Conceito também possui uma entrada na tabela `Conceito`, que é registrada no campo `concept_class_concept_id`;

Concept Relationships (Relacionamento de Conceitos):

- Os relacionamentos geralmente podem ser classificados como hierárquicos (pai-filho) ou não hierárquicos (laterais);
- Todos os Relacionamentos são direcionais e cada Relacionamento é representado duas vezes simetricamente na tabela `CONCEPT_RELATIONSHIP`;
- Há um registro para cada relacionamento de conceito;
- Relacionamentos de conceito definem relacionamentos diretos entre conceitos;

Relationship Table (Tabela de Relacionamentos):

- Existe um registro para cada Relacionamento;
- Os relacionamentos são classificados como hierárquicos (pai-filho) ou não hierárquicos (laterais);
- Eles são usados para determinar quais registros de relacionamento de conceito devem ser incluídos no cálculo da tabela `CONCEPT_ANCESTOR`;
- O campo `relacionamento_id` contém um identificador alfanumérico, que também pode ser utilizado como abreviação do Relacionamento;
- O campo `Relationship_name` contém os nomes não abreviados do Relacionamento;

- Cada Relacionamento também possui uma entrada equivalente na tabela Conceito, que é registrada no campo relacionamento_conceito_id;
- Relacionamentos hierárquicos são usados para construir uma árvore hierárquica a partir dos conceitos, que é registrada na tabela [CONCEPT_ANCESTOR](#);
- As relações, também hierárquicas, podem ser entre Conceitos dentro de um mesmo Vocabulário ou aqueles adotados de diferentes fontes de Vocabulários;

Concept Synonyms (Sinônimos de Conceitos):

- O campo concept_synonym_name contém um sinônimo válido de um conceito, incluindo a descrição no próprio concept_name. Cada conceito possui pelo menos um Sinônimo na tabela CONCEPT_SYNONYM;
- Somente sinônimos ativos, atuais em inglês, são armazenados na tabela CONCEPT_SYNONYM;

Concept Ancestor (Conceito Ancestral):

- Cada conceito também é registrado como um ancestral de si mesmo.
- Somente conceitos válidos e padrão participam da tabela [CONCEPT_ANCESTOR](#);
- Normalmente, apenas Conceitos de um mesmo Domínio são conectados através de registros da tabela [CONCEPT_ANCESTOR](#), mas pode haver exceções.

Source to Concept Map (Mapa de Conceitos):

- Esta tabela não é mais usada para distribuir informações de mapeamento entre códigos-fonte e Conceitos Padrão para os Vocabulários Padrão. Em vez disso, a tabela CONCEPT_RELATIONSHIP é usada para esse propósito, usando o relacionamento_id='Maps to';
- No entanto, esta tabela ainda pode ser usada para a tradução de códigos-fonte locais em Conceitos Padrão;
- Os campos valid_start_date, valid_end_date e invalid_reason são utilizados para definir o ciclo de vida das informações de mapeamento;

Drug Strength (Força da droga):

- A tabela DRUG_STRENGTH contém informações para cada conceito de medicamento padrão ativo;
- Um medicamento que contém vários ingredientes ativos resultará em vários registros DRUG_STRENGTH, um para cada ingrediente ativo;
- As informações sobre a concentração do ingrediente são fornecidas como quantidade absoluta;
- Todos os vocabulários de Medicamentos contendo Conceitos Padrão possuem entradas na tabela DRUG_STRENGTH;

5.6 Athena: consulta, seleção e download de vocabulários

Athena²⁴ é um aplicativo de software de código-fonte aberto, baseado na web, desenvolvido pela comunidade OHDSI, gratuito e disponível publicamente. Permite a pesquisa e navegação pelos códigos que compõem os diferentes vocabulários com o objetivo de selecionar aqueles que melhor representem os conteúdos da nossa base de origem. No site selecionamos o conjunto dos vocabulários para download. Esta seleção será usada junto ao CDM OMOP e na ferramenta USAGI para mapeamento dos códigos dos vocabulários locais para o vocabulário padrão. A Figura 5.14 apresenta a interface da ferramenta. No primeiro acesso, fazer cadastro e login (simples e rápido com um clique na opção mais à direita do aplicativo).

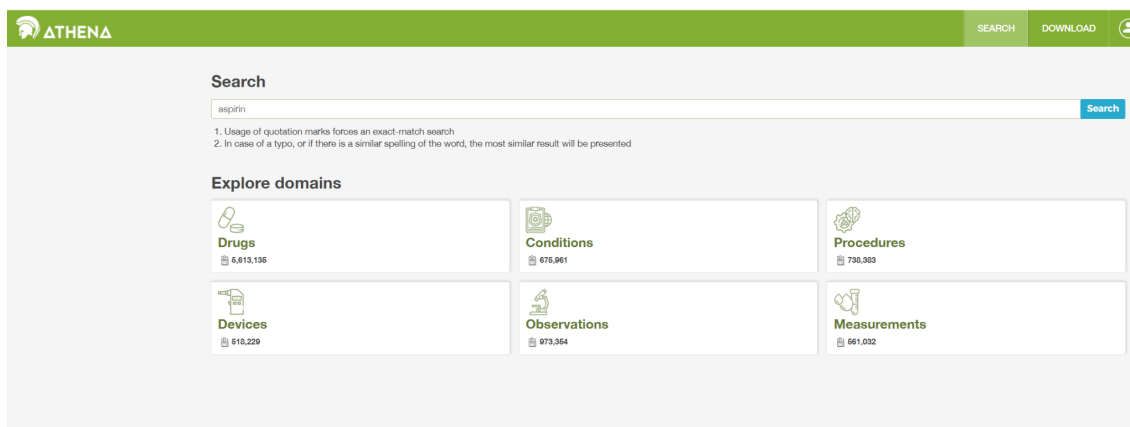


Figura 5.14 Interface do site do ATHENA

A seguir, um breve descritivo das funcionalidades da ferramenta:

A opção SEARCH, permite navegar nos:

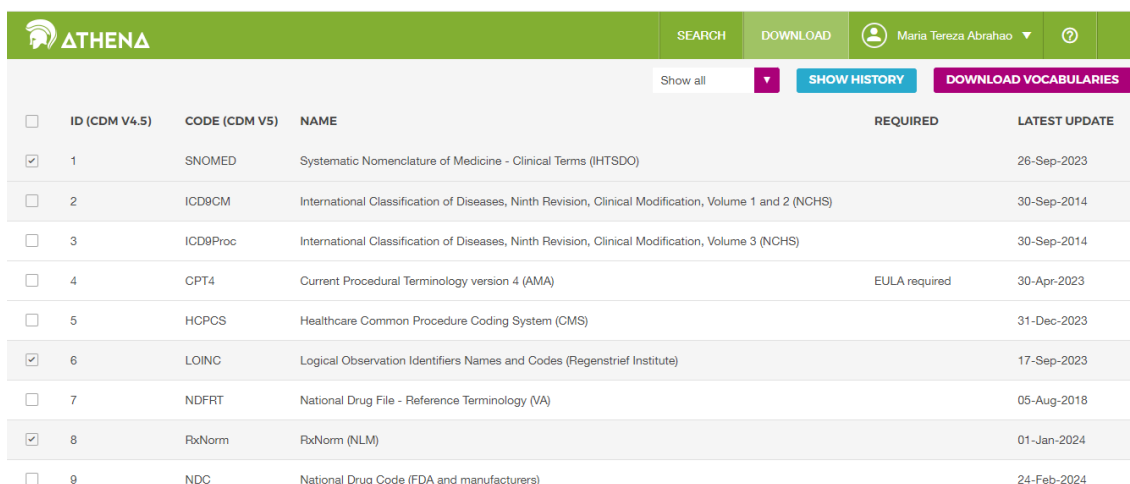
- **Domínios**, cada conceito é atribuído a um domínio, que agrupam códigos dos vocabulários, tipo: condições, procedimentos, visitas, dispositivos, medidas, etc. Os domínios também direcionam para qual tabela e campo do CDM OMOP um evento clínico ou atributo de evento deve ser registrado;
- **Conceitos**, representam o significado de cada evento clínico e se subdividem em:
 - Padrão, apenas os conceitos SNOMED são padrão e representam a condição nos dados. Os demais são designados como conceitos não padronizados ou de origem e mapeados para os padrões. Os Conceitos Padrão são indicados por um “S” no campo STANDARD_CONCEPT. E somente esses Conceitos Padrão são usados para registrar dados nos campos do CDM OMOP;

²⁴ <https://athena.ohdsi.org/search-terms/start>

- Não padrão, os conceitos não padronizados não são usados para representar os eventos clínicos, mas ainda fazem parte dos Vocabulários Padronizados e são frequentemente encontrados nos dados de origem. Por essa razão, eles também são chamados de “conceitos de origem”. A conversão de conceitos de origem em Conceitos Padrão é um processo chamado “mapeamento”;
- Classificação, esses conceitos não são padrão e, portanto, não podem ser usados para representar os dados. Mas eles estão participando da hierarquia com os Conceitos Padrão e, portanto, podem ser usados para realizar consultas hierárquicas
- **Classes**, alguns vocabulários classificam seus códigos ou conceitos univocamente. Por exemplo, o SNOMED tem 33 dessas classes de conceito, que o SNOMED chama de “tags semânticas”: achado clínico, contexto social, estrutura corporal, etc. Estas são divisões verticais dos conceitos. Outros, como MedDRA ou RxNorm, possuem classes de conceito que classificam níveis horizontais em suas hierarquias estratificadas;
- **Vocabulários**, existem 111 vocabulários atualmente suportados pelo OHDSI, dos quais 78 são adotados de fontes externas, enquanto o restante são vocabulários internos do OMOP. Esses vocabulários são normalmente atualizados trimestralmente. A fonte e a versão dos vocabulários são definidas no arquivo de referência VOCABULARY.
- **Validade**, se o conceito ainda está válido ou não para um vocabulário.

Os vocabulários necessários para o CDM OMOP:

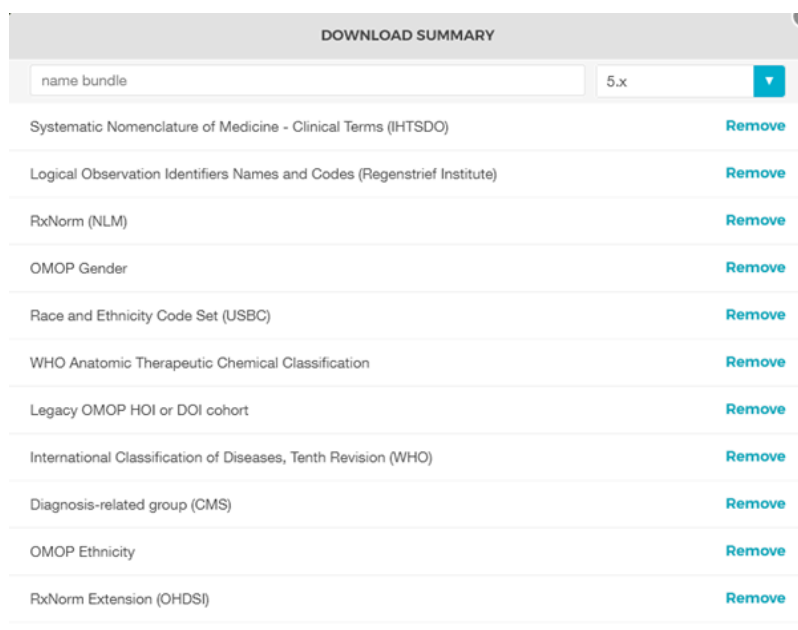
Use a opção DOWNLOAD e selecione todos os vocabulários necessários para o CDM OMOP de uso local. Vocabulários com conceitos padrão e uso muito comum são pré-selecionados. Adicione vocabulários que são usados em seus dados de origem. Os vocabulários proprietários não têm botão de seleção. Clique no botão “Licença necessária” para incorporar tal vocabulário em sua lista. O time OHDSI responsável pelos vocabulários entrará em contato e solicitará que justifique a necessidade da sua licença ou ajudará a se conectar com as pessoas certas para obter uma. A Figura 5.15 apresenta o detalhamento de uma seleção de vocabulários.



ID (CDM V4.5)	CODE (CDM V5)	NAME	REQUIRED	LATEST UPDATE
<input checked="" type="checkbox"/>	1	SNOMED	Systematic Nomenclature of Medicine - Clinical Terms (IHTSDO)	26-Sep-2023
<input type="checkbox"/>	2	ICD9CM	International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 1 and 2 (NCHS)	30-Sep-2014
<input type="checkbox"/>	3	ICD9Proc	International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 3 (NCHS)	30-Sep-2014
<input type="checkbox"/>	4	CPT4	Current Procedural Terminology version 4 (AMA)	EULA required 30-Apr-2023
<input type="checkbox"/>	5	HCPCS	Healthcare Common Procedure Coding System (CMS)	31-Dec-2023
<input checked="" type="checkbox"/>	6	LOINC	Logical Observation Identifiers Names and Codes (Regenstrief Institute)	17-Sep-2023
<input type="checkbox"/>	7	NDFRT	National Drug File - Reference Terminology (VA)	05-Aug-2018
<input checked="" type="checkbox"/>	8	RxNorm	RxNorm (NLM)	01-Jan-2024
<input type="checkbox"/>	9	NDC	National Drug Code (FDA and manufacturers)	24-Feb-2024

Figura 5.15 Rol de vocabulários selecionados

Ao terminar a seleção, a ferramenta apresenta um sumário dos vocabulários que foram selecionados e solicita um nome para o arquivo. Confirmando, dá-se início ao processo de geração das tabelas (pode demorar horas). No término, é enviado um e-mail de aviso e o arquivo estará pronto. A Figura 5.16 ilustra um exemplo do sumário.



DOWNLOAD SUMMARY	
name bundle	5.x
Systematic Nomenclature of Medicine - Clinical Terms (IHTSDO)	Remove
Logical Observation Identifiers Names and Codes (Regenstrief Institute)	Remove
RxNorm (NLM)	Remove
OMOP Gender	Remove
Race and Ethnicity Code Set (USBC)	Remove
WHO Anatomic Therapeutic Chemical Classification	Remove
Legacy OMOP HOI or DOI cohort	Remove
International Classification of Diseases, Tenth Revision (WHO)	Remove
Diagnosis-related group (CMS)	Remove
OMOP Ethnicity	Remove
RxNorm Extension (OHDSI)	Remove

Figura 5.16 Sumário dos vocabulários selecionados

Baixe o arquivo zip com todas as tabelas de vocabulários padronizados e carregue em seu banco CDM local e na ferramenta USAGI.

A continuação, segue um breve resumo:

- Todos os eventos e fatos administrativos são representados nos Vocabulários Padronizados do OMOP como conceitos, relacionamentos de conceito e hierarquia de conceitos ancestral
- A maioria é adotada a partir de esquemas de codificação ou vocabulários existentes, enquanto alguns deles são selecionados de novo pela equipe de vocabulário OHDSI
- Todos os conceitos são atribuídos a um domínio, que controla onde o fato representado pelo conceito é armazenado no CDM OMOP
- Conceitos de significado equivalente em diferentes vocabulários são mapeados para um deles, que é designado por Conceito Padrão. Os outros são conceitos de origem.
- O mapeamento é feito através das relações conceituais “Maps to” e “Maps to value”
- Há uma classe adicional de conceitos chamados conceitos de classificação, que não são padronizados, mas, em contraste com os conceitos de origem, eles participam da hierarquia
- Os conceitos têm um ciclo de vida ao longo do tempo.
- Os conceitos dentro de um domínio são organizados em hierarquias. A qualidade da hierarquia difere entre os domínios, e a conclusão do sistema de hierarquia é uma tarefa contínua.

5.7 Usagi: Mapeando vocabulários locais para o Padrão

Usagi [Schuemie 2023] é uma ferramenta de mapeamento utilizada no contexto da terminologia médica, especialmente em projetos que envolvem a implementação de padrões como SNOMED CT. Desenvolvido pela *International Health Terminology Standards Development Organisation (IHTSDO)*, o Usagi auxilia na criação de mapas entre terminologias médicas, facilitando a integração e interoperabilidade entre sistemas de saúde que utilizam diferentes conjuntos de códigos ou terminologias.

O Usagi automatiza parte do processo de mapeamento, sugerindo correspondências entre conceitos de diferentes terminologias com base em suas características semânticas e hierárquicas, garantindo uma integração mais eficiente e precisa entre os sistemas de informação de saúde.

Usando as funções do aplicativo, para importar códigos-fonte para Usagi, primeiro exporte os códigos-fonte do sistema de origem para um arquivo CSV ou Excel (.xlsx). Esse arquivo deve ter pelo menos as colunas contendo o código-fonte e uma descrição do código-fonte em inglês, porém informações adicionais sobre os códigos também podem ser trazidas (descrição no idioma original). Além das informações sobre os códigos-fonte, a frequência do código deve ser preferencialmente trazida, pois isso pode ajudar a priorizar quais códigos devem receber mais esforço no

mapeamento. Se alguma informação do código-fonte precisa ser traduzida para o inglês, o Google Translate pode ser usado.

Os extratos do código-fonte devem ser divididos por domínios (isto é, medicamentos, procedimentos, condições, observações) e não agrupados em um arquivo grande. Os códigos-fonte são carregados no Usagi no menu Arquivo -> Importar códigos. A partir daqui, será exibido “Códigos de importação...”, conforme visto na Figura 5.17.

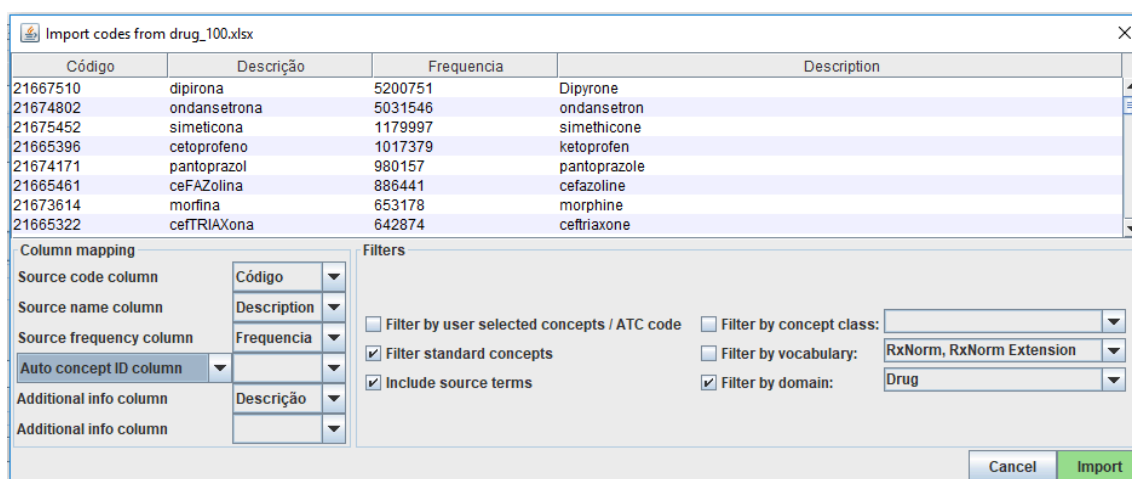


Figura 5.17 Tela de entrada dos códigos-fonte no Usagi

A figura apresenta os códigos (Code), os códigos-fonte que foram traduzidos para o inglês (English term), a frequência e os termos locais que foram adicionados. Usagi irá alavancar as traduções em inglês para mapear para o vocabulário padrão.

Pode-se definir algumas restrições para o Usagi ao mapear. Por padrão, o Usagi mapeia apenas para os conceitos padrão, mas se a opção 'Filtrar conceitos padrão' estiver desligada, o Usagi também irá considerar os conceitos de classificação. Assim que todas as suas configurações forem finalizadas, clique no botão “Importar” para importar o arquivo. A importação do arquivo levará alguns minutos, pois está executando o algoritmo de similaridade para mapear os códigos-fonte para os códigos padrão. Ao concluir a importação dá-se início ao processo de mapeamento. O Usagi é composto por 3 seções principais: uma tabela que permite a visão geral do mapeamento, a seção de mapeamento que foi selecionado e o local para realizar as pesquisas e alterar o mapeamento, se necessário. A Figura 5.18 ilustra essas seções.

The screenshot displays the Usagi v1.4.3 interface with three highlighted sections:

- Tabela de visão geral:** A table listing source terms and their mappings. Columns include Status, Source code, Source term, Frequency, Descrição, Match score, Concept ID, Concept name, Domain, Concept class, Vocabulary, Concept code, Standard concept, Parents, Children, Assigned To, Equivalence, Comment, and Status Pts. The first row shows a match for 'ampicillin' with a score of 1.00.
- Mapeamento selecionado:** A detailed view of a selected mapping for 'ampicillin' (Concept ID: 1717327) mapped to 'AMPicilina' (Concept ID: 21554335). It shows the source term, frequency, and description.
- Facilidades de busca:** A search interface with a query input field, filters (e.g., 'Filter by user selected concepts / ATC code'), and a results table. The results table lists various 'ampicillin' related terms with their scores and concept IDs.

Figura 5.18 Tela de correspondência do Usagi

Logo após a importação dos códigos-fonte, são apresentados os mapeamentos sugeridos gerados automaticamente com base na similaridade de termos e nas opções selecionadas. O Usagi compara as descrições do código-fonte com nomes de conceitos e sinônimos para encontrar a melhor correspondência. Se for selecionado 'Incluir termos de origem', o Usagi considera os nomes e sinônimos de todos os conceitos de origem no vocabulário que mapeiam para um conceito específico. Se o Usagi não puder fazer um mapeamento, ele será mapeado para `CONCEPT_ID = 0`.

É necessário experiência em sistemas de codificação para mapear os códigos-fonte para seu vocabulário padrão associado. Código por código na aba de visão geral deverá ser visto para aceitar o mapeamento que Usagi sugeriu ou escolher um novo mapeamento.

É possível adicionar comentários aos mapeamentos, que podem ser usados para documentar o porquê uma determinada decisão de mapeamento foi feita.

Melhores Práticas:

- O mapeamento sempre deve ser feito por alguém que tenha experiência com esquemas de codificação.
- Ao clicar no nome de uma coluna, você pode classificar as colunas na aba de visão geral. Pode ser valioso classificar por pontuação (Match score); revisar os códigos nos quais Usagi tem mais confiança primeiro pode eliminar rapidamente um pedaço significativo de códigos. Também é valioso classificar por “Frequência”, gastar mais esforço em códigos frequentes do que em códigos não frequentes.

- É normal mapear alguns códigos para `CONCEPT_ID = 0`, alguns códigos podem não valer a pena encontrar um bom mapa e outros podem apenas não ter um mapa adequado.
- É importante considerar o contexto de um conceito, especificamente seus ascendentes e descendentes.

Depois de criar um mapa no USAGI, para sua utilização é preciso exportá-lo e anexá-lo à tabela `SOURCE_TO_CONCEPT_MAP` do Vocabulário OMOP.

Requisitos de instalação:

- Requer Java 1.8 ou superior <http://www.java.com>
- Obtenha a versão mais recente do vocabulário em Athena <http://athena.ohdsi.org/>
 - É necessária uma conta
 - Quando seu pacote de vocabulário estiver pronto, baixe-o e descompacte
- Instalação Usagi <http://ohdsi.github.io/Usagi/installation.html>
- Versão mais recente <https://github.com/OHDSI/Usagi/releases/tag/v1.4.3>
- Após o download, o Usagi pode ser iniciado simplesmente clicando duas vezes no arquivo jar

Instalação e suporte: Todo o código-fonte e instruções de instalação estão disponíveis no site GitHub da Usagi: <https://github.com/OHDSI/Usagi>

Configuração única:

Na primeira vez que você iniciar o Usagi, será solicitado que indexe os vocabulários. O Usagi não vem com os índices, você deve fornecer os arquivos de vocabulário. Para fazer isso, siga as seguintes etapas:

1. Obtenha a versão mais recente do vocabulário de [Athena](#). É necessária uma conta. Você pode selecionar qualquer vocabulário que precisar.
2. Quando seu pacote de vocabulário estiver pronto, baixe-o e descompacte o pacote.
3. Acesse o Usagi conforme descrito acima.
4. Quando solicitado, especifique a localização dos arquivos de vocabulário baixados para criar o índice. A criação do índice de vocabulário é um processo caro do ponto de vista computacional e pode levar horas para ser concluído.
5. Ao terminar, a versão do vocabulário deve ser exibida no canto inferior direito do Usagi. As estatísticas do seu índice podem ser visualizadas em Ajuda -> Mostrar estatísticas do índice

5.8 Atlas: consulta, seleção e grupos de conceitos de vocabulários

A OHDSI oferece uma ampla variedade de ferramentas de código aberto para dar suporte a vários casos de uso de análise em dados observacionais no nível do paciente, que permitem interagir com um ou mais bancos de dados usando o modelo comum de dados, o CDM OMOP.

Existem três abordagens principais para a implementação de um estudo. A primeira é escrever código personalizado que não faça uso de nenhuma das ferramentas que a OHDSI tem a oferecer. A segunda abordagem envolve desenvolver a análise em R e fazer uso dos pacotes da [Biblioteca de Métodos OHDSI](#). A terceira abordagem baseia-se na plataforma de análise interativa [ATLAS](#), uma ferramenta que permite que não programadores realizem uma ampla variedade de análises com eficiência. O ATLAS faz uso das Bibliotecas de Métodos, mas fornece uma interface gráfica simples para projetar e executar análises.

O ATLAS é uma ferramenta gratuita, publicamente disponível e baseada na web, desenvolvida pela comunidade OHDSI que facilita o projeto e execução de análises em dados observacionais em nível de paciente, padronizados no modelo CDM OMOP. É a ferramenta usada para fenotipagem baseada em regras ou definição de coorte baseada em regras.

O ATLAS é implantado como um aplicativo da web em combinação com o OHDSI WebAPI. Para o desempenho de análises em tempo real requerer acesso a dados de nível de paciente no CDM. Para executar o Atlas, você deve estar atrás do Firewall de sua instituição.

É a ferramenta principal para a preparação das coortes e montagem dos estudos. Nela temos tópicos que nos permitem escolher os termos que vão formar, junto com outros critérios, a definição da nossa coorte (seleção dos pacientes que participam do estudo). Podemos navegar entre os conceitos, escolher os relacionamentos hierárquicos para incluir ou excluir determinados códigos, agrupar conceitos, exportá-los e importá-los.

À esquerda da ferramenta, temos a barra de navegação mostrando as várias funções fornecidas pelo ATLAS, sendo:

Fontes de dados: Fornecem relatórios padronizados e descritivos de caracterização das fontes de dados configuradas na plataforma Atlas. Esse recurso usa a estratégia de análise em larga escala, pré-computados ao término do mapeamento dos dados para o CDM OMOP.

Pesquisa de vocabulário: Para explorar o vocabulário padronizado OMOP para entender quais conceitos estão disponíveis e como aplicá-los em suas análises.

Conjuntos de Conceitos: Permite criar coleções de expressões lógicas que serão usados nas análises padronizadas. Um conjunto de conceitos é composto por vários conceitos do vocabulário padronizado em combinação com indicadores lógicos que permitem ao usuário incluir ou excluir conceitos relacionados na hierarquia do

vocabulário. Portanto, pesquisar o vocabulário, identificar os conceitos e especificar a lógica a ser usada para resolver um conjunto de conceitos, fornece um mecanismo poderoso para definir a linguagem médica usada nos planos de análise.

Definições de coorte: Para selecionar um conjunto de pessoas que satisfaçam um ou mais critérios por um período e são a base de entrada para todas as análises subsequentes.

Caracterizações: É um recurso analítico que permite examinar uma ou mais coortes e resumir características sobre essas populações de pacientes. Visa, através do uso de estatísticas descritivas, gerar hipóteses sobre os determinantes da saúde e da doença e para compreender os resultados clínicos de grupos específicos da população.

Caminhos de coorte: Possibilita observar a sequência de eventos clínicos que ocorrem em uma ou mais populações.

Taxas de incidência: Para estimar a incidência de desfechos em populações-alvo de interesse.

Perfis: Permite explorar dados observacionais longitudinais de pacientes para resumir o que está acontecendo com um determinado indivíduo.

Estimativa: Analisa os efeitos causais das exposições (por exemplo, intervenções médicas, tais como exposições a medicamentos ou procedimentos) sobre resultados de saúde específicos de interesse (por exemplo, a segurança ou eficácia de medicamentos ou outros tratamentos).

Predição: Realiza a previsão de resultados de saúde futuros a partir de dados existentes ao nível do paciente, a partir da aplicação de algoritmos de aprendizado de máquina, para apoiar a tomada de decisões clínicas, a avaliação de riscos e a validação de tais modelos de previsão.

Jobs: Essa opção é usada para acompanhar o estado dos processos que estão sendo executados. Os trabalhos geralmente são processos de longa execução, como gerar uma coorte ou computar relatórios de caracterização de coorte.

Configuração e Feedback: Ferramentas utilizadas pelo administrador do site.

A Figura 5.19 apresenta a interface da ferramenta Atlas.

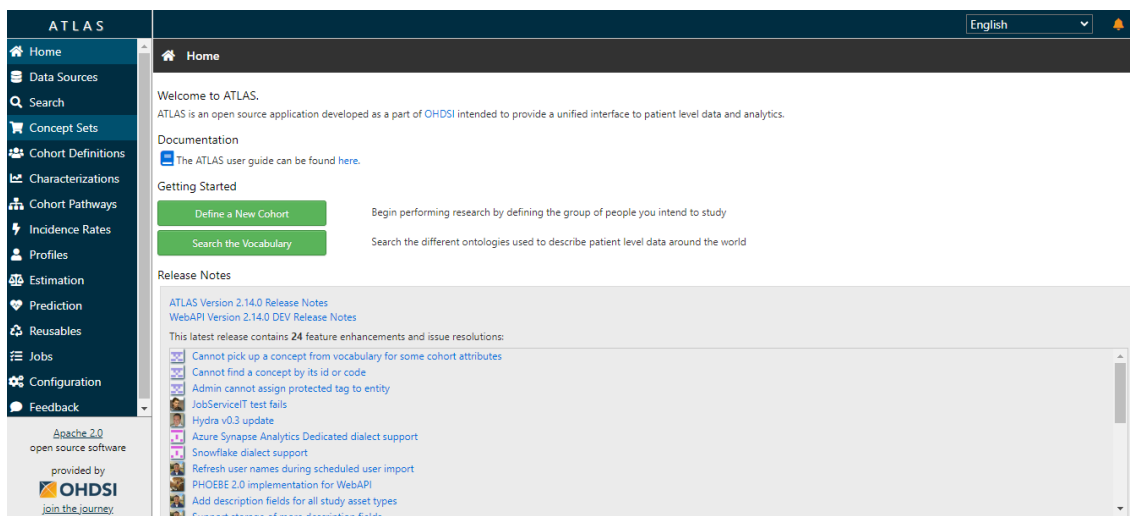


Figura 5.19 Interface web do Atlas

5.8.1 Conjunto de Conceitos

Os Conjuntos de Conceitos no contexto do CDM OMOP, são ferramentas essenciais para criar coleções de expressões lógicas utilizadas em análises padronizadas. Eles permitem a inclusão e exclusão de conceitos relacionados, baseando-se na hierarquia do vocabulário padronizado.

Funcionalidade e Utilidade:

1. Pesquisa do Vocabulário: Permite que os usuários pesquisem no vocabulário padronizado OMOP para encontrar os conceitos relevantes.
2. Identificação de Conceitos: Uma vez encontrados, os conceitos são identificados e selecionados para formar o conjunto de conceitos.
3. Especificação de Lógica: Os usuários podem definir a lógica de inclusão e exclusão, determinando como os conceitos relacionados serão tratados com base em suas relações hierárquicas.

Poder e Flexibilidade:

- Definição Precisa: Ao usar indicadores lógicos, como "incluir" ou "excluir", os usuários podem definir de maneira precisa e flexível os conjuntos de conceitos que serão aplicados em suas análises.
- Hierarquia do Vocabulário: A hierarquia do vocabulário OMOP facilita a inclusão ou exclusão de conceitos relacionados, garantindo que todas as variações relevantes de um termo sejam consideradas ou descartadas conforme necessário.
- Planos de Análise: Fornece um mecanismo poderoso para definir a linguagem médica usada nos planos de análise, permitindo uma padronização e consistência nas pesquisas.

Exemplo de Uso:

- Definição de Coorte: Um pesquisador pode criar um conjunto de conceitos para definir uma coorte de pacientes com diabetes, incluindo conceitos específicos de diabetes tipo 1 e tipo 2, enquanto exclui conceitos relacionados a outras formas de diabetes.
- Construção de Covariáveis: Pode ser usado para construir covariáveis específicas para análises, como incluir apenas medicamentos específicos dentro de uma classe terapêutica.

Benefícios:

- Eficácia na Pesquisa: Facilita a fenotipagem eficiente e a construção de covariáveis, melhorando a qualidade e a consistência dos dados utilizados em pesquisas observacionais.
- Interoperabilidade: A utilização de um vocabulário padronizado garante que diferentes grupos de pesquisa possam colaborar e compartilhar resultados de maneira eficiente.

Os Conjuntos de Conceitos OMOP são, portanto, uma ferramenta fundamental para pesquisadores que buscam gerar evidências robustas e confiáveis a partir de dados observacionais de saúde.

A Figura 5.20 apresenta o Atlas na aba de definições de conceitos, um exemplo de um conjunto de conceitos para COVID 19.

Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants
<input type="checkbox"/>	37311061	COVID-19	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	439676	Coronavirus infection	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/>	605554	Acute COVID-19	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Figura 5.20 Concept set para diagnóstico de COVID 19

5.8.2 Coortes na ferramenta Atlas

As coortes são usadas (e reutilizadas) nas ferramentas de análise OHDSI para definir, por exemplo, as exposições e os resultados de interesse. A estratégia para construir uma coorte dependerá do rigor clínico de como seu consenso de especialistas define a doença. Isso quer dizer que o design de coorte certo dependerá da pergunta que está tentando responder. Pode-se optar por criar uma definição de coorte que use tudo o que

puder obter, use o menor denominador comum para compartilhá-la com outras instituições. Em última análise, fica a critério do pesquisador qual limite de rigor é necessário para estudar adequadamente a coorte de interesse.

Uma definição de coorte é uma tentativa de inferir algo que gostaríamos de observar a partir dos dados registrados. Em geral, a validação de uma definição de coorte baseada em regras ou algoritmo probabilístico, pode ser pensada como um teste da coorte proposta em comparação com alguma forma de referência “padrão ouro” (por exemplo, revisão manual de gráficos de casos).

As coortes parametrizadas vão ser geradas e depois visualizadas em relatórios pré formatados no Atlas. Toda a definição, conceitos e parametrização de uma coorte pode ser exportada no formato JSON para ser importada por outra entidade que faça uso do CDM OMOP. Essa coorte pode ser gerada e analisada em parceiros de estudos e seus resultados podem ser agrupados e analisados em conjunto, sem compartilhamento de dados, somente dos scripts das análises.

A Figura 5.21 apresenta a parametrização de uma coorte para COVID 19 na ferramenta Atlas. A Figura 5.22 mostra um template para exportação dos parâmetros e definições da coorte.

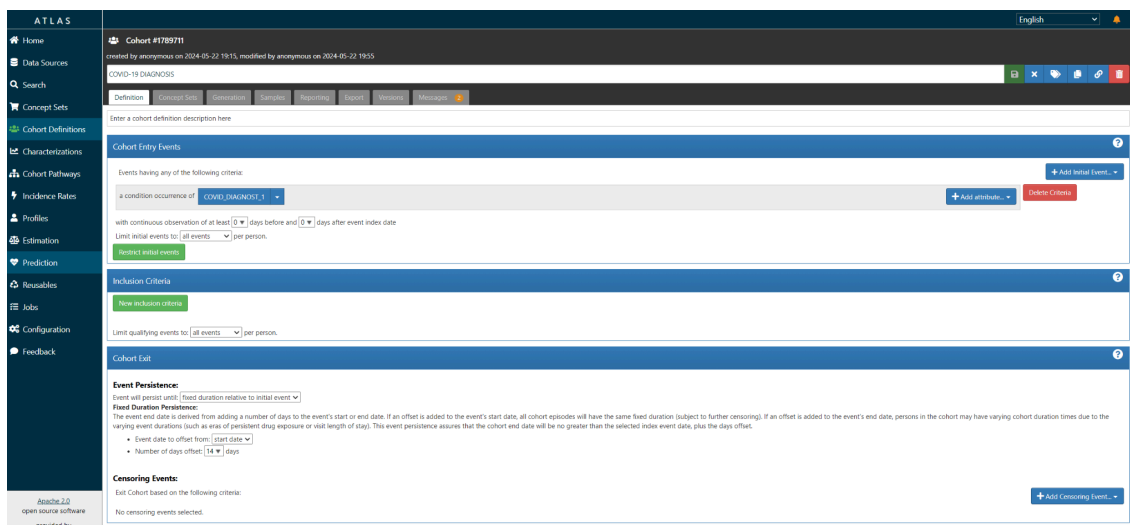


Figura 5.21 Exemplo de uma coorte de COVID 19

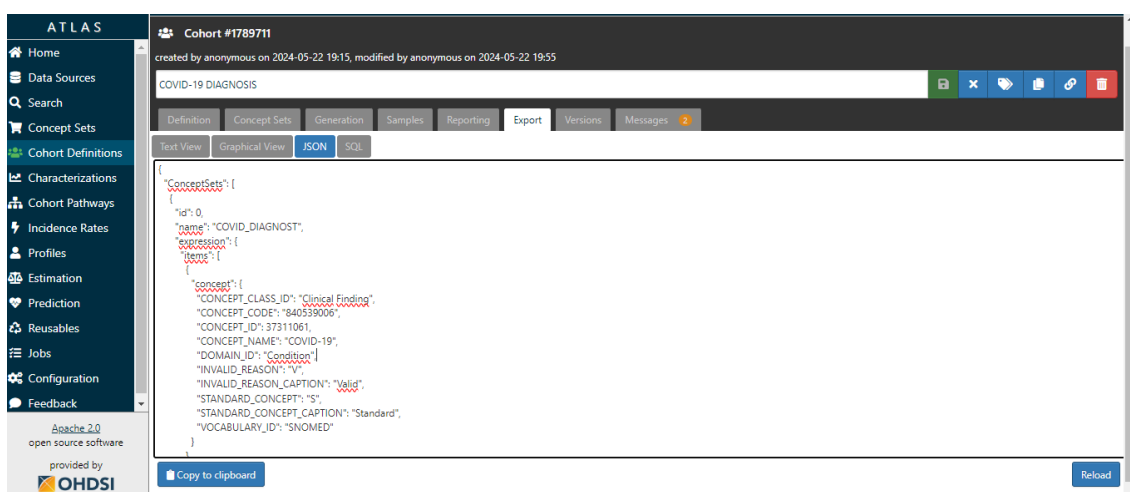


Figura 5.22 Exportação do JSON de uma coorte de COVID 19

Construir uma coorte é o bloco mais importante para responder a uma pergunta de pesquisa nas ferramentas OHDSI. As coortes são a base para realização de estudos de caracterização de populações, estimativas a nível populacional (*Population-Level Estimation - PLE*)²⁵ e predições a nível de paciente (*Patient-Level Prediction - PLP*)²⁶ [Abrahão 2022].

5.9 Considerações finais e conclusões

O vocabulário da OHDSI, definido como o conjunto dos códigos padrão (standard codes), cumpre um papel fundamental em limitar as variações possíveis na definição da seleção de pacientes que compõem o estudo (coorte) limitando o conjunto de termos e códigos plausíveis de serem utilizados nela. Com isso, cumpre a função de "*authority file*" dando consistência às análises estatísticas feitas em cima da seleção de pacientes definida pelo fenótipo.

A sua estrutura é de um vocabulário controlado plano, em particular uma classificação estatística mono-hierárquica e exaustiva. É uma lista de códigos e termos, dado que os termos de classificação (*classification codes*) não fazem parte da seleção da coorte, sendo apenas utilizados para pesquisa dos termos padrão pela ferramentas (ATLAS, ATHENA, etc).

O motivo disto é a necessidade de compatibilizar diversos sistemas de codificação de diferentes domínios (diagnósticos, drogas, procedimentos, etc), alguns dos quais ontologias (SNOMED), outras classificações (CID) ou tesouros (RxNorm), muitos dos quais se sobrepõem, complementam ou contradizem e evitar contagens duplicadas.

²⁵ <https://ohdsi.github.io/TheBookOfOhdsi/PopulationLevelEstimation.html>

²⁶ <https://ohdsi.github.io/TheBookOfOhdsi/PatientLevelPrediction.html>

Os vocabulários são a pedra angular desta padronização. Compreender a sua estrutura e vivenciar o seu uso em situações práticas permite ganhar critérios de discernimento e uma visão crítica na análise de outras ferramentas e de resultados de estudos.

Quais bases estamos comparando? A seleção de termos é condizente com o seu uso no conteúdo das bases? A comparação está semanticamente correta? Houve vieses de seleção pela diferença entre os vocabulários que descrevem os conteúdos das bases usadas no estudo? Estas são perguntas que devem ser sempre feitas ao analisar o resultado de qualquer estudo observacional.

Para ser possível mapear os diversos sistemas de registro de informações médicas com o objetivo de realizar comparações válidas entre fontes de dados diversas, foi necessário sacrificar a expressividade de conhecimento do vocabulário e escolher o mínimo denominador comum.

Como consequência, o trabalho da resolução das ambiguidades, compatibilização das terminologias e mapeamento dos diversos vocabulários entre si, foi deixada na mão de dois agentes: o processo de tradução e mapeamento do vocabulário local do sistema de registro eletrônico para o dicionário padrão durante a carga do banco CDM OMOP, e do próprio pesquisador durante a montagem do fenótipo, quer dizer, da definição dos critérios que vão selecionar a coorte do seu estudo.

Por isso, o mapeamento de vocabulário local para os termos padrão é o mais complexo de todo o processo de criação e carga de um banco CDM OMOP.

Não podemos deixar de mencionar as diversas iniciativas que estão em andamento para melhorar esta situação, compreendendo entre elas, a definição de fenótipos através de um processo probabilístico [Banda 2017] e o mapeamento do dicionário CDM-OMOP para estruturas ontológicas [Callahan, T J et al 2023] que suportem esquemas de inferências e raciocínio lógico na montagem dos fenótipos.

As decisões adotadas no sentido de abranger a maior quantidade possível de codificações e terminologias com uma estrutura plana pode parecer limitante, porém junto com um modelo de dados comum e análises estatísticas padronizadas, está permitindo pela primeira vez, fazer comparações entre fontes de dados diversas com segurança e confiabilidade sem compartilhar dados brutos.

Através dos exemplos do CID, MeSH e outros, vemos a importância de manter compatibilidade e estabilidade para que análises feitas através do tempo se mantenham válidas e consigam mostrar tendências que ultrapassam o período de vida dos sistemas de informação que as geraram. A definição dos vocabulários da OHDSI faz parte essencial deste processo.

A padronização do conteúdo das informações médicas é uma tarefa árdua e trabalhosa, porém, absolutamente vital para poder comparar resultados de estudos realizados em fontes diversas e obter evidências reprodutíveis e confiáveis do mundo real.

Referências bibliográficas

- Abrahão, M T., Nobre M R C. e Madril, P J. (2019) "O estado da arte em pesquisa observacional de dados de saúde: A iniciativa OHDSI". In: Artur Ziviani; Natalia Castro Fernandes; Débora Christina Muchaluat Saade. (Org.). Livro de Minicursos do 19o Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2019). 19o ed. Porto Alegre: Sociedade Brasileira de Computação-SBC, (2019), I SBN-13 (15) 978-85-7669-472-4, v. 1, p. 141-189.
- Abrahão, M T e Madril, P J. (2022) "Fenótipos no contexto da pesquisa observacional: OHDSI Phenotype February 2022". In: Sociedade Brasileira de Computação. SBC 2022. (Org.). Minicursos do XXII Simpósio Brasileiro de Computação Aplicada à Saúde. 22 ed. Porto Alegre: SBC, 2022, v. 6, p. 219-261.
- Banda, Juan., Halpern, Yoni., Sontag, David and Shah, Nigam. (2017) "Electronic phenotyping with APHRODITE and the Observational Health Sciences and Informatics (OHDSI) data network". AMIA Summits on Translational Science Proceedings. 2017. 48-57.
- Callahan, T J et al. (2023) "Ontologizing Health Systems Data at Scale: Making Translational Discovery a Reality", <https://arxiv.org/abs/2307.05727v2>
- Cornet, R, and C G Chute. (2016) "Health Concept and Knowledge Management: Twenty-five Years of Evolution." Yearbook of medical informatics vol. Suppl 1, Suppl 1 S32-41. 2 Aug. 2016, doi: 10.15265/IYS-2016-s037 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5171511/>
- Cimino, J J. (1998) "Desiderata for controlled medical vocabularies in the twenty-first century." Methods of information in medicine vol. 37,4-5: 394-403. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3415631/>
- DeMars, M M and Perruso C. (2022) "MeSH and text-word search strategies: precision, recall, and their implications for library instruction". J Med Libr Assoc. 2022 Jan 1;110(1):23-33. doi: 10.5195/jmla.2022.1283. PMID: 35210959; PMCID: PMC8830400.
- Ferreira, Miguel and Baptista, Ana Alice. (2005) "The use of taxonomies as a way to achieve interoperability and improved resource discovery in DSpace-based repositories". <http://repositorium.sdum.uminho.pt/handle/1822/873>

Gruber, Thomas R. (1993) "A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition", 5(2):199-220, 1993
<https://tomgruber.org/writing/ontolingua-kaj-1993.pdf>

Gruber, Thomas R. (1995). "Toward principles for the design of ontologies used for knowledge sharing". Originally in N. Guarino and R. Poli, (Eds.), International Workshop on Formal Ontology, Padova, Italy. Revised August 1993. Published in International Journal of Human-Computer Studies, Volume 43 , Issue 5-6 Nov./Dec. 1995, Pages: 907-928, special issue on the role of formal ontology in information technology. <https://tomgruber.org/writing/onto-design.pdf>

Heather Hedden. (2022) "The Accidental Taxonomist", Third Edition, Information Today, Inc. Medford, N.J. (November 8, 2022), ISBN: 978-157387-586-8

Karl Fast, Fred Leise and Mike Steckel. M. (2002) "What Is a Controlled Vocabulary?"
http://web.archive.org/web/20030811115443/http://www.boxesandarrows.com/archives/what_is_a_controlled_vocabulary.php

Library of Congress (2004) "Library of Congress Subject Headings (LCSH)" 27th edition. Washington (DC). Five volumes. Hardbound. Published annually. ISSN 1048-9711

Garshol L. M. (2004) "Metadata? Thesauri? Taxonomies? Topic Maps!" Ontopia.

Garshol L. M. (1986) "Guidelines for the establishment and development of monolingual thesauri", International Organization for Standardization (ISO)

Lomax, Jane and Wolf, Elizabeth. (2021) "The Evolution and Importance of Biomedical Ontologies for Scientific Literature"
<https://www.copyright.com/wp-content/uploads/2021/02/White-Paper-Evolution-Importance-of-Biomedical-Ontologies.pdf>

Madsen, B. N. and Erdman Thomsen, H. (2009) "Ontologies vs. Classification Systems". NEALT (Northern European Association of Language Technology) Proceedings Series, (4), 27-32.
http://dspace.utlib.ee/dspace/bitstream/10062/9840/1/13-4-Final-Ontologies_vs_classification_systems_NODALIDA-2009.pdf

Mazzocchi, Fulvio. (2018) "Knowledge organization system (KOS)". Knowledge Organization 45, no.1: 54-78. Also available in ISKO Encyclopedia of

Knowledge Organization, eds. Birger Hjørland and Claudio Gnoli,
<https://www.isko.org/cyclo/kos>

Mucheroni, Marcos and Modesto, Fernando. (2011) “A interoperabilidade dos sistemas de informação sob o enfoque da análise sintática e semântica de dados na web” DOI: [10.9771/1981-6766rpa.v5i1.3622](https://doi.org/10.9771/1981-6766rpa.v5i1.3622).

National Information Standards Organization. (2005) “Guidelines for the construction, format, and management of monolingual controlled vocabularies” (ANSI/NISO Z39.19-2005). Bethesda, MD: NISO Press.
<https://groups.niso.org/higherlogic/ws/public/download/12591/z39-19-2005r2010.pdf>

Pagel, M. "Q&A. What is human language, when did it evolve and why should we care?" BMC Biol 15, 64 (2017). <https://doi.org/10.1186/s12915-017-0405-3>
<https://bmcbiol.biomedcentral.com/articles/10.1186/s12915-017-0405-3#citeas>

Schuemie M. (2023) “Usagi” <https://github.com/OHDSI/Usagi>

Souza, Renato Rocha, Douglas Tudhope and Mauricio B. Almeida (2012). "Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems". Knowledge Organization 39, no. 3: 179–192.

Stocks, K.I., Neiswender, C., Isenor, A.W., Graybeal, J., Galbraith, N., Montgomery, E.T., Alexander, P., Watson, S., Bermudez, L., Gale, A., Hogrefe, K. (2010) "The MMI Guides: Navigating the World of Marine Metadata", https://uop.who.edu/techdocs/presentations/MMI_Guides.pdf

Chapter

6

Applications of Artificial Intelligence to Support the Diagnosis, Treatment, and Prognosis of Mental Disorders

Paulo Mann (UERJ), Elton H. Matsushima (UFF), Aline Paes (UFF)

Abstract

This theoretical short course aims to present the main challenges and trends in Artificial Intelligence applications to support the diagnosis, treatment, and prognosis of mental disorders. First, we approach the fundamental concepts of mental disorders that affect the general population the most — depressive disorders and anxiety disorders, with particular emphasis on the former. With this, we hope to provide greater visibility and knowledge to computing professionals about mental disorders' psychiatric and psychological aspects. Additionally, we will detail specific AI applications and techniques that support these disorders' diagnosis, treatment, and prognosis. Finally, we will discuss the main research challenges in this area, such as LGPD regulation and the ethical implications of automated systems that handle sensitive data. By the end of the short course, participants are expected to (i) understand the fundamentals of depressive disorders and anxiety disorders, (ii) know the main AI techniques and model architectures used by applications, (iii) understand the main methodologies of AI applications to deal with both disorders at different stages, (iv) be aware of the research trends in this area, and (v) comprehend the main ethical and legislative challenges that permeate the research and application of automated models to support the mental health field.

Resumo

O objetivo deste minicurso teórico é apresentar os principais desafios e tendências das aplicações de Inteligência Artificial para apoiar o diagnóstico, tratamento, e o prognóstico dos transtornos mentais. Em primeiro lugar, serão abordados os principais conceitos fundamentais dos transtornos mentais que mais afetam a população - os transtornos depressivos e os transtornos de ansiedade, com particular ênfase para o primeiro. Com isso, esperamos dar maior visibilidade e conhecimento para profissionais de computação

sobre aspectos psiquiátricos e psicológicos dos transtornos mentais. Ademais, iremos detalhar aplicações e técnicas específicas de IA que atuam no diagnóstico, tratamento e prognóstico destes transtornos. Por fim, trataremos dos principais desafios de pesquisa nessa área, como a regulamentação da LGPD e das implicações éticas de sistemas automatizados que lidam com dados sensíveis. Espera-se que ao final do minicurso os participantes sejam capazes de (i) conhecer os fundamentos sobre transtornos depressivos e transtornos de ansiedade, (ii) conhecer as principais técnicas e arquiteturas de modelos de IA mais utilizados pelas aplicações, (iii) conhecer as principais metodologias de aplicações de IA para lidar com ambos os transtornos em diferentes fases, (iv) conhecer as tendências de pesquisa nesta área, e (v) compreender os principais desafios éticos e legislativos que permeiam a pesquisa e aplicação de modelos automatizados para apoiar a área de saúde mental.

6.1. Introduction

Major Depressive Disorder (MDD) is the leading mental health disorder worldwide [Kupferberg et al., 2016; WHO, 2017], contributing significantly to global disability by driving dysfunctional behaviors that impair both social and professional functioning [Greenberg et al., 2021]. The socioeconomic impact of MDD is substantial, with costs in the United States alone exceeding \$300 billion, primarily due to workplace-related expenses [Greenberg et al., 2021]. Despite the stability in the number of US adults receiving treatment over the past decade, the prevalence of depression has been on the rise, indicating that many individuals remain untreated and continue to suffer from depressive symptoms [Greenberg et al., 2021].

Several barriers prevent individuals from seeking treatment for MDD. These include fear of social stigma, limited knowledge about mental health, and financial constraints. While educational initiatives can address the first two issues, and improved public health services can alleviate the third, a significant number of individuals remain undiagnosed and untreated. To address this, there are a few possible solutions: (1) implementing effective screening mechanisms; (2) creating effective methods for supporting the treatment; (3) providing supportive interventions, such as educational programs and access to psychological and psychiatric services. Effective screening can identify a maximum number of individuals with depression and guide them toward appropriate assistance, while effective methods for supporting treatment will help through episodes of mental disorder. Supporting educational programs equates to providing a better prognosis both in terms of the individual and communities as a whole.

In addition to depression, anxiety disorders are another major category of mental health issues that often co-occur with depression [Aina and Susman, 2006], exacerbating the overall burden on individuals and society. Anxiety disorders, like MDD, contribute to significant functional impairment and can further complicate the course and treatment of depression. Therefore, any comprehensive mental health initiative must also account for the prevalence and impact of anxiety disorders.

One way to support the diagnosis is to automatically screen individuals with depression and anxiety to raise awareness and knowledge about these disorders. Automated screening can help with early identification and intervention, improving mental health

literacy. Mental health literacy encompasses understanding how to achieve and maintain positive mental health, recognizing mental disorders and their treatments, reducing stigma, and enhancing help-seeking efficacy [Kutcher et al., 2016]. According to the World Health Organization (WHO), health literacy, defined as "the ability to gain access to, understand, and use information in ways that promote and maintain good health," is a crucial predictor of health quality [Jorm et al., 1997].

In line with this rationale, we focus on Artificial Intelligence (AI) applications that help not only in screening individuals with depression or anxiety but also support the therapeutic and prognostic phases as well. As AI models rely on data to be trained, we give particular emphasis to social media data from Social Media Platforms (SMPs) like Twitter, Reddit, Weibo, and Instagram. This approach allows for non-intrusive screening that does not disrupt individuals' daily lives.

Furthermore, the integration of supportive interventions following the screening process is essential. These interventions include educational programs designed to increase mental health literacy, reduce stigma, and provide information on where and how to seek help. Access to psychological and psychiatric services is another critical component, ensuring that individuals identified through screening receive the necessary professional care.

With that, we expect to show applications and trends that aim to identify individuals with MDD and anxiety disorders but also strive to foster a broader understanding and acceptance of mental health issues. By leveraging social media data and advanced ML techniques, we show, through examples, that it is possible to create a robust system that addresses both the screening and support aspects of mental health care. The diverse array of applications has the potential to reach a broad audience, offering a proactive solution to the growing mental health crisis.

This chapter is organized as follows. In Section 6.2, we present the formal definition of Major Depressive Disorder and Anxiety Disorder. In Section 6.3, we demonstrate the main applications for screening mental health disorders using social media data. In Section 6.4, we explore the main ethical implications of applications that automatically screen for mental health disorders; to do that, we create an analytical framework. Finally, we conclude in Section 6.5.

6.2. Major Depressive Disorder and Anxiety Disorder

Major Depressive Disorder (MDD) is a Mental Disorder defined as a persistent feeling of sadness and loss of interest, which negatively affects how you feel, the way you think and how you act, leading to a several emotional and physical problems that disrupt your ability to function in almost every context [Association et al., 2013]. To be diagnosed with a mental disorder, the behavior should reflect a severe dysfunction in the individual's cognition, emotions, and functioning, which ultimately causes one to suffer.

Specifically, MDD has nine associated symptoms: (1) depressed mood; (2) loss of interest or pleasure; (3) significant weight loss or gain; (4) insomnia or hypersomnia; (5) psychomotor agitation or retardation; (6) fatigue or loss of energy; (7) feelings of worthlessness; (8) impaired concentration, indecisiveness; (9) recurring thoughts of death

or suicide [Association et al., 2013]. To say that an individual is suffering from MDD, five or more of these symptoms have to be present nearly every day during a 2-week period and represent a change from previous functioning. Additionally, one of the five symptoms has to be at least the symptom (1) or (2) [Association et al., 2013]. The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning and are not attributable to the physiological effects of a substance or another medical condition [Association et al., 2013]. These criteria represent a Major Depressive Episode (MDE) with varying degrees of severity, ranging from one episode to a recurrent number of episodes and remission status that provide the diagnostic for the MDD¹.

On the other hand, Anxiety Disorders² are a group of mental disorders characterized by significant and excessive fear or anxiety that interferes with daily activities [Association et al., 2013]. While fear is related to perceived imminent threat, anxiety is related to the anticipation of future threat [Association et al., 2013]. These disorders are marked by persistent, intense, and often irrational worry that affects how individuals feel, think, and behave. It leads to various emotional and physical symptoms that can severely disrupt their ability to function in different contexts [Association et al., 2013]. According to the DSM-5, an anxiety disorder diagnosis requires that the behavior reflects severe dysfunction in cognition, emotions, and overall functioning, causing significant distress and impairment.

Specifically, Generalized Anxiety Disorder (GAD), one of the most common anxiety disorders, has several associated symptoms: (1) excessive anxiety and worry occurring more days than not for at least six months about several events or activities; (2) difficulty controlling the worry; (3) anxiety and worry associated with three or more of the following symptoms: restlessness or feeling keyed up or on edge, being easily fatigued, difficulty concentrating or mind going blank, irritability, muscle tension, and sleep disturbance (difficulty falling or staying asleep, or restless, unsatisfying sleep) [Association et al., 2013]. To diagnose GAD, anxiety and worry should be present for more days than not for at least six months and should be about several events or activities.

Additionally, the symptoms must cause clinically significant distress or impairment in social, occupational, or other important areas of functioning. They must not be attributable to the physiological effects of a substance (e.g., a drug of abuse, a medication) or another medical condition [Association et al., 2013]. These criteria define Generalized Anxiety Disorder, but similar criteria are used for diagnosing other anxiety disorders, such as Panic Disorder, Social Anxiety Disorder, and Specific Phobias, each with its specific symptom profile and duration requirements.

Anxiety disorders are often chronic and can fluctuate in severity, with periods of increased symptoms and times of relative calm. The persistence and recurrence of anxiety can severely impact an individual's quality of life, requiring comprehensive treatment

¹Depression is an overloaded term that is often used for short periods of distress or mourning. For this paper, we interchangeably use the terms “depression”, and “MDD” to refer to the Major Depressive Disorder.

²We refer simply as “anxiety” to the whole group of disorders characterized as an Anxiety Disorder, such as Generalized Anxiety Disorder (GAD), Agoraphobia, Panic Disorder, Social Anxiety Disorder, Specific Phobia (the most common Anxiety Disorder in the U.S), etc.

approaches, including psychotherapy, medication, and lifestyle modifications to manage symptoms effectively.

However, the symptoms of Anxiety and Depressive disorders may vary under different sociocultural norms, as one behavior is acceptable or encouraged by one society. Albeit there are well-defined criteria for diagnosing a mental disorder, the spectrum of sociocultural norms should always be considered for a proper diagnosis. For example, a sample of Japanese individuals who are not depressed might behave similarly to a sample of depressed Brazilian individuals. As culture evolves, clinicians should be prepared to differentiate normative behavior from an impairing symptom criterion for diagnosis. Significantly, the cultural etiquette formed in online social media is an organism that evolves even faster.

Moreover, screening depression and anxiety using social media cues becomes challenging as online language and identity are fluid over time. Social media constantly adopts new memes, terms, inside jokes, and new hashtags that provide a new way to interact, but at the same time, it is also a social space used by users with a mental disorder — or not. Individuals with depression and anxiety disorders use tools that do change over time — and frequently — which inevitably will make the manifestation of their symptoms take different forms: a new hashtag, a newly coined term; although different, it is still a manifestation of a depressive or an anxiety symptom.

Accordingly, it has become increasingly difficult for any human to keep up with the constant flow of new data, especially with social media; this is no different for clinicians. To cope with this fast pace and vast amounts of data, machine learning-aided diagnosis, usually referred to as high-performance medicine [Topol, 2019], is a tool that can help healthcare professionals in the diagnosis. In this way, automated methods can help analyze from a single piece of information, to the general holistic view. Thus, ML trained models on a specific sample can be swiftly adapted to new individuals and new social media terms.

Finally, the traditional way to screen depressed individuals is through a clinical interview. However, this is often costly and error-prone, and it requires an active role of the individual to look for help. To help general practitioners, psychometric tests can provide a second opinion on the intensity of depressive or anxiety symptoms. There are two most widely used psychometric tests in the literature for identifying the intensity of depressive symptoms: Beck's Depression Inventory (BDI) [Beck et al., 1996] and the Center for Epidemiologic Studies Depression Scale (CES-D) [Radloff, 1977].

The Beck Depression Inventory-II (BDI-II) is a self-report questionnaire that measures the severity of depression by evaluating all of its symptoms. It has 21 items responded in a four-alternative scale; each answer yields a score between 0–3, which indicates the severity of the symptom the question is evaluating. The final BDI-II score is the sum of all 21 questions answers, which might yield a score in the range 0–63. The final score is classified into four distinct categories: 0–13 is minimal; 14–19 is mild; 20–28 is moderate; and 29–63 is severe depression [Gorenstein et al., 2011].

The Generalized Anxiety Disorder 7-item (GAD-7) scale is a self-report questionnaire designed to measure the severity of generalized anxiety disorder (GAD) by as-

sessing its symptoms over the past 2 weeks. It consists of 7 items, each describing a symptom associated with GAD. Respondents rate how often they have been bothered by each symptom over the past two weeks on a 4-point scale, where 0 means “not at all,” 1 means “several days,” 2 means “more than half the days,” and 3 means “nearly every day.” The total GAD-7 score is calculated by summing the scores of all 7 items, resulting in a range of 0 to 21. The final score is classified into four distinct categories to indicate the severity of anxiety, where 0–4 is minimal anxiety, 5–9 is mild anxiety, 10–14 is moderate anxiety, and 15–21 is severe anxiety.

The GAD-7 scale is widely used in both clinical and research settings due to its brevity and effectiveness in screening for generalized anxiety disorder. It helps healthcare providers quickly identify the level of anxiety and make informed decisions about further evaluation and treatment options. The scale is also helpful in monitoring changes in anxiety symptoms over time, making it a valuable tool for both initial assessment and ongoing management of anxiety disorders.

6.3. Artificial Intelligence for Mental Disorders

In this section, we present a systematic literature review to summarize the main applications of AI for mental health disorders. The work we present here has been collected from the following digital libraries: Web of Science, IEEE, ACM Digital Library, and SCOPUS. When applicable, if the search yielded more than 300 studies, we get the 300 most relevant works. We sort the result by the number of citations, the relevance of the publishing medium, and the publication date — recent publications with higher priority.

Finally, the study selection process occurs as follows: (1) we first submit the search string to the mentioned repositories, which yielded 635 items across all repositories, including publications from journals and conferences. We saved and managed the references using the Mendeley Reference Manager³ tool; (2) next, the authors read all paper’s titles, keywords, and abstracts, removing the studies that met at least one of the exclusion criteria. This phase resulted in 28 articles; (3) next, the author read the introduction and conclusion of all remaining papers and removed those that meet the exclusion criteria, thus resulting in 15 articles, of which 4 are surveys. The remaining papers were completely read and did not fit any exclusion criteria.

Furthermore, we also applied the snowballing technique [Jalali and Wohlin, 2012] to find relevant references of a study (backward snowballing) or relevant works that mention the final selection of 23 articles (forward snowballing). We do not include the four surveys [Skaik and Inkpen, 2020; Ríssola et al., 2021; Dhelim et al., 2023; Mathur et al., 2023] in the list of selected works, demonstrated in Table 6.1; they are, instead, used as seeds for the forward and backward snowballing method, and as a reference to structure this section.

This section is organized as follows: we begin with Section 6.3.1 by exploring the aspects of data, such as which social media the data was collected and how digital data footprints can leave signs of mental health expressions. In Section 6.3.2, we explore different methods for obtaining data, such as asking users for permission — explicitly or

³<http://mendeley.com>

implicitly — to collect their data. In Section 6.3.3, we show different methods for extracting the features from different modalities of data and the main methods for classification.

6.3.1. Data

In the broad field of psychotherapy, understanding the nuances of how different methodologies process various input data can illuminate their practical applications in a therapeutic context. Different approaches to psychotherapy process the input data differently — for input data, consider the voice tone, for example. The Cognitive Behavioral Therapy (CBT) often relies on the theory of behavior and cognition to find and intervene in thought processes, behavior, and habits that originate dysfunctional ways of living [Rice, 2015]. For example, journaling is one of the cognitive techniques used in CBT. Taking notes in a stressful situation might reveal reinforcement mechanisms of dysfunctional outcomes [Rice, 2015], which helps both the individual suffering and the therapist with enhanced information. Cognitive restructuring is another technique to help individuals identify, evaluate and modify the faulty thoughts responsible for their psychological disturbance [Clark, 2013]. In that sense, the analyst collaborates with the patient to improve their quality of life. The input received by the analyst is processed in methodological ways to produce the desired outcome.

Psychoanalysis, on the other hand, was initially focused on the unconscious and free association. The psychoanalyst encourages patients to talk freely about anything. In the famous case of obsessional neurosis as written by Sigmund Freud, the “Rat man” said “[...] I used to have a morbid idea that my parents knew my thoughts [...] There were certain people, girls, who pleased me very much, and I had a very strong wish to see them naked. But wishing this I had an uncanny feeling, as though something must happen if I thought such things [...] that my father might die.” [Freud, 1909]. After conducting several sessions, Freud concluded that the “Rat Man”, as a child, was scolded by his father for some situation related to masturbation. Since then, he has created not only trauma from masturbation but also a terrible grudge against his father, which has motivated his thoughts of death against his father. Freud freely associated the “Rat Man”’s discourse with his creativity and the theory of unconsciousness to conclude that certain events in his childhood led to developing his obsession with his father’s death. Although one might argue that this methodology seems vague and opaque, it illustrates the inherently different methods of different psychotherapies and how they process input data.

Not too distant, nonetheless, is the data generated through social media. While the “environment” and objective of talk therapy are inherently different from social media, both share a common aspect: behavior footprint. While in talk therapy, patients, intentionally or not, let several clues about their mental state through the spoken content or facial expressions and posture. This behavior footprint is the raw data for therapy. While on social media, users leave digital footprints behind. Although there is room for fictional online identities, a body of evidence shows how the language and behavior of individuals with mental health issues using social media differ from control groups [De Choudhury et al., 2013; Nguyen et al., 2014; Pan et al., 2020; Chancellor and De Choudhury, 2020; Kelley and Gillan, 2022]. Furthermore, social media is available most of the time, users can post virtually any time. On the other hand, the interaction with the health care system, be it for a general practitioner or a regular psychotherapist, is usually infeasible or not fre-

Table 6.1. List of related works and their main characteristics. The features column shows techniques for extracting features from textual, visual, or metadata content, FE: feature engineering. The user-level column refers to studies that train models for user-level classification by aggregating the post-level features in any way (be it an LSTM or simply taking the average of the post embeddings). “Data” shows the dataset’s source, be it manually collected by the researchers from Twitter, Reddit, Instagram, Facebook, Weibo, or Dcard, or using a dataset created for a shared task, such as eRisk [Parapar et al., 2021], CLPsych [Coppersmith et al., 2015] and SMHD [Cohan et al., 2018]. The field column is related to a mental disorder or a specific symptom of a mental disorder. MDD: Major Depressive Disorder, AN: Anxiety, SD: Suicide Ideation; “Mental health” refers to general mental health expressions, whether the user disclosed suffering from mental health symptoms.

Reference	Features	Classifier	Label	User-level	Post-level	Data	Modality	Field	Data Gather Method
[Tsuigawa et al., 2015]	FE, LDA, BoW	SVM	CES-D	Yes	No	Twitter	Text	MDD	Explicit
[Bagroy et al., 2017]	N-grams	LR	Inductive Transfer Learning	No	Yes	Reddit	Text	Mental Health	Implicit
[Shen et al., 2018]	FE	NB, MSNL, WDL, MDL, RF	Self-report	Yes	No	Twitter	Various	MDD	Implicit
[Reece and Dantforth, 2017]	FE	RF	CES-D	No	No	Instagram	Image	MDD	Explicit
[Cheng et al., 2017]	SC-LIWC	SVM	SPS, DASS-21	Yes	No	Weibo	Text	MDD, AN, SD	Explicit
[Orabi et al., 2018a]	word2vec	Bi-LSTM, CNN	Self-report	Yes	No	CLPsych 2015	Text	MDD	Implicit
[Coppersmith et al., 2018]	GloVe	LSTM	Self-report	No	Yes	Twitter	Text	SD	Mixed
[Ricard et al., 2018]	FE	LR	PHQ-8	Yes	No	Instagram	Text	MDD	Explicit
[Troitzek et al., 2018a]	FE, GloVe, fastText	CNN, LR	Self-report	Yes	Yes	Facebook	Text	MDD	Implicit
[Wongkrohlap et al., 2018]	FE	SVM, LR, DT, NB	CES-D	Yes	No	eRisk 2017	Text	MDD	Explicit
[Aragón et al., 2019]	BoSE	SVM	Self-report	Yes	Yes	Facebook	Text	MDD	Implicit
[Liu et al., 2019]	N-grams, FE	SVM, DT, RF, LR	Expert	No	Yes	Weibo	Text	SD	Implicit
[dos Santos et al., 2020]	LIWC, TF-IDF, word2vec	MLP, LR	Self-report	Yes	No	Twitter	Text	MDD	Implicit
[Fu et al., 2021]	BERT, FE	MLP	Expert, Knowledge Graph	No	Yes	Weibo	Text	SD	Implicit
[de Souza et al., 2022]	GloVe, Word2Vec	LSTM, CNN	Lexicon	Yes	No	SMHD	Text	MDD, AN	Implicit
[Cha et al., 2022]	word2vec, BERT	Bi-LSTM, CNN, MLP	MDI, PSS, TPI, GSE	No	Yes	SMHD	Text	MDD	Implicit
[Mukta et al., 2022]	MpNet	KNN, RF, AdaBoost, LGBM	Self-report	Yes	No	Facebook	Text	MDD	Explicit
[Bacur et al., 2023]	CLIP, EmoBERTa	Transformer, LSTM, GRU	Expert	Yes	No	Reddit, Twitter	Text, Image	MDD	Explicit
[Wu et al., 2023]	BERT	LSTM	Expert	No	Yes	Dcard, open data	Text	SD	Implicit

quent enough for most individuals — e.g., either because of professional unavailability or the associated cost.

Individuals use Social Media Platforms (SMP) for several reasons: to connect with distant relatives, to meet new people, to engage in sociopolitical activities, to talk to their friends, or to benefit from social relationships. Benefitting from social relationships can be seen as generating what is called social capital [Pan et al., 2020]. Although this concept is general enough for any social interaction, be it digital or not, the increasing popularity of SMPs, their convenience, and the paucity of time for real-life interactions make SMPs a natural gateway for social interactions in the digital world, especially for young adults. Although users might participate in SMP for several motivations, there is a crucial reason for screening mental health on social media: contributing an original post and responding to posts by other users, which might generate social capital [Pan et al., 2020]. Analyzing the relationship between social capital gains — or lack thereof — with mental health expressions can offer a unique opportunity to help individuals. Although users are not always looking for social support in social media, the word choices, the topic they are discussing, which posts they interact with, or what users they are friends with — all reveal a behavioral footprint that will inevitably depend on the structure and features provided by the specific social media platform. Next, we will discuss how different applications capitalize on the digital footprint.

6.3.2. Data Collection

How do researchers obtain the data? In order to obtain data, they either (1) asked volunteers to participate through formal questionnaires and with explicit terms and conditions to guarantee participant’s privacy and knowledge; (2) scraped publicly available content without asking for the owner’s direct permission; (3) or used one dataset already collected using method 1 or 2, generally known as datasets for shared tasks, such as the eRisk, CLPsych, and SMHD datasets in Table 6.1.

For the first method, the study generally applied for the local Institute Review Board (IRB) and often stored the data anonymously and securely. We refer to this method as explicitly asking for the participant’s consent — as demonstrated in the Data Gather Method column in Table 6.1. Previous research asked for participants’ permission to scrape their social media content and answer psychometric tests. However, even when users thoroughly answered the psychometric test, several participants refused to share their social media data. For example, in Reece et al. (2017) [Reece and Danforth, 2017], 43% of individuals who completed the survey refused to share the Instagram data. In Wongkoblapp et al. (2018) [Wongkoblapp et al., 2018], for one dataset, only 18% provided access to their Facebook data (931 individuals). Perhaps, a consequence of survey questionnaires is the demonstration that privacy often matters for individuals, and they have more concerns about their data being analyzed, especially for mental health footprint.

Informed consent is often the desirable method for ethical guidelines in research, and a few studies compensated the participants with monetary payment to boost response rate [Reece and Danforth, 2017; Cheng et al., 2017; Ricard et al., 2018]. It also has the benefit of asking individuals to answer psychometric tests to serve as a gold label and to obtain sociodemographic statistics. However, asking for informed consent also has

downsides, such as selection bias in online surveys, when a particular sociodemographic class is more prone to answer and help than others. For example, in Wongkoblap et al. (2018) [Wongkoblap et al., 2018] 59% of participants are female. On the other hand, in Tsugawa et al. (2015) [Tsugawa et al., 2015], for a sample of Japanese individuals, 58% are male. While sociodemographic factors may influence, there is also the possibility of bias related to the communication channels used to disseminate the research invitation. Additionally, as the online survey invitation often already reveals the research focus, i.e., mental health, individuals who believe they are experiencing mental health issues could be more inclined to participate.

For this last issue, consider the prevalence of depression among the general population: 5,8% for Brazil and 5,9% for the US [WHO, 2017]. Next, consider the prevalence of depression for each explicitly collected dataset. The Instagram dataset [Mann et al., 2020] contains 60% of depressed individuals as measured by the BDI-II; in Wongkoblap et al. (2018) [Wongkoblap et al., 2018], they reported 76% of depressed individuals as measured by Center for Epidemiologic Studies Depression Scale (CES-D); Tsugawa et al. (2015) [Tsugawa et al., 2015] reported 39% of depressed individuals as measured by CES-D; Reece et al. (2017) [Reece and Danforth, 2017] reported 43% also measured by CES-D. Note the difference between the prevalence of depression in each sample and the prevalence for the general population.

Furthermore, in all the research mentioned earlier, participants are often young adults, which explicitly excludes a big part of society — such as older individuals or individuals without access to the internet and SMPs. Machine learning models trained on biased data often underperform when faced with a data distribution not seen during training (more on this subject in Section 6.4). The high amount of young adults with depression in all those datasets demonstrates the difficulty of accessing a balanced sample according to sociodemographic markers.

The most preferred method for gathering data is scraping the data publicly without explicit consent from owners. The consent is implicit because SMP users must agree to the social media data policy and terms of service, which usually inform users that the SMP has broad rights to use and distribute the data that users create and share. Collecting data implicitly is more straightforward, less time-intensive, and cheaper. Thus, 64% of the related works relied on data obtained implicitly, as demonstrated in Table 6.1.

However, data obtained implicitly often lack several benefits from explicitly asking permission. First, social media users are often anonymous, so it is hard to scrape sociodemographic statistics. Second, to label posts (or users) for mental health disorders, researchers usually rely on posts containing the self-report of a specific mental disorder. For example, users who wrote posts that match the pattern “(I’m/ I was/ I am/ I’ve been) diagnosed depression”, or the ones that loosely mention “depress”, or even through regular expressions, would be labeled as “depressed”, among other mental health disorders or symptoms [Shen et al., 2018; Trozsek et al., 2018a; Coppersmith et al., 2018; Orabi et al., 2018a; Aragón et al., 2019; dos Santos et al., 2020]. There is also a concern about the validity of the collected data and whether the self-report text is accurate, for example.

On the other hand, the benefit of scraping data without asking for explicit permission is the possibility of scraping much more data in a large-scale scenario. When

gathering public data, researchers can collect up to 152,834 unique Reddit users with 446,897 posts [Bagroy et al., 2017], or 36,993 depression-candidate Twitter users with over 35 million tweets [Shen et al., 2018]. In comparison, asking for permission yields a dataset with no more than 1000 individuals [Tsugawa et al., 2015; Reece and Danforth, 2017; Cheng et al., 2017; Coppersmith et al., 2018; Ricard et al., 2018; Wongkoblap et al., 2018; Mukta et al., 2022], where studies with bigger datasets, in general, provided monetary compensation for participation [Cheng et al., 2017; Ricard et al., 2018]. Relying on individuals' self-reporting rather than using a psychometric instrument comes with the benefit of a larger dataset. With modern deep learning techniques relying less and less on specific architectural innovations and more on large datasets, especially with relaxed inductive biases on Transformer architectures, gathering more data could lead to improved prediction scores [Bucur et al., 2023]. Considering the listed related works in Table 6.1, there is still work to be done to understand the impacts of different types of datasets both in performance and ethical guidelines.

The third method to obtain data is to ask for other researchers who either collected explicitly or implicitly or to use shared tasks datasets, often distributed for competition purposes. There are two widely used datasets in the literature: CLPsych and eRisk.

6.3.3. Feature Extraction and Classifiers

In this section, we will explore two essential aspects of automatic classification: (1) extracting features from multiple modalities of data; (2) and the common classification methods used by the related works.

6.3.3.1. Extracting Features

There is also a significant concern about how to generate the model's input. Several studies in the literature use textual features to detect mental health disorders, mainly based on the assumption that there are psychological traits in the text produced by individuals [Pennebaker et al., 2003]. Among them, using the taxonomy proposed by Dhelim et al. (2023) survey [Dhelim et al., 2023], there are three high-level categories of features: textual features, multimedia features, and behavioral features.

Textual Features

There are three main categories of textual features: linguistic, sentiment, and ideograms. One of the most common choices of features is obtaining the psychological categories of words. The Linguistic Inquiry Word Choice (LIWC) is one of the most widely used psychological dictionaries (under the linguistic category in the taxonomy). Previous studies used LIWC because of its simplicity, ease of use, and the expressivity of the features, such as obtaining meaningful categories for words: positive and negative emotions, anger, and personal pronoun usage, among others. Several studies use it as a baseline method [Cheng et al., 2017; Liu et al., 2019; Mukta et al., 2022; dos Santos et al., 2020; Shen et al., 2018; Trozsek et al., 2018a; Wongkoblap et al., 2018]. There are other psychological dictionaries, such as empath [Fast et al., 2016], a psychological dictionary based on deep learning techniques that produce categories based on a set of seed words, which one of the related works also use [Mukta et al., 2022].

On the other hand, topic modeling is used to compute a set of latent topics that indicates the general discussion or directions — or simply topics — from a set of textual contents. The set of textual content could be from one social media user or the entire dataset. Researchers apply supervised or unsupervised topic modeling in the entire dataset to understand the general topic of discussion among depressed individuals against control groups and use it as a feature vector to feed classification models [Tsugawa et al., 2015; Shen et al., 2018].

Another commonly used linguistic feature, especially for baseline methods, is the Bag of Words (BoW), sometimes used with the TF-IDF weighing scheme [Ricard et al., 2018; dos Santos et al., 2020]. It is a simple method based on the count of the frequency of words. The product of this technique is a matrix where the rows represent posts (in post-level classification) or users (in user-level classification), and the columns represent the vocabulary. However, the main issue with this representation is that the order of words in a post is not represented, and the temporal dimension of mental disorders will not be modeled as part of the problem. N-grams is a technique that considers the probability of occurrence of a word (w_i) given a history of words preceding the word to counterbalance this issue. For simplicity, the entire history is often not considered but only a small window; hence, that is why N-gram. If $N = 2$, we simplify the problem of determining the probability of occurrence of a word given its entire history by the probability of a word given its preceding word, also called the Markov assumption. The product of this process is often a matrix where the columns are the N-grams, and the rows are the posts (post-level classification) or users (user-level classification), where each cell contains the frequency or the likelihood ratio of the N-gram.

Unlike BoW and the N-grams approach, word embeddings are crucial in deep learning techniques. To that end, word2Vec [Mikolov et al., 2013] was one of the pioneer techniques to improve several NLP tasks by allowing words to capture multiple degrees of meaning through their low-dimensional latent representation. However, this technique has a few limitations that the other recent ones do not have. First, it can not represent polysemy because of the same vector representation for the word regardless of context. Second, all embeddings are trained to an entire corpus, which means that words not seen during training are not represented at test time. Third, it does not consider hierarchical representation for words, impairing the representation of syntax and semantics aspects. A few works use either word2vec, or similar variants with similar limitations [Orabi et al., 2018a; Coppersmith et al., 2018; Trotzek et al., 2018a; dos Santos et al., 2020; Cha et al., 2022].

Recent word embedding techniques were improved to represent polysemous words, richer representations with improved methods, and a more extensive training corpus with a better tokenization strategy. Among several variations, BERT, CLIP, and RoBERTa are often used for representing words, or sentences, as vectors [Fu et al., 2021; Cha et al., 2022; Wu et al., 2023; Bucur et al., 2023].

Sentiments are primarily obtained from LIWC but can also be obtained through sentiment analysis algorithms. A few works rely on specific emotion lexicons, such as the Affective Norms for English Words (ANEW), VADER, NRC, and the Opinion Lexicon [Trotzek et al., 2018a]. Other works elaborated their lexicon based on experts [Cha

et al., 2022]. One naive approach is to consider posts with words such as “pain”, “depression”, and “unhappy” from individuals with possible signs of depression. Although this is simple enough for text matching, it is difficult to know whether the word, in context, does not have an entirely different meaning, such as irony.

Regarding ideograms, a few works removed emoji from the text as they were considered “incompatible with many text processing algorithms” [Shen et al., 2018; Fu et al., 2021]; other works included an emoji sentiment scale to map emojis to a happiness score [Ricard et al., 2018]. However, several works need to describe what they do with emojis: adding, modifying, or removing them entirely. One of the exciting things about the recent — contextual — word embedding techniques, such as BERT, is that they support the representations of emojis without relying on tricks to overcome the limitations of the previous static word embeddings, such as word2vec.

Multimedia Features

A few works relied on images for multimedia features, especially given the rise of photo-oriented platforms in recent years. They either manually extracted features, such as the hue, value, brightness, and number of faces [Reece and Danforth, 2017; Shen et al., 2018], or they relied on visual representation learning techniques [Bucur et al., 2023]. It has been demonstrated that using more than one modality improves the performance scores in several experiments [Shen et al., 2018; Mann et al., 2020, 2022; Bucur et al., 2023]. The type of modality faced by multimodal data is particularly challenging as they are characterized by meaning multiplication [Bateman, 2014]: the textual and visual contents may refer to distinct contexts, but both modalities are essential to creating a new meaning that diverges from merely making a decision separately from the unimodal meanings. However, it still needs to be determined how much each modality contributes to the performance and exactly how. None of the selected related works investigate audio data. Convolutional Neural Network (CNN) and CLIP is the preferable method for extracting visual representations [Bucur et al., 2023]. Although a few works use CNN, they use it for sentence classification instead of obtaining visual features.

Furthermore, manually extracted features bring awareness to the bias problem, where the choice of features relies on the researcher’s knowledge and how their culture perceives the mental disorder. The way society perceives depression in one person is often dependent on the subjectivity in culture and environment, which ultimately is reflected upon the diagnosis criteria of psychiatric manuals [Association et al., 2013]. For example, while the posting time is a determinant for classifying depression in a sample of American individuals, it is not for Japanese individuals [Tsugawa et al., 2015]. By leveraging Representation Learning and Deep Learning [Bengio et al., 2013], the model automatically learns feature representations according to the task using generic priors, eliminating the need to extract features manually. Additionally, we can use the learned representations to transfer the knowledge with transfer learning [Pan and Yang, 2009] to other domain-related problems; such a procedure is not straightforward to replicate for handcrafted features.

Behavioral Features

Social Media Platforms offer various options for users to connect or interact with

the platform itself or other users. For behavioral features, there are several categories of behavior. For example, online activities involve interaction with other posts and users, such as liking, sharing, or commenting on a post, the number of followers, or how many users they follow. Another behavioral feature is associated with communities, where individuals can affiliate with subreddits (Reddit) or Facebook groups. The information on which groups and the kind of relationship the user has with the group is another rich source of behavioral information.

Several related works rely on behavioral features to feed their classification models [Tsugawa et al., 2015; Shen et al., 2018; Reece and Danforth, 2017; Ricard et al., 2018; Wongkoblaph et al., 2018; Fu et al., 2021]. One possible reason for not relying on such data is that the behavioral activity is inherently attached to the SMP, and it is hardly helpful for transferring the learned knowledge to another SMP — because of the inevitably different characteristics or structure of the SMP. Another reason is to effectively experiment with the classification only on the textual content created by individuals without relying on any external behavioral trace. If it is possible to distinguish depressed individuals from control based only on their produced textual content, then we expect that any added behavioral footprint will improve the performance scores.

More specifically, Ricard et al. (2018) [Ricard et al., 2018] was the only study to compare user-generated content with community-generated content. Their experiments show evidence that community-generated content does indeed help improve performance scores as opposed to only using linguistic features.

6.3.3.2. Classification

The extracted features are used to feed classification models. The classification method will automatically classify different entities based on how features are computed: a user, a post, or even a community. All related works deal with supervised learning approaches and label their dataset according to Section 6.3.2.

It is important to note that if they labeled the data using a psychometric test, the obtained score naturally results in a user-level labeling process [Tsugawa et al., 2015; Cheng et al., 2017; Ricard et al., 2018; Mukta et al., 2022]. As the score obtained is usually an integer number, the related works often consider a threshold to split users into two classes: depressed and non-depressed — or other mental health disorders. For example, the BDI splits the score into four categories of intensity of depressive symptoms: 0–13 for minimal intensity, 14–19 for mild intensity, 20–28 for moderate intensity, and 29–63 for severe intensity. Furthermore, in the psychiatric literature, moderate and severe categories are related to a depressed individual [Gorenstein et al., 2011]. It is similar to other psychometric tests — such as CES-D —, where they establish optimal cutoffs to distinguish depressed from non-depressed individuals.

On the other hand, several studies relied on individuals self-reporting depression (or other mental health conditions) to separate into two groups — one with the mental health disorder and the control group [Shen et al., 2018; Orabi et al., 2018a; Trotzek et al., 2018a; Aragón et al., 2019; dos Santos et al., 2020; Bucur et al., 2023]. When considering self-reporting the mental health condition, researchers use the premise that

the individual is suffering from that condition and label them as of the positive group. For the negative class, researchers often scrape data from random users across the SMP or collect data from forums that discuss general topics, such as subreddits about movies, food, or news [Orabi et al., 2018a; Trozsek et al., 2018a; Aragón et al., 2019].

One exciting approach to gathering control group data is to collect data from individuals who talk about the mental health disorder but do not suffer from any mental health disorder [dos Santos et al., 2020]. Using posts from the “depression” subreddit as the positive class and posts from subreddits such as news, movies, or food as the negative class might induce the model to learn simple correlations that result in high metric scores. As such, the model could learn to correlate the occurrence of words such as “depressed” and “unhappy” with depression simply because those words are more likely to occur in the depression subreddit. Collecting data from individuals who shall use the same — prominent — words that depressed individuals use but are not suffering from depression makes the task more challenging for classification models — and more akin to the real world.

Classifying the user directly (user-level) can attribute the mental disorder directly to the user. However, one issue is relying on aggregating tricks to obtain the user-level feature vector. When doing feature engineering, it is straightforward to aggregate data; for example, when using n-grams or BoW, it is enough to sum the frequencies across all posts of a single user to obtain the frequency vector for a single user. Additionally, when using LIWC, the researcher only needs to aggregate the frequency for each category for each post, resulting in a user-level feature vector. Aggregating manually engineered features across time might result in a loss of information for classification. Several related works used aggregated engineered features data to train user-level classifiers [Tsugawa et al., 2015; Shen et al., 2018; Cheng et al., 2017; Ricard et al., 2018; Trozsek et al., 2018a; Aragón et al., 2019; dos Santos et al., 2020].

However, aggregating features is less effective for low-dimensional neural distributed representations, such as word2vec or BERT. Previous works have presented a theoretical and practical framework for constructing a Multiple Instance Learning (MIL) methodology [Mann et al., 2022]. One study implements a user-level classification that leverages the MIL paradigm [Bucur et al., 2023] — i.e., without aggregation tricks. We note, however, that none of the mentioned works gives formal MIL specifications for their tasks.

In the context of post-level classification, it might seem too granular to classify a single post. However, there are a few applications wherein such granular classification is desirable, such as identifying a post written by a person at risk of suicide [Coppersmith et al., 2018]. In such circumstances, using the latest post written by the individual to measure suicide risk is necessary. However, depression is a condition whose diagnosis criteria require two weeks to observe symptoms. Consequently, a post-level classification will inevitably lose the temporal component when analyzing mental health conditions.

When examining user-level classification, there remain valid concerns regarding the model’s training process over the user-level feature vector. Specifically, the model learns to establish a linear or otherwise correlation with the class based on the aggregated feature-engineered vectors (user-level feature vector). However, it is worth noting

that this feature vector remains static — it is not learned during training, just statically computed before it. As a result, it does not consider local temporal variations, such as the change in usage of emotional words from one week to another. Even if we let the model adjust the user-level feature representation during training, it will not compute the user-level representation based on the post-level representations. In contrast, employing the Multiple Instance Learning (MIL) methodology [Mann et al., 2022] allows the model to ascertain the best user-level representation during the training phase, provided that the feature extractor can be fine-tuned. The learning procedure dynamically determines the most effective user-level feature vector based on post representations.

Regarding supervised machine learning algorithms, the most frequently used classical machine learning techniques are the Support Vector Machines (SVM) [Tsunami et al., 2015; Cheng et al., 2017; Wongkoblaph et al., 2018; Aragón et al., 2019; Liu et al., 2019], Logistic Regression [Bagroy et al., 2017; Ricard et al., 2018; Trozsek et al., 2018a; Wongkoblaph et al., 2018; Liu et al., 2019; dos Santos et al., 2020], Random Forests (RF) [Reece and Danforth, 2017; Liu et al., 2019; Mukta et al., 2022], Decision Trees (DT) [Wongkoblaph et al., 2018; Liu et al., 2019], Naïve Bayes (NB) [Shen et al., 2018; Wongkoblaph et al., 2018], or tree-based ensemble algorithms, such as AdaBoost and LightGBM [Mukta et al., 2022]. These methods have the advantage of being low resource intensive and often provide some degree of interpretability — such as using the coefficients of a linear SVM.

However, these methods often reach a performance saturation point, showing minor improvements even with bigger datasets. As such, deep learning algorithms are often used to overcome this limitation, such as Long Short-Term Memory (LSTM) [Orabi et al., 2018a; Coppersmith et al., 2018; Cha et al., 2022; Bucur et al., 2023; Wu et al., 2023], CNN [Orabi et al., 2018a; Trozsek et al., 2018a; Cha et al., 2022], and Transformers [Bucur et al., 2023]. Nevertheless, a few works do not use either LSTM or CNN for user-level classification, but as a sequence (of words) classification [Coppersmith et al., 2018; Trozsek et al., 2018a; Cha et al., 2022; Wu et al., 2023]. Only two of the selected related works use either LSTM or Transformer to classify individuals based on the sequence of posts [Orabi et al., 2018a; Bucur et al., 2023].

6.4. Enacting Change: Principles and Directives for Socio-Ethical Machine Learning Models for Screening Mental Disorders

To screen depression automatically, researchers often investigate Machine Learning (ML) methods that rely on social media publications to learn patterns associated with depression. Those models offer an alternative way to large-scale screening depression that could guide mental health administrators to create better policies. Although one could argue that such models have been successful to a degree, they are still experimental or impose a significant risk to use in the real world. As these models are inherently embedded in social systems, it is imperative to discuss the general role of technology and the existing issues around it that impact — and are impacted by — social aspects. Consequently, social and technological components are not isolated in society. We must understand the main social catalysts that lead to the innovation of technological solutions. This intertwined relationship between the said “social” and “technology” often results in sociotechnical

systems [Selbst et al., 2019]. We could perceive society as a continuously evolving sociotechnical system because of our increasing overreliance on technology.

Technology has imposed its presence in arguably all aspects of life: from leisure to work, gaming to video streaming, and augmented reality to shopping. Its obtrusive characteristics, nonetheless, challenge society in many ways, such as data privacy and insidious changes in socioeconomic structures. As these tools are present in many aspects of life, they act in elusive ways. For instance, it can change people's votes without their awareness [Epstein and Robertson, 2015]. The subtlety nature of some technologies acts purportedly to change our *psyche* in some direction. In this sense, they can be used as a mechanism to control opinion according to power relations, showing as political tools with moral implications.

AI, and more specifically, Machine Learning (ML), casts even more challenges with its high ability to address complex problems. With the recent advances in ML and subsequent implementations of those new technologies in production, we, as a society, are falling behind in understanding the real consequences these new tools bring, especially when they are taken for granted as purely beneficial and neutral. However, as researchers, we must understand the impacts of the technologies we create. Furthermore, we should understand the social and political consequences of the wide adoption of these technologies imbued with *psyche-changing* capacities.

The inquiry about how a technological advancement could potentially change inherent human behavior was first documented in Plato's *Phaedrus*. Plato's discourse investigated the influence of a now-incorporated and widely used technology in society: writing⁴. Despite its practical benefits, Plato asked himself whether writing would weaken an individual's memory capacity. Indeed, once, the ancient Greeks could recite a significant part — or entirely — of the Iliad by memory [Jaeger, 2001; Foley, 2007; Parry, 1933].

Individuals in our modern society rely more and more on new technological tools. The immediate consequence, as theorized by Plato, is the effect on the ability it will directly replace or partially replace — such as memory. For example, the very nature of pre-processing input data to standardized formats dehumanizes and deskills humans, removing any contingency and making everything predictable; the consequence: it promotes hegemonic behavior that might homogenize creativity [Burrell and Fourcade, 2021]. This is crucial because it does not happen individually, as technology and AI is widely used. This will inevitably impact society at all levels, not only the specialized individuals directly using the technology, which ultimately will disturb the socioeconomic framework. Are the creators of such tools, governments, or even the public sphere aware that those changes will profoundly impact society?

Still in Plato's *Phaedrus*, the old god Thoth, the creator of many arts, argues that his creation — the writing art — benefits individuals because it will improve their memory and wisdom. However, the other god discussing with Thoth argues that the "inventor of an art is not always the best judge of the utility or inutility of his inventions to the users of them" [Plato et al., 1952]. This inquiry is more modern than ever: Society as a whole — not only specialists but laypeople too — need to take part in the public discourse

⁴Analogy borrowed from [Ballesteros, 2020].

and take action about impactful technologies in their lives. Even though Plato's inquiry proved correct, the undeniable benefits of writing vastly outweigh its problems, as he left us many texts. However, we need to pose the same question that Plato once investigated, with a few more: Will this new technology change our behavior? If so, what are the consequences? Do the benefits far outweighs the new problems created by using such technology? In what context should the proposed technology be used to avoid any form of moral violation? If technology can influence users' psyches, how do we ensure those in power do not misuse it for their benefit? Are these technologies fair and just? Are they addictive?

Previous works have discussed the ethical impacts surrounding those questions for predicting mental health state. However, we stress that previous works discussed more on the perspective of conducting ethical and moral research guidelines [Conway, 2014; Benton et al., 2017; Chancellor et al., 2019], or understanding the general population's ethical opinion about using social media data for research [Mikal et al., 2016; Fiesler and Proferes, 2018]. Inspired by Selbst et al. (2019) [Selbst et al., 2019] ripple effect trap⁵, we take another route: we are particularly interested in the humanistic, social, political, and ethical concerns that impact, and are impacted by, the creation and usability of models to screen depressed individuals deployed on social contexts.

As such, we rely on social sciences, philosophy, economy, and politics studies to develop an analytical framework discussing three main challenges of deploying models to screen depressed individuals based on social media data. Therefore, we contribute with the following an analytical framework: (a) First, we explore the challenges of the social system where the ML model is embedded, understanding how the model impacts the sociotechnical system and vice versa (Section 6.4.1); (b) Second, we investigate how the ML models impact different stakeholders to varying degrees (Section 6.4.2); (c) Finally, we investigate how data inequity and misrepresentation create inherently biased models that are ultimately dangerous for screening depressed individuals (Section 6.4.3). Finally, by observing the existing tensions in the aforementioned analytical framework, we propose strategies to mitigate the presented issues, such as mental health, data, and digital literacy. Given the gap in this interdisciplinary research, we contribute to approximate computer scientists to social sciences, giving more emphasis on the social aspect.

6.4.1. Those who rule and who are ruled

In this Section, we focus on the social, political, economic, ethical, and philosophical aspects that underpin the application of ML models to screen for depressed individuals using social media data. By doing this, we expect computer scientists to have a broader view of the fundamental social aspects underlying the deployment of models in the wild. We do not expect this Section to be an exhaustive enumeration of all social issues but rather to shed light on existing issues that computer scientists often overlook. Hence, we start by discussing that commercial interests are aligned with collecting behavioral data, which is essential to virtually any technology today. Next, we discuss how data creates a power relation deepening surveillance and control. Finally, we discuss the relationship of

⁵Defined by Selbst et al. (2019) as "Failure to understand how the insertion of technology into an existing social system changes the behaviors and embedded values of the pre-existing system". Here, however, we argue that this change is bidirectional.

commercial interests and data to the task of screening for depressed individuals on social media platforms.

The technology industry had a turning point just after the dot-com bubble. After firms failed to deliver profitable business, they found what would be called the “new oil”: data [Burrell and Fourcade, 2021]. They found that the digital trace, or behavioral data, that users left behind using a system could generate profit. Companies started to use this data to push directed advertisements for their users while reaping a share of advertisers’ payment. Furthermore, some firms also started to sell the data they had to other firms⁶.

The new proposed way to monetize over digital trace data proved financially successful. Firms deliberately eased access to their systems to facilitate unimpeded entry, ultimately attracting a larger user base and generating more digital traces, which led to more profit. This powerful mechanism rapidly boosted firms’ growth, leading to higher profits and stabilizing their market share and dominance. Companies that arrived earlier in technology gained so much power that they are now accused of antitrust violations [Kolhatkar, 2021]. Nevertheless, the monopoly is not only financial: they hold information and computing power, equating to political power in the era of *infocracy*. The consolidation of the called “Big Techs” resulted in conglomerates of firms in which no other firm can compete, or if that is the case, they buy the competitors or force them to give up [Noble, 2018].

Consolidated companies have collected users’ digital traces for more than two decades. Previous AI methods have not effectively handled big data sets, be it for lack of hardware power or because the method’s performance did not scale well with data. After the rebirth of neural networks under the garment of deep learning [Aggarwal et al., 2018], the Big Techs could now put their vast data sets under deep learning to improve their predictive capabilities and explore new avenues they could not access. For example, the studies mentioned in Section 6.3 collect user-generated content through social media, be it with explicit consent or not. However, individuals frequently do not know their data could be used in such a way, although they agree with the social media platform terms, often without paying proper attention to the small letters and large texts. As a result, users’ data generate value for third-party entities without the *explicit consent* of the data creator — the data is being used to train models for screening depressed individuals. The implicit consent through accepting those platforms’ agreement terms is too vague and frequently ambiguous. Users only know that their data could be shared or used to train models without explicitly knowing the model type and task it solves. Many individuals are unaware that their data can be used for such endings or think deleting previous posts will solve the problem [Mikal et al., 2016]. Furthermore, ML models are known to be vulnerable to attacks that target extracting information from the training set [Salem et al., 2019; Hu et al., 2022; Huang et al., 2022], which imposes even more risks for users whose data are used for training such sensitive models.

The data available for those companies are mainly personal or behavioral. They use this data to create what Byung-Chul Han calls *psychometrics*: a method to gener-

⁶Interestingly, the practice of surveying and selling data to other interested firms is much older than expected. For example, during his first job in 1907 in Portugal, Fernando Pessoa was tasked to collect information from firms to sell to other firms around the world [Zenith, 2022].

ate a personality profile [Han, 2022]. The vast data feeds deep learning models trained to predict the users' behavior or predilections. Based on the profiles, the company can better discriminate users' interests. For instance, it is possible to determine a group of individuals with depressed-related behavior through clustering techniques or topic modeling [Resnik et al., 2015; Dipnall et al., 2017]. When implying that individuals' behavior is similar to those of a group, we incur the risk of determining the individual identity as being equivalent to those of the group.

Consequently, we negate two fundamental principles of self-determination for identity: justification and control [Engelmann et al., 2022]. When firms group individuals by their psychometrics, they often do not disclose or justify that automatic decision. Thus, users do not know what firms “think” about them and have no access or chance to modify that “opinion”. This problem is aggravated by the fact that firms arbitrarily choose the amount of data to determine whether an individual pertains to a group of individuals with depressed-related behavior. The amount of information that determines whether an individual is depressed or not is hardly written in stone, and it will be different for each specific case. By limiting an entire life that encloses unconscious and conscious actions through deterministic mathematical formulations, we crystallize human behavior to a set of mechanical rules by ignoring the contingency of life. Therefore, those systems must let individuals self-determine their identity by justifying and controlling how they are portrayed, especially for sensitive “classifications” such as screening for depressed individuals.

While firms are driven by commercial interests and market dominance, the “side-effects” of applying those technologies on a large scale should not be negligible. Even if the firm does not misuse their predictions, they could sell the psychometrics data to other firms — a standard practice in the market [Noble, 2018]. Furthermore, as users surf the web, their data could be retroactively associated with digital traces from other sites, yielding more power to those institutions that can collect and use this data to train more effective models. The ethical implication is, “Do they own the right to sell our information, which other digital systems could use to create enhanced psychometrics”?

In essence, the individual is no longer entitled to privacy, and firms use data they can collect for their commercial interests. Although digital systems often disclose how they use data in their agreement terms, technological tools transform individuals into hostages. Even if the individual wants to use something other than the system, they are left behind with only a few choices, which are frequently deficient [Noble, 2018; Burrell and Fourcade, 2021]. Thus, the most widespread tool is consistently improved because it is widely used: they generate more psychometrics data to train better models, generating better-directed advertisements, increasing profits and market dominance in a perpetual cycle. This is especially the case for Google products, where they transform its users in the *cybertariat*: “individuals that perform a continuum of unpaid, micro paid and poorly paid human tasks” [Burrell and Fourcade, 2021] on tasks such as verifying if there is a crosswalk in the given images.

Although companies' commercial interests sound distant from screening for depressed individuals, we contend that these elements form the core political, economic, and social fabric that underpins predicting depression status, among many other tasks.

Because companies and applications optimize to favor profits, they often rely on techniques that touch human cognition, emotion, and behavior — such as making addictive applications to keep users engaged. Bombarding users with microtargeted advertisements; exposing individuals to an ever-increasing competitive life; killing the alterity by creating Filter Bubbles⁷; excess of positivism — notably in social media platforms; societal norms and expectations such as hyper-productivity and self-optimization; selling behavioral data to interested third parties. These forces make individuals more exposed to stress and depression [Han, 2015]. While ML models for screening depressed individuals pose a solution to this increasing mental health issue, it is also inherently embedded into the commercial interests of the same society that generated this problem in the first place. As a result, we must first ask ourselves what is the best long-term strategy: create ML models to large-scale screen depressed individuals or fundamentally change societal structure to avoid stress, burnout, and depression?

Arguably, one common approach to both problems is education. From one side, society must be educated to understand insidious and predatory commercial practices to protect themselves. For example, in a focus group, despite individuals knowing that Twitter data is publicly available, they think that deleting posting history will protect them because they lack the knowledge that the data could be collected and saved in other personal databases [Mikal et al., 2016], or collected through paid API's and reselling services. In terms of the general population, individuals need to understand the principles and mechanisms that govern technology because they lack fundamental knowledge about data permanence. As such, individuals are surprised when they are told that their social media posts could be used to find the prevalence of depression [Mikal et al., 2016]. For the personal aspect, we will further discuss the education aspect in Section 6.4.2.

In our application of interest, screening for depressed individuals is not always a stand-alone model. Typically, it could be embedded into a larger system as its primary objective or a collateral effect of trace data collected through system usage. The embedded model is explicitly trained to screen possibly depressed individuals, often using standard psychological tests or self-report data as labels, as explored in Section 6.3. The objective and scope of such models are naturally restrained, and thus, it is easier to understand the ethical implications such models might pose. However, identifying possibly depressed individuals based on the digital trace they left behind poses a significant risk to privacy and the capacity of individuals to self-determine.

However, different environments create different opportunities and goals. Social or humanitarian goals are often absent in a firm where data is used for profit. On the other hand, in places that should be more welcoming, such as schools and universities, the goal of screening for depression is to benefit the students. Since the sample of undergraduate and graduate students is often three to six times more exposed to depression than the general population [Ibrahim et al., 2013; Evans et al., 2018], it is vital to find depressed students to conduct an effective intervention. In a welcoming, affectionate setting, where the university's goal is not to profit above everything else, there are legitimate social or

⁷Eli Pariser's concept about recommender systems that select ideas and news that individuals might like, or agree, based on their digital behavior patterns. As a result, individuals are predominantly exposed to information that aligns with their existing beliefs and relies less and less upon alternative viewpoints to shape opinion — alterity.

humanitarian objectives to help students instead of exploring their data for profit — although only sometimes valid, since there are private universities as well. Furthermore, in a controlled setting with clear goals to help students, we expect data collection and processing to follow strict ethical guidelines, such as the approval of the institutional review boards (IRBs). With that, volunteers to participate in the study (or deployed system) explicitly concede permission to use their data and that by providing the data, they explicitly benefited from the resulting system.

Unlike the university environment, firms are optimized for best matching the advertisement with their users. To do that, they need to discriminate well among various profiles, one of them possibly being the “depressed” profile. Depending on the technique they use, the profile is latent, which means that they do not know *a priori* if he or she is interested in “drama”, “technology”, “sports”, or “depressed” content. However, they know well enough to predict predilections — a proxy for identity. If they take the time to evaluate groups of individuals with similar interests or personalities, they might find a group tied to depressive content; in the end, they could further explore and benefit from this particularity. More specifically, a social media platform company aims to improve engagement by neglecting the consequences it brings. For example, social media platforms increase engagement based on provocative and divisive publications, which opens the door to several conspiracy theories with manipulative content and false information [Rauchfleisch and Kaiser, 2020; Fisher, 2022]. When creating a space that amplifies divisiveness, addiction, and envy, social media platform companies lay the fundamental triggers for causing and sustaining depressive symptoms. Hence, when screening depressed individuals on social media, we may want to solve a problem inherently amplified by how commercial interests are intertwined with social media platforms.

Furthermore, with the psychometrics information, firms can disclose their findings to other subsidiary companies — improving their services —, but can also sell the same information to interested third parties. Even when users know that the platform sells their information, they often need to learn that their digital trace is being processed to find much richer information, such as psychometrics. In this scenario, users must judge whether they want their personality screened by predatory practices, but before that, they should be educated on how this technology works. While the general population is not educated, firms mine information arguably freely⁸.

From another standpoint, a system incurs the risk of broadcasting inaccurate information. For example, one individual might be wrongly screened as depressed by the system. Next, driven by financial interests, the firm might commercialize identity information to other firms or redistribute it to its subsidiaries. It is clear that selling inaccurate information is not only prejudicial to the buyers — or the seller —, but it may cause harm to the individual whose misjudgment of the model directly affects. Furthermore, the user frequently has no mechanism to change how its identity is rendered nor control over its commercialization: they lack the justification and control capacities for self-determination [Engelmann et al., 2022]⁹. On the other hand, if the firm and its sub-

⁸Although we have General Data Protection Regulation (GDPR) and Lei Geral de Proteção de Dados Pessoais (LGPD) to protect individuals’ privacy and information, it is hard to keep a vast and unregulated space like the internet under control.

⁹The reader could argue that users can stop using the system. However, some systems are pervasive

sidiaries keep the false prediction, it might create an echo chamber of disinformation in all systems owned by the firm that initially committed the mistake.

This situation is undesirable by itself and aggravated when the general population widely uses a system — notably when the system is seen as a public resource¹⁰. The harm is not only on the individual; however, it can shift the perception of other users of the tool — which will inevitably promote hegemonic representations about an individual or a group of individuals. In that sense, Byung-Chul Han also argues that algorithmic operations have the potential to totalitarianism [Han, 2022]: as the classic totalitarianism thoroughly explains the past, present, and future through a straightforward truth to the detriment of multiple possibilities, ML models poses the same issues through black-box predictions. The model prediction is a new truth. By promoting a unitary truth — or worldview —, widely believed to be neutral and unbiased, we reach a state of totalitarianism by inadvertently burying other opinions and reinforcing the dominant standpoint, similar to what happens with recommendation systems that create Filter Bubbles. The dominant culture’s views and opinions, including race hierarchy and power relations, often surmount the opinions of the marginalized groups. Even worse, the dominant’s opinions of the marginalized groups often prevail over the opinions of the marginalized groups have about themselves¹¹. Thus, depressed individuals suffer from social stigma because the dominant view of the mental disorder is not portrayed nor publicly disseminated the way individuals who have this mental disorder expect. The constant disinformation about the mental disorder in conformity with the rise of political ideologies of meritocracy and neoliberalism¹² results in a dominant view that does not comply with how depressed individuals see themselves. The representation of information is rendered as a function of the dominant view. Popularity is vital because AI and generalized technology are heavily based on statistical processes. However, popularity does not equate to truth.

If groups of individuals are being misrepresented in technological systems, how should we, as a society, approach this problem? First, we should educate society about the underlying functions of such digital systems so they know what is happening with the digital traces they left behind. Next, accountability is of utmost importance for harm caused to marginalized groups. However, the harms such as “black girls” portrayed as pornographic and sexualized girls in Google Search are often said to be simply a “glitch” that they can “fix” [Noble, 2018]. However, the “glitch” culminates in predatory practices that misrepresent identity and the self. Despite the huge impacts, companies such as Google bypass by “fixing” the “glitches” and explain that they are not guilty because this is an “anomaly” in the system [Noble, 2018]. Even worse, a study has demonstrated how to manipulate people’s votes by changing the ranking in search results [Epstein and Robertson, 2015]. However, if the developers and owners of those systems are not accountable, since there is no policy to intervene, and public opinion perceives such systems as neu-

enough to cause social exclusion when users do not use them. Both social media and online search tools are examples.

¹⁰Google Search is an example of a widely used tool seen as a public resource.

¹¹See [Noble, 2018] for an example of how a search with the keywords “black girls” used to portray black girls as the stereotype of sexualized individuals in 2011.

¹²This is notably true if we consider the dominant ideologies of software engineers from Silicon Valley, which are responsible for the vast majority of systems used by individuals across the globe. See [Noble, 2018] and [Burrell and Fourcade, 2021] for more information.

tral [Burrell and Fourcade, 2021; Noble, 2018], who will be accountable for such actions? Depression is already stigmatizing and isolating [Chancellor et al., 2019]; thus, how do we deal with technologies that statistically infer the presence of depression that is not detrimental to social media users? Although the Samaritan's Radar app was created to prevent suicide, bad actors were bullying and stalking vulnerable individuals [Lee, 2014]. It is crucial to emphasize the social media platforms' responsibility to their users' mental health. Moreover, the suggestion that “glitches” are beyond human control equates to denying the problems originated from multiple sources, such as dataset bias, algorithmic malfunction, and unethical corporation attitudes.

While firms are optimized for commercial interests and market dominance, can they capitalize on public resources widely used by society? How do we make the creators of digital systems, such as software engineers, CEOs, and programmers, accountable for the impact of the systems they create? While we, as a society, do not create a better public policy to intervene in such questions, we are left behind with manipulative software that amplifies the hegemonic worldviews that users seldom have other options to replace. Furthermore, firms are structuring tacit overarching narratives through technology. Although implicit, they develop pernicious unexpected outcomes that are not socially or politically discussed.

Restricting the scope and application of the model that works on screening for depressed individuals is one step towards more fairness in such a sensitive problem. On the one hand, institutions that control the data and the resulting model use them to optimize profit. On the other hand, we have institutions whose primary goal is to create models from provided data (explicitly) and whose primary benefits are returned to the individuals who provided the data themselves. We advocate that the second case is crucial for benefiting those who help create the predictive system for screening mental health social media users. In contrast, the first case might create severe consequences for the individuals who provide the data: predatory neoliberal practices, such as pushing improved marketing to users or selling services to users who contributed with unpaid work to create those said services.

6.4.2. Technologically Mediated Behavior

This Section discusses the impacts of potentially deployed ML models to screen depressed individuals on stakeholders. The stakeholders are mental health administrators, public policymakers, individuals who are the target of ML models, social media platform administrators, psychologists, or psychiatrists. Thus, we first discuss the fundamental challenges of technology that change daily behavior and routine to discuss more deeply the role of ML models as mechanisms that change perception and behavior for the case of screening depression using social media data.

Since elementary school, we have received feedback based on our assignments, exams, and quizzes. The feedback is essential to improving students' performance, but it also influences how students perceive themselves in the world. For example, if a student frequently receives low-grade scores, it can significantly impact their opinion, ability, and perception of themselves in society [Festinger, 1954; Burrell and Fourcade, 2021].

Nonetheless, the feedback loop is not restricted to education. The presence of

feedback loops in day-to-day work and relationships has been a common element. However, with technology being more pervasive in people's lives, the feedback takes the shape of constant visualizations, assessments, scores, and recommendations [Burrell and Fourcade, 2021]. In other words, we can monitor how many steps we walk, how many hours we sleep, our body temperature, and our heart rate. While those metrics are often helpful, they promote a continuous and intermittent feedback loop that users are sometimes unprepared for.

Metrified systems also create another effect: users might rely on them to have constant feedback. However, as Byung-Chul Han argues, algorithms exclude “the possibility of the experience of contingency” [Han, 2022], which have always been an essential tool for many scientific discoveries in history. By relying on contingency and through trial and error, humans constantly arrive at surprising conclusions which leveraged unpredicted situations. On the other hand, metrified feedback encourages comparison to other individuals, often called “social comparison” [Festinger, 1954]. This phenomenon lowers the barrier to comparing oneself to friends or acquaintances, which once was subjective, but now takes place as — supposedly — objective and metrified comparisons. The *apparently* objective comparisons can severely impact individuals' self-perceptions, possibly aggravating self-esteem and subjective well-being, which might reverberate into depression and anxiety symptoms [Verduyn et al., 2017; Hwnag, 2019; Sharma et al., 2022].

Consider, for example, the use of Grammarly¹³. It often notifies users to change phrases to a more positive tone. Although it may provide some practical value for both experienced and inexperienced users, it has the potential to *nudge* our writing capacity in a particular direction: writing positively. It also has another component: its *opinion* is implicitly saying that the person is writing *negatively*. At the same time that the machine's *opinion* seems subjective (the text *seems* negative, write it more positively), it wears the mantle of objectivity because the suggestions are mathematically driven and thus seem objective. While our opinion of our ability sometimes depends on other opinions — human opinions —, unambiguous criteria, such as mathematical formulae and metrics, provide a clear path to comparison: “I am better because my score is 10 and yours is 5”. By suggesting to write positively, what are the impacts it could create? Will individuals see themselves as negative writers? Will people start writing more positively as a consequence? If society widely uses such a system, what are the consequences of suggesting everyone write positively? Another example, for instance, is when users are asked to make moral decisions. In ethical-sensitive domains, it is common to elicit the ethical values of stakeholders since any ML model will affect the stakeholders differently in varying degrees. Often, users guide their ethical values based on past and present information. However, when faced with predictive information, i.e., the opinion resulted from the prediction of an ML model or human expert, individuals' ethical preferences are directly impacted [Narayanan et al., 2022]. Moreover, they found evidence that humans prefer to rely more on the predictions of the ML model than a human expert [Narayanan et al., 2022]. In another experiment, even in light of the evidence of privacy and security issues, popular social media users did not delete nor change the privacy settings, which shows more evidence that individuals think they are immune to manipulation [Hinds et al.,

¹³A technological tool that assists users in detecting English errors while suggesting rewriting phrases for different tones.

2020]. Those experiments support the idea that ML copilots — although helpful for making decisions — have the potential to *nudge* behavior, i.e., manipulate opinion, possibly even about oneself. The worrying part is that this opinion is hegemonic and insidious. Given the same input, the model will take the same opinion (prediction).

Hannah Arendt argues that the sense of reality is mediated through shared knowledge among the community [Ballesteros, 2020]. She says that although we perceive the world through our point of view, understanding is reaffirmed because we have others perceiving the world the way we do. In that way, suppose the machine is the “other” — anthropomorphizing the machine —, by yielding constant feedback to users, the users might change the sense of reality to be coopted by the machine. General technology and AI are not only changing our perception of the world and sense of reality, but they also change how we are perceived — the said technologies of reputation, such as social media [Ballesteros, 2020]. As a result, users think that the technology they are using is innocuous, while the system acts on a pre-reflexive level [Han, 2022] to *nudge* behavior.

More formally, B. F. Skinner elaborated on the operant conditioning mechanism to explain how new behavior is learned in non-human animals. Based on positive and negative rewards, animals can keep or extinguish behavior accordingly. Animals learn new behavior through reinforcers (rewards, often food) and a reinforcement schedule, which delivers the reinforcers based on pre-defined rules. Skinner and his colleagues found that variable-time reinforcement schedules are the most effective for maintaining and strengthening behavior because the unpredictability of reward keeps the subject constantly engaged, anticipating the subsequent reinforcer [Staddon and Cerutti, 2003]. Although most experiments validate the theory for non-human animals, they also have been shown to have similar results for humans [Staddon and Cerutti, 2003].

Moreover, operant conditioning can be found in online gaming, traditional gambling games, and slot machines [Deibert, 2019]. When applying the operant conditioning theory to social media platforms, users are conditioned to expect reinforcers, such as likes, comments, or private messages — social rewards — in a variable time interval. The social rewards are assumed to share neural mechanisms with non-social rewards [Lindström et al., 2021]. Consequently, users engage in a continual expectation of new notifications and interactions (reinforcers), which in turn consolidates new addictive behavior. The immediate nature of online communication exposes individuals to immediate gratification, further strengthening addictive behavior. Social media platform addiction has been demonstrated to share neurological mechanisms with substance abuse [Turel et al., 2014; Kupferberg et al., 2016], and individuals with social media addiction experience similar symptoms to substance-related addiction symptoms [Kupferberg et al., 2016]. In a North American national survey, 32% of people with a substance use disorder also have a Major Depressive Disorder (MDD) [Carey, 2019; Xu et al., 2020]. More broadly, we advocate that addictive behavior is not only restricted to social relation expectations — such as likes, comments, and followers —, since social media platforms offer various services and experiences. For example, Facebook users might be addicted to the platform because of its games or buying and selling goods in Facebook groups.

Firms, through technology, explore fundamental laws of humans’ neurophysiology mechanisms by hacking our reward system. Based on that, firms create addictive

software designs to keep users engaged in their systems: with engagement, they generate more data to secure profits. The elaborated strategy acts upon the reward system's affective dimension because it is faster and more effective than a reasoned argument [Han, 2022]. Instigating excitement, emotions, and engagement through simple images (memes) is more straightforward than a 10-page essay.

While the system slowly directs the user behavior in a particular direction, users often believe these tools to be neutral and non-biased [Noble, 2018; Burrell and Fourcade, 2021]. This false perception creates an ominous problem: technology changes user behavior without the user's consent or awareness. Even more, it transforms the users into hostages because not using the digital systems equates to social isolation or unproductivity, which are deemed necessary in our fast and hyperconnected society. In the context of screening for depressed individuals, the same problems can happen. Individuals using any digital system today are prone to be tracked and submitted to psychometrics, as explained in Section 6.4.1. Therefore, in the context of commercial applications, firms have the potential to find latent variables connected to depressed individuals based on digital traces. When users engage with the system, they also receive feedback, such as a system that alerts them that their written text contains negative emotions or that they are frequently recommended to depressed-related content. The Filter Bubble is crucial to understanding how users are constantly exposed to no content other than the content recommended by the system. The reception of highly channeled content related to depression potentially reinforces the identity of the self as a depressed individual. Despite the efforts to leave the Filter Bubble, individuals will face other systems that also use their data to recommend the same content. This perpetual cycle provokes rumination, especially for negative thoughts, which have a significant role in worsening or maintaining depressive episodes [Cooney et al., 2010].

Moreover, even when individuals can see and control the predictions about their identity in social media systems, it could not be beneficial. Data transparency is ubiquitously considered a good feature; however, as we have been demonstrating through this study, the predictions of ML models have the power to shape opinion. Thus, there are two negative sides when explicitly showing predictive information to social media users: first, transparency paves the way for feedback loops that improve the firms' systems, which then return as paid services or enhanced marketing to users — resulting in unpaid work; second, data transparency could work against the capacity of individuals to self-determine once the predictive opinion is seen as objective and factual truth, thus overriding the individual's opinion about oneself [Engelmann et al., 2022]. In a focus group, one individual said “The fact that if it was an algorithm, and they were looking like, ‘Hey, we think you're feeling low right now.’ I feel like it might make me feel even more low” [Mikal et al., 2016]. From another angle, surveillance and data transparency are linked: “it is not people but information that is truly free” [Han, 2022]. The consequence is that people's data are truly transparent, while domination itself is not transparent, and neither the black-box nature of ML models is transparent [Han, 2022].

Differently, when screening for depression is not embedded into a larger system — such as a search tool or social media — the process will take a different principle. We argue that an ML model exclusively used to screen depressed users mediated through public entities and strict policies should produce a more secure outcome. The impact will

not be directly on the perception of the average user of digital systems but on who uses the model. The “user” (or stakeholder) will be general practitioners, psychologists, psychiatrists, healthcare administrators, and public policy makers. In that scenario, the problem associated with a lack of education in users is reduced to a few professionals whose education could include ethics in technology. Despite general practitioners being educated on the usage of ML models, they are still prone to changing their perception when using the model, which hints at the question: how will this model impact the perception of those professionals?

On the one hand, even if the average general practitioner is educated or trained to use the tool, they can blindly rely — often unintended — on the technology. In a hypothetical situation, the general practitioner did not expect the patient to be depressed, but the model has alerted about the possibility of depression. The general practitioner might be biased to agree with the model and treat the patient as depressed, although the model functions as a support mechanism and not as a decisive truth. Though the general practitioner gives the last word, the black-box nature of the model’s decision cuts any aspect of communicative rationality [Han, 2022]. For example, peer discussion about a patient’s diagnosis and treatment is essential to any medical residency. By relying on a decision without further explanation or dialogue, communicative rationality is eliminated from the process, which is so profoundly important to shape rationality and, thus, to reason about any object. Although the model might help raise awareness of the possibility of depression, it still can not offer any explanation or reason — at least not how they are conceivable today.

Another possible collateral effect is the *belief* in technology. With the generalized perception that any technology is improving in ways that challenge our rationality, technology can turn into artifacts that hold truthfulness. For example, a study demonstrates that American users often think search engines are an unbiased source of information and even believe that what is shown is true [Noble, 2018]. Once, those technologies were not perceived as trustworthy; however, as firms improve their services, perception slowly shifts to an ominous *belief* state. Currently, ML models to screen depressed individuals are limited by construction, both in terms of performance and in the input signals they leverage. However, better models will not only rely on textual or visual cues: they can also leverage the improvements in IoT, such as capturing data, but not limited to, for example, blood pressure, body temperature, voice tone, video, and audio. The future is uncertain, but as models improve performance scores, they might reach a state of *belief* where they are seen as inherently better than human capacity in the task. The risk is when general practitioners see their abilities overtaken by the model’s predictive performance. What is the impact of ML models tasked to screen depressed individuals if they are deemed better than humans? Though this is not happening now, preparing for what might come is vital.

On the other hand, a different opinion (from the AI model) might bring attention to an otherwise forgotten or unpracticed ability. General practitioners might understand they lack the expertise to diagnose depression while the tool helps them — because it might raise the possibility of depression while they are not seeing it. It might promote the inverse of what was discussed earlier: general practitioners will prepare themselves to better diagnose depression in light of evidence that their ability is deficient. The tool here is merely a support mechanism to help the general practitioner. The perception shift here

is beneficial to the user's ability.

Improving general practitioners' ability in primary care benefits society as a whole. In primary care, patients constantly complain about many symptoms, such as lack of concentration, sleep disturbances, forgetfulness, back pain, or headaches. These physical symptoms presented to general practitioners are often somatization processes associated with depression. Frequently, patients attribute emotional distress through physical manifestations of pain rather than psychological processes [Lipowski, 1990], which are called somatizations. Thus, allied to the fact that general practitioners often misdiagnose depression — only 10%–60% of cases are correctly diagnosed [Löwe et al., 2008] —, they prescribe unnecessary exams and medications based on these physical manifestations rather than the primary cause — depression. Therefore, inadequate treatment will bring risks to the patient's health without any benefit. At the same time, it will incur a waste of public resources, burdening the public health system. By relying on ML models to help identify the possibility of depression, the general practitioner will depend less on invasive diagnostic tests and treat the disorder accordingly without wasting public resources. Similar to what happened with Plato's writing inquiry, the benefits might vastly outweigh the issues.

In conclusion, we should worry about the entities who create the technology, but we also need to understand the impacts on users. One technology might cause significant harm to democracy while providing benefits to users, yet the users might be blind to such harm. Users should be educated, but more is needed to engage private corporations to include ethical processes in their pipeline.

6.4.3. Data, Inequity, and Misrepresentation

“Democracy is degenerating into infocracy” [Han, 2022]. Information — or data — has been commodified; consequently, all internet users have also been commodified since they are data producers. Although the internet is simple to join — you only must have a device and an internet connection —, it is far from a just and fair place. The internet is open, but often, those who use it and produce data are from specific hegemonic social hierarchies propagated through the digital medium. The evident problem of misrepresentation in traditional media is carried and amplified to the world of the internet [Noble, 2018]. The hegemonic representation (or opinion) will prevail over the representation of minorities' opinions due to firms' commercial interests and statistical processes. For example, the — English — data used to train GPT-2 and GPT-3 is heavily based on Reddit content, in which 67% of users are men between 18 and 29 years old [Bender et al., 2021]. Moreover, the content used to train GPT-3, for example, is highly filtered to remove unintelligible data or undesirable content, which might contain, among others, content created by marginalized communities that should not be filtered [Bender et al., 2021].

Such heavily skewed data is not only a matter of statistics but should open our eyes to the hegemonic data that originates hegemonic representations in technological systems. While general-purpose systems and ML models rely on statistics, in which case the most prevalent data will be emergent in any such system, the issue of misrepresentation will persist. Those said “neutral systems” are not neutral [Noble, 2018; Burrell and Fourcade, 2021]. Though exact and deterministic, the mathematics behind the systems is situated in a complex social body. Their mathematical processes are inherently political because

they are capable of changing behavior, and therefore they are also moral and ethical. This fight for representation hits directly into marginalized communities.

The excluded — or underrepresented — data is not only statistically insignificant to emerge but also opens another avenue to marginalization. The predominance of content related to marginalized communities on the internet is often not how marginalized communities self-perceive themselves. To illustrate this situation, a Google search with the keywords “black girls” in 2011 yielded several pages with porn content [Noble, 2018]. From hundreds of thousands of pages that could be returned to “black girls”, the Google algorithm decided to return porn content; this is not unintentional. This could be a direct consequence of commercial interests drawn from the broad appeal of the hegemonic opinions circulated online. In other words, Google search users intend to see this type of content by searching “black girls”. However, it could be a combination of both. Either way, this is not how black girls see themselves. As a consequence, historical issues of enslavement and portraying black women as sexual objects are still perpetuated through digital systems by hegemonic representations [Noble, 2018].

Users of social media platforms also need to be educated on how their data are being factually used. They lack knowledge about data permanence, thinking that deleting their posting history will render them immune to tracking, post-processing, or aggregate analysis [Mikal et al., 2016]. If users were to be educated on how their data are being used in post-processing mechanisms and how much profit is generated from their content production, they could safeguard themselves in a way that could significantly change how companies handle data.

Another issue related to social media data is the bias related to content production. As aforementioned, social media users tend to be skewed toward young adults, although the number of elderly individuals has been increasing [Center, 2021]. As such, models to screen depressed individuals trained with as much data are prone to perform better for young adults simply because they contain the most active individuals in online communication. In contrast, a sample of elderly individuals might not enjoy the same performance benefits from the model as young adults because they are not as present online. Another issue might be related to digital literacy, as elderly individuals still need to be educated on digital communication. As code (or programs) sets the range of usability, it enables and disables individuals and groups [Youmans and York, 2012]. Although social media platforms are open and accessible, underrepresented groups might not feel welcome in such places for various reasons, such as moderation practices or cyberbullying [Bender et al., 2021]. Hence, ML models relying on datasets where structural — coded — gatekeeping mechanisms impair a wide array of individuals to participate are inevitably prone to be biased. Therefore, using various data sources, looking particularly for sources where underrepresented groups are more prone to use, is one way of helping bridge the performance gap between different samples. Mainly, understanding where the sample of depressed individuals is posting or feel most comfortable posting is critical to accessing abundant sources of data to create better and robust ML models for screening depressed individuals.

While firms are optimized for commercial interests, they seldom reserve time or interest to understand their systems’ “bugs” or “glitches”. As data is the fundamental

commodity for those firms, they optimize their systems to keep users engaged and productive; in other words, they generate more data. For instance, racist, divisive, or conspiracy theory content promotes increased engagement, securing higher profits. Moreover, even when software engineers actively moderate such content, commercial interests inevitably stand in the way. It is easier and more profitable to keep users engaging freely — with possibly divisive content — than restricting them, resulting in less content [Roberts, 2016]. From interviews with commercial content moderators, Sarah T. Roberts found that despite the employees arguing that specific content — blackface, for instance — is undesired, the company decided to keep it [Roberts, 2016]. The threshold for keeping or removing content is a careful balance between profit and the company’s public image. Moreover, even entire teams on machine-learning ethics have been dismantled, as happened with Twitter in 2022 [Chayka, 2022].

As the prominent data will be created through addictive practices and well-engineered algorithms to keep users engaged, the online data will inevitably be biased. If researchers do not carefully evaluate the data they are feeding to ML models, it will carry the same bias and misrepresentations in the data. Even worse, the model could amplify the misrepresentation and cause unpredicted harm, mainly directed at marginalized communities.

For the task of screening depressed individuals, there are key elements that distinguish who will benefit more or less from such technology. For example, as data online is prominently in English, native English speakers will have the advantage of having models trained on more data than native Brazilian speakers. Not only quantity determines model performance, but also quality. Native speakers of not-so-famous languages are left with fewer and low-quality data samples, resulting in less desirable performant ML models. Even if we consider transferring the knowledge from one ML model trained on a more extensive dataset to start training another ML model to work with a small dataset, we suffer from the portability trap [Selbst et al., 2019]. It means the social context in which the ML model is trained, such as the data and its type (which media), where it was produced (which social media), and from which sample (e.g., median age, nationality) might mislead or harm when used to train models to be embedded in other social contexts. Thus, it is imperative to correctly model the social context and possible ethical implications when transferring knowledge from one model to another.

Furthermore, using pre-trained models — trained on large datasets of generalized text or images — to improve the performance of the downstream task might pose other risks. First, as we have argued, those said “generalized” texts or images often suffer from bias and misrepresentation issues. Second, the said “pre-trained” models, also named “foundation models” [Bommasani et al., 2021], are a standard block used in many classification systems. As foundation models are trained on a massive corpus of data through deep learning algorithms, they will not only eclipse less frequent content but can also amplify the hegemonic content. As the current ML models for screening depressed individuals heavily rely on foundation models to extract good textual or visual representations for classification [Trotzek et al., 2018b; Orabi et al., 2018b; Aragón et al., 2019; Mann et al., 2020, 2022; Bucur et al., 2023], they also inherit the bias and misrepresentations included in the foundation model. On top of that, fine-tuning the foundation model to specialize it to the domain data, a standard practice in many automatic systems that screen individuals’ data for depression, does not necessarily eliminate these issues — on the

contrary, it could perpetuate and amplify them.

As we approach a state of singularity in Computer Vision and Natural Language Processing (NLP) by using the same foundation models for various tasks, we risk perpetuating and spreading bias and misrepresentations to many aspects of the field. Furthermore, though unexpected emergent properties that arise in models are sometimes beneficial, there are times they may create undesirable collateral effects. For example, the emergent property of in-context learning that emerged from GPT-3 training, which is the capacity to adapt the language model to a downstream task based on a given context prompt as input, is attractive because it helps solve many tasks with higher accuracy and adaptability. However, it also might be the cause of hallucinations. As emergent properties are primarily unpredictable and need extensive empirical validation, it is too unstable to be reliable for screening depressed individuals using social media data.

Furthermore, GPT-3 and GPT-4 lays the groundwork for the vast improvement of Natural Language Generation (NLG). Differently from what is argued by [Bommasani et al., 2021], all data is not only created by people, nor only by people for other people. GPT-3 or GPT-4, a Large Language Model (LLM), can automatically generate data for diverse tasks with high capability. The internet, full of racist, homophobic, and divisive content, now faces another challenge: the content created by advanced NLG models, such as GPT-3 and GPT-4. Those models not only perpetuate the misrepresentations of the internet but also accentuate them by automatically creating content quicker and more effectively — though not always correct — than any human being. The risk of such stochastic parrots [Bender et al., 2021] for screening depressed individuals is that they could automatically generate “depressed-related” content. The boundaries for human-made content are blurry since we now have LLMs capable of creating human-like textual and visual content. Consequently, we advocate practitioners not using an LLM to “augment” datasets for screening depressed individuals, at least as conceivable today, because of the many inherent risks associated with the content they produce.

6.4.4. Enacting Change

From previous sections, we can observe the persistent tension between opposing interests or motivations that can explain the main challenges in screening MDD using social media data. In this Section, we explore two key areas to discuss further the above tensions for screening depressed individuals in social media: education and data protection laws. Finally, we conclude this Section by stressing a few suggestions highlighted from the discussed tensions and the proposed critical areas of discussion.

6.4.4.1. Education

There are two main ways of conceiving AI and education today: We can use AI to help educate individuals or educate individuals to prepare for the age of AI — here, we will focus on the latter. The same could be said for mental health disorders, where individuals still need to learn or improve their knowledge about them. For 193 member states in the United Nations, Quality Education (SGD-4) and Good Health and Well-Being (SGD-3) are among the 17 shared Sustainable Development Goals (SDG).

Education will be highly impacted by AI, be it by newly incorporated ways to learn and educate or by new ways of interacting and using technology that inevitably change the desired abilities when companies hire personnel. Nonetheless, the current United Nations SDG report states how the COVID-19 pandemic highly impacted education and advocates for higher investments in education and embracing technology to improve education [of Economic and Affairs, 2023]. While severe primary educational concerns remain, digital literacy, fundamental for benefiting and avoiding AI's dangers, is still largely lacking [of Economic and Affairs, 2023]. Digital literacy can be defined as “a set of skills required by 21st Century individuals to use digital tools to support the achievement of goals in their life situations” [Reddy et al., 2020].

We add to the above definition the necessity to understand how Information Technology (IT) companies are commercially motivated and how the data produced by individuals can be used for several purposes, as discussed in Section 6.4.1. Because companies explore ways to maximize profit at the expense of social and humanitarian goals, citizens must inform themselves that companies are consciously creating ways to explore human psychological vulnerabilities to expand engagement. Therefore, behavior is altered to promote engagement and addiction, which possibly culminates in anxiety and depression in expectation of social rewards, as discussed in Section 6.4.2.

Mental health disorders are still largely stigmatized and isolating [Chancellor et al., 2019]. As such, pushing the education boundaries to help citizens understand and recognize mental health disorders is crucial. To that, Mental Health Literacy (MHL) is essential to help promote recognizing, managing, and preventing mental health disorders [Jorm et al., 1997]. It can be defined as “understanding how to obtain and maintain positive mental health; understanding mental disorders and their treatments; decreasing stigma related to mental disorders; and enhancing help-seeking efficacy (knowing when and where to seek help and developing competencies designed to improve one's mental health care and self-management capabilities)” [Kutcher et al., 2016]. MHL shapes how lay individuals understand mental health disorders and seek appropriate help — otherwise, individuals might not even know they have depression.

For the tensions we have discussed in this text, we focus on how to educate students — or society, more broadly — to prepare for the age of AI, notably through the lens of applications to screen depressed individuals as one of the fundamental ways to enact change. Hence, digital and mental health literacies are essential for individuals to preserve autonomy and privacy and empower individuals in political contexts. We will explore two notable applications to understand better the potential of MHL and digital literacy.

For example, the Samaritans, a philanthropic organization that provides help to individuals with emotional distress and suicide ideation, created an app in 2014. This app, called “Radar”, was aimed at Twitter users to detect signs of suicide or depression [Hsin et al., 2016]. However, the app was designed to alert the user when one of their friends exhibits the said signs of suicide or depression, even if their friends did not opt-in to participate. The signs of depression were determined by text matching with pre-determined phrases such as “I am sad” and “I want to kill myself”, i.e., based on a keyword list. Such a simplistic approach to detection is easily prone to errors and misclassifications. This

raises serious concerns regarding privacy and accuracy because users are being analyzed without explicit consent, and the accuracy is questionable. Even worse, bad actors downloaded the app to encourage targeted individuals by the app — possibly suicidal — to take their lives.

Similarly, Facebook also offers the service of suicide prevention¹⁴. Facebook implemented two main strategies for suicide prevention, identifying potential individuals by (1) reactive reporting and (2) proactive reporting [Gomes de Andrade et al., 2018]. Reactive reporting is based on the collaboration of Facebook users that can “flag” whether a post demonstrates suicidal ideation, similar to reporting a post for violating community rules. The user who reported the post is then prompted to offer help directly or to delegate the responsibility to Facebook human evaluators. However, the Facebook staff noted that several posts containing suicidal ideation were not being reported. As such, they developed proactive reporting based on ML models. The model can identify posts from users at-risk, which, based on a threshold, a report is sent to the Community Operations to review the post and send resources, if applicable [Gomes de Andrade et al., 2018].

The primary issue learned from the Samaritans’ Radar app fiasco is that privacy is vital: Who can access the information about screening depressed individuals on social media? We argue that access to this information needs to be very carefully controlled by mental health professionals. The user did not explicitly opt-in (agree to give explicit consent) to participate and might not even know they have been screened for depression and suicide risk. Equally important, the Facebook staff can infer suicidal ideation based on human and automatic evaluations. Are Facebook users aware that their friends can report their posts for containing the risk of suicide? Moreover, as a consequence, Facebook can learn about users’ mental states by relying on the evaluation of Facebook users and the Community Operations personnel. While the Facebook service for suicide prevention is not as open to bad actors as happened to the Samaritans app, the service is embedded in a social media platform with commercial interests.

As discussed in Section 6.4.1, social media platforms, — and other mega-technological corporations — allied to advertisement strategies to influence individuals to purchase products may rely on vast psychometrics data. Based on this data, they could act like bad actors like the Samaritans’ Radar app. Not only that, they have the potential to trigger and induce depression, stress, and anxiety, as we discussed in Section 6.4.2. The question is, do they have the right to induce such information based on self-generated data by their users?

With rapid AI development, institutions have yet to create better public policies. However, one of the best ways to avoid these pernicious practices is to increase society’s digital and mental health literacy through education strategies. In a similar vein, the World Health Organisation has underscored the fact that health literacy¹⁵ is a better predictor of health than many other factors, such as income and employment status [Furnham and

¹⁴Although it is not explicitly designed to screen depressed users, we note that the comorbidity of suicide ideation and depression is reported to be the highest among several mental disorders [Henriksson et al., 1993]

¹⁵MHL was inspired by health literacy, which could broadly be defined as “the ability to gain access to, understand, and use information in ways which promote and maintain good health” [Jorm et al., 1997].

Swami, 2018]. Although there is still a need for a body of evidence to state the same thing for mental health literacy, it is clear that there is an urgency to simultaneously increase the digital and mental health literacies to improve over the many discussed challenges. Therefore, by understanding how commercial interests are aligned with practices to addict and promote engagement, citizens need to protect their interests in favor of their (mental) health.

6.4.4.2. Data and Legislation

Data protection laws came to help in a longstanding issue regarding IT companies exploiting personal data without transparency on how they were using it or by obtaining explicit consent from individuals. Two notable protection laws are GDPR for the European Union (EU) and the Brazilian LGPD. Now, companies must comply with their countries' data protection regulations and the data protection laws of the countries where their companies are partners. Although both regulations improve towards protecting social and human rights in the era of AI, their implementations happened very recently — GDPR in 2016, and LGPD in 2020 —, and the effect of these laws are in its infancy.

Both data protection laws share the concept of "controller" and "data subject" ("*titular*" in LGPD). The controller is "any legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of processing personal data" [Parliament and of the European Union, 2016]. In LGPD, the data subject is defined as "a natural person to whom the personal data that are the object of processing refer to". Furthermore, personal data can be classified as sensitive, which refers to data concerning health, ethnic origin, political opinion, and genetic or biometric data when related to a natural person, among others.

Both LGPD and GDPR assume that any processing of personal data must be directly communicated and only conducted over explicit approval from the data subject. At the same time, processing personal sensitive data in GDPR and LGPD are forbidden except when the data subject provided explicit consent; however, there is a legal basis for processing sensitive data even without explicit consent from the data subject. Notably, among the seven exceptions in LGPD, we stress two of them most related to mental health disorders: (1) "to protect the health, exclusively, in a procedure carried out by health professionals, health services or sanitary authorities"; (2) protecting life or physical safety of the data subject or a third party. Processing personal sensitive data for health interests is also supported by GDPR but only under the responsibility of a professional obliged to professional secrecy [Parliament and of the European Union, 2016].

The data related to the mental sphere includes the possibility and risks of identifying mental disorders and the identity of data subjects. As such, when the controller can obtain psychometrics data to single out individuals, it violates privacy under the GDPR and LGPD laws. However, as Ienca et al. [Ienca and Malgieri, 2022] pointed out related to GDPR legal basis, "the data revealing thoughts or memories are not automatically sensitive data just because they refer to the 'mental sphere'". When the data is not classified as sensitive, the possibilities to process and commercialize the data are less restricted than sensitive data. Mainly, non-sensitive data in LGPD can be processed for new purposes

— such as financial gains —, while subsequent processing for sensitive health data is explicitly prohibited for commercial gains.

Although GDPR explicitly provides a legal basis for mental ill-health issues, based on the “data concerning health” that explicitly includes mental health, there is no particular mention of mental health in LGPD, only to the broader “health” term. Moreover, while both laws provide specific guidelines to guarantee the data subject’s privacy and self-determination, there is still a gap, especially for LGPD and mental health disorders, that legal entities can explore on a legal basis, even if such actions may not be ethically or morally sound. As suggested by other authors, this gap could be explained by the lack of more specific categories in processing sensitive data, such as including “emotions”, “desires”, “thoughts” (textually) as data that could be used to identify a natural person uniquely or infer several other personal dimensions [Ienca and Malgieri, 2022].

Therefore, as discussed in Sections 6.4.1 and 6.4.2, legal entities are in a position of obtaining, processing, and inferring depression and identity lawfully. Although other obligations apply, such as asking for consent, there is a clear gap in the legislation that allows treating such data as non-sensitive. Even when psychometrics data might be considered sensitive, both GDPR and LGPD allow processing them under the responsibility of a professional subject (in the case of health-related services) or by acquiring explicit consent from the data subject as defined in Article 9(2)(a) and 9(2)(h) in GDPR.

For the case of social media platforms, while legal entities can capture psychometrics data to infer other attributes lawfully, such as mental health states, users agree to Terms of Services based on what is called “weak consent regimes” [Ienca and Malgieri, 2022]. Social media platform users consent without reading the Terms of Services, or even when they read, it often contains complex language and specialized terminology that individuals do not understand [Obar and Oeldorf-Hirsch, 2020]. Consequently, obtaining consent from social media platform users under LGPD’s Article 7 (and 11 for personal sensitive data) and GDPR’s Article 6 (and 9 for personal sensitive data) is straightforward.

Although there might be legitimate interest from data subjects to self-track depression or suicide ideation, we note that the gap in treating textual thoughts and emotions allows legal entities to subsequently process the information for commercial interests, such as using psychometrics for advertising. Legal entities can proceed under this path lawfully, especially under LGPD, which is a legitimate interest only to the legal entities [Ienca and Malgieri, 2022]. However, we note several risks associated with collecting and inferring depression in an individual, as discussed in Sections 6.4.1 and 6.4.2. Besides these risks, Chancellor et al. (2019) [Chancellor et al., 2019] mentions the risks of advertisement for prescription drugs, credit score based on mental health state, and health insurance raising premiums because of existing mental disorders. Are there enough reasons for legal entities to lawfully share and process self-generated data, especially under the risks of manipulating behavior and opinion and amplifying existing prejudice?

While data protection laws are still improving, especially in light of the rapid AI development, individuals also need to understand how their data could be used without the scope of their legitimate interests. Data protection laws enforce privacy, but we argue that it is still insufficient to protect mental data, as it could be used to infer mental health disorders or other mental states [Ienca and Malgieri, 2022]. Remarkably, the two

mentioned laws have different approaches to mental health disorders, which is another relevant point of discussion — while GDPR is more robust to mental health data, LGPD is more flexible.

Thus, similarly to digital and mental health literacies, there is still an urgency to educate individuals to understand how legal entities are commercially motivated to use their data. Therefore, data literacy is another important aspect of today's digital world, where general-purpose technology, specifically social media, is present in virtually all aspects of life. Individuals must understand and be able to protect their self-determination and legitimate interests and reliably reject abusive data collection practices. Another way to help individuals in this direction is to promote legislation and Terms of Services in a simplified (plain language) way so that they can further comprehend their privacy rights and how legal entities are using their data.

6.4.4.3. Discussion

Based on the three main areas of tension we explored in Sections 6.4.1, 6.4.2, and 6.4.3, and the two main approaches to mitigate the tensions in Sections 6.4.4.1 and 6.4.4.2, here we propose practical suggestions to mitigate the political, social and ethical challenges of screening depressed individuals in social media.

As we have discussed, data protection laws need to be more comprehensive to protect individuals in terms of inferring mental states from raw visual and textual self-generated content. Consequently, policymakers, public administrators, and legislators should consider including mental data as personal sensitive data. More importantly, we advocate explicitly including mental health and mental data into LGPD to prevent misuse and promote the privacy and safety of individuals. In this way, the mental data used to predict MDD, for example, would be strictly used either (1) under the explicit consent of the individual as a tool for self-assessment and with the legitimate use only for that purpose; or (2) to use for purposes of research or medical diagnosis, the provision of health care or treatment conducted by a professional obliged to a professional oath.

Although current data protection laws offer some security and privacy to personal sensitive data, using the data for only the two scenarios is very restrictive. As such, we call for a public debate to discuss privacy versus the greater social good. On the one hand, less privacy would improve the performance of ML models since they rely on more data. On the other hand, more privacy restricts the performance — and generalizability — of ML models or concentrates the power to a few legal entities with power, as currently happens with the Big Techs. Furthermore, there is the public versus private debate: Should the government oversight and regulate more or less? Many restrictions and oversight would result in an Orwellian monitoring [Mikal et al., 2016], but too few restrictions could lead to predatory and unethical practices.

One way to enforce restrictions without creating an Orwellian society is to change the dynamics underpinning predatory practices that alter behavior and, consequently, opinion and thoughts. As such, Governments could focus on the main strategies legal entities are using that impose several risks to the mental health of individuals. This suggestion relies on the fact that depression and other mental health disorders in the era of

technology might correlate with technological addiction [Han, 2015; Lin et al., 2016; Primack et al., 2017]. Consequently, dedicating efforts to changing the dynamics that motivate corporations might improve collective mental health more extensively than dedicating resources to screen depressed individuals using ML models. We call for public opinion and more research towards this open debate.

Another suggestion is that Governments need to ensure the right actors are held responsible for their actions. Accountability is essential to ensure that legal entities comply with the legislation and data protection laws. When Big Tech giants say it is “a bug” or “a glitch” in the system, it is not. It is a critical problem in the system that relies on the cacophony of prejudice or errors in human-generated data — and now, the LLM data.

There is also the concern of using personal data to infer depression. Users of social media platforms might agree that using individual data to infer mental disorders, such as depression, is undesirable. However, if the platform could create a mechanism to create population-based algorithms instead of user-based algorithms, we could advance research by understanding the behavior of individuals suffering from depression, for example. Thus, such findings can help public policymakers create better policies and strategies to improve services and social welfare. When using a population-based method and ensuring that no individual can be singled out, we incur fewer risks of violating privacy. We advocate creating models that can work on aggregate-level data without access to individual-level data in ways that would be impossible to identify individuals.

When implementing the ML model, the debate of intervention versus observation must also be considered. For research, intervention is complex and limited, given the unstable nature of the developed models and limited resources. For commercial or governmental interests, deploying models on real-life situations might help if aligned with intervention strategies. Consequently, training the relevant personnel on the inherent risks, benefits, and optimal utilization of observation tools for effective intervention is crucial. The goal of this entire process should be focused on the appropriate intervention strategies that will ultimately benefit society.

Education is fundamental to the professionals involved in the observation and intervention aspects of screening depressed individuals on social media. However, society must also be educated. On the one hand, individuals lack fundamental knowledge about data literacy; on the other hand, they also lack appropriate knowledge about mental health literacy. While technology consistently relies on data to create improved services, micro-targeted advertisements, and addictive interfaces and mechanisms, individuals often lack the knowledge of minimum safeguards. As such, they use predatory services at the expense of their mental health. We advocate for improving mental health and data literacy.

6.5. Final Considerations and Future Perspectives

The use of artificial intelligence (AI) in supporting the diagnosis, treatment, and prognosis of mental disorders, especially depressive and anxiety disorders, presents significant potential and challenges. It is crucial to reflect on the current achievements and future directions in this interdisciplinary field.

AI’s capacity to analyze large amounts of data from social media and other digital

sources offers a unique opportunity for early detection and intervention of mental health issues. Integrating machine learning models with psychological and psychiatric practices can improve the accuracy and efficiency of screening mental disorders, leading to more personalized and timely treatment plans. Additionally, AI-driven tools can provide continuous monitoring and support, aiding individuals in managing their conditions more effectively.

However, several challenges remain. The ethical implications of using AI in mental health care are vital. Privacy, data security, and informed consent issues are especially crucial when dealing with sensitive personal information. Researchers and practitioners must navigate these concerns carefully to ensure that the deployment of AI tools respects individuals' rights and complies with regulatory standards such as the General Data Protection Regulation (GDPR) and the *Lei Geral de Proteção de Dados* (LGPD) in Brazil.

Another critical area for future research is the development of more robust and inclusive AI models. Current models often face limitations due to biased training data, which can result in inaccurate predictions and worsen health disparities. Efforts should be directed towards creating diverse and representative datasets and implementing fairness-aware algorithms that mitigate biases and promote equitable outcomes across different demographic groups.

Interdisciplinary collaboration is essential to advance the field of AI in mental health care. Partnerships between computer scientists, mental health professionals, ethicists, and policymakers can promote the development of innovative solutions that are both technically sound and ethically responsible. Furthermore, involving patients and the broader public in the design and implementation of AI tools can enhance their acceptance and effectiveness.

Looking ahead, integrating AI with other emerging technologies such as wearable devices, virtual reality, and telehealth platforms holds great promise. These technologies can complement AI applications by providing real-time data, immersive therapeutic experiences, and remote care options, making mental health services more accessible and responsive to individual needs.

In conclusion, the application of AI in mental health care represents a transformative approach to addressing the growing burden of mental disorders. By continuing to explore new methodologies, address ethical and technical challenges, and foster interdisciplinary collaboration, we can harness the full potential of AI to improve mental health outcomes and enhance the well-being of individuals and communities worldwide.

References

- Charu C Aggarwal et al. Neural networks and deep learning. *Springer*, 10(978):3, 2018.
- Yemi Aina and Jeffrey L Susman. Understanding comorbidity with depression and anxiety disorders. *The Journal of the American Osteopathic Association*, 106(5 Suppl 2):S9–14, 2006.
- Mario Ezra Aragón, Adrian Pastor López-Monroy, Luis Carlos González-Gurrola, and Manuel Montes. Detecting depression in social media using fine-grained emotions. In *Proc. of the 2019 NAACL-HLT, Volume 1 (Long and Short Papers)*, pages 1481–1486, 2019.

- American Psychiatric Association et al. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.
- Shrey Bagroy, Ponnuram Kumaraguru, and Munmun De Choudhury. A social media based index of mental well-being in college campuses. pages 1634–1646. Association for Computing Machinery, 2017. ISBN 9781450346559. doi: 10.1145/3025453.3025909. URL <https://doi.org/10.1145/3025453.3025909>.
- Alfonso Ballesteros. Digitocracy: Ruling and being ruled. *Philosophies*, 5(2):9, 2020.
- John Bateman. *Text and image: A critical introduction to the visual/verbal divide*. Routledge, 2014.
- Aaron T Beck, Robert A Steer, and Gregory Brown. Beck depression inventory–ii. *Psychological Assessment*, 1996.
- Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. On the dangers of stochastic parrots: Can language models be too big? In *Proc. of the 2021 ACM FAccT*, pages 610–623, 2021.
- Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1798–1828, 2013.
- Adrian Benton, Glen Coppersmith, and Mark Dredze. Ethical research protocols for social media health research. In *Proceedings of the first ACL workshop on ethics in natural language processing*, pages 94–102, 2017.
- Rishi Bommasani, Drew A Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, et al. On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*, 2021.
- Ana-Maria Bucur, Adrian Cosma, Paolo Rosso, and Liviu P. Dinu. It’s just a matter of time: Detecting depression with time-enriched multimodal transformers. In Jaap Kamps, Lorraine Goeriot, Fabio Crestani, Maria Maistro, Hideo Joho, Brian Davis, Cathal Gurrin, Udo Kruschwitz, and Annalina Caputo, editors, *Advances in Information Retrieval*, pages 200–215, Cham, 2023. Springer Nature Switzerland. ISBN 978-3-031-28244-7.
- Jenna Burrell and Marion Fourcade. The society of algorithms. *Annual Review of Sociology*, 47: 213–237, 2021.
- Theadia L Carey. Use of antidepressants in patients with co-occurring depression and substance use disorders. *Antidepressants: From Biogenic Amines to New Mechanisms of Action*, pages 359–370, 2019.
- Pew Research Center. Social media use in 2021, 2021. URL <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/>. Accessed: 2023-08-14.
- Junyeop Cha, Seoyun Kim, and Eunil Park. A lexicon-based approach to examine depression detection in social media: the case of twitter and university community. *HUMANITIES & SOCIAL SCIENCES COMMUNICATIONS*, 9, 6 2022. doi: 10.1057/s41599-022-01313-2.

- Stevie Chancellor and Munmun De Choudhury. Methods in predictive techniques for mental health status on social media: a critical review. *NPJ digital medicine*, 3(1):1–11, 2020.
- Stevie Chancellor, Michael L Birnbaum, Eric D Caine, Vincent MB Silenzio, and Munmun De Choudhury. A taxonomy of ethical tensions in inferring mental health states from social media. In *Proceedings of the conference on fairness, accountability, and transparency*, pages 79–88, 2019.
- Kyle Chayka. A twitter employee’s account of surviving layoff day. <https://www.newyorker.com/culture/infinite-scroll/a-twitter-employees-account-of-surviving-layoff-day>, 2022. Accessed: 2023-01-25.
- Qijin Cheng, Tim M H Li, Chi-Leung Kwok, Tingshao Zhu, and Paul S F Yip. Assessing suicide risk and emotional distress in chinese social media: A text mining and machine learning study. *JOURNAL OF MEDICAL INTERNET RESEARCH*, 19, 6 2017. ISSN 1438-8871. doi: 10.2196/jmir.7276.
- David A Clark. Cognitive restructuring. *The Wiley handbook of cognitive behavioral therapy*, pages 1–22, 2013.
- Arman Cohan, Bart Desmet, Andrew Yates, Luca Soldaini, Sean MacAvaney, and Nazli Goharian. SMHD: a large-scale resource for exploring online language usage for multiple mental health conditions. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1485–1497, Santa Fe, New Mexico, USA, August 2018. Association for Computational Linguistics. URL <https://aclanthology.org/C18-1126>.
- Mike Conway. Ethical issues in using twitter for public health surveillance and research: developing a taxonomy of ethical concepts from the research literature. *Journal of medical Internet research*, 16(12):e290, 2014.
- Rebecca E Cooney, Jutta Joormann, Fanny Eugène, Emily L Dennis, and Ian H Gotlib. Neural correlates of rumination in depression. *Cognitive, Affective, & Behavioral Neuroscience*, 10(4): 470–478, 2010.
- Glen Coppersmith, Mark Dredze, Craig Harman, Kristy Hollingshead, and Margaret Mitchell. Clpsych 2015 shared task: Depression and ptsd on twitter. In *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, pages 31–39, 2015.
- Glen Coppersmith, Ryan Leary, Patrick Crutchley, and Alex Fine. Natural language processing of social media as screening for suicide risk. *Biomedical informatics insights*, 10: 1178222618792860, 2018.
- Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. Predicting depression via social media. In *Proc. of the ICWSM*, volume 7, 2013.
- Vanessa Borba de Souza, Jéferson Campos Nobre, and Karin Becker. DAC stacking: A deep learning ensemble to classify anxiety, depression, and their comorbidity from reddit texts. *IEEE J. Biomed. Health Informatics*, 26(7):3303–3311, 2022. doi: 10.1109/JBHI.2022.3151589. URL <https://doi.org/10.1109/JBHI.2022.3151589>.
- Ronald J Deibert. Three painful truths about social media. *J. Democracy*, 30:25, 2019.

- Sahraoui Dhelim, Liming Chen, Sajal K Das, Huansheng Ning, Chris Nugent, Gerard Leavey, Dirk Pesch, Eleanor Bantry-White, and Devin Burns. Detecting mental distresses using social behavior analysis in the context of covid-19: A survey. *ACM Comput. Surv.*, 6 2023. ISSN 0360-0300. doi: 10.1145/3589784. URL <https://doi.org/10.1145/3589784>.
- Joanna Frith Dipnall, JA Pasco, Michael Berk, LJ Williams, Seetal Dodd, FN Jacka, and D Meyer. Why so glumm? detecting depression clusters through graphing lifestyle-environs using machine-learning methods (glumm). *European Psychiatry*, 39:40–50, 2017.
- Wesley Ramos dos Santos, Amanda M. M. Funabashi, and Ivandré Paraboni. Searching brazilian twitter for signs of mental health issues. In *Proceedings of The 12th Language Resources and Evaluation Conference, LREC 2020, Marseille, France, May 11-16, 2020*, pages 6111–6117. European Language Resources Association, 2020. URL <https://aclanthology.org/2020.lrec-1.750/>.
- Severin Engelmann, Valentin Scheibe, Fiorella Battaglia, and Jens Grossklags. Social media profiling continues to partake in the development of formalistic self-concepts. social media users think so, too. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, pages 238–252, 2022.
- Robert Epstein and Ronald E Robertson. The search engine manipulation effect (seme) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences*, 112(33):E4512–E4521, 2015.
- Teresa M Evans, Lindsay Bira, Jazmin Beltran Gastelum, L Todd Weiss, and Nathan L Vanderford. Evidence for a mental health crisis in graduate education. *Nature biotechnology*, 36(3):283, 2018.
- Ethan Fast, Binbin Chen, and Michael S Bernstein. Empath: Understanding topic signals in large-scale text. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 4647–4657, 2016.
- Leon Festinger. A theory of social comparison processes. *Human relations*, 7(2):117–140, 1954.
- Casey Fiesler and Nicholas Proferes. “participant” perceptions of twitter research ethics. *Social Media+ Society*, 4(1):2056305118763366, 2018.
- Max Fisher. *The Chaos machine: the inside story of how social media rewired our minds and our world*. Hachette UK, 2022.
- John Miles Foley. “reading” homer through oral tradition. *College Literature*, pages 1–28, 2007.
- Sigmund Freud. Notes upon a case of obsessional neurosis. *Standard edition*, 10, 1909.
- Guanghai Fu, Changwei Song, Jianqiang Li, Yue Ma, Pan Chen, Ruiqian Wang, Bing Xiang Yang, and Zhisheng Huang. Distant supervision for mental health management in social media: Suicide risk classification system development study. *JOURNAL OF MEDICAL INTERNET RESEARCH*, 23, 6 2021. ISSN 1438-8871. doi: 10.2196/26119.
- Adrian Furnham and Viren Swami. Mental health literacy: A review of what it is and why it matters. *International Perspectives in Psychology*, 7(4):240–257, 2018.

- Norberto Nuno Gomes de Andrade, Dave Pawson, Dan Muriello, Lizzy Donahue, and Jennifer Guadagno. Ethics and artificial intelligence: suicide prevention on facebook. *Philosophy & Technology*, 31:669–684, 2018.
- C Gorenstein, WY Pang, IL Argimon, and BSG Werlang. Manual do inventário de depressão de beck–bdi-ii. *São Paulo: Editora Casa do Psicólogo*, 2011.
- Paul E Greenberg, Andree-Anne Fournier, Tammy Sisitsky, Mark Simes, Richard Berman, Sarah H Koenigsberg, and Ronald C Kessler. The economic burden of adults with major depressive disorder in the united states (2010 and 2018). *Pharmacoeconomics*, 39(6):653–665, 2021.
- Byung-Chul Han. *The burnout society*. Stanford University Press, 2015.
- Byung-Chul Han. *Infocracy: Digitization and the crisis of democracy*. John Wiley & Sons, 2022.
- Markus M Henriksson, Hillevi M Aro, Mauri J Marttunen, Martti E Heikkinen, ET Isometsa, Kimmo I Kuoppasalmi, JK Lonnqvist, et al. Mental disorders and comorbidity in suicide. *American journal of psychiatry*, 150:935–935, 1993.
- Joanne Hinds, Emma J Williams, and Adam N Joinson. “it wouldn’t happen to me”: Privacy concerns and perspectives following the cambridge analytica scandal. *International Journal of Human-Computer Studies*, 143:102498, 2020.
- Honor Hsin, John Torous, and Laura Roberts. An adjuvant role for mobile health in psychiatry. *JAMA psychiatry*, 73(2):103–104, 2016.
- Hongsheng Hu, Zoran Salcic, Lichao Sun, Gillian Dobbie, Philip S Yu, and Xuyun Zhang. Membership inference attacks on machine learning: A survey. *ACM Computing Surveys (CSUR)*, 54(11s):1–37, 2022.
- Jie Huang, Hanyin Shao, and Kevin Chen-Chuan Chang. Are large pre-trained language models leaking your personal information? In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 2038–2047, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics. URL <https://aclanthology.org/2022.findings-emnlp.148>.
- Ha Sung Hwnag. Why social comparison on instagram matters: Its impact on depression. *KSI Transactions on Internet and Information Systems (TIIS)*, 13(3):1626–1638, 2019.
- Ahmed K Ibrahim, Shona J Kelly, Clive E Adams, and Cris Glazebrook. A systematic review of studies of depression prevalence in university students. *Journal of psychiatric research*, 47(3):391–400, 2013.
- Marcello Ienca and Gianclaudio Malgieri. Mental data protection and the gdpr. *Journal of Law and the Biosciences*, 9(1), 2022.
- Werner Jaeger. A formação do homem grego. *São Paulo: Fontes*, 2001.
- Samireh Jalali and Claes Wohlin. Systematic literature studies: database searches vs. backward snowballing. In *Proceedings of the ACM-IEEE international symposium on Empirical software engineering and measurement*, pages 29–38, 2012.

- Anthony F Jorm, Ailsa E Korten, Patricia A Jacomb, Helen Christensen, Bryan Rodgers, and Penelope Pollitt. “mental health literacy”: a survey of the public’s ability to recognise mental disorders and their beliefs about the effectiveness of treatment. *Medical journal of Australia*, 166(4):182–186, 1997.
- Sean W Kelley and Claire M Gillan. Using language in social media posts to study the network dynamics of depression longitudinally. *Nature communications*, 13(1):870, 2022.
- Sheelah Kolhatkar. Lina khan’s battle to rein in big tech. <https://www.newyorker.com/magazine/2021/12/06/lina-khans-battle-to-rein-in-big-tech>, 2021. Accessed: 2023-09-06.
- Aleksandra Kupferberg, Lucy Bicks, and Gregor Hasler. Social functioning in major depressive disorder. *Neuroscience & Biobehavioral Reviews*, 69:313–332, 2016.
- Stan Kutcher, Yifeng Wei, and Connie Coniglio. Mental health literacy: Past, present, and future. *The Canadian Journal of Psychiatry*, 61(3):154–158, 2016.
- Dave Lee. Samaritans pulls ‘suicide watch’ radar app. *BBC News*, 11 2014. URL <https://www.bbc.com/news/technology-29962199>. Accessed: 2023-08-12.
- Liu Yi Lin, Jaime E Sidani, Ariel Shensa, Ana Radovic, Elizabeth Miller, Jason B Colditz, Beth L Hoffman, Leila M Giles, and Brian A Primack. Association between social media use and depression among us young adults. *Depression and anxiety*, 33(4):323–331, 2016.
- Björn Lindström, Martin Bellander, David T Schultner, Allen Chang, Philippe N Tobler, and David M Amodio. A computational reward learning account of social media engagement. *Nature communications*, 12(1):1311, 2021.
- Zbigniew J Lipowski. Somatization and depression. *Psychosomatics*, 31(1):13–21, 1990.
- Xingyun Liu, Xiaoqian Liu, Jiumo Sun, Nancy Xiaonan Yu, Bingli Sun, Qing Li, and Tingshao Zhu. Proactive suicide prevention online (pspo): Machine identification and crisis management for chinese social media users with suicidal thoughts and behaviors. *JOURNAL OF MEDICAL INTERNET RESEARCH*, 21, 6 2019. ISSN 1438-8871. doi: 10.2196/11705.
- Bernd Löwe, Robert L Spitzer, Janet BW Williams, Monika Mussell, Dieter Schellberg, and Kurt Kroenke. Depression, anxiety and somatization in primary care: syndrome overlap and functional impairment. *General hospital psychiatry*, 30(3):191–199, 2008.
- Paulo Mann, Aline Paes, and Elton H Matsushima. See and read: detecting depression symptoms in higher education students using multimodal social media data. In *Proc. of the ICWSM*, volume 14, pages 440–451, 2020.
- Paulo Mann, Elton H Matsushima, and Aline Paes. Detecting depression from social media data as a multiple-instance learning task. In *Proc. of the ACII*, pages 1–8. IEEE, 2022.
- Priya Mathur, Amit Kumar Gupta, and Abhishek Dadhich. Mental health classification on social-media: Systematic review. Association for Computing Machinery, 2023. ISBN 9781450399937. doi: 10.1145/3590837.3590946. URL <https://doi.org/10.1145/3590837.3590946>.

- Jude Mikal, Samantha Hurst, and Mike Conway. Ethical issues in using twitter for population-level depression monitoring: A qualitative study. *BMC Medical Ethics*, 17, 4 2016. ISSN 14726939. doi: 10.1186/s12910-016-0105-5.
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.
- Md. Saddam Hossain Mukta, Salekul Islam, Swakkhar Shatabda, Mohammed Eunus Ali, and Akib Zaman. Predicting academic performance: Analysis of students' mental health condition from social media interactions. *BEHAVIORAL SCIENCES*, 12, 6 2022. doi: 10.3390/bs12040087.
- Saumik Narayanan, Guanghui Yu, Wei Tang, Chien-Ju Ho, and Ming Yin. How does predictive information affect human ethical preferences? In *ACM Conference on AI, Ethics, and Society*, 2022.
- Thin Nguyen, Dinh Phung, Bo Dao, Svetha Venkatesh, and Michael Berk. Affective and content analysis of online depression communities. *IEEE Transactions on Affective Computing*, 5(3): 217–226, 2014.
- Safiya Umoja Noble. Algorithms of oppression. In *Algorithms of Oppression*. New York University Press, 2018.
- Jonathan A Obar and Anne Oeldorf-Hirsch. The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, 23(1):128–147, 2020.
- United Nations. Department of Economic and Social Affairs. *The Sustainable Development Goals: Report 2023*. UN, 2023.
- Ahmed Husseini Orabi, Prasadith Buddhitha, Mahmoud Husseini Orabi, and Diana Inkpen. Deep learning for depression detection of twitter users. pages 88–97, 2018a.
- Ahmed Husseini Orabi, Prasadith Buddhitha, Mahmoud Husseini Orabi, and Diana Inkpen. Deep learning for depression detection of twitter users. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, pages 88–97, 2018b.
- Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2009.
- Wenjing Pan, Bo Feng, and Cuihua Shen. Examining social capital, social support, and language use in an online depression forum: social network and content analysis. *Journal of Medical Internet Research*, 22(6):e17365, 2020.
- Javier Parapar, Patricia Martín-Rodilla, David E. Losada, and Fabio Crestani. Overview of erisk 2021: Early risk prediction on the internet. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12880 LNCS:324 – 344, 2021. doi: 10.1007/978-3-030-85251-1_22. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85115867066&doi=10.1007%2f978-3-030-85251-1_22&partnerID=40&md5=a5a71fb7eea7c4abea8e8208da6f1649. Cited by: 19.

- European Parliament and Council of the European Union. General data protection regulation (gdpr), 2016. URL <https://gdpr-info.eu/>. Accessed: 2023-08-17.
- Milman Parry. Whole formulaic verses in greek and southslavic heroic song. In *Transactions and Proceedings of the American Philological Association*, pages 179–197. JSTOR, 1933.
- James W Pennebaker, Matthias R Mehl, and Kate G Niederhoffer. Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, 54(1):547–577, 2003.
- Plato et al. *Phaedrus*, volume 275. Bobbs-Merrill Indianapolis, 1952.
- Brian A Primack, Ariel Shensa, César G Escobar-Viera, Erica L Barrett, Jaime E Sidani, Jason B Colditz, and A Everette James. Use of multiple social media platforms and symptoms of depression and anxiety: A nationally-representative study among us young adults. *Computers in human behavior*, 69:1–9, 2017.
- Lenore Sawyer Radloff. The ces-d scale: A self-report depression scale for research in the general population. *Applied psychological measurement*, 1(3):385–401, 1977.
- Adrian Rauchfleisch and Jonas Kaiser. The german far-right on youtube: An analysis of user overlap and user comments. *Journal of Broadcasting & Electronic Media*, 64(3):373–396, 2020.
- Pritika Reddy, Bibhya Sharma, and Kaylash Chaudhary. Digital literacy: A review of literature. *International Journal of Technoethics (IJT)*, 11(2):65–94, 2020.
- Andrew G Reece and Christopher M Danforth. Instagram photos reveal predictive markers of depression. *EPJ Data Science*, 6, 2017. doi: 10.1140/epjds/s13688-017-0110-z. URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85027011738&doi=10.1140%2fepjds%2fs13688-017-0110-z&partnerID=40&md5=12591ab9e3de234208ab344a688f4cec>.
- Philip Resnik, William Armstrong, Leonardo Claudino, Thang Nguyen, Viet-An Nguyen, and Jordan Boyd-Graber. Beyond lda: exploring supervised topic modeling for depression-related language in twitter. In *Proceedings of the 2nd workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality*, pages 99–107, 2015.
- Benjamin J Ricard, Lisa A Marsch, Benjamin Crosier, and Saeed Hassanpour. Exploring the utility of community-generated social media content for detecting depression: An analytical study on instagram. *JOURNAL OF MEDICAL INTERNET RESEARCH*, 20, 6 2018. ISSN 1438-8871. doi: 10.2196/11817.
- Robert H Rice. Cognitive-behavioral therapy. *The Sage encyclopedia of theory in counseling and psychotherapy*, 1:194, 2015.
- Esteban A. Ríssola, David E. Losada, and Fabio Crestani. A survey of computational methods for online mental state assessment on social media. *ACM Trans. Comput. Heal.*, 2(2):17:1–17:31, 2021. doi: 10.1145/3437259. URL <https://doi.org/10.1145/3437259>.
- Sarah T Roberts. Commercial content moderation: Digital laborers’ dirty work. 2016.

- Ahmed Salem, Yang Zhang, Mathias Humbert, Pascal Berrang, Mario Fritz, and Michael Backes. MI-leaks: Model and data independent membership inference attacks and defenses on machine learning models. In *26th Annual Network and Distributed System Security Symposium, NDSS 2019, San Diego, California, USA, February 24-27, 2019*. The Internet Society, 2019.
- Andrew D Selbst, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*, pages 59–68, 2019.
- Aparna Sharma, Kavish Sanghvi, and Prathamesh Churi. The impact of instagram on young adult’s social comparison, colourism and mental health: Indian perspective. *IJIM Data Insights*, 2(1):100057, 2022.
- Tiancheng Shen, Jia Jia, Guangyao Shen, Fuli Feng, Xiangnan He, Huanbo Luan, Jie Tang, Thanassis Tiropanis, Tat-Seng Chua, and Wendy Hall. Cross-domain depression detection via harvesting social media. In *Proc. of the 27th Int. Joint Conf. on Artificial Intelligence (IJCAI-2018)*, pages 1611–1617, 2018.
- Ruba Skaik and Diana Inkpen. Using social media for mental health surveillance: A review. *ACM Comput. Surv.*, 53, 6 2020. ISSN 0360-0300. doi: 10.1145/3422824. URL <https://doi.org/10.1145/3422824>.
- John ER Staddon and Daniel T Cerutti. Operant conditioning. *Annual review of psychology*, 54(1):115–144, 2003.
- Eric J Topol. High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine*, 25(1):44–56, 2019.
- Marcel Trotzek, Sven Koitka, and Christoph M Friedrich. Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences. *IEEE Trans. on Knowledge and Data Engineering*, 2018a.
- Marcel Trotzek, Sven Koitka, and Christoph M Friedrich. Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences. *IEEE Trans. on Knowledge and Data Engineering*, 2018b.
- Sho Tsugawa, Yusuke Kikuchi, Fumio Kishino, et al. Recognizing depression from twitter activity. In *Proc. of the ACM CHI*, pages 3187–3196, 2015.
- Ofir Turel, Qinghua He, Gui Xue, Lin Xiao, and Antoine Bechara. Examination of neural systems sub-serving facebook “addiction”. *Psychological reports*, 115(3):675–695, 2014.
- Philippe Verduyn, Oscar Ybarra, Maxime Résibois, John Jonides, and Ethan Kross. Do social network sites enhance or undermine subjective well-being? a critical review. *Social Issues and Policy Review*, 11(1):274–302, 2017.
- WHO. Depression and other common mental disorders: global health estimates. 2017.
- Akkapon Wongkoblaph, Miguel A Vadillo, and Vasa Curcin. A multilevel predictive model for detecting social network users with depression. pages 130–135, 2018. ISBN 978-1-5386-5377-7. doi: 10.1109/ICHI.2018.00022.

En-Liang Wu, Chia-Yi Wu, Ming-Been Lee, Kuo-Chung Chu, and Ming-Shih Huang. Development of internet suicide message identification and the monitoring-tracking-rescuing model in taiwan. *JOURNAL OF AFFECTIVE DISORDERS*, 320:37–41, 6 2023. ISSN 0165-0327. doi: 10.1016/j.jad.2022.09.090.

Le Xu, Jun Nan, and Yan Lan. The nucleus accumbens: A common target in the comorbidity of depression and addiction. *Frontiers in neural circuits*, 14:37, 2020.

William Lafi Youmans and Jillian C York. Social media and the activist toolkit: User agreements, corporate interests, and the information infrastructure of modern social movements. *Journal of Communication*, 62(2):315–329, 2012.

Richard Zenith. *Pessoa: uma biografia*. Schwarcz S.A, 2022.