SBC
Sociedade Brasileira de Computação

GRANDES DESAFIOS

Sociedade Brasileira de Computação

IV SEMINÁRIO DOS GRANDES DESAFIOS DA COMPUTAÇÃO NO BRASIL: TRABALHOS APRESENTADOS

ANDRÉ SANTOS E FLÁVIO RECH WAGNER

Organizadores
Andre Luis de Medeiros Santos
Flávio Rech Wagner

IV Seminário dos Grandes Desafios da Computação no Brasil: Trabalhos Apresentados

Sociedade Brasileira de Computação
Porto Alegre
2025

Dados Internacionais de Catalogação na Publicação (CIP)

S471 Seminário dos Grandes desafios da computação no Brasil (4. : 27 – 28 novembro 2024 : São Paulo)
Trabalhos apresentados [recurso eletrônico] / organização: André Luis de Medeiros Santos e Flávio Rech Wagner. – Dados eletrônicos. – Porto Alegre: Sociedade Brasileira de Computação, 2024.
123 p. : il. : PDF ; 7,5 MB

ISBN 978-85-7669-640-7

1. Computação – Brasil – Evento. 2. Desafios da computação. I. Santos, André Luis de Medeiros. II. Wagner, Flávio Rech. III. Sociedade Brasileira de Computação. IV. Título.

CDU 004(063)

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

Sociedade Brasileira de Computação
Av. Bento Gonçalves, 9500
Setor 4 | Prédio 43.412 | Sala 219 | Bairro Agronomia
Caixa Postal 15012 | CEP 91501-970
Porto Alegre - RS
Fone: (51) 99252-6018
sbc@sbc.org.br

IV Seminário dos Grandes Desafios da Computação no Brasil: Trabalhos Apresentados

Índice

Apresentação

A Sociedade Brasileira de Computação organizou seu primeiro evento com o objetivo de definir seus Grandes Desafios em 2006, em São Paulo. Foi uma iniciativa pioneira em Computação no país, no sentido de planejar e direcionar a pesquisa em Computação para um período de 10 anos (de 2006 a 2016). Neste evento foram definidos cinco grandes desafios, que se mostraram ao mesmo tempo precisos e abrangentes em sua visão do futuro da Computação, tendo servido de base para os eventos posteriores, realizados em 2008, 2009, 2013 e 2014.

Em novembro de 2024 a SBC realizou uma nova edição dos Seminários dos Grandes Desafios da Computação, com a participação de mais de 80 pesquisadores e representantes da indústria, com o objetivo de fortalecer a pesquisa da comunidade científica de Computação em torno dos desafios para a próxima década. Foram propostos novos Grandes Desafios, tendo em vista os muitos avanços científicos ocorridos desde o seminário de 2006, com foco especial nos impactos socioeconômicos da Computação, diante do acelerado processo de transformação digital pelo qual passa a sociedade e da utilização de soluções computacionais em virtualmente todos os aspectos de nossa vida.

Este eBook contém os artigos apresentados neste seminário, submetidos por pesquisadores de nossa comunidade acadêmica, que ajudaram a definir os novos grandes desafios, nas áreas de Inteligência Artificial e Ciência de Dados; Computação Sustentável; Cibersegurança; Ubiquidade da Internet; Computação Quântica; e Computação e Sociedade. Esperamos que estes trabalhos influenciem o desenvolvimento de projetos que avancem os estado da arte nestas áreas e gerem impactos sociais e tecnológicos no Brasil.

Ética na Inteligência Artificial: como apoiar os desenvolvedores?

Geber L. Ramalho

Centro de Informática – Universidade Federal de Pernambuco (UFPE)
Av. Jornalista Aníbal Fernandes, s/n – Cidade Universitária – Recife - PE – Brasil

glr@cin.ufpe.br

Abstract. *This article examines the primary ethical challenges associated with Artificial Intelligence (AI), particularly in the development of Machine Learning systems. Although ethical principles have been increasingly consolidated, a significant gap remains between these principles and their practical implementation by developers. The study reviews the state of the art and emphasizes the "ethics by design" approach, underscoring the need to rethink the training of professionals in the field, as well as the development of tools, processes, and best practices aligned with their day-to-day activities.*

Resumo. *Este artigo discute os principais desafios éticos da Inteligência Artificial (IA), especialmente no desenvolvimento de sistemas de Aprendizado de Máquina. Apesar da consolidação de princípios éticos, persiste uma lacuna entre esses princípios e sua aplicação prática por desenvolvedores. O texto analisa o estado da arte e destaca a abordagem "ethics by design", que evidencia a necessidade de repensar a formação desses profissionais, bem como o desenvolvimento de ferramentas, processos e boas práticas orientadas para suas atividades cotidianas.*

1. Introdução

A Inteligência Artificial (IA) tem emergido como uma força transformadora na sociedade moderna, trazendo revoluções em áreas tão diversas quanto saúde, finanças, educação e indústria. No entanto, juntamente com seus benefícios notáveis, a IA também traz uma série de questões éticas que precisam ser enfrentadas [Crawford 2021; Huang et al. 2022]. Enfrentá-las é fundamental para assegurar que a IA seja desenvolvida e utilizada de maneira que respeite os valores humanos e sociais, evitando consequências indesejadas. Há um consenso hoje sobre os princípios éticos da IA, mas *"the gap between principles and practice is large, and widened by complexity, variability, subjectivity, and lack of standardization"* [Morley et al. 2020].

Apesar dos avanços em termos de leis e normas para traduzir os princípios em instrumentos práticos, o **desafio** que permanece aberto é **como formar e apoiar os profissionais de computação a criarem sistemas de IA que adiram a princípios éticos em todo seu ciclo de vida, do design ao uso, passando pelo desenvolvimento e implantação?** Este é o desafio por excelência atualmente dentro da área de ética na IA, que cresceu muito nos últimos 4 anos, junto com a disseminação dos sistemas de IA. É um desafio crucial para garantir que as inovações tecnológicas promovam o bem-estar social e não amplifiquem problemas socio-econômicos existentes.

2. Questões éticas da IA

No desenvolvimento e uso de sistemas de IA, em particular os de Aprendizagem de Máquina (AM), várias questões éticas emergem, e elas podem ser organizadas em três categorias principais.

Há questões relacionadas ao próprio funcionamento dos modelos de AM. A eficácia de um sistema de AM depende de sua confiabilidade técnica, da sua precisão em previsões e decisões. Falhas podem causar danos graves, especialmente em áreas críticas como veículos autônomos [Thadani et al. 2023]. A justiça é outra preocupação, pois os algoritmos de AM podem amplificar preconceitos presentes em dados históricos, perpetuando discriminação com base em raça, gênero ou status socioeconômico, como visto no sistema COMPAS [Kirkpatrick 2017]. A autonomia humana também não pode ser posta em risco com a crescente automação de decisões [Tiribelli 2023], e garantir o alinhamento dos objetivos das máquinas com os dos seres humanos problema complexo [Yudkowsky 2016]. Além disso, muitos modelos de AM, especialmente os de aprendizagem profunda, são consideradas caixas-pretas, o que dificulta a explicabilidade e transparência [Dwivedi 2023]. Questões de responsabilização também existem, como no caso dos veículos autônomos, onde a culpa por acidentes pode ser difícil de atribuir entre desenvolvedores e usuários [Mueller et al. 2021].

Os dados, insumos fundamentais para os sistemas de AM, levantam preocupações éticas sobre privacidade e segurança. A coleta, armazenamento e uso de dados pessoais devem proteger os direitos dos indivíduos [Aldboush et al. 2023]. A segurança dos dados contra vazamentos é crucial para manter a confiança nos sistemas de IA [Mothukuri et al. 2021]. A transparência sobre o uso dos dados também é importante para garantir maior responsabilização [McCoy et al. 2023]. Questões de direitos autorais estão em pauta quando dados protegidos são usados sem autorização [Chesterman 2024].

O impacto social dos sistemas de AM inclui preocupações com o desemprego devido à automação, a proteção dos direitos humanos [Ximenes, Salcedo & Ramalho

2020], e o impacto na democracia por meio da disseminação de desinformação e publicidade política direcionada [Łabuz & Nehring 2024]. Além disso, a pegada de carbono dos modelos de AM, especialmente os de aprendizado profundo, é significativa, exigindo novas estratégias para reduzir emissões [CNBC 2024]. Por fim, há potenciais riscos de longo prazo envolvendo sistemas superinteligentes em relação aos interesses humanos [Bengio, 2023].

3. Fundamentos éticos para IA

Não há uma definição precisa de IA ética, entretanto, para este documento vamos adotar que **um sistema de IA ético é aquele que segue princípios éticos em todo seu ciclo de vida: design, desenvolvimento, implantação e uso**. Cobrir todo o ciclo de vida é crucial, já que questões éticas podem surgir em qualquer etapa do desenvolvimento e impactar componentes além dos algoritmos e modelos de AM [Morley et al. 2020; Lu et al. 2024]. Por exemplo, um sistema de AM que usa dados protegidos sem autorização para treinar o modelo não deve ser considerado ético.

São cinco os princípios éticos para IA, amplamente aceitos na literatura e entre especialistas [Floridi et al. 2018; Jobin et al. 2019]: (1) Beneficência, que visa promover o bem-estar, preservar a dignidade e sustentar o planeta; (2) Não-maleficência, que determina que a ação (mesmo benéfica) deve causar o menor dano possível, abordando questões como privacidade, confiabilidade técnica e segurança; (3) Justiça, que promove a prosperidade e preserva a solidariedade, estabelecendo a equidade como uma condição fundamental; (4) Autonomia, que preserva o poder de decidir (ou não decidir) e equilibra o poder de decisão que mantemos para nós mesmos e o que delegamos aos agentes artificiais; e (5) transparência, que envolve explicabilidade, inteligibilidade e responsabilização, garantindo a compreensão e a responsabilidade pelos processos de decisão dos sistema de IA.

Embora haja um consenso geral sobre a importância desses princípios, a tradução prática deles em diretrizes operacionais para o desenvolvimento de IA ainda é um desafio significativo, sobretudo para os desenvolvedores. De fato, "*as consensus across the various published AI ethics principles is approached, a gap remains between high-level principles and practical techniques that can be readily adopted to design and develop responsible AI systems"* [Sanderson et al. 2023]

4. Estado da arte

Diante dos desafios éticos decorrentes da rápida adoção e desenvolvimento de sistemas de IA, diversas comunidades de pesquisa surgiram para estudar o tema. Essas comunidades estão organizadas sob as denominações de "IA responsável", "IA confiável", "IA ética" e "IA segura". Apesar de pequenas diferenças em suas abordagens, elas compartilham o objetivo comum de garantir que as tecnologias de IA beneficiem a sociedade sem causar danos.

As inciativas para traduzir princípios gerais em mecanismos práticos podem ser classificadas em quatro vertentes, seguindo o paradigma das forças de regulação de Lessig [Lessig 2006].

Primeiro, nas leis, houve inúmeros avanços nos últimos anos. O Regulamento Geral de Proteção de Dados (GDPR) da União Europeia é um marco na regulamentação de dados pessoais, estabelecendo novos padrões que influenciaram outros países, como os do BRICS e da América Latina [Belli & Doneda 2023]. Recentemente, a Europa propôs o AI Act, um marco regulatório para garantir o uso seguro e ético da IA, classificando sistemas de acordo com seus níveis de risco [European Union 2024]. O Brasil, depois de lançar a LGPD, está trabalhando em um Projeto de Lei sobre IA no senado [Senado 2023].

Nas normas, segunda força regulatória, não-obrigatória, houve também avanços. A ISO 27000, focada em segurança da informação e privacidade, é uma das primeiras normas relevantes, mas há hoje 16 certificações ISO/IEC relacionadas a sistemas de IA éticos, incluindo ISO/IEC 24028 (Confiabilidade em IA), 24027 (Viés em IA), 38507 (Governança da IA por organizações) e 23894 (Gestão de riscos) [Janaćković et al. 2024]. Há trabalhos acadêmicos relevantes, com inclusive participação de brasileiros, sobre governança da IA de forma geral [Gasser & Almeida 2017].

O mercado, terceira força, via boicotes, auto-regulação e regulação imposta a fornecedores, pode impactar na regulação como as lojas de aplicativos da Apple e Google Play que estabelecem regras para classificar jogos voltados para crianças e adolescentes [Google 2024]. Diversos manuais de ética internos às empresas também surgiram [Sony Global 2024; Microsoft 2024].

O problema é que as regulações citadas são ou abstratas, ou complexas ou focadas em gestores, de forma que os desenvolvedores (arquitetos, cientistas de dados, engenheiros de AM, engenheiros de dados, engenheiros de software e operadores) não sabem como proceder no seu dia a dia para construir sistemas de IA éticos [Sanderson 2023, Lucaj et al. 2023]. De fato, *"dozens of proposals for addressing ethical aspects of artificial intelligence (AI) have been published. However, many of them are too abstract for being easily translated into concrete designs for AI systems"* [Prem 2023].

Para tanto, precisamos entrar no que Lessig chama de regulação por arquitetura, traduzida hoje por *"ethics by design"* e que é particularmente poderosa, pois busca desenvolver sistemas que sejam éticos desde a concepção, em vez de depender de leis, normas ou forças de mercado para mitigar problemas posteriores, sobretudo porque leis e normas nunca chegam na velocidade do avanço da tecnologia. É nesse quesito que se encontram o desafio para a computação que propomos e que se traduze na questão: como formar e apoiar os profissionais de computação a criarem sistemas de IA que adiram a princípios éticos em todo seu

ciclo de vida, do design ao uso, passando pelo desenvolvimento e implantação? Dentro dessa força de regulação, podemos ver as iniciativas atuais em **quatro eixos: educação; ferramentas e técnicas; processos e boas práticas.** Infelizmente, ainda estamos em estágios preliminares, com pouca validação, e praticamente nenhuma padronização.

Sobre a educação, os desenvolvedores geralmente não recebem treinamento ético extensivo. Educar desenvolvedores de sistemas de AM em princípios éticos, ferramentas e melhores práticas é uma etapa essencial para que eles priorizem justiça, transparência e responsabilidade, e que antecipem e mitiguem possíveis problemas ao desenvolver sistemas. Iniciativas de educação em IA ética estão surgindo [Tubella 2024; Alam 2023], identificando objetivos de aprendizado concretos e casos de uso implementáveis, mas ainda são muito poucas e longe de uma padronização ou adoção ampla. No CIn-UFPE, temos oferecido uma disciplina de ética em IA para a pós-graduação desde 2019, e algumas outras correlatas na graduação, mas elas ainda não formam um corpo coerente e coeso que abordem todas as competências necessárias sobretudo para os alunos da graduação.

Quanto às ferramentas e técnicas, houve muitos avanços nos últimos anos particularmente em segurança de dados, não discriminação e explicabilidade. Novas ferramentas estão surgindo a cada dia para ajudar os desenvolvedores a criar sistemas de AM éticos. Recentemente, foi apresentada uma compilação dessas ferramentas por meio de uma revisão sistemática [Prem 2023]. Exemplos incluem o AI Fairness 360 da IBM [IBM 2024], que oferece métricas e algoritmos para detectar e mitigar viés em modelos. Ferramentas de explicabilidade, como o SHAP [SHAP 2024], que ajudam a entender e explicar as decisões de modelos complexos, melhorando a transparência. Além disso, técnicas de preservação de privacidade, como a privacidade diferencial [Hassan 2019] e o aprendizado federado [Li et al. 2020], estão sendo integradas aos sistemas de IA para proteger os dados dos usuários. Claro, essas ferramentas tratam apenas de alguns pontos no ciclo de vida de sistemas de IA.

O terceiro eixo é o dos processos. Existem processos para criar sistemas de AM, desde o clássico CRISP-DM. No entanto, não há um sequer processo maduro para ajudar os desenvolvedores a criar sistemas de AM ou IA éticos. Alguns frameworks e metodologias emergentes estão sendo propostos para atender a essa necessidade na indústria, como o da Google [Google 2024b], mas ainda estão em estágios iniciais de adoção. Esforços acadêmicos propuseram processos para o desenvolvimento ético de IA, como o de Hundt [Hundt 2024] e o de Prem [Prem 2023], mas nenhum deles foi validado ainda. Outra iniciativa, o "Trustworthy Development Process" [Hohma & Lütge 2023], é uma coleção de diretrizes para desenvolvedores, mas também tem limitações de validação.

Enfim, sobre boas práticas, há diversas proposições nesse sentido, já que é mais fácil começar por elas do que por processos abrangentes. Mas há vários

problemas: (a) poucas das recomendações podem ser consideradas boas práticas no sentido de "medidas acionáveis". A maioria é de "recomendações concretas", com a vantagem de serem menos abstratas do que os princípios, porém menos claras do que boas práticas [Madaio et al. 2020]; (b) a maioria é focada em governança da IA voltada para gestores das empresas e no máximo gerentes de projetos, mas não para desenvolvedores [Chancellor 2023; Baldassarre et al. 2024]; (c) a maioria não tem uma validação empírica ou qualitativa por especialistas [Zhang et al. 2021; Lu et al. 2024]. Uma das poucas proposições de boas práticas validada é a que desenvolvemos no CIn-UFPE anos atrás [Ximenez & Ramalho 2021] e que está para ser publicada com mais profundidade e amplitude após um período de implantação de mais de um ano em uma empresa de IA com mais de 35 milhões de usuários ativos mensais.

5. Possíveis métricas de avanço no desafio

O desafio proposto reside em **transformar os princípios éticos da IA em práticas concretas e aplicáveis para desenvolvedores**, desafio que pode ser abordado a partir de quatro eixos fundamentais: educação dos desenvolvedores, processos de desenvolvimento, ferramentas e técnicas, e recomendações concretas e boas práticas. Assim, podemos tentar definir métricas para o avanço nesse desafio. Claro, as métricas já são um desafio em si dada a atualidade e a velocidade da adoção de IA, assim como sua versatilidade de aplicação, mas algumas métricas possíveis são:

- Quantidade de empresas que adotam ferramentas, processos e boas práticas éticas para sistemas de IA, em todo seu ciclo de vida;
- Número de frameworks, processos, e boas práticas propostas e validadas solidamente;
- Processos de auditoria adotados e validados;
- Número de cursos e treinamentos disponíveis em ética para IA, que deve ser amplo.

6. Conclusões e conexão com os outros desafios

A ética na IA é um campo emergente e vital para o futuro da tecnologia. Apesar dos avanços realizados, ainda há uma lacuna significativa entre os princípios éticos e sua aplicação prática, em especial para desenvolvedores de sistemas IA que não sabem como aderir no dia a dia aos princípios éticos para IA, se é que conhecem esses princípios. A criação de diretrizes concretas e boas práticas, acompanhadas por ferramentas, técnicas, processos e métricas adequados e eficazes, é essencial para garantir que a IA desenvolvida beneficie a sociedade de maneira ampla e equitativa. Avançar nesse desafio requer uma abordagem abrangente, que inclua também educação dos desenvolvedores.

Além desses, há desafios mais profundos na agenda da pesquisa em ética na IA. O primeiro, de caráter epistemológico, envolve a necessidade de se estabelecer um diálogo permanente e estruturado entre as comunidades da computação e das

humanidades. Não é possível falar de antecipar ou mitigar impactos negativos na sociedade sem envolver os pesquisadores que estudam a sociedade. Nós da computação habitualmente conversamos somente com nós mesmos.

O segundo, de caráter ontológico, envolve repensar o próprio conceito de algoritmo. Em tese recente na UFPE, é defendida a ideia de que estamos em uma transição de "closed algorithms" para "open algorithms", dos quais os sistemas de aprendizado de máquina fazem parte [Falcão 2025]. Os "algoritmos abertos" são menos "controláveis" por diversas razões, entre elas porque não são apenas código, mas um amálgama de código, dados, pessoas e interfaces. Não controlamos as pessoas e por consequência os dados que elas provêm. Reconhecer esse novo "animal" pode nos ajudar a melhorar e criar ferramentas, processos, métricas e práticas que levem em conta princípios éticos.

A proposta de desafio está fortemente conectada aos desafios tecnológicos e socioeconômicos atuais da SBC. Em termos de desafios tecnológicos, o tema "IA ética" dialoga com a "IA" em si, "ciência de dados" (porque dados são o principal insumo dos modelos de IA), "ciber-segurança" (também por conta do uso dados) e "computação sustentável" (porque o treinamento dos modelos consume muita energia). Quanto aos desafios socioeconômicos, a IA ética se relaciona com praticamente todos também: diretamente com os "aspectos éticos da computação", mas também com o "combate à desinformação" (pois IA ajudar a gerar ou supervisionar esse tipo de conteúdo), o "fortalecimento da inovação e do empreendedorismo" (visto que a IA será motor de muitas inovações), e "impactos socioeconômicos" que são a principal motivação da proposição desse desafio.

7. Referências

Alam, Ashraf. Developing a Curriculum for Ethical and Responsible AI: A University Course on Safety, Fairness, Privacy, and Ethics to Prepare Next Generation of AI Professionals. *Intelligent Communication Technologies and Virtual Mobile Networks*. Singapore: Springer Nature Singapore, 2023. 879-894

Aldboush et al. Building Trust in Fintech: An Analysis of Ethical and Privacy Considerations in the Intersection of Big Data, AI, and Customer Trust. *International Journal of Financial Studies* 11.3 (2023): 90.

Baldassarre et al. Fostering Human Rights in Responsible AI: A Systematic Review for Best Practices in Industry. *IEEE Transactions on Artificial Intelligence* (2024).

Belli, L., and Doneda, D. Data protection in the BRICS countries: legal interoperability through innovative practices and convergence. *International Data Privacy Law* 13.1 (2023): 1-24.

Bengio, Y. (2023). AI and catastrophic risk. *Journal of democracy*, *34*(4), 111-121.

Chancellor, S. (2023). Chancellor, Stevie. Toward practices for human-centered machine learning. *Communications of the ACM* 66.3 (2023): 78-85.

Chesterman, Simon. Good models borrow, great models steal: intellectual property rights and generative AI. *Policy and Society* (2024): puae006.

Crawford K. *Atlas of AI : power, politics, and the planetary costs of artificial intelligence*, 1st edn. (2021). Yale University Press
CNBC - Consumer News and Business Channel. Google's carbon emissions surge nearly 50% due to AI energy demand (2024). Available in https://www.cnbc.com/2024/07/02/googles-carbon-emissions-surge-nearly-50percent-due-to-ai-energy-demand.htAM
Dwivedi et al. Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM Computing Surveys* 55.9 (2023): 1-33.
European Union - The EU Artificial Intelligence Act. Available in https://artificialintelligenceact.eu (2024)
Falcão, J. (2025). *Open Algorithms: an interdisciplinary inquiry of artificial intelligence systems*. Tese de doutorado. Centro de Informática – UFPE.
Floridi L, Cowls J, Beltrametti M, et al (2018) AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds Mach* 28:689–707.
Gasser, U. and Virgilio Almeida. A layered model for AI governance. *IEEE Internet Computing* 21.6 (2017): 58-62.
Google. Google Play Families Policies. Available in https://support.google.com/googleplay/android-developer/answer/9893335?hl=en&sjid=13527497315176022638-SA#1&2&3&4&5&6&7&87&9&zippy=%2Cexamples-of-common-violations. (2024)
Google. Responsible AI. Available in https://cloud.google.com/responsible-ai. (2024b)
Hassan, M., Mubashir Rehmani, & Jinjun Chen. Differential privacy techniques for cyber physical systems: A survey. *IEEE Communications Surveys & Tutorials* 22.1 (2019): 746-789.
Hohma, E. & Christoph Lütge. From trustworthy principles to a trustworthy development process: The need and elements of trusted development of AI systems. *AI 4.4* (2023): 904-925.
Huang et al. An overview of artificial intelligence ethics. *IEEE Transactions on Artificial Intelligence* 4.4 (2022): 799-819.
Hundt, A., Schuller, J. & Severin Kacianka. Towards Equitable Agile Research and Development of AI and Robotics. arXiv preprint arXiv:2402.08242 (2024).
IBM. AI Fairness 360. Available in https://aif360.res.ibm.com (2024)
Janaćković, G., Vasović, D. and Bojan Vasović. Artificial Intelligence standardisation efforts. *Engineering management and competitiveness* (EMC 2024) (2024): 250.
Jobin et al. The global landscape of AI ethics guidelines. *Nat Mach Intell* 1:389–399. (2019)
Kirkpatrick, K. It's not the algorithm, it's the data. *Communications of ACM* 60:21–23. (2017)
Łabuz, M. & Nehring, C. On the way to deep fake democracy? Deep fakes in election campaigns in 2023. European Political Science (2024): 1-20.
Lessig L *Code: And Other Laws of Cyberspace*, Version 2.0, 2nd edn. Basic Books (2006)
Li, Tian, et al. Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine* 37.3 (2020): 50-60.
Lu et al. Responsible AI pattern catalogue: A collection of best practices for AI governance and engineering. *ACM Computing Surveys* 56.7 (2024): 1-35.

Lucaj, L., Van Der Smagt, P. and Djalel Benbouzid. "AI regulation is (not) all you need." Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. (2023).

Madaio et al. Co-designing checklists to understand organizational challenges and opportunities around fairness in AI. Proceedings of the 2020 CHI conference on human factors in computing systems. (2020)

McCoy et al. Ethical responsibilities for companies that process personal data. *The American Journal of Bioethics* 23.11 (2023): 11-23.

Microsoft. Empowering responsible AI practices. Available in https://www.microsoft.com/en-us/ai/responsible-ai. (2024)

Morley, J., Floridi, L., et al. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Sci. Eng. Ethics* 26, 2141–2168 (2020).

Mothukuri et al. A survey on security and privacy of federated learning, *Future Generation Computer Systems*, vol. 115, pp. 619–640, (2021).

Mueller et al. Addressing driver disengagement and proper system use: human factors recommendations for level 2 driving automation design. *Journal of cognitive engineering and decision making* 15.1 (2021): 3-27.

Prem, E. From ethical AI frameworks to tools: a review of approaches. *AI and Ethics* 3.3 (2023): 699-716.

Sanderson et al. AI Ethics Principles in Practice: Perspectives of Designers and Developers, *IEEE Transactions on Technology and Society*, pp. 1–1, (2023).

Senado Federal Brasileiro. Projeto de Lei n° 2338. Disponível em https://www25.senado.leg.br/web/atividade/materias/-/materia/157233 (2023)

SHAP. Available in https://shap.readthedocs.io/en/latest. (2024)

Sony Global. Sony Group's Initiatives for Responsible AI. Available in https://www.sony.com/en/SonyInfo/sony_ai/responsible_ai.html. (2024)

Thadani, Trisha, et al. Tesla drivers run Autopilot where it's not intended-with deadly consequences. *The Washington Post* (2023): NA-NA.

Tiribelli, S. The AI ethics principle of autonomy in health recommender systems. *Argumenta* 16 (2023): 1-18.

Tubella, A. Marçal Mora-Cantallops, A. and Juan Carlos Nieves. How to teach responsible AI in Higher Education: challenges and opportunities. *Ethics and Information Technology* 26.1 (2024): 3.

Ximenes, B. Salcedo, D. & Ramalho, G. Towards broadening the perspective on lethal autonomous weapon systems ethics and regulations. *Rio Seminar on Autonomous Weapons Systems.* (Rio de Janeiro, Naval War College) - Brasília: FUNAG, 2020

Ximenes B & Ramalho G. Concrete Ethical Guidelines and Best Practices in Machine Learning Development. In: *Proceedings of IEEE International Symposium on Technology and Society*. (2021)

Yudkowsky, Eliezer. "The AI alignment problem: why it is hard, and where to start." *Symbolic Systems Distinguished Speaker* 4.1 (2016).

Zhang, T., Qin, Y., & Li, Q. (2021). Trusted artificial intelligence: technique requirements and best practices. *International Conference on Cyberworlds (CW).* IEEE (2021).

A Big Challenge: Tools to Guarantee Robust and Controlled Behavior of Large Language Models

Fabio G. Cozman, Sarajane M. Peres, Marcelo Finger, Renata Wassermann, Anna H. Reali Costa, Edson S. Gomi, Artur J. L. Correia, Anarosa A. F. Brandão, Karina V. Delgado, Denis D. Mauá, Thiago A. S. Pardo, Fátima L. S. N. Marques

[1]Universidade de São Paulo (USP)
Av. Profº Lúcio Martins Rodrigues, 370 – 05508-020 — São Paulo — SP – Brazil

{fgcozman,sarajane,anna.reali,gomi,arturjordao,kvd}@usp.br

{anarosa.brandao,denis.maua,fatima.nunes}@usp.br

{mfinger,renata}@ime.usp.br, taspardo@icmc.usp.br

Abstract. *Large language models (LLMs) have changed Artificial Intelligence, and in fact greatly affected Computer Science and its applications. Even though the capabilities of LLMs are impressive, they offer responses without any solid guarantees of rationality, are prone to hallucinations, are relatively weak when faced with long reasoning chains, and offer a limited degree of controllability. Despite the impressive performance, ensuring robust and controlled output is a major challenge. That is, the big challenge is to produce tools that make these models insensitive to irrelevant features, resistant to unexpected failures, amenable to control and in accordance to performance requirements, as well as tools to formally verify their success and failure modes, and to evaluate them in meaningful ways.*

Resumo. *Grandes modelos de linguagem transformaram a área da Inteligência Artificial e, de fato, impactaram profundamente a Ciência da Computação e suas aplicações. Embora suas capacidades sejam notáveis, esses modelos produzem respostas sem garantias sólidas de racionalidade, são suscetíveis a alucinações, demonstram fragilidade diante de cadeias longas de raciocínio e oferecem um grau limitado de controlabilidade. Apesar de seu desempenho impressionante, assegurar que suas saídas sejam robustas e controladas continua sendo um desafio. Nesse contexto, portanto, o grande desafio é desenvolver ferramentas que tornem esses modelos insensíveis a características irrelevantes, resistente a falhas inesperadas, passíveis de controle e conforme com requisitos de desempenho, além de instrumentos capazes de verificar formalmente seus modos de acerto e falha e de avaliá-los de maneiras que façam sentido.*

1. Introduction

Large language models (LLMs)[1] have had extraordinary impact. They were developed at first to convey statistical properties of languages [Jurafsky and Martin 2024], but re-

[1]For a curated and updated list of recent work, see https://github.com/Hannibal046/Awesome-LLM

cent developments have demonstrated that LLMs can mimic human conversation to astonishing levels, leading some to think that they offer a direct path to Artificial General Intelligence [Bubeck et al. 2023].

Alas, whatever mechanisms are at play within LLMs, the research community can hardly say they are fully understood. For instance, the community does not know when LLMs will fail, and how to make them withstand changes; it knows LLMs hallucinate, but does not know exactly when they do it, how they do it, and how this behavior can be blocked or at least how it can be guaranteed to stay within given required bounds. Moreover, while LLMs often succeed in seemingly human ways, for example by correctly summarizing documents, they fail in surprising non-human ways, for example by missing easy questions while nailing hard questions in a given exam [Locatelli et al. 2024].

At this point the research community does not have the right tools to analyze and synthesize them so as to guarantee they satisfy a set of requirements; does not know how to guarantee that their responses follow given rules; and does not know how to make them reason in formally guaranteed ways, for instance by staying within prescribed logical schemes. Moreover, there is little understanding on how to design LLMs, other than by repeating a limited set of architectures that have been successful in the past.

This context presents us with a number of challenges, ranging from technical to ethical issues. However, we wish here to focus on one key challenge that is directly related to the theme of this meeting: ***the development of theoretical and practical computational tools that can ensure, or at least significantly improve, the robustness and control of LLMs' behaviour.***

2. LLMs with robust and controlled behaviour

First, there is a need for basic research on the mathematical tools to understand the inner workings of LLMs. To be able to control something, it is important to understand it; hence there must be a better grasp of the relationship between complexity and expressivity, of optimization algorithms and performance, of architecture size/structure and robustness to failure. Besides, it is necessary to determine some paradigmatic problems in this effort, much as resolution for first-order logic or completeness for automata offer a guiding path to investigation. These aspects depend on connections with research in statistics, statistical physics, and mathematics; similar multi-disciplinary efforts have been pursued throughout the world (for instance, take the recent call for proposals by the US-NSF.[2])

There is also a need for a systematic study of LLM architectures that guarantee correctness and assertiveness of outputs, both in the context of their fundamental task of language generation (in an intrinsic perspective) and in well-established contexts of downstream applications (in an extrinsic perspective) [Bommasani et al. 2021]. These architectures must be studied with respect to metrics such as accuracy, but also with respect to their robustness and reliability.

One particular strategy to enhance LLMs has been Retrieval-Augmented Generation (RAG) [Gao et al. 2024], where a LLM may query external databases. This sort of strategy has been expanded to include queries to reasoning engines, often supported

[2]Program Solicitation NSF 24-569, https://new.nsf.gov/funding/opportunities/mfai-mathematical-foundations-artificial-intelligence

by complex prompting techniques that ask for the LLM to expose its chain of thought — indicating steps that may require external reasoners. On one hand, reliance on formal reasoning is a promising approach to minimizing hallucinations and maximizing control through explicit rules. On the other hand, beyond the complexity of designing such formal reasoning, efforts to enhance robustness and control inevitably introduce trade-offs — particularly in terms of flexibility, creativity, and computational efficiency. Stricter control mechanisms may constrain model expressiveness or increase inference latency, potentially undermining usability in dynamic, real-time applications. Quantifying and balancing these dimensions remains a significant and specific research challenge that depends heavily on the intended application domain. Alas, current RAG systems are still far from guaranteeing specific levels of performance.

Actually, RAG offers one possible strategy within *neuro-symbolic* approaches, where the goal is to mix the data-centric power of neurally inspired architectures with the knowledge-centric power of formal symbolic reasoning [Garcez and Lamb 2023]. Several different paths are possible here [Lamb et al. 2020]. One path consists of systems where a neural module calls a symbolic engine (RAG and its variants fit here). Another path explores the reverse idea: a symbolic engine that calls a neural module (say, an LLM) and processes its output. Yet another strategy is one where symbolic rules and formal constraints are used to help build a neural module so as to guarantee given requirements; a related scheme embeds rules and deductive patterns into numerical spaces, so as to enforce them in the same spaces where embeddings operate. All such approaches, and their possible combinations, deserve more study as their potential is still unfilled.

Another challenge to solve to make LLMs more applicable is to reduce their size and their complexity, while pursuing robust and controlled behavior. Current LLMs are very large and energy-hungry at both learning and inference time. To what extent is so much flexibility needed? A related question is whether LLMs could be more modular and hence easier to analyze and to design. The recent work on distillation, pruning, quantization and similar techniques to reduce the size of LLMs, make the resulting models even more opaque and do not add any formal guarantees or modularity. Note that the challenge posed herein goes beyond mere miniaturization. The goal is to achieve reductions in size and complexity while embedding guarantees — whether statistical, logical, or otherwise — that render smaller models as trustworthy and predictable as their larger counterparts.

Yet another key element in this challenge is the verification of LLMs. While most computing systems can today be verified by ever more powerful formalisms, there are few formal ways to verify neurally inspired systems, and they do not generally scale up [Preto and Finger 2023]. There is a sore need for guaranteed LLM verification by appropriate algorithms.

The concretization of the theoretical directions outlined above is essential and constitutes a fundamental part of the proposed challenge. Their technical feasibility hinges on advancements in model instrumentation, modular training pipelines, and access to explainability layers. Practical implementation may also depend on the development of middleware capable of enforcing symbolic constraints or verifying runtime behavior through formal methods — or through systematic and robust heuristic procedures.

3. Technological Autonomy and Local Restrictions

Brazil has already seen cases of LLM use in public sector conversational agents, legal document summarization, and policy analysis. These examples reinforce the feasibility and urgency of developing locally designed, controlled, and trusted solutions that provide performance guarantees. Furthermore, strengthening domestic R&D fosters workforce development and reduces dependence on foreign platforms with opaque mechanisms.

In this context, the development of robust and controllable LLMs is especially critical for Brazil, as the country's digital infrastructure, regulatory frameworks, and sociocultural conditions may differ substantially from those of highly digitized nations. Off-the-shelf foreign models often fail to meet local demands – whether due to language limitations, legal incompatibilities, or a lack of contextual alignment. To be effective, LLMs in Brazil must be adapted to local legislation (e.g., the LGPD), cultural specificities, and public service needs. Thus, any formalism for the verification, validation, and monitoring of LLMs should be effective both as a general framework and in local applications.

4. Evaluation

Of course, a challenge only makes sense when it is possible to determine whether it is met. There are two possible standards to which respect success can be evaluated here. First, our challenge will be met at a theoretical level when it becomes possible to analyze and synthesize LLMs against prescribed requirements. Second, our challenge will be met at an empirical level when concrete LLMs are able to reach prescribed performance on benchmarks and meet the usage requirements determined by the technical, organizational, and social contexts in which they will be applied.

Some requirements are easily expressed: we might ask a system never to return different answers if a question is formulated in distinct natural languages. And some metrics are obvious: the probability of returning the correct answer is important in a system that answers questions — there are, in fact, dozens of metrics that apply to natural language processing in general and to LLMs in particular [Liang et al. 2023]. However, we submit that finding precise formalisms so as to express requirements and metrics is *part of the challenge*, given the lack of guidance concerning requirements, and the need for more nuanced metrics that really capture semantics. It is necessary to compare existing formalims and metrics; agreement on how to evaluate LLMs is itself a tangible victory.

It is also important to have some agreed-upon testing scenarios. For instance, we believe an interesting scenario of extraordinary social impact is the *generation of correct arguments and the detection of false arguments* in public discourse. Arguments are complex objects that can be formally analyzed and validated. Other concrete applications include educational tutoring systems capable of verifying mathematical reasoning, legal assistants that must comply with jurisdiction-specific constraints, and healthcare conversational agents required to operate within ethical and regulatory boundaries. These scenarios offer fertile ground for controlled evaluations and the specification of formal requirements. Responding to these demands with guaranteed levels of performance is an important activity that can test the robustness and control of LLMs. We expect that a plethora of new scenarios will be developed in the next years.

Finally, understanding and verification within the theoretical or experimental scope will not be sufficient to ensure that society adequately perceives and benefits from

the scientific progress eventually achieved. In this sense, real-world and holistic evaluations will serve as a definitive measure to determine to what extent the proposed challenge and the solutions presented are necessary and sufficient to position LLMs as beneficial tools for society.

5. Conclusion

This paper sets a challenge focused on the development of theoretical and practical computing tools that *guarantee* levels of robustness and control for Large Language Models (LLM). We have commented on a number of specific research directions that may lead us to meet this challenge; it is possible that they also lead to other unanticipated strategies. But the goal is clear: the research community must develop novel LLMs that can satisfy formal requirements and follow formal rules with guarantees, perhaps of statistical nature. Once such guaranteed behavior is available, we can collectively find ways to use LLMs responsibly and to impose ethical rules. Also, closer work with industry can also speed up progress toward the proposed challenge. Industry partners often operate under stringent constraints, such as latency, compliance, cost, and risk minimization. That calls for robust control requirements, realistic testbeds, and for system-level monitors that can evaluate performance both during development and during real-world deployment scenarios.

Acknowledgements

F. G. Cozman is partially supported by CNPq grant Pq 305753/2022-3. M. Finger is partially supported by CNPq grant Pq 302963/2022-7. A. H. R. Costa is partially supported by CNPq grant Pq 312360/23-1. D. D. Mauá is partially supported by CNPq grant Pq 305136/2022-4 and FAPESP grant 2022/02937-9. The authors would like to thank the Center for Artificial Intelligence (C4AI-USP) and the support from the São Paulo Research Foundation (FAPESP) grant 2019/07665-4) and from the IBM Corporation.

Referências

Bommasani, R. et al. (2021). On the opportunities and risks of foundation models. *ArXiv*.

Bubeck, S. et al. (2023). Sparks of artificial general intelligence: Early experiments with GPT-4.

Gao, Y. et al. (2024). Retrieval-augmented generation for large language models: A survey.

Garcez, A. d. and Lamb, L. C. (2023). Neurosymbolic AI: the 3rd wave. *Artificial Intelligence Review*, 56(11):12387–12406.

Jurafsky, D. and Martin, J. H. (2024). *Speech and Language Processing*.

Lamb, L. C. et al. (2020). Graph neural networks meet neural-symbolic computing: A survey and perspective. In *Int. Joint Conf. on Artificial Intelligence*.

Liang, P. et al. (2023). Holistic evaluation of language models. *arXiv preprint arXiv:2211.09110*.

Locatelli, M. S. et al. (2024). Examining the behavior of LLM architectures within the framework of standardized national exams in Brazil. *arXiv preprint arXiv:2408.05035*.

Preto, S. and Finger, M. (2023). Proving properties of binary classification neural networks via łukasiewicz logic. *Log. J. IGPL*, 31(5):805–821.

Adversarial Machine Learning: Aprendizado de Máquina em Contextos Inseguros

Oscar Zibordi de Paiva[1,3], Marcos Antonio Simplício Jr[1], Charles Christian Miers[2]

[1]Escola Politécnica da Universidade de São Paulo
São Paulo, SP – Brasil

[2]Universidade do Estado de Santa Catarina
Joinville, SC – Brasil

[3]Banco Bradesco S.A.
Osasco, SP – Brasil

{ozpaiva,mjunior}@larc.usp.br, charles.miers@udesc.br

Abstract. *The ubiquity of Machine Learning (ML) applications in a wide range of computing systems raises worrying questions about the resilience of ML models and algorithms in the face of malicious actions. Such questions concern a sub-area of cybersecurity called Adversarial Machine Learning (AML) - a name that refers to ML in adversarial contexts, that is, in which there are attackers. The number of publications on AML has grown significantly in the last five years, a fact that is indicative of the importance and the number of open questions in the area. This proposal details some of the reasons why these open questions will constitute research challenges for the next decade and highlights the reasons why traditional cybersecurity fails when applied to ML.*

Resumo. *A ubiquidade de aplicações de Machine Learning (ML) nos mais variados sistemas computacionais suscita questões concernentes à resiliência dos modelos e algoritmos de ML frente às ações maliciosas. Questões desse tipo inserem-se numa subárea da cibersegurança denominada Adversarial Machine Learning (AML) - nome que faz referência a ML em contextos adversariais, i.e., em que há atacantes. O número de publicações sobre AML cresceu significativamente nos últimos cinco anos, fato que é indicativo da importância e do número de questões ainda em aberto na área. Esta proposta detalha algumas das razões pelas quais essas questões em aberto serão desafios de pesquisa para a próxima década e destaca as razões pelas quais a cibersegurança tradicional falha ao ser aplicada no âmbito de ML.*

1. Introdução

A Cibersegurança é uma área da Computação que envolve tecnologia, pessoas, informações e processos para possibilitar operações com garantias de segurança. A área tem natureza multidisciplinar, contemplando a criação, operação, análise e teste de sistemas computacionais seguros, ao mesmo tempo que envolve aspectos legais, políticos, fatores humanos, ética e gestão de risco, com o objetivo de considerar contextos adversariais. O atendimento a requisitos de cibersegurança é relevante em uma multitude de cenários, ganhando ainda maior importância e necessidade de esforços de pesquisa no

contexto de tecnologias emergentes, as quais muitas vezes não são projetadas com proteções em mente. Uma das tecnologias atuais que apresenta esse aspecto emergente consiste em sistemas de inteligência artificial (IA). Historicamente, a sinergia entre IA e cibersegurança já é bastante explorada para construir sistemas de segurança inteligentes. Nesse sentido, o papel da IA, especificamente da subárea de Aprendizado de Máquina (*Machine Learning* - ML), é similar àquele apresentado em diversas outras áreas: uma ferramenta útil para lidar com diversas tarefas rotineiras e intensivas em dados, permitindo que especialistas humanos se concentrem em aspectos estratégicos, criativos e eticamente complexos (CSA, 2024). Por outro lado, é menos comum observar uma exploração de sinergias no sentido inverso, de proteger os algoritmos, práticas e ferramentas de IA contra potenciais adversários maliciosos. Com a crescente adoção de sistemas inteligentes em uma expressiva variedade de cenários, entretanto, têm se tornado cada vez mais importantes os esforços para sanar essa lacuna. A área de estudo que se dedica a essa questão se denomina *Adversarial Machine Learning* (AML) (Vassilev et al, 2025), e tem por objetivo a proteção de sistemas inteligentes baseados em *Machine Learning* (ML). Algumas ameaças nesse cenário incluem o eventual mau uso de ferramentas inteligentes para atacar outros sistemas, o envenenamento de bases de dados para treinamento, a violação de privacidade por meio da extração de dados de usuários incluídos nesta base, a violação de direitos autorais por meio da extração de modelos de terceiros, entre vários outros. Não por acaso, essa é uma área que tem recebido especial atenção por órgãos de segurança ao redor do mundo, como o *National Institute of Standards and Technology* (NIST) (Vassilev et al, 2025) e a *European Union Agency for Cybersecurity* (ENISA) (Malatras et al., 2021). Ao mesmo tempo, o número de trabalhos científicos sobre AML cresceu significativamente nos últimos anos, tendência que deve se intensificar com a adoção cada vez maior de sistemas inteligentes em diversos cenários de aplicação.

2. Fatores Diferenciativos de AML e Questões de Pesquisa Subjacentes

Em comparação a sistemas de software tradicionais, a utilização de ML introduz riscos de cibersegurança bastante peculiares. Uma das razões para isso está relacionada à característica inata destes algoritmos, grosso modo, de automática ou semi-automaticamente, transformar dados em programas. Tal transformação ocorre de forma indireta e convoluta, dividindo-se em duas etapas: uma primeira de treinamento, em que dados (denominados de dados de treinamento) são utilizados para buscar, em um espaço de representação, um modelo matemático que aproxima uma determinada função-alvo parcialmente representada nos dados de treinamento; e uma etapa posterior de inferência, em que o modelo aprendido é utilizado para inferir o valor da função-alvo para valores novos, ausentes do conjunto de dados de treinamento. Este arranjo, baseado em dados e não em um processo de desenvolvimento clássico, tende a dificultar a incorporação de propriedades básicas de segurança, a aplicação de boas práticas consagradas de programação defensiva e, até mesmo, a modelagem de requisitos para a correta proteção do sistema (Cazzaniga et. al., 2024). Dentre as propriedades e boas práticas consagradas violadas pelo ML, podem-se listar:

- **não-interferência**: propriedade de segurança de sistemas segundo a qual as ações de um programa ou usuário não devem afetar outros programas ou usuários além do que é explicitamente necessário. O treinamento de modelos em diversos sistemas inteligentes atuais frequentemente usa dados provenientes de usuários e não-usuários,

gerando um modelo global efetivamente criado e utilizado por todos: torna-se, portanto, difícil garantir propriedades de não-interferência. Um problema análogo é observado em cenários nos quais dados de vários usuários são usados também na etapa de inferência, influenciando o comportamento global das saídas do sistema.

- **codificação segura**: diz respeito às práticas de desenvolvimento de software com superfície de ataque reduzida, resistente a entradas possivelmente maliciosas. Dentre as práticas preconizadas pela área, podem-se listar desde a verificação completa de adequação de entradas (programação defensiva) até a utilização de linguagens com gestão segura de memória. No âmbito de ML, uma vez que os "programas" aprendidos são na realidade modelos estatísticos consultados em tempo de inferência, o campo de conhecimento da programação segura oferece pouca ajuda à tarefa de proteger sistemas usando ML. Consequentemente, o efeito de entradas maliciosas enviadas a esses sistemas costuma ser imprevisível.
- **mínimo privilégio**: princípio de projeto segundo o qual um programa ou usuário não deve ter acesso a mais dados ou recursos de sistema além do que estritamente necessário para a execução de suas tarefas. Especificamente no âmbito de ML, modelos são treinados a partir de conjuntos de dados globais, potencialmente sigilosos (e, por vezes, até mesmo maliciosos). Ainda, durante a etapa de inferência, usuários comumente têm acesso à totalidade da capacidade preditiva dos modelos, violando o princípio de mínimo privilégio. Com isso, o efeito de entradas maliciosas durante a etapa de inferência pode, por exemplo, dar origem a divulgação de dados de forma indevida, comprometendo o sigilo dos dados de treinamento ou inferência, e possivelmente a privacidade das pessoas às quais os dados estão associados.
- **simplicidade**: mecanismos de segurança devem ser tão simples quanto possível, no intuito de minimizar as chances de erros de projeto ou implementação. Os modelos e algoritmos de ML costumam violar essa propriedade de duas formas: (i) Modelos e algoritmos de ML são frequentemente gigantescos e incompreensíveis (e.g., redes neurais multicamadas). Com isso, a inspeção humana (como um mecanismo de segurança, mesmo que rudimentar) torna-se inviável. Soma-se a essa incompreensão o fato de que pesquisas na área de ML encontram-se frequentemente em estágio pré-científico, contendo dados empíricos sem explicações e resultados irreprodutíveis. (ii) Projetar políticas de segurança torna-se muito mais difícil no contexto de ML. Tradicionalmente, as políticas de segurança costumam se restringir a regras de controle de acesso bastante “cartesianas”, versando regras como “determinado programa só pode acessar um dado recurso em certas circunstâncias”. Em sistemas empregando ML, as políticas de segurança usualmente não podem ser descritas em tamanha simplicidade. Um exemplo aparece na construção de modelos usando dados sigilosos: nesse caso, é necessário definir políticas de segurança em termos de privacidade, uma propriedade de segurança relacionada à confidencialidade, porém muito mais sutil e, por isso, muito mais difícil de proteger do que esta útima mesmo em sistemas tradicionais. Outro exemplo se dá no contexto da definição de quais assuntos um chatbot deve ou não abordar em LLMs: tais filtros envolvem conceitos de linguagem natural e, além disso, critérios de julgamento bastante subjetivos.

Frente a essa dificuldade de aplicar boas práticas e propriedades de segurança em ML, não é surpreendente que técnicas de ataque diversas venham sendo demonstradas na literatura (Liu et al., 2024). Essa superfície de ataque só vem crescendo com a expansão

nos tipos de algoritmos e modelos de ML elaborados na última década, sobretudo nos últimos cinco anos. Dentre essas categorias de ataque, podem-se listar: evasão de modelo, envenenamento de modelos, extração de dados de modelo (ou inversão de modelo), inferência de dados de treinamento do modelo, e, mais recentemente, subversão de políticas ("*jailbreak"*) e injeção de *prompt* em *chatbots* baseados em LLMs. Embora alguns mecanismos de proteção tenham sido propostos para combater esses ataques, ainda é difícil determinar se estes operam de forma satisfatória, i.e., se são eficazes e abrangentes, e se introduzem compromissos aceitáveis com relação a custos computacionais, carga de trabalho de analistas, e redução de acurácia preditiva. Justamente pelas características de cibersegurança únicas de modelos ML, as quais ainda se encontram em fase de estudo pela comunidade acadêmica especializada em proteção de sistemas, o corpo de conhecimento da área de cibersegurança fica aquém de fornecer alternativas práticas, em prontidão, para o enfrentamento das questões de AML. Para sanar essa lacuna, há ainda muita pesquisa científica a ser realizada, ao mesmo tempo que modelos e algoritmos de ML continuam sendo colocados em operação nas mais diversas aplicações. Assim sendo, o desafio que se coloca é ainda uma verdadeira corrida contra o tempo para garantir que todo o potencial transformador da IA não se torne sinônimo de riscos de cibersegurança. (ENISA, 2023).

3. Indicadores de Sucesso de Pesquisa em AML

Como subárea de cibersegurança, o principal indicador de sucesso da área de ML é a descrição de mecanismos de segurança que possam ser utilizados no mundo real, de forma prática. Isso se traduz em métricas para aferir a efetividade desses mecanismos contra ataques conhecidos, e aferir também o quão aceitáveis são os impactos causados em termos de funcionalidade, custo e desempenho. Adicionalmente, a área de AML carece ainda de modelos capazes de descrever os funcionamentos internos de algoritmos e modelos de ML. Essa propriedade é importante para avaliações diversas, inclusive de segurança: usando esses modelos, seria possível fazer previsões, por exemplo, do efeito de perturbações em dados de treinamento ou de inferência, atestando-se o correto funcionamento de sistemas em diferentes cenários (o que é, afinal, um requisito fundamental de engenharia). Essa questão tem forte correlação com a área de IA Explicável, que é também emergente e ainda pouco madura. Por fim, em termos práticos, é necessário o estabelecimento de modelos de análise de riscos e de modelamento de ameaças de ML. Isso não só direcionará a pesquisa científica a adotar modelos que sejam realistas, como poderá orientar empresas e demais organizações na adoção de mecanismos de segurança de forma priorizada.

4. AML e Questões Socioeconômicas

Os cenários adversariais a que estão sujeitas aplicações de ML podem influenciar e até agravar possíveis impactos socioeconômicos naturalmente causados por essas aplicações. Alguns desses impactos e seus agravamentos são:

- **Vieses preconceituosos**: em aplicações cujas predições possam ter qualquer influência direta ou indireta sobre grupos sociais, usuários maliciosos podem explorar vulnerabilidades dos algoritmos de ML para, por exemplo, introduzir vieses que prejudiquem determinados grupos. Sistemas de concessão de crédito, prognóstico de propensão a crimes, diagnóstico médico, e até mesmo *chatbots* são exemplos de sistemas nos quais esse tipo de influência pode ocorrer.

- **Crimes cibernéticos**: sistemas de autenticação biométrica atuais são, em grande parte, dependentes de ML. Logo, nesse cenário, as vulnerabilidades inerentes a algoritmos e modelos de ML podem levar, quando exploradas, a violações de acesso. Em algumas aplicações, pode haver perdas materiais de alvos específicos ou de grupos sociais inteiros (e.g., subtração de benefícios de vulneráveis). A autenticação entre humanos também pode ser prejudicada ao se usar sistemas de ML em ataques de engenharia social, gerando textos para *spear phishing* ou áudio e vídeo que personificam familiares ou colegas de trabalho de possíveis vítimas.
- **Violação de direitos autorais**: a IA generativa pode reproduzir ideias e estilos de artistas cujas obras forem usadas para treinar os modelos subjacentes. A proteção de direitos autorais nesse caso suscita questões adversariais, como a inserção de perturbações nas obras (envenenamento) para evitar que os modelos treinados reproduzam o estilo nelas presente.
- **Geração automática de conteúdo prejudicial**: modelos generativos de áudio, texto e vídeo podem ser usados maliciosamente para gerar notícias falsas e outros conteúdos socialmente prejudiciais em escalas sem precedentes. A proteção desses modelos, prevenindo tal utilização, e o uso de modelos para detecção de conteúdo sintético, relacionam-se a aspectos de AML.

Referências

(Cazzaniga et. al., 2024). Cazzaniga et. al., “Gen-AI: Artificial Intelligence and the Future of Work.” IMF Staff Discussion Note SDN2024/001, International Monetary Fund, Washington, DC.

(CSA, 2024) “AI Model Risk Management Framework”. Cloud Security Alliance. (Julho/2024).

(ENISA, 2023). European Union Agency for Cybersecurity. "Artificial Intelligence and Cybersecurity Research".

(Liu et al., 2024) Liu, Y. "A hitchhiker’s guide to jailbreaking chatgpt via prompt engineering". 4th International Workshop on Software Engineering and AI for Data Quality in Cyber-Physical Systems/Internet of Things, SEA4DQ 2024. ACM.

(Malatras et. al., 2021) European Union Agency for Cybersecurity, Malatras, A., Agrafiotis, I., and Adamczyk, M. "Securing machine learning algorithms". Publications Office of the European Union.

(Vassilev et al., 2025) Vassilev et. al., “Adversarial machine learning: A taxonomy and terminology of attacks and mitigations.”

Agentes Conversacionais Inteligentes para a Inclusão Digital

Vasco Furtado[1,2], **Elizabeth S. Furtado**[1,3]

[1]Universidade de Fortaleza, UNIFOR
Av. Washington Soares, 1321, Fortaleza, CE, Brasil

[2]ETICE - Empresa de Tecnologia da Informação do Ceará
Av. Pontes Vieira 220, Fortaleza, CE, Brasil

[3]Universidade Estadual do Ceará - UECE
Campus do Itaperi, Fortaleza, CE, Brasil

{vasco,elizabet}@unifor.br

Abstract. *In Brazil, where socioeconomic inequality and low literacy rates hinder access to technology, the emergence of intelligent conversational agents based on natural language — especially those powered by Large Language Models (LLMs) — holds the potential to include historically marginalized populations in the digital ecosystem. However, for this scenario to unfold responsibly, it is essential to recognize that most widely used LLMs have been trained predominantly on foreign data, detached from Brazil's linguistic and cultural realities. Beyond linguistic barriers, there are significant ethical and social risks associated with the use of AI agents in direct interaction with people. The tendency toward anthropomorphism may lead vulnerable users to develop emotional bonds or excessive trust in intelligent agents, thereby increasing the risk of manipulation, misinformation, or malicious use. In this context, one of the greatest challenges for Brazilian computing in the coming years will be to design and implement technologies for human–conversational agent interaction that foster equitable access, taking into account the country's social, linguistic, and cultural particularities. This challenge encompasses efforts ranging from the development and/or fine-tuning of language models adapted to Brazilian contexts, to the evaluation of ethical implications and risks associated with the widespread adoption of intelligent conversational agents, and the formulation of public policies and inclusive strategies that leverage these technologies as tools for digital citizenship.*

Resumo. *No Brasil, onde a desigualdade socioeconômica e os baixos índices de letramento comprometem o acesso à tecnologia, a emergência de agentes conversacionais inteligentes baseados em linguagem natural – especialmente os impulsionados por Modelos Amplos de Linguagem (LLMs) – pode viabilizar a entrada de populações historicamente marginalizadas no ecossistema digital. Contudo, para que esse cenário se concretize de forma responsável, é essencial reconhecer que os LLMs amplamente utilizados hoje foram treinados majoritariamente com dados estrangeiros, distantes da realidade linguística e cultural brasileira. Além das barreiras linguísticas, há riscos éticos e sociais substanciais associados ao uso de agentes de IA em interação direta com pessoas. A tendência ao antropomorfismo pode levar usuários vulneráveis a desenvolver vínculos emocionais ou confiança excessiva nos agentes inteligentes,*

agravando riscos de manipulação, desinformação ou uso malicioso. Nesse contexto, o grande desafio da computação brasileira nos próximos anos será projetar e implementar tecnologias para interação entre agentes inteligentes conversacionais e humanos que promovam acesso com equidade, considerando as singularidades sociais, linguísticas e culturais do país. Esse desafio incorpora ações que vão desde o desenvolvimento e/ou refinamento de modelos de linguagem ajustados à realidade brasileira, avaliação dos impactos éticos e os riscos da adoção massiva de agentes conversacionais inteligentes e formulação de políticas públicas e estratégias inclusivas que utilizem essas tecnologias como instrumento de cidadania digital.

1. Introdução

A interação entre humanos e sistemas de Inteligência Artificial (IA) tornou-se um eixo central para os avanços tecnológicos e sociais contemporâneos, especialmente em contextos marcados por desigualdades estruturais, como o brasileiro. Em um país onde disparidades socioeconômicas e baixos níveis de letramento ainda comprometem o acesso equitativo ao mundo digital, as interfaces baseadas em linguagem natural — viabilizadas pelos Modelos Amplos de Linguagem (LLMs, do inglês *Large Language Models*) — emergem como promissoras ferramentas de inclusão. Tais modelos permitem o desenvolvimento de agentes conversacionais inteligentes capazes de mediar o acesso à informação de maneira mais acessível e fluida, ampliando o alcance de serviços e conhecimentos a grupos historicamente marginalizados.

Embora essas características dos agentes conversacionais representem um cenário promissor, elas também trazem complexidades que motivam o desafio central deste artigo: como tornar assistentes conversacionais inteligentes vetores importantes para a inclusão digital. Isso exige uma perspectiva multidisciplinar que vá além da computação (especialmente IA e IHC), envolvendo saberes da filosofia, sociologia, comunicação, direito, psicologia e outras áreas do conhecimento. É por meio dessa articulação que será possível compreender e antecipar os impactos sociais, cognitivos e culturais desses agentes na vida das pessoas e propor soluções que tornem a tecnologia inclusiva.

Para isto, propomos uma abordagem que se inspira e complementa alguns desafios propostos na área de IHC [Pereira et al. 2024]. Enfatizamos que, para que essas tecnologias sejam de fato inclusivas, é fundamental que estejam ancoradas na realidade brasileira. Isso inclui treinar os agentes com dados representativos dos contextos cultural e educacional do país. Estudos demonstram que o vocabulário e as estruturas frasais utilizadas por pessoas em situação de analfabetismo funcional ou baixa escolaridade divergem substancialmente da norma culta empregada por modelos linguísticos convencionais [Tarallo 1985, Monte 2019]. Tais diferenças, se ignoradas, podem comprometer a eficácia das interações [da Silva et al. 2022]. Portanto, os riscos e oportunidades trazidos pela adoção de sistemas de IA devem ser analisados sob uma ótica crítica e situada, considerando as especificidades sociais e culturais da sociedade brasileira.

2. Justificativa

A presente proposta justifica-se pela urgência crescente de tornar os serviços digitais mais acessíveis, inclusivos e socialmente adequados ao contexto brasileiro. Em um país

onde a exclusão digital ainda atinge milhões de pessoas, a adoção de soluções baseadas em agentes conversacionais inteligentes surge como uma estratégia promissora para ampliar o acesso à informação e aos serviços públicos e privados. Interações em linguagem natural — tanto escritas quanto faladas — podem representar uma via crucial de aproximação para indivíduos com baixa escolaridade, permitindo que esses usuários acessem conteúdos e recursos que, de outro modo, permaneceriam inacessíveis.

Contudo, a efetividade dessas interações é condicionada por variáveis demográficas, regionais e educacionais, que impactam a forma como os usuários se comunicam com os sistemas. No caso da língua portuguesa falada no Brasil, estudos demonstram que o desempenho dos assistentes conversacionais por voz pode variar conforme os padrões regionais e os níveis de escolaridade dos interlocutores [Lima et al. 2019, da Silva et al. 2022]. Essas variações afetam aspectos acústicos, cognitivos e estruturais da linguagem, criando obstáculos à compreensão e à personalização da interação.

No campo da Interação Humano-Computador (IHC), já se discute amplamente aspectos éticos, de privacidade e de apoio à diversidade no desenvolvimento de sistemas interativos. A iniciativa de Pereira et al. (2024) propõe seis desafios para a área de IHC no período de 2025 a 2035, incluindo o desafio específico das 'Implicações da IA', tratado como uma questão à parte. No entanto, ao abordar agentes conversacionais inteligentes, torna-se necessário adotar uma perspectiva mais ampla. Isso implica incorporar e expandir os seis desafios propostos, de modo que orientem tanto o desenvolvimento, através de técnicas de treinamento e engenharia de *prompt*, quanto a análise crítica desses agentes. Pois, esses agentes não apenas simulam comportamentos humanos com crescente naturalidade, mas também atuam com relativa autonomia, assumindo decisões, emitindo juízos e, por vezes, extrapolando as intenções originalmente projetadas por seus desenvolvedores.

Essas dificuldades são amplificadas quando se considera que os principais modelos de linguagem disponíveis no mercado são treinados com corpora estrangeiros, que não refletem com precisão as diversidades linguística e cultural brasileiras. Isso tem levado à reprodução de vieses e limitações no desempenho dos agentes, como apontado em estudos recentes [Silva et al. 2024]. Como resultado, as interações podem se tornar frustrantes ou ineficazes — e, em casos mais graves, até manipuladoras, comprometendo a segurança informacional e a autonomia dos usuários.

Além disso, é preciso considerar o elevado grau de receptividade do brasileiro a novas tecnologias, o que, embora positivo sob o ponto de vista da adoção, pode facilitar fenômenos de antropomorfismo e confiança excessiva em sistemas automatizados. Interfaces que simulam características humanas, como nos assistentes corporificados (ECAs) [Yasavur and Rishe 2014], ou agentes projetados para atuar como companheiros pessoais [Oliveira et al. 2025], tendem a gerar vínculos emocionais que podem agravar a vulnerabilidade cognitiva de certos grupos populacionais. Esses riscos levantam questões cruciais sobre a segurança, a privacidade, a credibilidade e a ética na utilização de tecnologias interativas [Slattery et al. 2024, Shneiderman 2020].

Diante desse cenário, é imprescindível que a comunidade científica nacional se debruce sobre os temas desse desafio, desenvolvendo estratégias e soluções para agentes conversacionais inteligentes que dialoguem com a realidade brasileira de forma ética,

crítica e socialmente comprometida.

3. Objetivos

O objetivo desta proposta é identificar e analisar propostas relacionadas ao desafio do desenvolvimento de agentes conversacionais inteligentes no contexto brasileiro que promovam a inclusão digital, a fim de planejar como mitigar os riscos que lhes são associados e alavancar a introdução da tecnologia. As propostas se concentram em:

1. Desenvolvimento de modelos de linguagem e técnicas de engenharia de *prompt* ajustados ao Brasil: Criar e treinar modelos que considerem as particularidades linguísticas e culturais do Brasil, especialmente focados em populações marginalizadas;

2. Desenvolvimento de abordagens multidisciplinares; Elaborar estratégias e diretrizes, com contribuições de áreas multidisciplinares, para que os *stakeholders* mais relevantes possam aplicar a fim de compreenderem e anteciparem os impactos e perigos desses agentes na vida das pessoas;

3. Políticas e Ações de Inclusão Digital: Explorar o potencial de agentes conversacionais inteligentes como ferramenta para reduzir a exclusão digital no Brasil, promovendo a inclusão de pessoas com pouca ou nenhuma familiaridade com o mundo digital; e

4. Avaliação Ética da Interação com agentes conversacionais inteligentes: Analisar os impactos éticos do uso de agentes conversacionais inteligentes, identificando potenciais perigos, como manipulação e dependência emocional, além de riscos de segurança relacionados ao uso malicioso dessas tecnologias.

4. Descrição do Desafio

A interação entre humanos e agentes conversacionais inteligentes representa, atualmente, um importante desafio da computação no contexto brasileiro. Tal complexidade decorre, em grande medida, da necessidade de adaptar essas tecnologias às especificidades linguísticas, culturais e educacionais da população. O desafio adquire contornos ainda mais críticos quando se considera que muitas dessas interações acontecerão com pessoas em situação de baixa escolaridade, inseridas em um cenário de profunda diversidade sociocultural.

Sem diretrizes claras que orientem o desenvolvimento e a implementação desses sistemas, corre-se o risco de desperdiçar uma oportunidade estratégica de promoção da inclusão digital. Em vez de reduzir desigualdades, a tecnologia mal calibrada pode perpetuar ou até mesmo ampliar barreiras existentes, tornando a experiência digital excludente para os públicos que mais necessitam de apoio.

Outro aspecto preocupante diz respeito à confiança desmedida que usuários com baixa cognição podem depositar em agentes conversacionais. Atribuir a esses sistemas um grau elevado de autoridade ou humanidade — especialmente em contextos marcados por baixa literacia digital — pode gerar efeitos colaterais graves, como dependência emocional, manipulação involuntária e vulnerabilidade a fraudes. Esses riscos são intensificados pela já precária infraestrutura de segurança digital no país, agravando o potencial de danos individuais e coletivos.

Diante desse cenário, torna-se imprescindível compreender em profundidade as múltiplas dimensões envolvidas na interação das pessoas com esses agentes. Essa compreensão deve fundamentar o desenvolvimento de agentes conversacionais mais seguros, éticos [Gero et al. 2025] e inclusivos, apoiando a criação de normas, protocolos e diretrizes que orientem seu uso responsável [Neves Camêlo and Alves 2023]. Esses instrumentos devem ser passíveis de incorporação tanto em sistemas de governança institucional — como políticas públicas e regulamentações — quanto em práticas de engenharia de *prompt* e de governança distribuída, lideradas pela sociedade civil e comunidades usuárias da tecnologia.

5. Metodologia

- Coleta de Dados: Obter dados sobre interações de usuários brasileiros com modelos de linguagem amplos, identificando falhas na compreensão linguística e nas respostas fornecidas. Isso incluirá um enfoque em populações com baixo nível educacional;
- Interdisciplinaridade: Promover uma forte interação entre as áreas de IA e IHC bem como entre essas áreas e outros ramos do saber como comunicação, filosofia, linguística, psicologia e sociologia.
- Desenvolvimento de Modelos: Criar versões personalizadas de modelos de linguagem, com ajustes baseados nos dados coletados, visando melhorar a compreensão e a utilidade das respostas para o público-alvo brasileiro;
- Avaliação de Riscos: Implementar avaliações contínuas dos riscos associados à interação humano-IA, como manipulação emocional e insegurança digital. Essas avaliações devem ser feitas em colaboração com especialistas em ética e segurança da informação;
- Políticas de Mitigação: Desenvolver diretrizes para o desenvolvimento e uso de agentes conversacionais inteligentes no Brasil, com foco em minimizar os riscos e maximizar os benefícios da inclusão digital;
- Desenvolvimento de serviços digitais: Atender a grandes parcelas da população excluídas digitalmente usando agentes inteligentes apropriados; e
- Educação: integrar teoria e prática na educação, promovendo uma reflexão contínua sobre sociedade e regulamentações, princípios éticos, processos técnicos e contemplando-os na engenharia de *prompt* para a interação do usuário com agentes inteligentes conversacionais [Porto and Furtado 2024] e [Silva et al. 2024].

6. Potenciais Métricas de Avaliação

Para avaliar o progresso das soluções propostas, as seguintes métricas serão consideradas:

- Precisão Cultural-Linguística: Percentual de interações bem-sucedidas entre usuários brasileiros e agentes conversacionais inteligentes, com base em parâmetros específicos de linguagem e contexto cultural;
- Nível de Satisfação do Usuário: Avaliação da experiência do usuário com agentes conversacionais inteligentes, medida através de pesquisas de satisfação e feedback direto;
- Redução na Exclusão Digital: Percentual de novos usuários digitais incluídos com sucesso por meio de interações com agentes conversacionaisi inteligentes, especialmente em regiões com baixos índices de acesso à tecnologia;

- Risco de Manipulação: Incidências de manipulação ou mal-uso da IA, identificadas em testes controlados e monitoramento contínuo; e
- Conformidade Ética: Grau de aderência às diretrizes éticas desenvolvidas, para apoiar o letramento dos dados e a privacidade acessível, com auditorias regulares para garantir conformidade.

7. Conclusão

A interação entre humanos e agentes conversacionais inteligentes é uma fronteira relevante e desafiadora para o futuro da computação no Brasil. Trata-se de uma área que, ao mesmo tempo em que oferece um potencial transformador para ampliar o acesso à informação, facilitar o uso de serviços públicos e privados, e promover a inclusão digital de milhões de brasileiros, também traz consigo riscos substanciais que não podem ser ignorados.

Tais riscos — que vão desde a exclusão involuntária de populações com baixa escolaridade até a indução à dependência emocional e à manipulação — exigem uma abordagem crítica e comprometida com os princípios da equidade, da transparência e da responsabilidade social. No contexto brasileiro, onde coexistem intensa desigualdade social, rica diversidade cultural e alta adesão às tecnologias digitais, o desafio de desenvolver sistemas que respeitem as singularidades da linguagem, da cognição e das formas de vida locais é particularmente urgente.

Acreditamos que, por meio de uma abordagem ética, interdisciplinar e culturalmente situada, é possível conceber e implementar agentes conversacionais inteligentes que não apenas ampliem a participação digital, mas que o façam de maneira segura, inclusiva e respeitosa às condições reais de seus usuários. Isso implica o desenvolvimento de modelos de linguagem treinados com dados representativos do Brasil, o estabelecimento de protocolos para mitigação de riscos, e a construção de políticas públicas orientadas à promoção de uma cidadania digital plena.

Nesse esforço, destaca-se ainda a importância de desenvolver técnicas de engenharia de *prompt* que estejam em conformidade com as diretrizes estabelecidas para a construção de interações responsáveis. Essas técnicas devem considerar as variações linguísticas e culturais dos usuários, bem como os objetivos éticos e funcionais dos agentes, garantindo que os modelos operem de forma transparente, previsível e alinhada às necessidades e limitações da população atendida.

Enfrentar esse desafio com seriedade e visão estratégica permitirá ao Brasil não apenas avançar internamente em termos de justiça tecnológica e inclusão social, mas também posicionar-se como referência internacional na formulação de soluções de inteligência artificial que aliem inovação técnica a compromisso ético.

References

da Silva, T. H. O., Furtado, V., Furtado, E., Mendes, M., Almeida, V., and and, L. S. (2022). How do illiterate people interact with an intelligent voice assistant? *International Journal of Human–Computer Interaction*, 40(3):584–602.

Gero, K. I., Desai, M., Schnitzler, C., Eom, N., Cushman, J., and Glassman, E. L. (2025). Creative writers' attitudes on writing as training data for large language models. CHI '25, New York, NY, USA. Association for Computing Machinery.

Lima, L., Furtado, V., Furtado, E., and Almeida, V. (2019). Empirical analysis of bias in voice-based personal assistants. In *Companion Proceedings of The 2019 World Wide Web Conference*, pages 533–538.

Monte, A. (2019). A influência da escolaridade e do sexo/gênero no uso variável da concordância verbal de terceira pessoa do plural. *Revista Diálogos*, 7:89–104.

Neves Camêlo, M. and Alves, C. (2023). G-priv: A guide to support lgpd compliant specification of privacy requirements. *iSys - Brazilian Journal of Information Systems*, 16:2:1 – 2.

Oliveira, J., Silva, T., Oliveira, R., and Furtado, E. (2025). Recommendations of embodied conversational agents to healthcare applications.

Pereira, R., Darin, T., and Silveira, M. S. (2024). Grandihc-br: Grand research challenges in human-computer interaction in brazil for 2025-2035. In *Proceedings of the XXIII Brazilian Symposium on Human Factors In Computing Systems*.

Porto, A. V. F. and Furtado, M. E. S. (2024). Framework to specify dialogues for natural interaction with conversational assistants applied in prompt engineering. In Arai, K., editor, *Intelligent Systems and Applications*, pages 231–253, Cham. Springer Nature Switzerland.

Shneiderman, B. (2020). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered ai systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 10(4):1–31.

Silva, V., Furtado, E., Oliveira, J., and Furtado, V. (2024). Engenharia de prompts em assistentes conversacionais para promoção de autocuidado baseados em modelos amplos de linguagem. In *Simpósio Brasileiro de Computação Aplicada à Saúde*, Goiânia.

Slattery, P., S. A. K., Grundy, Emily A. C.; Graham, J., Noetel, M.; Uuk, R. D. J. P. S. C. S., and Thompson, N. (2024). The ai risk repository: A comprehensive meta-review, database, and taxonomy of risks from artificial intelligence.

Tarallo, F. (1985). *A pesquisa sociolinguística*. Ática.

Yasavur, U.; Lisetti, C. and Rishe, N. (2014). Let's talk! speaking virtual counselor offers you a brief intervention. *Journal on Multimodal User Interfaces*, 8(4):381–398.

A Supercomputação na Era da Inteligência Artificial

Alba C. M. A. Melo[1], Philippe O. A. Navaux[2], Lucia M. A. Drummond[3], Cristina Boeres[3], Vinod Rebello[3], Alfredo Goldman[4], Márcio Castro[5]

[1]Departamento de Ciência da Computação – Universidade de Brasília (UnB)
Brasília – Brasil

[2]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Porto Alegre – Brasil

[3]Instituto de Computação – Universidade Federal Fluminense (UFF)
Niterói – Brasil

[4]Instituto de Matemática e Estatística – Universidade de São Paulo (USP)
São Paulo – Brasil

[5]Depto. de Informática e Estatística – Universidade Federal de Santa Catarina (UFSC)
Florianópolis – Brasil

Abstract. *This paper discusses the main challenges for the evolution of supercomputers in the era of Artificial Intelligence (AI) over the next 10 years. It analyzes issues related to architecture, storage, integration with cloud computing, and energy sustainability. Considering the exponential growth trend in computational demand driven by AI models, the need for new approaches in hardware, software, and infrastructure to meet this demand is also highlighted.*

Resumo. *Este artigo discute os principais desafios para a evolução dos supercomputadores na era da Inteligência Artificial (IA) para os próximos 10 anos. São analisadas questões relacionadas à arquitetura, armazenamento, integração com a computação em nuvem e sustentabilidade energética. Considerando-se a tendência de crescimento exponencial da demanda computacional por modelos de Inteligência Artificial, destaca-se ainda a necessidade de novas abordagens em hardware, software e infraestrutura para atender a esta demanda.*

1. Introdução e Contextualização

Os supercomputadores são máquinas utilizadas para resolver problemas extremamente complexos, inviáveis de serem resolvidos por computadores de uso comum, pois exigem Computação de Alto Desempenho (*High Performance Computing* - HPC). São considerados supercomputadores as máquinas mais rápidas existentes em um dado momento no tempo [Bell 2015]. O Cray-1, lançado em 1976, foi um dos primeiros supercomputadores, o qual tinha capacidade de processamento vetorial, atingindo 250 milhões de operações de ponto flutuante por segundo (MFLOPS).

Desde então, temos observado um crescimento impressionante no desempenho dos supercomputadores, rompendo as barreiras de GFLOPS (10^9 FLOPS), TFLOPS (10^{12} FLOPS), PFLOPS (10^{15} FLOPS) e EFLOPS (10^{18} FLOPS) a aproximadamente cada 11

anos. Em Novembro/2024, o *El Capitan*, contendo nodos de computação compostos por CPUs e GPUs (aceleradores), localizado nos Estados Unidos, é o supercomputador listado como mais rápido do mundo. Combinando processadores AMD EPYC de 24 núcleos (4ª geração) e GPUs AMD Instinct MI300A (11.039.616 núcleos de processamento no total), este supercomputador é capaz atingir 1,7 EFLOPS no *High-Performance Linpack* benchmark usado para classificar os sistemas. Considerando-se a taxa de evolução dos computadores recente, espera-se portanto que ultrapassemos a barreira dos ZFLOPS (10^{21} FLOPS) na década de 2030 [Matsuoka et al. 2023].

Tradicionalmente, os supercomputadores tem sido utilizados de forma efetiva para resolver problemas complexos, tais como dobramento acurado de proteínas [Streit et al. 2024], simulação de materiais para eletrônica de alto desempenho [Lyu et al. 2024], clima e previsão do tempo [Govett et al. 2024], simulação de energias renováveis [Veers et al. 2023], dentre outros. Mas, na presente década, temos observado um crescimento sem precedentes da Inteligência Artificial. Isso se dá principalmente devido à disponibilidade de conjuntos imensos de dados, que possibilitam respostas acuradas para modelos com uma quantidade enorme de parâmetros baseados em Redes Neurais do tipo *transformer* [Guo et al. 2019]. Tais arquiteturas deram origem à área de Inteligência Artificial Generativa e diversas soluções existentes, tais como o ChatGPT (OpenAI), são utilizadas milhões de vezes a cada dia ao redor do mundo. O treinamento de modelos com 1 trilhão de parâmetros (GPT-5) requer supercomputadores, com alta capacidade de poder de processamento, memória, interconexão e sistema de armazenamento em disco.

As empresas da área de Inteligência Artificial estão trabalhando para produzirem supercomputadores mais adequados a esse tipo de processamento. A empresa *Cerebras*, por exemplo, desenvolveu o *Cerebras Wafer-Scale Cluster*, com *design* otimizado, desacoplando memória e processamento, com acesso remoto à memória, interconexão de rede com nova topologia, dentre outros recursos. Com isso, foi lançado no final de 2023 o supercomputador Condor Galaxy 1, com capacidade de processamento de 4 EFLOPS (meia precisão) [Lie 2023] e o 8-EFLOPS Galaxy 3 já está em desenvolvimento.

Os próximos Modelos de Linguagem de Grande Escala (*Large Language Models* - LLMs) necessitarão de ainda mais os avanços no projeto de supercomputadores e, além de processamento, memória, interconexão, sistema de armazenamento secundário, o consumo de energia deverá ser considerado. Sendo assim, novas arquiteturas serão necessárias para que os supercomputadores do futuro possam processar tais modelos sem aumentar significativamente o consumo de energia. O consumo de energia associado a um supercomputador é enorme, sendo que quatro dos equipamentos entre os mais rápidos do mundo tem um consumo de mais de 10 MegaWatts, o que corresponde a cerca de três vezes o consumo de uma cidade como Campinas ou Guarulhos[1].

Para abordar o desafio da supercomputação na era da IA, devemos determinar uma maneira viável e sustentável de se ultrapassar a barreira dos ZFLOPS na década de 2030. As métricas objetivas e claras para avaliar o progresso das soluções são desempenho em FLOPS e a eficiência energética (GFLOPS/Watts) atingidas por cada supercomputador.

[1] https://dadosenergeticos.energia.sp.gov.br/PortalCEv2/Municipios/Eletricidade/M_Eletricidade.asp

2. Desafios para os Próximos 10 Anos

Nesta seção, partiremos dos supercomputadores mais avançados atualmente (1 a 4 EFLOPS) e discutiremos as principais questões para endereçarmos o desafio de supercomputadores na era da IA para os próximos dez anos.

2.1. Qual Deverá ser a Arquitetura do Supercomputador do Futuro?

Serão necessários grandes avanços em nível arquitetural para ultrapassar a barreira dos ZFLOPS, tais como: novas tecnologias de interconexão entre os nodos, novas tecnologias para resfriamento de supercomputadores de maneira efetiva e mais sustentável, novos mecanismos em software e hardware para tratar a falha de componentes, entre outros. Defendido por vencedores do Prêmio Turing [Hennessy and Patterson 2019], uma alternativa é projetar arquiteturas adaptadas a um domínio de problema específico. Arquiteturas específicas de domínio (DSAs) podem fornecer ganhos significativos de desempenho (e eficiência) e, dado que os acessos à memória se tornaram muito mais caros do que os cálculos aritméticos, podem ajudar fazer uso mais eficaz da hierarquia de memória. Portanto, supercomputadores do futuro deverão possuir tanto arquiteturas convencionais (processadores de propósito geral e aceleradores, incluindo aceleradores específicos para IA), como arquiteturas alternativas do tipo *Processing in Memory* (PIM). Há um interesse crescente em plataformas de hardware que implementem diferentes arquiteturas (ISAs) alternativas com desempenhos superiores por watt, por exemplo, a ISA aberta proposta pela iniciativa RISC-V[2] (da qual o Brasil faz parte), e que por ser aberta evita a dependência do fabricante do *chip*. Em termos de montagem, os fabricantes esperam que *chiplets* permitirão que eles abordem melhor três dos aspetos de eficiência computacional: energia computacional, energia de comunicações e energia de memória. Em um futuro próximo, a computação quântica deverá se aliar à computação de alto desempenho para, em conjunto, viabilizar a ultrapassagem da barreira dos ZFLOPS.

2.2. Como Será o Armazenamento dos Dados na Era da Supercomputação para IA?

O volume de dados utilizados em algoritmos de IA em larga escala é imenso. O primeiro desafio a ser superado é como armazenar, de maneira compacta e eficiente, volumes de dados na ordem de PetaBytes em constante (e rápido) crescimento. Técnicas de compressão de dados e de redução de dimensionalidade devem ser consideradas, dentre outras. Tendo os dados armazenados, o segundo desafio é recuperá-los de maneira rápida para processamento. Ou seja, deve ser definido como será a interconexão entre os nodos e com o sistema de arquivos e também como se dará o processamento (e.g., processamento com o dado comprimido).

2.3. Como Será a Interação entre Computação em Nuvem e Computação de Alto Desempenho na Era da IA?

Provedores de nuvem têm oferecido arquiteturas especializadas para computação paralela, tais como GPUs, FPGAs e aceleradores específicos para IA (e.g., TPUs da Google e Trainium e Inferentia da AWS e o MAIA da Microsoft) [Reed et al. 2023]. Portanto, a infraestrutura de computação em nuvem tem sido vista como uma forma de estender

[2]https://riscv.org/

uma infraestrutura de supercomputação instalada localmente. Nesse contexto, há diversos desafios relacionados a interoperabilidade entre nuvem e supercomputadores, o compartilhamento de recursos entre diversos usuários na nuvem e a dependência dos pesquisadores em certos serviços oferecidos por provedores de nuvem. Novas estratégias que permitam a redução do custo financeiro no uso de infraestruturas de nuvem também precisarão ser desenvolvidas.

2.4. Como Tratar o Problema de Consumo de Energia e da Sustentabilidade de Maneira Transversal?

Para sustentar os avanços na IA, são necessárias medidas urgentes para resolver o conjunto de gargalos enfrentados pela computação de alto desempenho. Os supercomputadores modernos são construídos usando dezenas de milhares de aceleradores, mas as limitações nas abordagens usadas para agregá-los significam que o desempenho próximo do pico raramente, ou nunca, é alcançado na prática. Isto implica que a quantidade de energia necessária para obter os ganhos de desempenho do ZFLOPS provavelmente excederá proporcionalmente os dos avanços anteriores. Se o desenvolvimento tecnológico continuar em sua trajetória atual, é estimado que um supercomputador da classe ZFLOPS precisaria em torno de 500 megawatts de potência[3], pouco menos da quantidade de energia gerada por uma das 20 turbinas da barragem de Itaipu e comparável ao consumo atual dos centros de dados de tamanho médio. A sede, aparentemente insaciável, de poder da IA ganhou atenção considerável dos provedores de nuvem. Num esforço para encontrar fontes de energia para novos centros de dados, os fornecedores de nuvens estão atualmente recorrendo à energia nuclear[4]. Embora não produza poluição do ar (ou dióxido de carbono durante sua operação), esta solução levanta a questão do tratamento de resíduos radioativos. Com os centros de dados de nuvem hoje consumindo de 2% a 3% de toda a energia utilizada, espera-se que o crescimento contínuo da IA faça com que este valor duplique. Isto está causando preocupação aos governos. Por exemplo, a União Europeia (UE) estabeleceu uma meta para reduzir o consumo de energia na Europa em 11,7% até 2030. Em relação à previsão para 2030 feita em 2020, ela está exigindo que as organizações que operam centros de dados nos países da UE apresentem relatórios anuais detalhando o consumo de água e de energia, bem como as medidas tomadas para sua redução[5]. Durante a próxima década, a eficiência computacional se tornará a prioridade número um. O desafio é descobrir como melhorá-lo significativamente numa forma economicamente e ambientalmente sustentável.

3. CVs dos Proponentes

Alba Cristina Magalhaes Alves de Melo é Professora Titular do Departamento de Ciência da Computação da UnB e atualmente é Chefe do Departamento. Atua na área de supercomputação e computação de alto desempenho desde 1996, quando concluiu seu doutorado no INPG, França. A Prof. Alba é pesquisadora Nível 1C do CNPq, Conselheira da SBC e Membro Senior da Sociedade IEEE. É editora associada de vários periódicos de prestígio e organizou diversas conferências internacionais na área de computação de alto desempenho. Coordenou também projetos de cooperação

[3]https://www.nextbigfuture.com/2023/02/intel-and-amd-path-to-zettaflop-supercomputers.html

[4]https://https://www.datacenterknowledge.com/energy-power-supply/going-nuclear-a-guide-to-smrs-and-nuclear-powered-data-centers

[5]https://www.cio.com/article/2100517/eu-moves-toward-regulating-data-center-energy-and-water-use.html

com o Canadá e Espanha e participou/participa de projetos com a Comunidade Européia, França e Alemanha.

Alfredo Goldman é professor da Universidade de São Paulo e bolsista de produtividade do CNPq. Atualmente é conselheiro da Sociedade Brasileira de Computação e coordenador da comunidade de processamento paralelo na *IEEE Computer Society* (TCPP). Possui doutorado em Informática e Sistemas pelo Instituto Nacional Politécnico de Grenoble. Trabalha com programação paralela e distribuída e desenvolvimento ágil de software. Coordena um projeto temático da FAPESP.

Cristina Boeres é professora do Instituto de Computação da UFF, e PhD em Ciência da Computação pela Universidade de Edimburgo (Reino Unido). Realiza pesquisas em HPC e gerenciamento de execução em ambientes paralelos e distribuídos. Tem participado de projetos financiados por CNPq, Petrobras e Faperj e coordena projeto CNPq/AWS.

Lúcia M. A. Drummond é Professora Titular da Universidade Federal Fluminense. É Pesquisadora Nível 1 do CNPq. É Cientista do Estado pela FAPERJ. Foi coordenadora da Comissão Especial de Arquitetura de Computadores e Processamento de Alto Desempenho da Sociedade Brasileira de Computação de 2019 a 2023. Atualmente é coordenadora de projetos financiados pela FAPERJ, CNPQ, CAPES e PETROBRAS na área de Computação de Alto Desempenho, envolvendo pesquisadores das Universidades de Sorbonne e Bordeaux (INRIA).

Márcio Castro é professor da Universidade Federal de Santa Catarina (UFSC) e PhD em Ciência da Computação pela *Université Grenoble Alpes* (França). É pesquisador Nível 2 do CNPq, coordenador da CE-ACPAD da SBC e do Programa de Pós-Graduação em Ciência da Computação da UFSC. Realiza pesquisas em HPC e Cloud em projetos financiados pelo CNPq, CAPES, FAPESC e MPSC.

Philippe O. A. Navaux é professor emérito do Instituto de Informática da UFRGS, Doutorado em Computação, INPG, 1979. Pesquisador nível 1 do CNPq nas áreas de Arquitetura de Computadores e HPC. Orientou mais de 100 estudantes de doutorado e mestrado e publicou mais de 400 artigos em periódicos e conferências. Sócio da SBC, SBPC, ACM e IEEE. Foi coordenador do Comitê da Computação da Capes, CNPq e Fapergs. Foi conselheiro da UFRGS, da Fapergs, do CATI/ MCTI, do LNCC. Atualmente membro do conselho do CIEE-RS e presidente do Conselho do SCALAC.

Vinod Rebello é Professor Associado da Universidade Federal Fluminense, tem um PhD em Ciência da Computação pela Universidade de Edimburgo (Reino Unido), e realiza pesquisas na área de HPC, nuvens e sistemas paralelos e distribuídos. Tem participado em varias projetos financiados por agências nacionais e internacionais e na organização de conferências da área.

Referências

Bell, G. (2015). Supercomputers: The amazing race. Microsoft Technical Report MSR-TR-2015-2, Microsoft Research.

Govett, M., Bah, B., Bauer, P., et al. (2024). Exascale computing and data handling: Challenges and opportunities for weather and climate prediction. *Bulletin of the American Meteorologic Society*, early access.

Guo, Y., Zheng, Y., Tan, M., Chen, Q., Chen, J., Zhao, P., and Huang, J. (2019). NAT: Neural architecture transformer for accurate and compact architectures. In *Advances in Neural Information Processing Systems (NeurIPS 2019)*, volume 32, pages 1–12. Curran Associates, Inc.

Hennessy, J. L. and Patterson, D. A. (2019). A new golden age for computer architecture. *Commun. ACM*, 62(2):48–60.

Lie, S. (2023). Inside the cerebras wafer-scale cluster: Cerebras systems. In *IEEE Hot Chips Symposium (HCS)*, volume 35, pages 1–41. IEEE.

Lyu, B., Chen, J., Wang, S., et al. (2024). Graphene nanoribbons grown in hbn stacks for high-performance electronics. *Nature*, 628:758–764.

Matsuoka, S., Domkel, J., Wahib, M., Drozd, A., and Hoefler, T. (2023). Myths and legends in high-performance computing. *Int. Journal of High Performance Computing and Applications*, 37(3-4):245–259.

Reed, D., Gannon, D., and Dongarra, J. (2023). HPC Forecast: Cloudy and uncertain. *Commun. ACM*, 66(2):82–90.

Streit, J., Bukvin, I., Chan, S., et al. (2024). The ribosome lowers the entropic penalty of protein folding. *Nature*, 633:232–239.

Veers, P., Bottasso, C. L., Manuel, L., et al. (2023). Grand challenges in the design, manufacture, and operation of future wind turbine systems. *Wind Energy Science*, 8(7):1071–1131.

This Future Without SQL

Eduardo C. de Almeida, Eduardo H. M. Pena, Altigran S. da Silva

[1] Departamento de Informática – Universidade Federal do Paraná (UFPR)
Curitiba, PR, Brazil

[2]Universidade Tecnológica Federal do Paraná (UTFPR)
Brazil

[3]Instituto de Computação – Universidade Federal do Amazonas (UFAM)
Manaus, Amazonas, Brazil

{eduardo@inf.ufpr.br, eduardopena@utfpr.edu.br, alti@icomp.ufam.edu.br}

Resumo. *O futuro da gerência de dados está se voltando para sistemas que processam consultas diretamente em linguagem natural, eliminando a necessidade de SQL. Diferente dos sistemas NL-para-SQL, que traduzem a linguagem natural para SQL, essa nova abordagem permite a interação direta com bancos de dados, impulsionada pelos avanços em processamento de linguagem natural (NLP) e inteligência artificial (IA). Propomos que a gerência de dados do futuro aproveite o PLN para lidar com consultas complexas e sensíveis ao contexto, com foco em aprendizado adaptativo para refinar a interpretação das consultas, além de enfrentar os desafios de segurança e privacidade. Nossa visão de "Linguagem Natural para Bancos de Dados" (NL-para-DB) integra soluções de NLP com estruturas robustas, possibilitando uma adaptação intuitiva e personalizada das consultas. Essa mudança repensa a forma como acessamos informações, promovendo inclusão e novas possibilidades na gerência de dados.*

Abstract. *The future of data management is shifting towards systems that process queries directly in natural language, eliminating the need for SQL. Unlike NL-to-SQL systems, which translate natural language into SQL, this new approach allows direct interaction with databases, driven by advances in NLP and AI. We propose that future data management leverage NLP for complex, context-aware queries, focus on adaptive learning to refine query interpretation, and address security and privacy challenges. Our vision for a "Natural Language to Databases" (NL-to-DB) integrates NLP solutions with robust frameworks, enabling intuitive, user-specific query adaptation. This shift reimagines how we access information, promoting inclusivity and innovation in data management.*

1. Introduction

The evolution of data management has long relied on Structured Query Language (SQL) as the main interface for interacting with databases. SQL introduced a declarative approach to querying and manipulating relational data with intuitive table-based structures, establishing itself as the industry standard [Abadi et al. 2022]. While SQL provides powerful and precise query capabilities, it imposes a steep learning curve for non-technical users, limiting broader accessibility. As data becomes more complex and widespread, the need for more intuitive ways to interact with databases has grown. Natural Language

Processing (NLP) and Artificial Intelligence (AI) offer promising alternatives by enabling users to query databases in natural language, bypassing the SQL syntax constraints.

This paper discusses the potential of a SQL-free future, where natural language interfaces democratize access to data, allowing users across various industries to interact with databases more naturally and efficiently. It discusses the advantages of this data access shift, such as improved accessibility and query efficiency, while addressing the technical, security, and adoption challenges. We are not proposing yet another form of NoSQL database. Instead, this vision seeks to redefine declarative data interaction within the relational model, which remains the dominant data model [Abadi et al. 2022], making it more inclusive and intuitive for a wider audience.

Consider a lawyer or an accountant who wants to check the database of a financial system to verify whether a transaction complies with specific regulations. This professional may not understand the underpinnings of the relational model required to explore a large schema with several hundred relations correctly or be able to program a very long SQL code with multiple operators. For example, in the TPC-DS benchmark[1], many queries reflect such real-world, high-level questions. The textual specification for query #88 outlines a complex business rule in 3 lines of plain English. The implementation of the corresponding SQL code spans almost 100 lines with aggregations, inline views, and dozens of joins, among other operators. Programming such queries requires a database background, which limits non-experts' flexibility in exploring the data.

This paper overviews some of the limitations of SQL for database interactions and proposes natural language interfaces as a more accessible solution. It reviews the current state of SQL and Natural Language to SQL (NL-to-SQL) systems and then discusses the advantages of transitioning to a SQL-free future. Then, it explores the technical, security, and adoption challenges involved, offering strategies for gradual implementation. Finally, it concludes with key takeaways, emphasizing the potential for innovation and providing recommendations for stakeholders to support this transition.

2. Current State of Art and Practice

As data becomes increasingly democratic and complex, SQL's intricate syntax poses barriers for many non-technical users who now need to work with databases. Reflecting on SQL's original vision, Don Chamberlin, its co-inventor, recently emphasized, "Database queries should not look like programs that tell the computer what to do. We wanted to express queries in a high-level, non-procedural language," [Chamberlin 2024]. With recent advances in NLP, it might be the right time to revisit SQL and explore more intuitive ways for users to interact with data.

Text-to-SQL systems aim to bridge the gap between user proficiency and SQL-based data retrieval by translating natural language queries into SQL statements [Katsogiannis-Meimarakis and Koutrika 2021, Li et al. 2024]. This approach has gained traction recently, driven by advances in large language models like GPT and specialized benchmarks [Gkini et al. 2021]. These lines of research approach the text-to-SQL task as a form of language translation, training neural networks on large datasets of paired natural language questions and their cor-

[1] https://www.tpc.org/tpcds/

responding SQL queries. However, such systems must overcome several challenges [Katsogiannis-Meimarakis and Koutrika 2023], including ambiguity in natural language, difficulty with complex queries, and mismatches between user terms and database schema. They often struggle with advanced SQL functions like aggregates and domain-specific terminology. Additionally, they lack effective error handling and feedback mechanisms, making it hard to capture the user's intent in SQL form accurately.

Another alternative approach is to move beyond NL-to-SQL systems and focus directly on NL-to-DB systems, also known as natural language interfaces to databases (NLIDB) [Li and Jagadish 2014]. It resembles search engine interfaces and often overlaps with research in keyword search interfaces [Yu et al. 2010]. An orthogonal line of research is table question answering (TQA), which focuses on extracting information from structured data in response to NLP queries [Pasupat and Liang 2015]. TQA represents a multifaceted challenge that requires a blend of language comprehension, logical analysis, and data interpretation skills. The goal is to process a user's query, understand the underlying tabular structure, and provide precise responses through reasoning and data extraction [Zhang et al. 2023]. In industry, NLIDBs like Tableau's Ask Data, Power BI, and Cognos Assistant exemplify a relatively new wave of tools to allow non-technical users to ask questions and explore data sets through simple, conversational queries.

3. A Claim for This Future Without SQL

SQL as the primary database interface poses challenges for non-technical users, particularly as data complexity grows. While NL-to-SQL systems translate natural language into SQL, they still depend on SQL, limiting intuitive data interactions. We propose eliminating SQL as an intermediary, allowing users to query databases directly through natural language. Advances in NLP and AI enable this shift, making data access more seamless and intuitive for non-experts, reducing complexity, and enabling context-aware responses.

A SQL-free approach democratizes data access by allowing non-technical users to interact with databases naturally without needing to learn SQL. This broadens access to data-driven decision-making and reduces project delays in real-time environments. Thanks to advancements in NLP and AI, these systems provide more accurate, context-sensitive responses, capturing user intent more effectively. Additionally, they are adaptable, learning from user interactions and evolving to meet specific domain needs, improving accuracy in areas like healthcare and finance.

Adopting a SQL-free system requires retraining and infrastructure changes, which can be both costly and time-consuming. Users may view natural language interfaces as less reliable than SQL, necessitating a gradual transition to build trust in the new system. A hybrid approach is a practical solution, allowing organizations to retain SQL while gradually adopting natural language interfaces. This phased transition minimizes disruption and prepares the groundwork for a fully SQL-free future.

Implementing SQL-free systems transforms data access across healthcare, finance, and law industries. Non-technical users can interact with data directly, focusing on insights rather than query mechanics. This shift could create new services in customer support, personalized recommendations, and faster decision-making processes. As these systems mature, new applications for natural language-driven interactions will emerge. Indeed, a SQL-free future fosters interdisciplinary collaboration by making data systems

more accessible to professionals from diverse fields. This democratization of data access breaks traditional silos, encouraging more collaborative, data-driven projects. Designing these systems will require collaboration across AI, linguistics, and domain experts, enhancing the system's contextual accuracy.

The shift to a SQL-free paradigm drives innovation in databases, NLP, and AI. This challenges traditional database interaction, opening possibilities for more flexible and adaptive systems and pushing advances beyond databases into conversational agents and context-aware computing. SQL-free innovations will influence other areas like user interfaces and decision support systems. Natural language could be embedded into productivity tools, creating more intuitive environments where data insights are readily accessible, ultimately expanding the scope of AI-powered assistants and automation.

Key performance indicators are essential to successfully transitioning to an SQL-free system. These include improved accessibility for non-technical users, query precision in interpreting complex natural language, and operational efficiency regarding response time and resource usage. Measuring the success of SQL-free techniques can also be evaluated by query accuracy and response relevance in aligning generated results with user intent. Specific metrics include query success rate, user satisfaction, execution efficiency, and resource consumption. The system's adaptability and continuous learning from user interactions, measured by reduced error rates and improved handling of complex queries, are crucial for long-term effectiveness. Recent database management and information retrieval research has already started investigating and implementing some of these metrics [Xing et al. 2024, Afonso et al. 2024].

4. Challenges Ahead

Our vision aligns with the Seattle Report on Database Research on declarative language abstractions [Abadi et al. 2022]. However, we see numerous challenges ahead.

Technical Challenges. While advances in NLP have made it possible to write code in various programming languages, including SQL, writing complex queries that capture business rules and regulations remains an open challenge. Complex rules require new sophisticated compilers and substantial computing resources for execution. Even state-of-the-art DBMSs can take hours to process rules with just a few predicates due to inefficient intermediate memory representations [Pena et al. 2021]. A new NLP engine will face the same scalability problem. The intermediate representations to support NLP queries, such as Candidate Joining Networks(CJNs) and Query Matches(QMs) are still in initial development stages [Martins et al. 2023]. A key metric for evaluating progress in SQL-free systems is their increasing adoption.

Privacy and Security Concerns. Differential privacy and homomorphic encryption are the state-of-the-art approaches in SQL engines. The former ensures that statistical properties remain intact even without individual data [de Farias et al. 2020]. The latter supports unbounded database aggregation queries. However, recent research shows that the homomorphic encryption in relational databases is yet slower than plaintext processing in magnitudes [Ren et al. 2022], leaving room for new initiatives.

Databases are increasingly exposed to sophisticated cyberattacks. SQL injection is a well-known form of attack, but we are still in the early stages of understanding "NLP

injection" vulnerabilities like prompt injection or malicious queries. These challenges are new for relational databases if NLP becomes the query language. Assessing the progress of new solutions includes addressing new types of attacks that will certainly appear.

Adoption and Usability Challenges. Many data scientists rely on programming interfaces, such as Jupyter and Zeppelin, in their daily work despite their limitations in collaborative coding and operationalizing code. Non-technical users may not be familiar with these programming interfaces. Therefore, it is essential to design intuitive NLP interfaces to improve usability, similar to industry initiatives such as Cognos and Tableau.

We must also ensure usability with reliable performance to increase adoption, particularly for critical applications. NLP-to-DB can be initially implemented as extensions of SQL engines utilizing new compilation paradigms [Jungmair et al. 2022] to benefit from the internal representations of SQL engines, such as vectorization and data-centric code generation [Kersten et al. 2018], without requiring translation of NLP to SQL. For example, the DuckDB team introduced a database extension that allows querying CSV files using vectorization without uploading them to a database. Building pure NLP query engines is still being determined, but progress toward our vision can be initially measured by the development of NLP database extensions that benefit from the high performance and maturity of modern SQL engines.

Regulatory and Legal Challenges. Brazilians are engaged in a significant debate on artificial intelligence's legal and regulatory aspects, as outlined in Congress in Bill # 2338/2023. There are at least three key challenges: what should be regulated, who will be responsible for regulating, and how to implement the regulation. As for *what* to regulate, the following problems should be considered [Rodrigues 2020]: algorithmic transparency, cybersecurity vulnerabilities, unfairness, bias and discrimination, lack of contestability, legal personhood issues, intellectual property issues, adverse effects on workers, privacy and data protection issues, liability for damage and lack of accountability.

Concerning *who* and *how* to regulate, recent evidence shows that automated systems can produce unfair outcomes and exacerbate existing inequalities [Ghasemaghaei and Kordzadeh 2024, Sunyé 2020]. Additionally, the global rise of AI-generated fake news shows that much more work is needed. In Brazil, Bill # 2338/2023 formalized the National Data Protection Authority (ANPD) as the National System for Regulation and Governance of Artificial Intelligence (SIA) coordinating body. In our NLP vision, significant effort is required to create data collection and disposal frameworks that align with regulations and policy constraints.

5. Conclusions, Remarks, and Takeouts

A SQL-free paradigm offers a transformative shift in data management, making databases more accessible and intuitive. While SQL has been practical, it poses barriers for non-technical users. Natural language interfaces can democratize data access and improve efficiency. Still, significant challenges remain, as technical limitations in handling complex queries and security and privacy concerns still need to be appropriately addressed.

The SQL-free transition will require collaboration between researchers, industry leaders, policymakers, and technologists. A hybrid approach combining SQL and natural language can ease this shift. Though challenging, this future promises more inclusive, efficient data interactions, reshaping how we access and use information.

About the Proponents

Eduardo C. de Almeida is an Associate Professor at UFPR. He holds a Ph.D. in Computer Science from the University of Nantes, INRIA GDD Team, France. His research areas include data structures, query processing, and data management. He has experience in both industry and academia, having worked as a database engineer and held research positions in Luxembourg and Switzerland.

Eduardo H. F. Pena is a professor at UTFPR and a permanent faculty member at the UEM graduate program in Computer Science. His Ph.D. on data quality and metadata extraction earned the CAPES award for Best Thesis in Computing in 2021. He is currently a postdoctoral researcher at NYU, developing tools for biomedical data integration.

Altigran S. da Silva is a Full Professor at UFAM. He earned his Ph.D. from UFMG in 2002. His interests include Data Management, Information Retrieval, Data Mining, Machine Learning, and Language Models. He has coordinated and participated in dozens of research projects resulting in over 150 publications in journals and conference. He served as Dean for Research and Graduate Studies at UFAM (2007/2009), coordinator of CA-CC at CNPq (2023/2024), and adjunct coordinator of the computing area at CAPES (2011/2013). He was also a board member (2005/2015) and council member (2016/2019) of SBC. He co-founded companies such as Akwan (acquired by Google in 2005), Neemu (acquired by Linx Systems in 2015), and Teewa (acquired by JusBrasil in 2019).

References

Abadi, D., Ailamaki, A., Andersen, D. G., Bailis, P., Balazinska, M., Bernstein, P. A., Boncz, P. A., Chaudhuri, S., Cheung, A., Doan, A., Dong, L., Franklin, M. J., Freire, J., Halevy, A. Y., Hellerstein, J. M., Idreos, S., Kossmann, D., Kraska, T., Krishnamurthy, S., Markl, V., Melnik, S., Milo, T., Mohan, C., Neumann, T., Ooi, B. C., Ozcan, F., Patel, J. M., Pavlo, A., Popa, R. A., Ramakrishnan, R., Ré, C., Stonebraker, M., and Suciu, D. (2022). The seattle report on database research. *Commun. ACM*, 65(8):72–79.

Afonso, A., Martins, P., and da Silva, A. (2024). Sereia: document store exploration through keywords. *Knowledge and Information Systems*.

Chamberlin, D. (2024). 50 years of sql | don chamberlin computer scientist and co-inventor of sql. Accessed: 2024-08-31.

de Farias, V. A. E., Brito, F. T., Flynn, C. J., Machado, J. C., Majumdar, S., and Srivastava, D. (2020). Local dampening: Differential privacy for non-numeric queries via local sensitivity. *Proc. VLDB Endow.*, 14(4):521–533.

Ghasemaghaei, M. and Kordzadeh, N. (2024). Understanding how algorithmic injustice leads to making discriminatory decisions: An obedience to authority perspective. *Information and Management*, 61(2):103921.

Gkini, O., Belmpas, T., Koutrika, G., and Ioannidis, Y. (2021). An in-depth benchmarking of text-to-sql systems. In *Proceedings of the 2021 International Conference on Management of Data*, SIGMOD '21, page 632–644, New York, NY, USA. Association for Computing Machinery.

Jungmair, M., Kohn, A., and Giceva, J. (2022). Designing an open framework for query optimization and compilation. *Proc. VLDB Endow.*, 15(11):2389–2401.

Katsogiannis-Meimarakis, G. and Koutrika, G. (2021). A deep dive into deep learning approaches for text-to-sql systems. In Li, G., Li, Z., Idreos, S., and Srivastava, D., editors, *SIGMOD '21: International Conference on Management of Data, Virtual Event, China, June 20-25, 2021*, pages 2846–2851. ACM.

Katsogiannis-Meimarakis, G. and Koutrika, G. (2023). A survey on deep learning approaches for text-to-sql. *The VLDB Journal*, 32(4):905–936.

Kersten, T., Leis, V., Kemper, A., Neumann, T., Pavlo, A., and Boncz, P. A. (2018). Everything you always wanted to know about compiled and vectorized queries but were afraid to ask. *Proc. VLDB Endow.*, 11(13):2209–2222.

Li, F. and Jagadish, H. V. (2014). Constructing an interactive natural language interface for relational databases. *Proc. VLDB Endow.*, 8(1):73–84.

Li, H., Zhang, J., Liu, H., Fan, J., Zhang, X., Zhu, J., Wei, R., Pan, H., Li, C., and Chen, H. (2024). Codes: Towards building open-source language models for text-to-sql. *Proc. ACM Manag. Data*, 2(3):127.

Martins, P., Afonso, A., and da Silva, A. S. (2023). Pylathedb - A library for relational keyword search with support to schema references. In *39th IEEE International Conference on Data Engineering, ICDE 2023, Anaheim, CA, USA, April 3-7, 2023*, pages 3627–3630. IEEE.

Pasupat, P. and Liang, P. (2015). Compositional semantic parsing on semi-structured tables. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China, Volume 1: Long Papers*, pages 1470–1480. The Association for Computer Linguistics.

Pena, E. H. M., de Almeida, E. C., and Naumann, F. (2021). Fast detection of denial constraint violations. *Proc. VLDB Endow.*, 15(4):859–871.

Ren, X., Su, L., Gu, Z., Wang, S., Li, F., Xie, Y., Bian, S., Li, C., and Zhang, F. (2022). HEDA: multi-attribute unbounded aggregation over homomorphically encrypted database. *Proc. VLDB Endow.*, 16(4):601–614.

Rodrigues, R. (2020). Legal and human rights issues of ai: Gaps, challenges and vulnerabilities. *Journal of Responsible Technology*, 4.

Sunyé, M. S. (2020). A quem servem os dados? *SBC Horizontes*, 4.

Xing, J., Wang, X., and Jagadish, H. V. (2024). Data-driven insight synthesis for multidimensional data. *Proc. VLDB Endow.*, 17(5):1007–1019.

Yu, J. X., Qin, L., and Chang, L. (2010). Keyword search in relational databases: A survey. *IEEE Data Eng. Bull.*, 33(1):67–78.

Zhang, L., Zhang, J., Ke, X., Li, H., Huang, X., Shao, Z., Cao, S., and Lv, X. (2023). A survey on complex factual question answering. *AI Open*, 4:1–12.

“Mais com Menos”- Processamento de Linguagem Natural Inteligente e Sustentável baseado em Engenharia de Dados e Inteligência Artificial Avançada

Marcos André Gonçalves[1], Leonardo Rocha[2], Washington Cunha[1], Guilherme Dal Bianco[3]

[1]Departamento de Ciência da Computação – Universidade Federal de Minas Gerais

[2]Departamento de Ciência da Computação – Universidade Federal de São João del-Rei

[3]Universidade Federal da Fronteira Sul

mgoncalv@dcc.ufmg.br, lcrocha@ufsj.edu.br

washingtoncunha@dcc.ufmg.br, guilherme.dalbianco@uffs.edu.br

Resumo. *Grandes Modelos de Linguagem (GMLs), baseados em técnicas de Inteligência Artificial (IA), tem revolucionado o Processamento de Linguagem Natural (PLN), sendo considerados o estado-da-arte em diversas tarefas práticas de PLN tais como classificação de texto, análise de sentimentos, sumarização de textos, e sistemas de perguntas-e-respostas. No entanto, tanto a construção desses modelos pré-treinados, quanto a sua adaptação a tarefas específicas (i.e. ajuste fino), demandam um custo computacional elevadíssimo, exigindo hardware e infraestrutura energética especializados, com impactos ambientais negativos em termos de emissão de* CO_2*. Mais ainda, a infraestrutura necessária para trabalhar nesses desafios no modelo hoje utilizado pelos big players, o qual denominamos de “Lei do Mais” (mais dados, mais hardware, mais energia) é não apenas insustentável, mas também desinteressante do ponto de vista de competitividade nacional – não temos os recursos financeiros e humanos para competir em pé de igualdade. Nesse contexto, nossa proposta de desafio é confrontar a “Lei do Mais” por meio do desenvolvimento de soluções inovadoras baseadas em engenharia de dados e IA Avançada que aumentem a efetividade dos modelos enquanto reduzem os custos computacionais e o consumo energético, com impactos ambientais positivos.*

Abstract. *Large Language Models (LLMs), based on Artificial Intelligence (AI) techniques, have revolutionized Natural Language Processing (NLP) and are considered state-of-the-art for many practical NLP tasks such as text classification, sentiment analysis, text summarization, and question-answering systems. However, constructing and adapting these pre-trained models to specific tasks (i.e., fine-tuning) requires extremely high computational costs, demanding specialized hardware and energy infrastructure, with negative environmental impacts regarding* CO_2*emissions. Furthermore, the infrastructure needed to address these challenges under the current paradigm used by big players – which we refer to as the ”More is Better Law”(more data, more hardware, more energy) – is unsustainable and unappealing from a national competitiveness perspective. We lack the financial and human resources to compete on equal footing. In this context, our challenge proposal is to confront the ”More is Better Law”by*

developing innovative solutions based on data engineering and Advanced AI that enhance model effectiveness while reducing computational costs and energy consumption, yielding positive environmental impacts.

Desafio

Desenvolver novas soluções *sustentáveis* que reduzam os *altos custos* computacionais, financeiros e de consumo energético, e consequente o impacto ambiental, associados à criação e ao ajuste-fino de Grandes Modelos de Linguagem (GMLs) utilizados em tarefas de PLN por meio de soluções avançadas de Engenharia de Dados e de Inteligência Artificial. Este desafio conecta 3 dos grandes temas da chamada da SBC: (1) IA; (2) Ciência dos Dados; e (3) Computação Sustentável. Potencialmente conecta também o tema (4) Computação Quântica pois existem soluções preliminares para este desafio (e.g. seleção de instâncias) implementadas por nosso grupo com esse paradigma computacional [7].

Argumentação

A Web permitiu que os usuários desempenhassem um papel crucial, não apenas no consumo, mas também na criação ativa de conteúdo. Isso vem resultando em volumes cada vez maiores de dados disponibilizados na Web (incluindo as Redes Sociais), dificultando a tarefa de encontrar de forma eficaz informações específicas. Organizar automática e adequadamente essas informações é portanto uma tarefa fundamental. Particularmente, no caso de dados textuais, ainda o principal tipo de dado encontrado na Web e nas redes sociais, tarefas de Processamento de Linguagem Natural (PLN), tais como Classificação Automática de Documentos e Análise de Sentimento tem se mostrado essenciais.

Técnicas de PLN testemunharam um avanço significativo na última década com propostas de estratégias inspiradas em arquiteturas de Redes Neurais (Aprendizado Profundo), incluindo as arquiteturas *Transformers* de 1ª e 2ª geração, tais como *RoBERTa* e *BART* [3, 5], e mais recentemente os Grandes Modelos de Linguagem (GMLs) (*Large Language Models* - LLMs) como *GPT* e *LLama* [10]. Os GMLs constituem o atual estado-da-arte em PLN, tendo alcançado resultados notáveis em diversas tarefas afim, incluindo recuperação de informação, ranqueamento (*ranking*), e perguntas-e-respostas (Q&A). Todavia, para alcançar esse notável desempenho, essas estratégias necessitam de grandes quantidades de dados no estágio de treinamento, muitas das vezes ruidosos e enviesados, impactando significativamente o custo computacional.

De fato, uma questão fundamental relacionada ao treinamento ou adaptação para tarefas específicas (i.e. ajuste-fino) de um modelo de aprendizagem profunda são os altos custos computacionais e de infraestrutura energética que, por sua vez, geram um grande impacto ambiental em termos de emissão de carbono. A quantidade de energia e tempo necessários para realizar o ajuste dos parâmetros para otimizar a eficácia dos modelos pode variar dependendo de diversos fatores, incluindo: (i) o tamanho e a complexidade do modelo; (ii) o hardware especializado utilizado, como GPU e TPU (consumo energético); e (iii) a quantidade de dados. Como a maior parte da eletricidade mundial é gerada a partir de combustíveis fósseis [9] o processo acima pode ser considerado diretamente responsável pela liberação de dióxido de carbono no meio ambiente. Nesse contexto, Patterson et al. [13] destacam que a fase de pré-treinamento do GPT-3 consumiu 1.287 MW/h. Esse consumo de energia resultou diretamente na emissão de pelo menos 552 toneladas de CO_2, equivalente ao consumo de um carro por cerca de 2,1 milhões de quilômetros.

Na busca constante por soluções mais efetivas, duas alternativas se destacam: (1) tornar os modelos mais complexos (e.g., mais parâmetros) ou (2) aumentar a quantidade e diversidade dos dados de treinamento, esbarrando por vezes em questões éticas, tais como o uso de dados pessoais ou com direitos preservados (*copyrighted*). O comum às duas alternativas é que ambas implicam no aumento de custos computacionais, energéticos e ambientais. Atualmente, esse é o modelo mais utilizado pelo *big players*, o qual denominamos de "Lei do Mais": mais dados, mais hardware, mais energia. Trata-se de um modelo que não é apenas insustentável, mas também desinteressante do ponto de vista de competitividade do Brasil a nível internacional, uma vez que não dispomos dos mesmos recursos financeiros e humanos que nações como os EUA e China, tão pouco das grandes companhias de TI do Vale do Silício. Em suma, o cenário que ilustramos envolve volumes de dados em crescente expansão, requisitos constantes de retreinamento, orçamentos multi-milionários e modelos de custo computacional e de energia de alta demanda.

A alternativa que advogamos em nossa proposta de Desafio vai na contra-mão da "Lei do Mais". De fato propomos como Grande Desafio para a comunidade nacional o investimento no desenvolvimento de soluções *sustentáveis* e de *baixo custo (financeiro, computacional e ambiental)* para a criação e ajuste-fino de Grandes Modelos de Linguagem utilizados em tarefas de Processamento de Linguagem Natural e de Recuperação de Informação. Particularmente, advogamos soluções baseadas em (1) engenharia de dados, incluindo pré-processamento [2, 14] e seleção de instâncias [3, 4, 7, 6] , destinadas a melhorar a qualidade e reduzir o volume dos dados de treinamento; e (2) estratégias de IA Avançadas, tais como: (i) compressão de modelos [12] aplicadas para reduzir a complexidade dos modelos de aprendizado profundo e (ii) aprendizado ativo [1], que também foca na redução de treinamento com "humanos-no-circuito"(humans-in-the-loop).

Direcionamentos

Para enfrentar o desafio proposto, apontamos como direcionamento para a comunidade nacional o desenvolvimento de novas soluções de engenharia de dados, e.g. seleção de instâncias [4], aplicadas na construção ou ajuste-fino de GMLs que sejam capazes de otimizar *três restrições simultaneamente*: (i) reduzir o volume de dados de treinamento; (ii) manter (ou mesmo melhorar) a eficácia removendo ruídos e redundâncias do treinamento; e (iii) reduzir o tempo total de aplicação de um modelo de ponta a ponta (que inclui desde as etapas tradicionais de pré-processamento até a etapa de treinamento do modelo) e, consequentemente, a emissão total de CO_2. Em situações específicas podemos incluir um quarto requisito, que seria: (iv) permitir que os modelos sejam mais explicáveis.

Generalizando, é fato que basicamente todas as tarefas de PLN fundamentadas em processos de aprendizado de máquina supervisionado e IA enfrentam as questões acima. Dessa forma, soluções para o desafio deverão também beneficiar diversas outras áreas, tais como sistemas de busca e recuperação da informação, sistemas de perguntas-e-respostas e sistemas de recuperação baseados em aprendizado de ranqueamento.

Do ponto de vista de técnicas de inteligência artificial avançadas, destacamos três abordagens principais para o desafio: duas relacionadas a compressão de modelos [12]: (i) quantização, um conjunto de técnicas que reduz a precisão numérica dos pesos dos modelos, geralmente convertendo valores em ponto flutuante para inteiros de 8 ou 4 bits; e (ii) poda (*pruning*), um conjunto de técnicas que elimina partes menos relevantes

do modelo, como neurônios e suas conexões, funcionando como uma generalização do *dropout*. Ambas as abordagens resultam na redução do tamanho dos modelos, otimizando seu desempenho e eficiência. A terceira alternativa que acreditamos que tem bastante potencial no contexto de reduzir a "Lei do Mais" é o Aprendizado Ativo (AA) que no contexto supervisionado busca selecionar um conjunto mínimo de instâncias não-rotuladas que, quando rotuladas pelo usuário, produzem um modelo com eficácia igual (ou até mesmo superior) se comparado a um conjunto completo de dados [1]. Diferentemente da SI, o AA foca em dados não rotulados e utiliza conhecimento especializado do usuário para a criação de conjuntos de treinamento mínimos para o ajuste-fino de modelos.

Questões Objetivas

1 - Por que o desafio proposto é relevante para o Brasil? Um exemplo prático imposto pelo uso de GMLs é o de empresas e grupos de pesquisa com restrições orçamentárias, especialmente em países em desenvolvimento. Diferente de nações como os EUA e a China, ou de grandes companhias do Vale do Silício, que dispõem de vastos recursos financeiros e humanos, no Brasil enfrentamos limitações significativas em ambas as dimensões. Isso afeta a capacidade de aplicar esses modelos em larga escala. Além disso, é comum a necessidade de realizar milhares de experimentos para alcançar avanços científicos ou inovações práticas. Portanto, este projeto tem o potencial de expandir significativamente a utilização desses modelos, tornando-os mais acessíveis, ao mesmo tempo em que promove uma abordagem mais sustentável e computacionalmente eficiente.

2 - Por que é viável considerar a proposta um desafio? A redução dos custos de treinamento e ajuste de GMLs é uma questão de relevância global considerando seu impacto ambiental. Com o compromisso político adequado e a alocação de recursos necessários, o projeto é realizável dentro do período de dez anos, viso a tendência mundial tanto na busca pela redução de custos quanto no aumento da adoção dessas tecnologias.

Nosso grupo vem trabalhando em algumas das direções apontadas alcançando relevantes resultados que demonstram a viabilidade de enfrentar este desafio com tecnologia nacional. Destaca-se a tese de doutorado de Felipe Viegas, vencedora do Prêmio Capes 2024. A tese propõe modelo um modelo simples, barato e criativo de representações de texto – Cluwords – que aproveita a eficiência e a interpretabilidade das representações (matriciais) tradicionais baseadas em frequências de palavras, ao mesmo tempo que explora as capacidades semânticas de modelos modernos baseados em *embeddings* de palavra. Esse modelo se mostrou tão ou mais eficaz que modelos do estado-da-arte baseado em Transformers, muito mais caros e complexos, em tarefas como Modelagem de Tópicos e Análise de Sentimentos [11, 15, 17, 16, 18]. Outro exemplo notável é a tese de doutorado de Washington Cunha [8], focada em *seleção de instâncias*, permitindo treinamento de modelos com conjuntos mínimos de dados, mantendo a efetividade. Mais especificamente, propomos e orquestramos novas etapas de pré-processamento baseadas em engenharia de dados aplicadas à PNL, oferecendo resultados tão eficazes quanto o estado-da-arte com um tempo de treinamento até 6,1 vezes menor. Esses exemplos demonstram a viabilidade de enfrentar o desafio, com fortes evidências de que o pré-processamento dos dados pode ser mais importantes que trabalhar em modelos de IA para alcançar resultados mais eficazes a menor custo.

3 - Quem deverá estar envolvido na sua solução? Nossa proposta convida toda a

comunidade científica brasileira a estar diretamente envolvida, com destaque especial para Computação. Pesquisadores, engenheiros e profissionais dessa área fornecerão as soluções tecnológicas essenciais para o desenvolvimento do projeto, desde a otimização de algoritmos até a criação de infraestruturas computacionais mais sustentáveis. Além disso, colaborações com outras áreas do conhecimento além da computação, como Matemática e Ciências Ambientais, são necessárias para integrar aspectos como eficiência energética e impacto ambiental nas soluções propostas. Além disso, a colaboração entre universidades, centros de pesquisa, empresas e órgãos governamentais será essencial para que a solução proposta se concretize e para que o Brasil se destaque no cenário global, não só pela adoção de novas tecnologias, mas também por sua capacidade de propor soluções sustentáveis para os desafios da computação de alta performance.

4 - Quais as iniciativas paralelas em outros Países? Diversos países ao redor do mundo já estão investindo em iniciativas para reduzir os custos e o impacto ambiental do treinamento de GMLs. Nos Estados Unidos, empresas como OpenAI, Google e Microsoft estão desenvolvendo técnicas de otimização que visam tanto a redução do consumo de energia quanto a melhoria na eficiência dos modelos. A criação de metodologias mais sustentáveis e a adoção de energias renováveis em data centers são algumas das medidas adotadas para mitigar os efeitos das emissões de carbono. Na Europa, países como Itália, Alemanha e França estão à frente em pesquisas que buscam alternativas mais verdes para o desenvolvimento de IA. Projetos como FAIR AI[1] incluem novos métodos de pré-processamento sustentáveis (*Green-aware AI*), oferecendo suporte ao desenvolvimento de IA de forma mais eficiente e com menor impacto ambiental.

Breve Curriculum Vitae dos Proponentes

M. A. Gonçalves é graduado pela UFC (1995), mestre pela UNICAMP (1997) e doutor pela Virginia Tech (2004), em Ciência da Computação. É professor titular da UFMG, pesquisador 1-B do CNPq e ex-membro afiliado da Academia Brasileira de Ciências.

L. Rocha é doutor em Ciência da Computação pela UFMG (2009), professor associado do Dept. de Ciência da Computação da UFSJ e Pesquisador 2 do CNPq.

W. Cunha é doutor em Ciência da Computação pela UFMG (2024) onde também obteve seu título de Mestre (2019) com Menção Honrosa no CTDBD-SBBD, possuindo publicações em importantes conferências e periódicos (e.g. *ACM SIGIR* e *ACM CSUR*).

G. Dal Bianco é doutor em C. Computação pela UFRGS (2014), e é professor adjunto na Univ. Fed. da Fronteira Sul (UFFS, 2015). Atualmente é pós-doutorando na UFMG.

Agradecimentos

Este trabalho contou com apoio do CNPq, CAPES, INCT-TILD-IAR, FAPEMIG, AWS, Google, NVIDIA, CIIASaúde, and FAPESP.

[1] `https://fondazione-fair.it/en/`

Referências

[1] G. D. Bianco, D. Duarte, and M. A. Gonçalves. Reducing the user labeling effort in effective high recall tasks by fine-tuning active learning. *IIS*, 61(2):453–472, 2023.

[2] W. Cunha, S. Canuto, F. Viegas, T. Salles, C. Gomes, V. Mangaravite, E. Resende, T. Rosa, M. Gonçalves, and L. Rocha. Extended pre-processing pipeline for text classification: On the role of meta-feature representations, sparsification and selective sampling. *IP&M*, 2020.

[3] W. Cunha et al. A comparative survey of instance selection methods applied to nonneural and transformer-based text classification. *ACM Comput. Surv.*, 2023.

[4] W. Cunha, C. França, G. Fonseca, L. Rocha, and M. A. Gonçalves. An effective, efficient, and scalable confidence-based instance selection framework for transformer-based text classification. In *ACM SIGIR*, pages 665–674, 2023.

[5] W. Cunha, V. Mangaravite, C. Gomes, S. Canuto, C. Nascimento, F. Viegas, C. França, W. S. Martins, J. M. Almeida, et al. On the cost-effectiveness of neural and non-neural approaches and representations for text classification: A comprehensive comparative study. *IP&M*, 58(3):102481, 2021.

[6] W. Cunha, A. Moreo Fernández, A. Esuli, F. Sebastiani, L. Rocha, and M. A. Gonçalves. A noise-oriented and redundancy-aware instance selection framework. *ACM Trans. Inf. Syst.*, 43(2), Jan. 2025.

[7] W. Cunha, A. Pasin, M. Goncalves, and N. Ferro. A quantum annealing instance selection approach for efficient and effective transformer fine-tuning. In *ACM ICTIR*, 2024.

[8] W. L. M. da Cunha. *A comprehensive exploitation of instance selection methods for automatic text classification*. PhD thesis, Aug. 2024. Available at `http://hdl.handle.net/1843/76441`.

[9] B. Dudley et al. Bp statistical review of world energy 2016. *British Petroleum Statistical Review of World Energy, Bplc. editor, Pureprint Group Limited, UK*, 2019.

[10] P. Liang et al. Holistic evaluation of language models. *Transactions on Machine Learning Research*, 2023. Featured Certification, Expert Certification.

[11] W. Luiz, F. Viegas, R. Alencar, F. Mourão, T. Salles, D. Carvalho, M. A. Gonçalves, and L. Rocha. A feature-oriented sentiment rating for mobile app reviews. In *WWW'18*, pages 1909–1918, 2018.

[12] F. M. Nardini, C. Rulli, S. Trani, and R. Venturini. Neural network compression using binarization and few full-precision weights. *arXiv preprint arXiv:2306.08960*, 2023.

[13] D. Patterson et al. The carbon footprint of machine learning training will plateau, then shrink. *Computer*, 55(7):18–28, 2022.

[14] M. Siino, I. Tinnirello, and M. La Cascia. Is text preprocessing still worth the time? a comparative survey on the influence of popular preprocessing methods on transformers. *Inf. Sys.*, 121:102342, 2024.

[15] F. Viegas, S. Canuto, C. Gomes, W. Luiz, T. Rosa, S. Ribas, L. Rocha, and M. A. Gonçalves. Cluwords: Exploiting semantic word clustering representation for enhanced topic modeling. In *Proc. of the 12th ACM Int. Conf. on Web Search and Data Mining*, WSDM '19, page 753–761, 2019.

[16] F. Viegas, S. D. Canuto, W. Cunha, C. França, C. M. V. de Andrade, G. Fonseca, A. Machado, L. Rocha, and M. A. Gonçalves. Pipelining semantic expansion and noise filtering for sentiment analysis of short documents - clusent method. *J. Interact. Syst.*, 15(1):561–575, 2024.

[17] F. Viegas, W. Cunha, C. Gomes, A. Pereira, L. Rocha, and M. Goncalves. CluHTM - semantic hierarchical topic modeling based on CluWords. In *ACL*, pages 8138–8150, 2020.

[18] F. Viegas, A. Pereira, W. Cunha, C. França, C. Andrade, E. Tuler, L. Rocha, and M. A. Gonçalves. Exploiting contextual embeddings in hierarchical topic modeling and investigating the limits of the current evaluation metrics. *Computational Linguistics*, pages 1–41, 03 2025.

Desenvolvimento e Automação de Software de Baixa Energia
Refinamento do Tópico de Computação sustentável
Luigi Carro - INF - UFRGS

Resumo
A crescente utilização de recursos computacionais pela sociedade em geral pode significar 21% do consumo mundial de energia em 2030 [1]. Neste cenário, as opções para redução do consumo de energia podem vir por mudanças no hardware ou software. Infelizmente, a diminuição do consumo energético pela melhoria do hardware encontra-se no limite, e a solução para redução do consumo terá de vir pelo desenvolvimento de software de baixo consumo energético. Contudo, as técnicas de software exigidas para redução de energia não são óbvias, e demandam tempo de projeto inaceitável para a velocidade de produção de software requerida pela sociedade atual. Nesta proposta discute-se o desafio de diminuir a energia somente via software, e o desenvolvimento de ferramentas que impeçam um aumento no tempo de produção de software.

Abstract
The increasing use of computing resources by society in general could account for 21% of global energy consumption by 2030 [1]. In this scenario, the options for reducing energy consumption may come from changes in hardware or software. Unfortunately, reducing energy consumption through hardware improvements is limited, and the solution to reducing consumption will have to come from developing low-energy software. However, the software techniques required to reduce energy are not obvious, and demand unacceptable design time for the software production speed required by today's society. This proposal discusses the challenge of reducing energy through software alone, and the development of tools that prevent an increase in software production time.

1)Alinhamento deste desafio com computação sustentável

O consumo de CO2 dos datacenters só aumenta. Pelos últimos 50 anos a solução para diminuição de energia sempre foi aguardar a próxima versão da tecnologia, que com o fim da lei de Moore não pode mais ser usada. Haverá portanto pressão da sociedade para que a computação diminua seu *footprint* de carbono, e também porque o rápido crescimento da demanda energética não poderá ser coberto pela expansão do sistema elétrico brasileiro. Sendo assim, mudanças no software serão a maneira de se conseguir alto desempenho com baixa energia, em diferentes plataformas (mobile e datacenters). As propostas deste documento lidam não somente com um problema brasileiro, mas mundial.

2)Os limites do hardware

Atualmente, software é executado em dois extremos: plataformas mobile, com severas restrições de potência, onde a duração da bateria é um prêmio; e nos datacenters, com CPUs ou GPUs, onde o volume de computadores exige enorme quantidade de recursos para resfriamento, e a energia é extremamente custosa em relação ao preço de funcionamento do

datacenter. Dados recentes apontam que 40% da energia elétrica consumida num datacenter é voltada à refrigeração [2].

Soluções de hardware poderiam ser utilizadas, e o são no telefone celular, por exemplo. Tem-se, pelo limite legal de 3W de potência, um conjunto de aceleradores especializados de vídeo e áudio que são ligados somente quando necessário. Contudo, a arquitetura de um celular é extremamente heterogênea, com diversas CPUs internas que funcionam como aceleradores, e seu desenvolvimento é custoso. Portanto, esperar que o hardware seja a solução é inviável, já que o fim da lei de Moore está muito presente. Para novos nós tecnológicos a potência máxima tem se mantido constante, e os ganhos energéticos são cada vez mais difíceis, como mostra o dark silicon.

A lei de Moore mais fraca significa que o tempo para se dobrar o número de transistores na mesma área deve passar de 2 anos para um período maior. Na figura 1, especulativamente, tem-se curvas de 2, 3 e 4 anos. A realidade estará provavelmente entre as curvas de 2 a 4 anos. Contudo, a desaceleração da lei de Moore é apenas uma variável. Outra variável importante é a velocidade de crescimento de software.

O crescimento do software é uma variável mais difícil de se obter. Por exemplo, há o crescimento em custo, estimado em 12% ao ano. Mas isto não reflete o número de linhas executadas, somente seu impacto. Dados da Nasa sobre o crescimento do número de linhas de código em seus projetos indicam um crescimento de 25% ao ano durante os anos de 1990 até 2010 [3]. Outra fonte facilmente observável é o crescimento do sistema operacional Windows, que passou de 10 milhões de linhas em 1993 para 50 milhões de linhas em 2003 [4]. Nestes 10 anos tem-se também 25% de crescimento composto ao ano.

Como a métrica de número de linhas de código é imperfeita, sobretudo com o uso corrente de linguagens mais abstratas, e portanto com maior capacidade de expressão, para melhor refletir realidade utilizou-se a equação clássica de arquitetura de computadores,
$T=\#I*CPI*Tciclo$ (1).

Em (1) o tempo de execução depende do número de instruções a serem executadas, do número de ciclos médio por instrução, e do tempo de ciclo. Partindo-se de 1 bilhão de instruções, com um ciclo de relógio de 1 GHz (1 segundo de execução numa CPU), pode-se analisar como o impacto do crescimento no número de instruções executadas (CAGR 25%aa) e das melhorias eventuais do hardware (baixa no Tciclo pela lei de Moore) são combinados. Na figura 1 assumiu-se um CPI de 1,5, levando-se em conta os efeitos de cache miss e diferença de velocidade de memórias DRAM, cada vez mais impactantes nos problemas atuais movidos por enorme massa de dados.

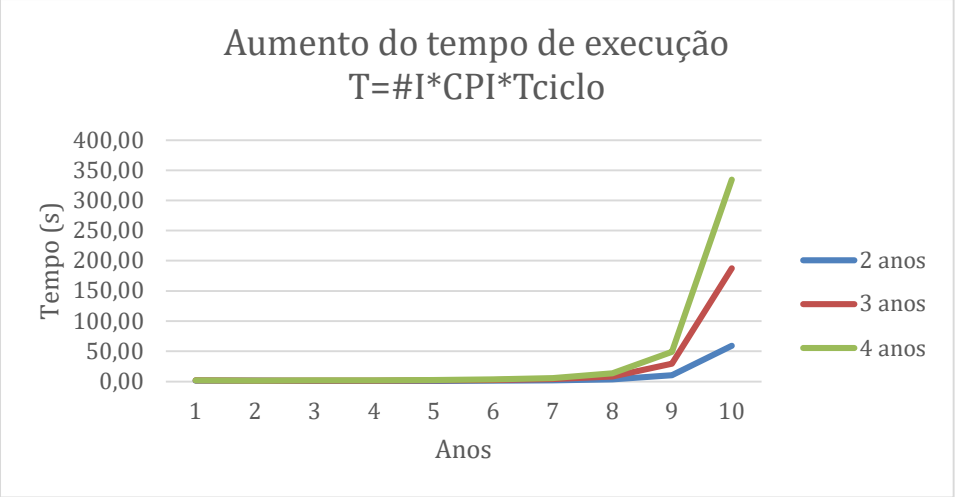

Figura 1 - os efeitos de uma lei de Moore mais fraca junto com o crescimento do número de instruções executada. Fonte: o autor

Na figura 2 tem-se como ambas as variáveis, crescimento do software e velocidade do hardware crescendo mais lentamente, afetarão a energia. Dado que a Energia é expressa por E = P*T (2), e assumiu-se a potência constante de 50W, como limite máximo de dissipação, e usou-se o tempo de computação da figura 1 (3 anos). O efeito é claramente exponencial, já que o crescimento do tempo também o é.

Mesmo arquiteturas novas como GPUs estão submetidas ao mesmo processo. A diferença aqui é que existem *n* processadores a serem usados, com o efeito de se diminuir drasticamente o tempo de execução, desde que a aplicação seja massivamnete paralela, mas com um excesso de potência. Neste contexto tecnológico, a responsabilidade para se obter baixa energia será do software.

3)As necessidades do software e o gap para produção de software de baixa energia

Como já mencionado neste artigo, o crescimento do software implica em milhões de linhas de código. Tem-se hoje o recurso de linguagens mais abstratas de programação, e o uso indiscriminado de APIs, tornando mais produtivo o programador. Na outra ponta, linguagens mais abstratas são obviamente mais distantes do hardware, e portanto é mais difícil para o programador controlar os aspectos que consomem energia (ponto flutuante, cache misses, acesso às memórias DRAM, predição de saltos, etc). Os dados que devem ser processados também sofreram um aumento exponencial na última década, e a tendência é que continuem aumentando. O consumo energético do ChatGPT-3, por exemplo, foi igual ao de uma cidade dinamarquesa de 75 mil pessoas em janeiro de 2023.

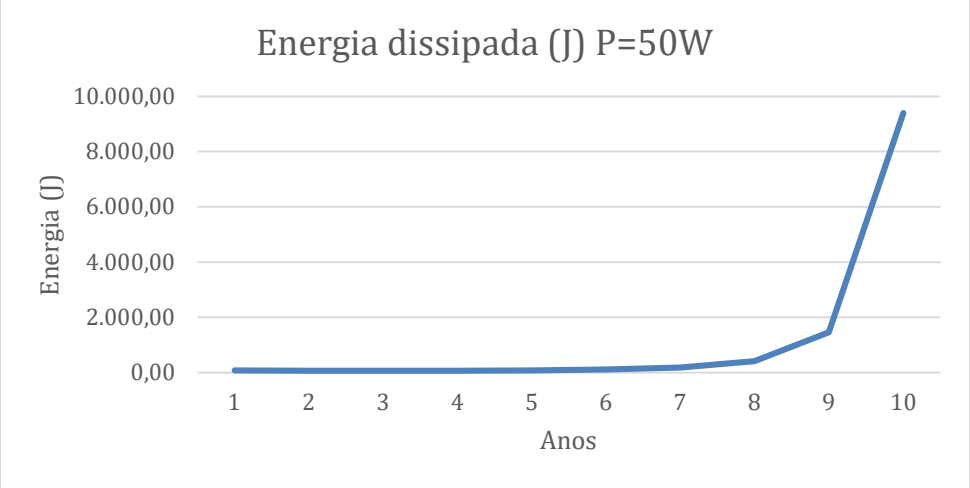

Figura 2 - impacto na energia de uma CPU. Fonte: o autor

A produção de software hoje em dia expõe o programador a uma enorme quantidade de APIs, mas o programador não tem noção do efeito de seu uso no consumo. Além disto, a velocidade exigida para entrega dos produtos de software não diminui, e portanto a pressão por velocidade de entrega não permite uma reflexão, por parte do desenvolvedor, de como atingir mínimos energéticos.

As necessidades de abstração e manutenção do software moderno acarretam mais consumo. Por exemplo, em [5] demonstra-se como uma ferramenta que auxilia na abstração provoca aumento de consumo de energia de até 65%. Embora existam algumas estratégias para diminuição de consumo em programas, largamente estudadas no âmbito de sistemas embarcados, todas elas implicam em alto tempo de projeto. O resultado então é uma impossibilidade: precisa-se de mais ferramentas para se desenvolver software de milhões de linhas de código, mas estas ferramentas aumentam o consumo, e ainda não há estratégias consolidadas de aplicação de técnicas de redução de consumo por parte do programador, ou de ferramentas que ajudem na diminuição do consumo energético sem prejuizo do tempo de projeto.

Some-se a este cenário a falta de treinamento de programadores. Como reduzir em 50% o consumo de um programa? Esta é uma pergunta para a qual os estudantes não foram treinados adequadamente, e há no mercado mundial poucos profissionais capazes de responde-la.

4)O desafio em si

O desafio pode ser dividido em duas partes. Na primeira, tem-se de desenvolver técnicas de redução de energia em software para aplicações-chave. Por exemplo, em [6] obteve-se uma redução de 96 vezes no consumo de CNNs grandes, somente usando técnicas de software. Em [7] foram estudados diferentes implementações de árvores e *forest-trees*, e observou-se que árvores diferentes demandam soluções diferentes, com ganhos de energia de até 4x, puramente pela escolha do algoritmo correto. Tem-se então a segunda parte do desafio: além da produção de algoritmos que possam economizar energia, tem-se de automatizar seu uso, que será função

do tipo de dado com que se está trabalhando. Este, na verdade, é um desafio mundial da Computação, não apenas brasileiro.

5)Possíveis soluções e métricas e avaliação

Algumas soluções já estão em andamento. Por exemplo, os banco de dados vetoriais partem de um modelo diferente, para melhor usar o hardware massivamente paralelo que se encontra à disposição na forma de GPUs. Embora a motivação inicial tenha sido desempenho, pela extrema velocidade de execução da multiplicação de matrizes a energia fica reduzida.

Além dos estudos em [6-7], há 2 anos o proponente oferece uma disciplina opcional de Software de Baixa energia, onde diferentes estratégias são apresentadas aos alunos. Pretende-se ampliar o esforço na produção de artigos, livros e material didático, que possa ser distribuído pelo país, para que mais alunos tenham contato com o tema. O ponto de "pensar sobre o assunto" já foi atingido, mas deve-se escalar a pesquisa na área. A implementação de datacenters no Brasil, e a chamativa economia pelo uso adequado de software podem ser ótimos catalisadores.

Quanto a métricas de avaliação, estas seriam relativamente fáceis de serem obtidas, pelo número de pessoas capazes de pensar o software de baixa energia, e pela implementação do mesmo, medindo-se o desempenho de datacenters no país.

Referências

[1] N. Jones. How to stop data centres from gobbling up the world's electricity. Nature, 561(7722):163–167, 2018.

[2] https://www.digitalrealty.com/resources/articles/future-of-data-center-cooling Último acesso em 08/09/2024

[3] Butler, R.; Pennotti, M. The evolution of software and its impact on complex system design in robotic spacecraft embedded systems. Dec 2013, Procedia Computer Science.

[4] https://en.wikipedia.org/wiki/Source_lines_of_code Último acesso em 08/09/2024

[5] Calero. C.; Polo, M.; Moraga, A. Investigating the impact on execution tiem and energy consumption of developing with Spring. Sustainable Computing: Informatics and Systems. 2021, Elsevier.

[6] Gonçalves, L.R.; Moura, R.F.; Carro, L. Aggressive energy reduction for video inference with software-only strategies. ACM transactions Embedded Computer Systems, v.18, issue 5, p. 1-20, 2019.

[7] Bergozzi, V.; Carro, L. Low energy decision trees. Relatório interno.

Computação Sustentável e Energeticamente Eficiente

Daniel Cordeiro[1], Emilio Francesquini[2], Alessandro Santiago dos Santos[3], Alvaro Luiz Fazenda[4], Silvana Rossetto[5] *

[1]Universidade de São Paulo (USP), São Paulo, Brasil

[2]Universidade Federal do ABC (UFABC), Santo André, Brasil

[3]Instituto de Pesquisas Tecnológicas (IPT), São Paulo, Brasil

[4]Universidade Federal de São Paulo (Unifesp), São José dos Campos, Brasil

[5]Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brasil

daniel.cordeiro@usp.br, e.francesquini@ufabc.edu.br, alesan@ipt.br
alvaro.fazenda@unifesp.br, silvana@ic.ufrj.br

Resumo. *Apresentamos refinamentos sobre os desafios sugeridos nas temáticas relacionadas com Computação Sustentável e os impactos sociais e econômicos da computação. Discutimos as demandas relacionadas com problemas computacionalmente complexos que requerem infraestruturas de processamento de alto desempenho e, consequentemente, geram demandas por soluções de hardware e software energeticamente eficientes. Destacamos a necessidade de incorporar essas questões na formação dos profissionais da área, incentivando e dando condições para se adotar abordagens de desenvolvimento de soluções computacionais cientes do consumo energético e ambientalmente sustentáveis. Por fim, propomos potenciais métricas para avaliação do progresso das soluções propostas.*

Abstract. *We present refinements on the challenges suggested in the themes related to sustainable computing and the social and economic impacts of computing. We discuss the demands related to complex computational problems that require high-performance processing infrastructures and, consequently, generate demands for energy-efficient hardware and software solutions. We highlight the need to incorporate these issues into the training of professionals in the field, encouraging and enabling them to adopt energy-conscious and environmentally sustainable approaches to developing computing solutions. Finally, we propose potential metrics for evaluating the progress of the proposed solutions.*

1. Contexto atual

Os desafios ambientais contemporâneos são de escala global e exigem uma abordagem interdisciplinar. Entre eles, destacam-se a drástica redução da biodiversidade em diversos biomas, a poluição atmosférica, o aquecimento global, a alta pegada energética e os danos causados por desastres naturais. Os avanços significativos em tecnologias de informação

*O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), processos #2023/00811-0 (“EcoSustain–Ciência de Dados e Computação para o Meio Ambiente”) e #2024/01115-0 (“CCD–Cidades Carbono Neutro”).

e comunicação (TIC), ciência de dados e inteligência artificial nas últimas décadas oferecem oportunidades promissoras para aplicar essas inovações em prol do meio ambiente. No entanto, o progresso da computação e a crescente complexidade dos problemas abordados têm gerado uma demanda substancial por energia. Um estudo de [Jones et al. 2018] aponta que atualmente cerca de 12% da energia global é consumida por *data centers*, computadores pessoais, celulares e TVs, com previsão de aumento para 20,9% até 2030. Para conter ou reduzir o consumo de energia, é necessário um esforço em várias frentes.

Um primeiro aspecto envolve o desenvolvimento de soluções computacionais — tanto de hardware quanto de software — que sejam energeticamente eficientes e de alto desempenho. Outra, de caráter mais humano, refere-se à revisão dos currículos de formação dos profissionais da área, permitindo que cultivem uma abordagem de desenvolvimento consciente em relação ao consumo energético, e que seja possível lhes dar um suporte adequado para atender às demandas sem sobrecarga excessiva de trabalho. O trabalho de Carro e Nazar [Carro and Nazar 2023] oferece uma visão abrangente sobre os desafios da computação energeticamente eficiente no contexto de aplicações com uso intensivo de dados. Os autores destacam que, para garantir a continuidade da evolução da computação, é fundamental buscar soluções mais eficazes para transformar algoritmos complexos em hardware capaz de executá-los de maneira eficiente e com economia de energia. Eles argumentam que a simples ampliação da quantidade de hardware para melhorar o desempenho computacional não tem se mostrado uma solução escalável nem ambientalmente sustentável. No âmbito do software e suas aplicações, um exemplo relevante é a computação aproximada, que propõe reduzir controladamente a precisão dos resultados em troca de uma economia significativa de energia durante o processamento. Alguns exemplos dessa abordagem incluem desde cálculos numéricos [Sudo et al. 2022] até simulações de tráfego urbano em cidades verdes [Rocha et al. 2022].

Outro aspecto diz respeito ao modelo de computação em nuvem que hoje é praticamente o padrão para novos desenvolvimentos e implantação de software escalável. Tecnologias vinculadas à virtualização e à computação em nuvem têm transformado a indústria, ao reduzir os custos de infraestrutura de TI, facilitar a implantação de novos produtos, diminuir os gastos com manutenção e permitir o ajuste rápido dos recursos conforme as variações de carga de trabalho [Oda et al. 2018]. Grandes empresas, como Netflix, LinkedIn e Facebook, adotam a computação em nuvem para oferecer soluções eficientes e escaláveis. No entanto, na base desse modelo estão os *data centers* de grande escala, que concentram um elevado consumo de energia em áreas relativamente pequenas. Além disso, para garantir alta disponibilidade e tolerância a falhas, esses *data centers* costumam ser geograficamente distribuídos, gerando impactos ambientais em várias regiões. À medida que o uso de soluções em nuvem cresce, aumenta também a necessidade de melhorar a eficiência energética e reduzir a pegada de carbono dessas grandes instalações [Cordeiro et al. 2023]. *Computação Verde* é uma área consolidada da Computação, mas repensar a Computação para torná-la adaptável aos seus impactos ambientais (ou seja, torná-la sustentável) em tal escala é um novo problema em si. O desafio que se apresenta é, portanto, desenvolver estratégias que tornem as aplicações que dependem de infraestruturas de nuvem em larga escala mais sustentáveis, por meio do projeto, gerenciamento e otimização de sua execução de forma eficiente. Esse esforço envolverá não apenas a academia e a indústria, mas também profissionais de ciência da computação de diversas áreas [Cordeiro et al. 2023]. Como mencionado

anteriormente, essa questão transcende a tecnologia, englobando também a formação humana. Os currículos acadêmicos em computação precisarão se adaptar para refletir os avanços nas pesquisas dessa área, promovendo uma nova cultura de desenvolvimento de soluções computacionais que considere não apenas a complexidade de processamento, o uso de memória e a comunicação, mas também o consumo de energia.

Tais exemplos destacam a relevância da computação sustentável e como esta se torna cada vez mais evidente e urgente. A criação de soluções computacionais que considerem tanto o consumo energético quanto o impacto ambiental é uma necessidade crítica. Trata-se de uma área estratégica, pois possibilita o desenvolvimento de sistemas capazes de resolver problemas complexos e processar grandes volumes de dados de forma eficiente, robusta e sustentável, tanto do ponto de vista ambiental quanto social. O gerenciamento de recursos ciente de energia traz desafios desde o dimensionamento do hardware, considerando a pegada de carbono (manufatura, uso e descarte) até o escalonamento de recursos ciente de energia com garantia de qualidade de serviço, levando-se em conta a intermitência imposta pelo uso de energias renováveis. O uso de soluções de software e hardware especializado (ex., FPGAs, ASICs, TPUs, VPUs) pode tornar mais eficientes — em termos de consumo de energia, tempo de processamento e custo financeiro — a resolução de problemas de domínio específico. Nesse sentido, faz-se necessário a construção de ambientes de desenvolvimento de software mais acessíveis para os desenvolvedores.

2. Desafio

O desafio que se coloca pode ser subdividido nas seguintes questões: (i) como desenvolver soluções computacionais — tanto em termos de hardware como de software — energeticamente eficientes e ambientalmente sustentáveis? (ii) como tornar as infraestruturas de computação em larga escala — necessárias para lidar com problemas complexos — mais sustentáveis? (iii) como definir e inserir métricas de consumo de energia na avaliação de algoritmos e aplicações? (iv) como motivar e dar condições para que os programadores desenvolvam soluções cientes do consumo de energia e dos impactos ambientais?

3. Soluções possíveis

[Cordeiro et al. 2023] apontam algumas possíveis direções para lidar com as questões colocadas, organizadas em três temáticas apresentadas nas subseções seguintes:

Gerenciamento de Recursos Ciente de Consumo de Energia. As plataformas de nuvem estão investindo no uso de energia renovável, mas ainda cabe ao programador utilizar eficientemente os recursos disponíveis (incluindo hardware especializado), garantir o balanceamento de carga e minimizar a movimentação de dados. Os futuros *data centers* com baixo impacto ambiental e alta eficiência energética precisarão aumentar ainda mais o uso de fontes de energia renováveis. Será necessário fornecer aos desenvolvedores de aplicações ferramentas para manter o escopo de seu desenvolvimento dentro do problema investigado e para minimizar o esforço necessário para executar suas aplicações na nuvem, de forma eficiente em termos de desempenho e energia. **Hardware e Software Especializados**. A criptomineração é um caso de sucesso de hardware especializado para fornecer soluções com melhor desempenho e menor consumo de energia para um problema específico (com transição de CPUs para GPUs, FPGAs e ASICs). Melhorias adicionais podem ser alcançadas por software especializado que considera as características

intrínsecas de uma aplicação. A Computação Aproximada é um exemplo de abordagem na qual a eficiência energética e o ganho de desempenho podem ser alcançados com soluções não exatas. Os futuros *data centers* podem se tornar mais eficientes ao fornecer hardware especializado ou soluções de Software como Serviço (SaaS) que considerem as características de cada aplicação, em vez de uma solução única. **Computação de Alto Desempenho**. No contexto da computação de alto desempenho, os benefícios de usar recursos de nuvem pública a tornam uma alternativa atraente aos caros clusters locais. No entanto, o ecossistema de software necessário para tornar possível uma plataforma de nuvem sustentável para computação de alto desempenho ainda não está maduro. Serão necessárias novas soluções para gerenciar e executar com eficiência as aplicações e fluxos de trabalho científicos em nuvens. As soluções atuais deverão ser adaptadas para prever o consumo de energia e a pegada de carbono das aplicações, ajudando os pesquisadores a construir novos escalonadores e soluções de forma geral, cientes da energia.

4. Potenciais métricas de avaliação

Como potenciais métricas de avaliação, podemos considerar as listadas a seguir. Essas métricas podem ajudar a fornecer uma avaliação abrangente, permitindo identificar áreas de melhoria e assegurar que os objetivos de sustentabilidade sejam alcançados. **Consumo de Energia**: (i) *Energia Total Consumida*, medida em kilowatt-hora (kWh) durante a execução de uma aplicação ou consumida por um *data center*; e (ii) *Eficiência Energética*, relação entre os resultados obtidos e a energia consumida. **Impacto Ambiental**: (i) *Pegada de Carbono*, emissões de CO_2 associadas à energia consumida, incluindo a análise do ciclo de vida do hardware; e (ii) *Redução de Resíduos*, avaliação do volume de resíduos eletrônicos gerados e estratégias de reciclagem ou reutilização. **Desempenho Computacional**: (i) *Makespan*, duração total para executar tarefas específicas ou processar grandes volumes de dados; e (ii) *Vazão*, quantidade de dados processados por unidade de tempo. **Escalabilidade**: *Capacidade de Expansão*, avaliação de como o sistema pode ser ampliado para atender a aumentos na carga de trabalho sem comprometer o desempenho ou a eficiência energética. **Custo Financeiro**: (i) *Custo Operacional*, total gasto com energia e recursos; e (ii) *Custo por Tarefa/Processamento*, custo médio associado à execução de tarefas específicas. Importante destacar que para que uma proposta sustentável tenha atratividade, é necessário que o custo financeiro seja compatível com o custo atual. **Sustentabilidade Social**: *Acessibilidade*, medida de quão acessíveis são as soluções desenvolvidas para diferentes usuários e desenvolvedores. **Utilização de Hardware Especializado**: *Desempenho do Hardware*, comparação entre hardware especializado (ex. FPGAs, ASICs) e hardware genérico em termos de eficiência energética e desempenho.

Referências

Carro, L. and Nazar, G. L. (2023). Desafios para a computação energeticamente eficiente. *Sociedade Brasileira de Computação*.

Cordeiro, D., Francesquini, E., Amarís, M., Castro, M., Baldassin, A., and Lima, J. (2023). Green cloud computing: Challenges and opportunities. In *Anais Estendidos do XIX Simpósio Brasileiro de Sistemas de Informação*, pages 129–131, Porto Alegre, RS, Brasil. SBC.

Jones, N. et al. (2018). How to stop data centres from gobbling up the world's electricity. *nature*, 561(7722):163–166.

Oda, R., Cordeiro, D., and Braghetto, K. R. (2018). Dynamic resource provisioning for scientific workflow executions in clouds. In *2018 IEEE International Conference on Services Computing (SCC)*, pages 291–294. IEEE.

Rocha, F. W., Fukuda, J. C., Francesquini, E., and Cordeiro, D. (2022). Accelerating smart city simulations. In Gitler, I., Barrios Hernández, C. J., and Meneses, E., editors, *High Performance Computing*, pages 148–162, Cham. Springer International Publishing.

Sudo, M. A., Fazenda, A. L., and Souto, R. P. (2022). Mixed precision applied on common mathematical procedures over gpu. In *Simpósio em Sistemas Computacionais de Alto Desempenho (SSCAD)*, pages 265–275. SBC.

Sobre os proponentes:

Daniel Cordeiro é Professor Doutor na Escola de Artes, Ciências e Humanidades da USP. É Doutor em *Mathématiques et en Informatique* pela Université de Grenoble, França, e Mestre e Bacharel em Ciência da Computação pela USP. Suas áreas de pesquisa incluem Computação de Alto Desempenho, Teoria do Escalonamento e Computação Sustentável. Seus projetos de pesquisa atuais estão relacionados à execução sustentável (de baixo carbono) de aplicações de alto desempenho em plataformas de Computação em Nuvem geo-distribuídas (ID Lattes: 5322325760113475).

Emilio Francesquini é professor adjunto do Centro de Matemática, Computação e Cognição da UFABC. Possui pós-doutorado na UNICAMP, doutorado em dupla titulação pela USP (Doutor em Ciências) e pela Universidade de Grenoble-Alpes, França (UGA, *Docteur Spécialité Informatique*). É também Mestre e Bacharel em Ciência da Computação pela USP. Seus projetos de pesquisa atuais incluem o emprego de computação aproximada para redução do impacto ambiental de sistemas de alto desempenho; além do estudo e implantação de soluções de escalonamento verdes em plataformas de computação em nuvem geo-distribuídas (ID Lattes: 8949216028517727).

Alessandro Santiago possui pós-doutorado na Universidade de Lisboa no tema de cidades neutras de carbono, doutorado e mestrado pela USP e bacharelado em Ciência da Computação pela UFMT. Atualmente é coordenador do mestrado profissional em computação aplicada e gerente de apoio de negócios de tecnologias digitais do IPT, lidando com desafios de pesquisa em cidades inteligentes, internet das coisas e monitoramento digital em sociedades de baixo carbono. Foi membro do comitê de contingência de combate ao coronavírus em São Paulo atuando no Sistema de Monitoramento Inteligente, em projetos de rodovias inteligentes pelo Estado de SP, e cooperação internacional com redes de pesquisas europeias (ID Lattes: 9738704704763672).

Álvaro Fazenda possui graduação em Computação Científica pela Universidade de Taubaté, mestrado e doutorado em Computação Aplicada pelo INPE e Pós-Doutorado pela University of Illinois at Urbana-Champaign. Atualmente é Professor Associado na UNIFESP. Atuou como Pesquisador Visitante no CPTEC/INPE e como professor assistente na Universidade de Taubaté. Tem experiência na área de Ciência da Computação e Computação Científica, com ênfase em: processamento de alto desempenho, sistemas distribuídos, ciência cidadã, dinâmica de fluídos computacional, modelos numéricos de

previsão de tempo/clima e geoprocessamento (ID Lattes: 7606159905559544).

Silvana Rossetto possui graduação em Ciência da Computação e mestrado em Informática pela UFES, e doutorado em Informática pela PUC-Rio. Realizou o programa de doutorado sanduíche no exterior, pela Politecnico di Milano (2004/2005). Atua na área de Ciência da Computação, com ênfase em computação concorrente, paralela e distribuída. Atualmente exerce o cargo de Professor Associado no Instituto de Computação da UFRJ (ID Lattes: 0054098292730720).

Os Desafios de Cibersegurança dos Referenciais do BCS/SBC

Aldri Santos (UFMG)[1], Altair Santin (PUCPR)[1], André Grégio (UFPR)[1], Carlos Raniery (UFSM)[1], Daniel Batista (USP)[1], Dianne Medeiros (UFF)[1], Diego Kreutz (UNIPAMPA)[1], Diogo Mattos (UFF)[1], Edelberto Franco (UFJF)[1], Igor Moraes (UFF)[1], Lourenço Pereira Jr. (ITA)[1], Marcos Simplicio (USP)[1], Michele Nogueira (UFMG)[1], Michelle Wangham (Univali/RNP)[1], Natalia Fernandes (UFF)[1], Rodrigo Miani (UFU)[1]

[1]Comissão Especial de Cibersegurança - Sociedade Brasileira de Computação (CESeg)

Abstract. *This document describes the main research and innovation challenges in the area of cybersecurity. The content builds upon education as a key element: it is organized into eight pillars, comprising the main areas listed in the Cybersecurity Curricular Guideline, published by the Brazilian Computer Society as a guide to inspire the creation of courses to train qualified personnel. The result is a comprehensive set of research topics addressing crucial cybersecurity needs in contemporary computer systems.*

Resumo.
Este documento descreve os principais desafios de pesquisa e inovação na área de cibersegurança. O conteúdo toma por base a educação como elemento-chave: ele se organiza em oito pilares, cobrindo as principais áreas elencadas nos Referenciais de Formação do Curso de Bacharelado em Cibersegurança (RF-CS), documento publicado pela Sociedade Brasileira de Computação como um guia para inspirar a criação de cursos para a formação de pessoal qualificado na área. O resultado é um conjunto abrangente de tópicos de pesquisa que abordam necessidades cruciais de cibersegurança em sistemas computacionais contemporâneos.

1. Introdução

O *déficit* estimado de profissionais de Cibersegurança apenas no Brasil é de 750 mil[1]. Em 2023, a comissão de educação da Comissão Especial de Cibersegurança (CE-Seg) da Sociedade Brasileira de Computação (SBC) publicou os Referenciais de Formação do Curso de Bacharelado em Cibersegurança (RF-CS) [SBC, 2023], que busca guiar a formação de profissionais em Cibersegurança, área crítica na Era da Informação atual, por meio de 8 eixos de formação: Segurança de Dados, Segurança de Sistemas, Segurança de Conexão, Segurança de Software, Segurança de Componentes, Segurança Organizacional, Fatores Humanos em Segurança, e Segurança e Sociedade. Dado que os eixos do RF-CS cobrem amplamente a área de Cibersegurança, este documento resume como os desafios da área podem ser divididos em temas de pesquisa, a fim de direcionar eventuais ações de colaboração entre academia, indústria e sociedade para lidar com aspectos relevantes de pesquisa em cibersegurança no Brasil e no mundo.

[1]Disponível em `https://www.fortinet.com/content/dam/fortinet/assets/reports/pt_br/2024-cybersecurity-skills-gap-report.pdf`.

2. O Desafio da Cibersegurança

O grande desafio da ciberseguraça no Brasil é detalhado em oito eixos temáticos, conforme definido pelo RF-CS da SBC. Os desafios de cada eixo são detalhados a seguir.

2.1. Segurança de Dados

Mitigar a ameaça da computação quântica é um dos principais desafios no eixo de Segurança de Dados, que trata da proteção de dados armazenados, em processamento e em trânsito. A razão é que computadores quânticos de grande porte podem obter chaves privadas dos principais esquemas criptográficos usados na Internet atualmente. Uma solução para esse problema é a chamada criptografia pós-quântica (*post-quantum cryptography* - PQC). O NIST lançou um processo de padronização de algoritmos criptográficos pós-quânticos. Apesar desse avanço, desafios importantes permanecem tanto para os esquemas já padronizados como para possíveis alternativas, considerando métricas de segurança, desempenho e flexibilidade. Em especial, destacam-se como áreas de pesquisa: (1) análise de segurança dos esquemas, cobrindo aspectos teóricos e de implementação (2) a otimização dos algoritmos, considerando técnicas de projeto e de implementação, para reduzir o uso de recursos a níveis mais próximos dos algoritmos clássicos; (3) a adaptação de soluções existentes para uso de PQC. Uma alternativa consiste no uso de tecnologias de comunicação quântica. Especificamente, aproveitando as propriedades físicas dos meios de comunicação, há a possibilidade de se fazer a distribuição de chaves secretas (*quantum key distribution* – QKD) entre usuários com garantias de que elas não foram capturadas durante a transmissão [Cao et al., 2022]. Nesse caso, embora métricas de segurança, desempenho e flexibilidade também se apliquem, um desafio da QKD em comparação a PQC é a ausência de padrões de QKD bem definidos e amplamente adotados.

2.2. Segurança de Sistemas

O eixo de Segurança de Sistemas é abrangente e inclui a integração entre componentes de hardware, software, infraestrutura de comunicação e sub-sistemas. Os principais desafios atuais são o uso de Inteligência Artificial (IA) e Ciência de Dados para auxiliar neste eixo. O uso de IA para cibersegurança é um tema explorado há tempos na automação de tarefas de monitoramento e detecção de ataques, geração de traços de auditoria, imposição de controles de segurança, análise de vulnerabilidades em software etc. Além disso, técnicas de ciência de dados são úteis para lidar com processamentos complexos envolvendo diversas fontes, volumes e velocidades de dados. Por outro lado, a aplicação de cibersegurança para sistemas de IA é uma área emergente, levando a conceitos como Aprendizado de Máquina em Contexto Adversário e Aprendizado de Máquina Seguro [Vassilev et al., 2024]. Desafios da área incluem: (i) mitigação dos efeitos do uso de IA por parte dos atacantes (por exemplo, para evadir sistemas inteligentes, roubar modelos, envenenar bases de dados públicas usadas na modelagem de IA, e praticar violações éticas, de privacidade de usuários e de direitos autorais); (ii) aplicação correta (eficiente, eficaz, ética e adequada) das técnicas de IA considerando a heterogeneidade dos dados de segurança (*streams*, dados discretos, múltiplos formatos etc.); (iii) consolidar os resultados do processamento inteligente distribuído de grandes massas de dados provenientes de redes de altíssima velocidade, com criptografia e formatos/fontes/protocolos variados para detecção e prevenção de ataques em tempo quase real.

2.3. Segurança de Conexão

Este eixo contempla as interligações lógicas e físicas entre componentes computacionais, cujos desafios de cibersegurança são em grande parte fruto da constante evolução das tecnologias de comunicação.Enquanto a implantação das redes além do 5G (*Beyond* 5G - B5G) prometem melhorias contínuas em métricas como taxas de transmissão e latência, isso também exige a tomadas de decisões de segurança autônomas, capazes de lidar com a conexão de bilhões de dispositivos e aplicações de forma transparente, resiliente e com garantia de desempenho. Este desafio arquitetural pode ser endereçado por meio de abordagens altamente flexíveis, baseadas em virtualização e softwarização de redes, como as redes de acesso via rádio abertas (Open RAN). Contudo, a abertura de interfaces e o uso de *software* de código-fonte de diversas fontes distintas aumenta a superfície de ataque a esses sistemas. Uma linha de pesquisa relevante nesse contexto tem como foco a proteção contra abuso e o acesso indevido a informações em redes B5G, usando princípios de segurança por projeto. Comumente, soluções na camada de controle dessas redes envolvem tecnologias descentralizadas e com elevado grau de transparência, buscando promover o modelo de "arquitetura de confiança zero" (*Zero-Trust Architecture* – ZTA). Logo, é importante medir a dependência das soluções por entidades confiáveis e a existência de pontos críticos de falha, e se são capazes de prover mecanismos eficientes para autenticação de entidades e preservação de privacidade dos dados. Na camada física, o gerenciamento de espectro e de recursos é fundamental para assegurar disponibilidade e permitir a detecção e isolamento eficiente de ameaças [Ramezanpour e Jagannath, 2022].

2.4. Segurança Organizacional, Segurança e Sociedade & Fatores Humanos

O eixo de Segurança Organizacional lida com a proteção de organizações contra ameaças e gestão de riscos de segurança. O eixo de Segurança e Sociedade aborda cibercrimes, privacidade e aspectos legais, éticos e políticos. Finalmente, eixo de Fatores Humanos em Segurança concentra-se na proteção de dados em contexto pessoais e corporativos. O ponto em comum entre esses eixos é o foco na interação de indivíduos com sistemas: as decisões e comportamentos por eles adotados tornam-se um ponto crítico para a cibersegurança e, portanto, compreendê-los é essencial para construir mecanismos de proteção eficientes [Pollini et al., 2022]. Por exemplo, ataques de engenharia social exploram diretamente as emoções e instintos humanos para ludibriá-los, afetando alvos específicos ou toda uma cadeia de suprimentos, como o incidente ocorrido com a biblioteca *XZ Utils* [Buchanan, 2024] em que atacantes se infiltraram como mantenedores do repositório para inserir portas dos fundos em duas versões da biblioteca. Lidar com tais desafios requer uma abordagem holística que combine educação, conscientização e tecnologia. A pesquisa na área envolve o desenvolvimento de ações e programas para melhorar o letramento digital dos usuários e sua postura de segurança. A efetividade de soluções deve ser medida pela sua capacidade de levar usuários a reconhecer e evitar cenários de ameaça, como tentativas de engenharia social, e diferenciar fatos de dados falsos. Para desenvolver esse pensamento crítico, são necessárias estratégias de comunicação eficientes, que, além de mostrar como implementar medidas de segurança, expliquem o racional envolvido, permitindo ao usuário lidar com cenários de ataque cada vez mais complexos.

2.5. Segurança de Software e Segurança de Componentes

O eixo de Segurança de Software se concentra no desenvolvimento e o uso de software de modo a garantir propriedades de segurança na aplicação-alvo, incluindo con-

ceitos de programação segura. O eixo de Segurança de Componentes trata do projeto, aquisição, teste, análise e manutenção de componentes de hardware. Em ambos, um tema desafiador é a proteção contra ataques na cadeia de suprimentos. O pano de fundo desse desafio é o fato de que sistemas computacionais modernos geralmente envolvem componentes de vários fornecedores especializados. No cenário com diferentes fornecedores, a falha em um elo da cadeia pode levar a impactos negativos expressivos. No caso de hardware, a potencial adulteração de componentes por meio de modificação de firmware permite ataques lógicos e físicos de difícil detecção. Esse risco é ilustrado pela explosão remota de dispositivos de comunicação usados pelo grupo Hezbollah, ação que envolveu a inserção de elementos estranhos ao sistema na sua fabricação [BBC News, 2024]. Questões similares aparecem na cadeia de produção de software, como ilustra o já mencionado caso da biblioteca *XZ Utils* [Buchanan, 2024]. Ainda, para evitar falhas não propositais, é necessária a incorporação de segurança no processo contínuo de desenvolvimento, integração e implantação.Para isso, é importante modelar adversários e usar adequadamente arcabouços que ajudam a retratar fidedignamente a realidade de cada aplicação. Para aferir se esse princípio está sendo respeitado, algo especialmente necessário em sistemas de controle industrial e de missão crítica, deve-se considerar o desenvolvimento de técnicas e ferramentas que facilitem verificabilidade e auditabilidade. No contexto de hardware, algumas estratégias incluem a adoção de protocolos de atestação de componentes [Cremers et al., 2023] e do conceito de *open hardware* na construção de sistemas complexos [Emanuilov, 2020]. Para os casos de computação de propósito geral, como em nuvem computacional, componentes de hardware confiáveis podem fazer parte da solução, ao permitir a construção de ambientes de execução confiável (TEE) [Muñoz et al., 2023]. No cenário de software, há a necessidade de soluções que permitam validar atualizações, melhorando a capacidade de detecção de vulnerabilidades para além do que fazem hoje ferramentas de análise estática e dinâmica de código. Métricas para avaliar o quão próximo se está da solução para esse desafio incluem: variedade e grau de adoção de soluções que melhorem a auditabilidade de hardware e software, permitindo a proteção de um amplo conjunto de sistemas; a facilidade de integração de mecanismos de proteção e verificação na cadeia de suprimentos; e a maturidade das tecnologias disponíveis, avaliado pelo número e gravidade das vulnerabilidades descobertas ao longo do tempo.

3. Considerações Finais

Este documento apresentou os grandes desafios da Cibersegurança organizado conforme os eixos de formação definidos no RF-CS da SBC. O eixos têm focos distintos, mas o desafio comum a todos é a Educação em Cibersegurança. É cada vez mais urgente atrair para a área profissionais de diferentes perfis, com formação desde ensino médio e técnico até mestrado e doutorado. A definição do RF-CS e o atual processo de sua transformação em Diretrizes Curriculares Nacionais pelo Conselho Nacional de Educação, que possibilita a abertura de cursos de graduação na área, é uma ação de médio e longo prazo para tratar a carência de profissionais em cibersegurança. Organizar competições temáticas com viés educativo e disponibilizar cursos sobre cibersegurança nos canais da SBC e CESeg podem ser medidas de curto prazo para atrair e capacitar pessoas. Ações de curto prazo, como o programa Hackers do Bem, devem ser uma prioridade e ajudam a desenvolver a área. Porém, políticas de Estado com planejamento de médio e longo prazo, especialmente para a educação em cibersegurança, devem ser um alvo de investimentos do país, incluindo incentivos para parcerias com a indústria. O custo das certificações

profissionais é bastante elevado e o Estado brasileiro poderia estabelecer parceiras sociais e de inclusão para estes casos também. É preciso formar professores em cursos de licenciatura visando a educação dos jovens nas boas práticas de segurança e privacidade no uso da tecnologia. A não priorização de cibersegurança pode causar prejuízos financeiros e à imagem, ameaçando a democracia, o direito à privacidade e as liberdades civis.

Referências

BBC News (2024). Como pagers do Hezbollah explodiram, em ataque atribuído ao Mossad. https://www.bbc.com/portuguese/articles/ce8vxg1xgx3o.

Buchanan, B. (2024). The social engineering of XZ. https://medium.com/asecuritysite-when-bob-met-alice/the-social-engineering-of-xz-f905084438fc.

Cao, Y., Zhao, Y., Wang, Q., Zhang, J., Ng, S. e Hanzo, L. (2022). The evolution of quantum key distribution networks: On the road to the Qinternet. *IEEE Communications Surveys & Tutorials*, 24(2):839–894.

Cremers, C., Dax, A. e Naska, A. (2023). Formal analysis of SPDM: Security protocol and data model version 1.2. Em *USENIX Security Symposium*, p. 6611–6628.

Emanuilov, I. (2020). Security through transparency and openness in computer design. Em *Int. Conf. on Risks and Security of Internet and Systems*, p. 105–116. Springer.

Muñoz, A., Rios, R., Román, R. e López, J. (2023). A survey on the (in) security of trusted execution environments. *Computers & Security*, 129:103180.

Pollini, A., Callari, T. C., Tedeschi, A., Ruscio, D., Save, L., Chiarugi, F. e Guerri, D. (2022). Leveraging human factors in cybersecurity: an integrated methodological approach. *Cognition, Technology & Work*, 24(2):371–390.

Ramezanpour, K. e Jagannath, J. (2022). Intelligent zero trust architecture for 5G/6G networks: Principles, challenges, and the role of machine learning in the context of O-RAN. *Computer Networks*, 217:109358.

SBC (2023). Referenciais de formação para o curso de bacharelado em cibersegurança. Relatório técnico, Sociedade Brasileira de Computação (SBC).

Vassilev, A., Oprea, A., Fordyce, A. e Anderson, H. (2024). Adversarial machine learning: A taxonomy and terminology of attacks and mitigations. Relatório técnico, National Institute of Standards and Technology.

Desafios Computacionais para uma Internet Quântica Brasileira

Antônio Abelém[1], Alberto Schaeffer-Filho[2], Gabriel Nazar[2], Jéferson Nobre[2], Juliano Wickboldt[2], Lisandro Granville[2], Luciano Gaspary[2], Weverton Cordeiro[2]

[1]Universidade Federal do Pará (UFPA)

[2]Universidade Federal do Rio Grande do Sul (UFRGS)

Resumo. *A Computação Quântica tem emergido como uma revolução nas tecnologias computacionais, prometendo resolver problemas complexos, como fatoração de grandes números e desenvolvimento de medicamentos. No entanto, surgem preocupações, como, por exemplo, a vulnerabilidade da criptografia atual a ataques quânticos. Além disso, a Computação Quântica está impulsionando a criação de redes quânticas, com impacto significativo em interconectividade e aplicações distribuídas. No Brasil, há uma demanda crescente por uma evolução da Internet Clássica nacional para sua integração com essas redes, sendo o principal desafio o desenvolvimento de um ecossistema robusto, focado em pesquisa, inovação e formação de talentos. Neste documento, desafios e oportunidades de pesquisa em Computação para o desenvolvimento da Internet Quântica Brasileira são apresentados. Tal desenvolvimento pode posicionar o país como líder regional e um importante ente global em Comunicação Quântica.*

Abstract. *Quantum Computing has emerged as a revolutionary force in computational technologies, promising to solve complex problems such as large-number factorization and drug development. However, concerns have arisen, for example, the vulnerability of current cryptographic systems to quantum attacks. In addition, Quantum Computing is driving the creation of quantum networks, which are expected to significantly impact interconnectivity and distributed applications. In Brazil, there is a growing demand for the evolution of the national Classical Internet to enable integration with these emerging networks. The main challenge lies in developing a robust ecosystem focused on research, innovation, and talent development. This document presents the key research challenges and opportunities in computing for the development of the Brazilian Quantum Internet. Such endeavor could position the country as a regional leader and a significant global player in Quantum Communication.*

1. Introdução

A Computação Quântica e a Comunicação Quântica emergem como campos de pesquisa que estão redefinindo os fundamentos da Computação e das Tecnologias de Comunicação. A Computação Quântica promete uma revolução ao permitir a resolução de problemas que são intratáveis para computadores clássicos, como a fatoração de grandes números (AHMED, 2024) ou o desenvolvimento de novos fármacos (SANTAGATI et al., 2024). Ao mesmo tempo, a Comunicação Quântica oferece a possibilidade de comunicações mais

seguras, baseadas nos princípios do entrelaçamento quântico e da criptografia quântica (e.g., protocolo BB84 (BENNETT; BRASSARD, 2014)). Nos últimos anos, grandes avanços foram feitos tanto na teoria quanto na experimentação dessas tecnologias, como evidenciado por pesquisas publicadas em eventos internacionais de renome, como a IEEE *International Conference on Quantum Computing and Engineering* (QCE) e em periódicos de alto impacto, como o *Nature* e o *Physical Review Letters*.

As diversas vantagens advindas da Computação Quântica são acompanhadas de algumas preocupações. Por exemplo, a possibilidade da utilização de computadores quânticos em ataques a esquemas de criptografia assimétrica (MAVROEIDIS et al., 2018). Apesar de haver novos esquemas criptográficos resistentes à Computação Quântica, o desenvolvimento, a adaptação e a implantação desses esquemas não é uma tarefa trivial (WANG et al., 2022). Além da resistência dos próprios algoritmos criptográficos, também é necessário assegurar as propriedades de segurança do hardware e do software utilizados. Finalmente, ainda há o desenvolvimento de esquemas criptográficos quânticos, os quais precisam também de garantias para que sua utilização seja difundida.

A emergência da Computação Quântica permite que sejam estabelecidas redes que operam com tecnologia quântica, o que traz impactos especialmente para a Internet. Comunicações quânticas já estão sendo implementadas em redes experimentais (e.g., Rede Quântica de Pequim-Shanghai (CHEN et al., 2021)). Tais redes representam um passo significativo para a criação de uma Internet Quântica, incorporando tecnologia quântica na infraestrutura da Internet. Nos últimos anos, pesquisas e experimentos têm se concentrado no desenvolvimento de protocolos para a Internet Quântica (ABELEM; TOWSLEY; VARDOYAN, 2020) (LI et al., 2024).

Os avanços na Computação e Comunicação Quântica não apenas sublinham a importância estratégica dessa área, mas também destacam a necessidade urgente do desenvolvimento de uma Internet Quântica Brasileira. Tal necessidade tem sido reconhecida por instituições científicas e governamentais. A Sociedade Brasileira de Computação (SBC) criou em 2024 o Grupo de Interesse em Computação e Comunicação Quântica (GICQ), e o Ministério da Ciência, Tecnologia e Inovação (MCTI) do Governo Federal criou internamente um Grupo de Trabalho (GT) dedicado às tecnologias quânticas. Ademais, na 5ª Conferência Nacional de Ciência e Tecnologia, realizada em Brasília em 2024, as questões de tecnologias quânticas foram debatidas em uma sessão plenária, refletindo o reconhecimento do tema como uma prioridade estratégica para o futuro do país.

Neste contexto, surge o grande desafio de desenvolver o ecossistema da Internet Quântica Brasileira de forma integrada. Esse ecossistema se faz imprescindível para posicionar o Brasil na vanguarda desta nova era tecnológica. Nesse documento, defendemos que um dos Grande Desafios a serem superados é a capacidade de desenvolver e utilizar mecanismos e técnicas computacionais com desenvolvimento nacional em conjunto com a formação de talentos e o desenvolvimento de pesquisas de ponta. O fomento a um ambiente acadêmico e inovador possibilitará o estabelecimento de parcerias internacionais e financiamento de projetos inovadores. O esforço dispendido contribuirá para o avanço global da ciência e para o fortalecimento da posição da Brasil como um centro de excelência em Comunicação Quântica.

A presenta proposta está descrita como segue. Na Seção 2, será apresentada a fundamentação em relação à Computação Quântica e a Comunicação Quântica. Na Seção 3, serão discutidos os desafios para o estabelecimento da Internet Quântica Brasileira.

2. Fundamentação Teórica

Um sistema quântico é descrito por seu estado quântico, o qual representa uma descrição completa do sistema em um momento específico. O equivalente clássico de um estado quântico seria uma sequência de valores binários. Assim como um bit individual (0, 1), um sistema quântico pode consistir em um único qubit, o qual existe como uma superposição de 2 possibilidades dentro do mesmo espaço binário ($a \mid 0\rangle + b \mid 1\rangle$) . No entanto, como os estados quânticos são regidos pelas leis da mecânica quântica, seu comportamento é significativamente diferente de um bit. Nesta seção, resumiremos alguns conceitos-chave da Computação Quântica. Em seguida, explicaremos suas implicações para as redes quânticas, em especial para a Internet Quântica.

Alguns conceitos da Computação Quântica são particularmente importante para as redes quânticas: emaranhamento, par de Bell, teletransporte, e troca de emaranhamento. O emaranhamento é o bloco de construção fundamental das redes quânticas. Qubits emaranhados têm propriedades não-locais interessantes, as quais podem ser usadas para comunicação. Pares de Bell são um tipo especial de estado quântico, no qual dois qubits exibem emaranhamento máximo. Por meio do teletransporte, a transmissão de um estado quântico desconhecido pode ser realizada. Finalmente, a troca de emaranhamento é o processo de distribuir um emaranhamento entre duas partes distantes por meio de alguns nós intermediários, o que é feito através do compartilhamento de um par de Bell.

A comunicação clássica (i.e., não-quântica) é um elemento fundamental das redes quânticas. As tarefas realizadas nesses redes também necessitam de conectividade clássica. Dentre essas tarefas, pode ser citado o gerenciamento da geração e troca de emaranhamentos. Protocolos distribuídos clássicos de sinalização e controle possibilitam tal gerenciamento. Dessa forma, uma rede quântica possui dois planos de dados, um clássico e um quântico. Além disso, as informações entre esses planos precisam ser correlacionadas (KOZLOWSKI et al., 2023).

A Internet Clássica é uma rede de redes clássicas, as quais podem ser definidas como um conjunto de nós que é capaz de trocar bits através de pacotes. De forma análoga, a Internet Quântica é uma rede de redes quânticas, as quais são definidas como um conjunto de nós que é capaz de trocar qubits e distribuir estados emaranhados entre si. A Internet Quântica será executada em conjunto com a Internet Clássica.

A emergência da Computação Quântica tem propiciado melhorias em aplicações tradicionais e o surgimento de aplicações quânticas. Os computadores quânticos podem ajudar no desenvolvimento de novos materiais e acelerar a descoberta de medicamentos e as decisões tomadas sobre transações financeiras. A Internet Quântica também oferece promessas para diversas novas aplicações (WANG et al., 2024). Este documento foca em aplicações quânticas que têm mais impacto na Internet Quântica, como configuração de comunicação segura e computação quântica distribuída.

3. Desafios, Áreas Habilitadoras e Oportunidades

A área da Computação tem muito a contribuir no desenvolvimento da Internet Quântica Brasileira. O processo de integração da Internet Quântica com a Internet Clássica é semelhante ao processo de introdução de novos paradigmas de comunicação na Internet já existente, mas com impactos mais significativos. A seguir, são discutidos desafios e oportunidades de pesquisa em um conjunto amplo de áreas da Computação.

Infraestrutura Nacional para a Internet Quântica. *Desafio:* Muitas infraestruturas nacionais e regionais têm sido desenvolvidas para a Internet Quântica. Podem ser citados os seguintes exemplos: Pequim-Shanghai (China), Jinan–Qingdao (China), DARPA (EUA), SwissQuantum (Suíca), Tóquio (Japão), entre outras. No Brasil, recentemente foi lançada a Rede Rio Quântica. No entanto, tais redes operam principalmente para experimentos relacionados com criptografia, não necessariamente usando uma pilha de protocolos compatíveis com a Internet (e.g., sem seguir padrões do *Internet Engineering Task Force* - IETF). Além disso, considerando o caso brasileiro, a rede possui abrangência apenas metropolitana. Dessa forma, o desenvolvimento de uma Internet Quântica nacional se impõe como uma desafio.

Desenvolvimento de Hardware e Software para Equipamentos para Internet Quântica. *Desafio:* Diversas necessidades de hardware e software são necessárias para o estabelecimento da Internet Quântica. Em relação à hardware, podem ser mencionados repetidores quânticos (e.g., responsáveis pela troca de emaranhamento) e roteadores quânticos (e.g., responsáveis pelo plano de controle da rede quântica). Já em relação ao software, podem ser mencionadas tanto aplicações especializadas a serem executadas nos nós quânticos finais quanto software de controle específicos para os nós que implementam a infraestrutura da rede. Apesar do hardware e software para redes quânticas ter potencialmente abrangência mundial, seria de grande importância a participação brasileira. Tal participação poderia posicionar o país como um dos líderes nessas redes, além de garantir a soberania quântica nacional. Além disso, no contexto desse desenvolvimento, espera-se o incentivo ao ecossistema de inovação no entorno da Internet Quântica Brasileira (e.g., com o surgimento de startups).

Desenvolvimento de Abordagens para o Gerenciamento e Operação da Internet Quântica *Desafio:* Diversos iniciativas de pesquisa têm sido propostas para concretizar a Internet Quântica (e.g., propondo melhorias em sua tolerância a erros). No entanto, pouco trabalhos enfrentam os desafios computacionais relacionados com abordagens para o gerenciamento e operação de rede até o presente momento. Diversas tarefas relacionadas com essas abordagens serão necessárias, como, por exemplo, descoberta da topologia e serviços da rede, gerenciamento de recursos, etc. Além de protocolos padronizados para o plano de controle de redes quânticas, também serão necessárias melhores práticas que assegurem um funcionamento adequado das Internet Quântica. Esses protocolos podem se utilizar apenas de comunicação clássica ou também necessitar de processos quânticos para sua operação.

Cenários de Transição da Internet Clássica Brasileira para sua Evolução Quântica *Desafio:* Durante a transição da Internet Clássica Brasileira para sua evolução quântica são esperados diversos desafios relacionados com tal evolução. Por exemplo, a escolha da utilização de criptografia resistente à computação quântica ou de criptografia quântica pode estar relacionada com requisitos específicos de cibersegurança das aplicações. Além disso, essa escolha também pode estar relacionada com as funcionalidades presentes nos nós participantes de uma determinada tarefa. Os cenários de transição potencialmente necessitam de uma coordenação entre um plano de controle clássico em conjunto com um quântico.

Desenvolvimento de Aplicações relacionadas à Realidade Brasileira para a Internet Quântica *Desafio:* A infraestrutura esperada para a Internet Quântica Brasileira pressupõe o suporte a diferentes tipos de aplicações quânticas distribuídas, heterogeneidade em relação a hardware e software, e respeito a padrões internacionais. No entanto, as particularidades da realidade brasileira precisam ser consideradas, especialmente no que tange ao desenvolvimento de aplicações. Dessa forma, espera-se que a comunidade acadêmica participe ativamente no projeto de aplicações com alto impacto nas questões mais significativas para a população brasileira

Referências

ABELEM, A.; TOWSLEY, D.; VARDOYAN, G. Quantum internet: The future of internetworking. In: ____. *Shortcourses' Book of the XXXVIII Brazilian Symposium on Computer Networks and Distributed Systems (SBRC 2020)*. Porto Alegre, RS, Brasil: Brazilian Computing Society SBC), 2020. cap. 2.

AHMED, N. Quantum computing algorithms for integer factorization: A comparative analysis. *Modern Dynamics: Mathematical Progressions*, v. 1, n. 1, p. 6–9, 2024.

BENNETT, C. H.; BRASSARD, G. Quantum cryptography: Public key distribution and coin tossing. *Theoretical computer science*, Elsevier, v. 560, p. 7–11, 2014.

CHEN, Y.-A. et al. An integrated space-to-ground quantum communication network over 4,600 kilometres. *Nature*, Nature Publishing Group UK London, v. 589, n. 7841, p. 214–219, 2021.

KOZLOWSKI, W. et al. *RFC 9340 Architectural principles for a quantum internet*. Wilmington, DE, USA: RFC Editor, 2023.

LI, Y. et al. A survey of quantum internet protocols from a layered perspective. *IEEE Communications Surveys & Tutorials*, IEEE, 2024.

MAVROEIDIS, V. et al. The impact of quantum computing on present cryptography. *International Journal of Advanced Computer Science and Applications*, The Science and Information Organization, v. 9, n. 3, 2018. Disponível em: ⟨http://dx.doi.org/10.14569/IJACSA.2018.090354⟩.

SANTAGATI, R. et al. Challenges and opportunities for applying quantum computers to drug design. *Bulletin of the American Physical Society*, APS, 2024.

WANG, C. et al. *RFC 9583 Application Scenarios for the Quantum Internet*. Wilmington, DE, USA: RFC Editor, 2024.

WANG, J. et al. Quantum-safe cryptography: crossroads of coding theory and cryptography. *Science China Information Sciences*, Springer, v. 65, n. 1, p. 111301, 2022.

Internet Disponível para Todos: Desafios do Acesso Ubíquo

Anelise Munaretto[1], Carlos Kamienski[2], Eduardo Cerqueira[3], Leobino Sampaio[4], Miguel Elias M. Campista[5] e Rossana M. C. Andrade[6]

[1]Universidade Tecnológica Federal do Paraná (UTFPR)
[2] Universidade Federado do ABC (UFABC)
[3] Universidade Federal do Pará (UFPA)
[4] Universidade Federal da Bahia (UFBA)
[5] Universidade Federal do Rio de Janeiro (UFRJ)
[6] Universidade Federal do Ceará (UFC)

anelise@utfpr.edu.br, carlos.kamienski@ufabc.edu.br, cerqueira@ufpa.br
leobino@ufba.br, miguel@gta.ufrj.br, rossana@ufc.br

Abstract. *This paper provides a historical overview of the Internet's evolution, highlighting its multifaceted role in computing and society. It examines the primary barriers to achieving true ubiquity and highlights the substantial technological and societal challenges that lie ahead. Realizing the Internet's potential as an equitable and inclusive infrastructure demands addressing not only issues like scalability, interoperability, and resilience but also critical concerns such as sustainability, privacy, and the ethical implications of the widespread adoption of artificial intelligence.*

Resumo. *Este artigo oferece uma visão histórica da evolução da Internet, enfatizando sua função transversal na computação e na sociedade. Ele explora as principais barreiras para alcançar a verdadeira ubiquidade e descreve os importantes desafios tecnológicos e sociais que estão por vir. Concretizar o potencial da Internet como uma infraestrutura equitativa e inclusiva exige a abordagem não apenas de questões como escalabilidade, interoperabilidade e resiliência, mas também de preocupações essenciais como sustentabilidade, privacidade e as implicações éticas da adoção generalizada da inteligência artificial.*

1. Um Breve Histórico da Internet

Criada em 1969 por acadêmicos como uma rede experimental, a Internet sempre teve a flexibilidade como característica central [Leiner et al. 2009]. Ao longo de sua história, a Internet passou por importantes evoluções, como a introdução da pilha TCP/IP após 15 anos de criação, e da Web, que somente surgiu na década de 1990. Hoje, após 55 anos de funcionamento, a rede ainda se adapta a novas tecnologias e serviços avançados.

Desde que a Internet se estabeleceu como uma rede de escala mundial no início dos anos 2000, muitas transformações impactaram o seu nível de popularidade, dentre as quais destacam-se os avanços nos meios de comunicação, na pluralidade de participantes e dispositivos e nas necessidades por serviços digitais inteligentes. Em relação ao acesso, é possível destacar algumas tecnologias de maior prevalência e impacto, como é o caso das redes sem fio (e.g., Wi-Fi e redes móveis 4G/5G), que oferecem suporte ao surgimento de

equipamentos como *laptops*, *tablets*, *smartphones*, *smartwatches* e, mais recentemente, os *smartrings* que possuem a mobilidade como destaque.

Além do acesso, a Internet vivencia mudanças na geração de conteúdos, sendo mais e mais centrada nos próprios usuários, que nos últimos anos se tornaram produtores multimídia. A distribuição da geração de conteúdo é também promovida pela maior facilidade de acesso à rede, já que permite que qualquer usuário assuma o papel de produtor, caso disponha de um dispositivo interconectado. Os dispositivos também merecem destaque na transformação da Internet, já que a redução de custos ocorreu em proporção inversa ao poder computacional. Por exemplo, os *smartphones*, de uso pessoal, possuem poder de processamento e armazenamento ordens de grandeza superiores àquele existente nos servidores dos primórdios da Internet. Tal poder distribuído culmina em dispositivos (ou coisas) conectados que monitoram o ambiente e atuam em sistemas ciberfísicos, gerando grandes massas de dados.

Por fim, o surgimento de serviços baseados em inteligência artificial é uma tendência, determinando preferências e expondo os usuários a conteúdos customizados. O impacto dessa transformação ainda não está totalmente compreendido, mas se ramifica nos vários aspectos da vida em sociedade, como nas compras, uso de serviços, trabalho, lazer e interações sociais. A digitalização social, ressaltada mais recentemente pela pandemia da COVID-19, impulsionou o processo de transformação digital, já que evidenciou a necessidade de serviços mediados por tecnologia. Os impactos dessa transformação são positivos, podendo oferecer simplicidade, agilidade e comodidade, mas também impõem consequências desafiadoras, como o isolamento social, alienação, desemprego, e invasão de privacidade, que devem ser combatidas.

2. Os Desafios da Internet Atual e as Perspectivas Futuras

Uma das realizações mais notáveis da Internet não é necessariamente o que ela é capaz de oferecer hoje, mas o fato de ter assumido as dimensões atuais, comparada aos seus princípios que também não são imutáveis:

> "*O princípio da mudança constante talvez seja o único princípio da Internet que deveria sobreviver indefinidamente.*" [Carpenter 1996]

A popularização da Internet está entre as grandes transformações possibilitadas pelo princípio da mudança constante e a sua simplicidade permite que grandes transformações se acomodem naturalmente em sua estrutura. Já o aumento do poder computacional, das taxas de transmissão e da inteligência aplicada aos serviços elevam o patamar de popularização da Internet. Nesse contexto, é essencial ter uma Internet ubíqua[1] que fornece serviços computacionais integrados e contínuos em todos os lugares, por meio da interconexão de uma vasta gama de dispositivos através de ecossistemas de redes de comunicação.

Entretanto, no Brasil, nem o acesso da maioria da população à Internet é uma realidade. Por exemplo, em 2023, apenas 84% dos domicílios brasileiros tinha acesso à Internet, como mostram as estatísticas do CGI.br. Embora esse número seja maior que os obtidos na mesma pesquisa em 2022, que foram de 80%, os desafios de alcançar a Internet plenamente ubíqua ainda persistem [Cetic.br 2023]. A presença de computadores nos

[1] O termo ubíqua surgiu na Computação com o artigo do Mark Weiser [Weiser 1999].

domicílios brasileiros revela desigualdade socioeconômica, sendo esse um dos entraves para o acesso à Internet. Computadores são encontrados em 10% dos domicílios das classes D e E, enquanto na classe A essa porcentagem é de até 97%. Portanto, além dos avanços tenológicos, o maior desafio é tornar a Internet ubíqua e igualmente acessível para todos. A seguir listamos os principais desafios e perspectivas futuras da Internet Ubíqua, tanto sob o ponto de vista técnico quanto social.

Infraestrutura e conectividade: Um dos principais desafios para a Internet móvel e ubíqua é garantir conectividade confiável e rápida em diferentes localidades. Dessa forma, é importante assegurar o acesso global e de qualidade à Internet, inclusive em áreas florestais e subdesenvolvidas, através de redes móveis de novas gerações, comunitárias ou mesmo satelitais. A preocupação com a cobertura tem sido perseguida através da concessão de espectros de frequência, como forma de contrapartida das concessionárias. No entanto, a construção da infraestrutura necessária para assegurar a conectividade tem custo elevado de implementação e operação. Então, investir em novas soluções para baratear o custo de implantação e operação da Internet ubíqua é importante para a universalização do acesso. As perspectivas em pesquisa visam alocar e gerenciar o espectro de frequências de forma a aumentar o número de dispositivos conectados e serviços disponíveis.

Escalabilidade e interoperabilidade: Com a crescente demanda de dispositivos IoT (*Internet of Things*), há um aumento massivo no tráfego de rede e nas interconexões de dispositivos. Dessa forma, é importante manter a compatibilidade de dispositivos heterogêneos tanto em relação aos diferentes fabricantes quanto aos recursos computacionais e interfaces de comunicação. O aumento no número de dispositivos leva à necessidade de escalabilidade em rede para que se possa acomodar bilhões desses dispositivos com baixa latência e alta disponibilidade, sobretudo em cenários de mobilidade. Além disso, a rede de comunicação deve assegurar a transferência das grandes quantidades de dados gerados, muitas vezes em tempo real, minimizando congestionamentos e consumo de energia. Novas formas de endereçamento, inclusive baseado em nomes, vem sendo exploradas como alternativas ao atual baseado na transição IPv4/IPv6. Como perspectiva futura está a softwarização das redes acesso de rádio através de interfaces abertas que permitem a programação do plano de controle.

Eficiência energética e sustentabilidade: A sociedade digital precisa de comunicação ubíqua, sustentável e energeticamente eficiente. Dessa forma, os procedimentos de transferência de dados devem ser energeticamente eficientes. Uma alternativa é trazer a computação mais próxima dos usuários através de estratégias de computação em borda. Além disso, os dispositivos devem ser mais resistentes e devem manter a disponibilidade mesmo que técnicas de economia de energia sejam aplicadas, como por exemplo, através da redução do ciclo de trabalho. Uma outra forma é através de estratégias inteligentes que podem aplicar comunicação semântica para identificar apenas aquilo que é importante para transferência, processamento e armazenamento.

Segurança, privacidade e neutralidade: A Internet ubíqua deve garantir a segurança e a privacidade dos dados e serviços, bem como a neutralidade da rede. Para isso, os dados gerados por dispositivos e sistemas devem ser protegidos, já que podem conter informações sensíveis de cunho pessoal, como os de saúde e localização; ou estratégicos, como os segredos indústrias. É importante o desenvolvimento de protocolos e sistemas

que tenham a segurança como um pilar nativo e assim possam detectar, mitigar e responder ameaças em tempo real e com alta acurácia. No contexto da IoT, torna-se ainda mais desafiador a criação de modelos mais escaláveis de *bootstrapping* de segurança em dispositivos remotos, como forma de garantir a identidade e legitimidade dos usuários. Uma forma de combater os ataques cibernéticos e torná-los menos atraentes economicamente é através de leis que de fato punam o mau uso da Internet ou a falta de neutralidade no tráfego de dados. O desafio da ubiquidade é proporcionar maior acesso e ao mesmo tempo garantir maior segurança e a equidade de oportunidades.

Padronização, regulamentação e governança: A geração distribuída de dados por uma quantidade massiva de dispositivos e serviços exige uma maior regulamentação e padronização da Internet de forma a garantir conectividade, interoperabilidade e qualidade das aplicações. É essencial criar e definir padrões globais para protocolos de comunicação, sintaxe e semântica de dados, interfaces abertas e seguras para que diferentes plataformas e arquiteturas interajam. Além disso, governos e órgãos reguladores devem estabelecer leis que abordem o uso ético da IA na Internet e cuja governança inclua leis relacionadas ao combate à desinformação.

Preocupações éticas e sociais: À medida que a Internet ubíqua se torna mais integrada à vida cotidiana, os desafios éticos e sociais se tornam mais proeminentes. Dessa forma, a ubiquidade pode levar à vigilância em massa, levantando a preocupações sobre privacidade, ética digital e liberdades individuais. A Internet inteligente oferece novos serviços digitais em diferentes verticais da sociedade, o que pode gerar a substituição de empregos destinados a humanos por máquinas e sistemas inteligentes. Então, a discussão sobre o papel do humano na sociedade digital e a promoção do emprego e da manutenção da qualidade de vida é essencial. A expansão da Internet ubíqua deve evitar a exclusão e o aletramento digital.

Resiliência e tolerância a falhas: A complexidade e a importância dos serviços e operações da Internet exigem que a rede continue operando mesmo em situações de falhas ou interrupções. Dessa forma, a Internet deve garantir que o ecossistema ubíquo continue a funcionar mesmo se alguns componentes falharem, particularmente em verticais críticas como saúde, indústria e direção autônoma. De maneira adicional, a redução das interações entre usuários e máquinas deve ser minimizada para que o uso da Internet se torne mais trivial e adaptativo. O desafio então está em desenvolver estratégias para recuperação de falhas na Internet causadas por diferentes fontes, inclusive falhas de equipamentos ou ataques cibernéticos, ou mau uso por parte dos usuários.

Ambientes de ensino, formação e experimentação: Um dos grande desafios do Brasil é assegurar o desenvolvimento de ambientes de teste e experimentação em larga escala para fomentar o desenvolvimento de protocolos, serviços e aplicações para a Internet de forma aberta e com alto impacto. As mudanças rápidas nas tecnologias, serviços, aplicações e protocolos inteligentes na Internet obriga a adequação ou criação de novas disciplinas ou metodologias de ensino em nível técnico, superior e pós-graduação. O uso de laboratórios virtuais pode eliminar barreiras, oferecendo uma experiência de aprendizado prática e mais inclusiva.

3. Comentários Finais

Uma Internet disponível para todos apresenta desafios em dimensões sociais, tecnológicas, científicas, econômicas, éticas e operacionais. Superar esses desafios exige avanços em tecnologias de conectividade e também em outras transversais como a inteligência artificial, a computação, a cibersegurança e a eficiência energética. A colaboração entre indústrias, governos, academia e sociedade é essencial para definição de requisitos de acesso que promovam a ubiquidade com equidade de oportunidades na Internet. Além disso, uma discussão urgente que resulte em estratégias de ensino, qualificação e conscientização de uso deve estar presente para que a Internet seja um motor do saber e não o oposto.

4. Autores

Anelise Munaretto é professora da Universidade Tecnológica Federal do Paraná (UTFPR). Foi pesquisadora visitante no INRIA-Saclay França e membro da CE-ReSD da SBC. Anelise é diretora executiva do LARC e membro do CADM da RNP. Seus principais interesses de pesquisa são Internet das Coisas e Cidades Inteligentes.

Carlos Kamienski é professor de Ciência da Computação na Universidade Federal do ABC (UFABC). Foi pesquisador visitante no IoT-Prism Lab da Universidade de Bolonha entre 2023 e 2024. Seus interesses de pesquisa atuais incluem a Internet das Coisas, agricultura inteligente, cidades inteligentes, computação em névoa, softwarização de rede e Internet do Futuro.

Eduardo Cerqueira, CNPq 1-C, é professor da Universidade Federal do Pará (UFPA), possui 4 patentes com registro internacional e mais de 280 artigos em conferências e periódicos nacionais e internacionais. Seus interesses de pesquisa atuais incluem a Internet das Coisas, aprendizado de máquina aplicado, Realidade Estendida e Internet ubíqua.

Leobino N. Sampaio é professor da Universidade Federal da Bahia (UFBA), PQ-2 do CNPq. Foi pesquisador visitante no IRL (*Internet Reserarch Lab*) da Universidade da Califórnia em Los Angeles (UCLA). Atualmente é vice-diretor técnico-científico do LARC. As suas pesquisas atuais estão voltadas para Redes de computadores e Cibersegurança.

Miguel Elias M. Campista é professor da Universidade Federal do Rio de Janeiro (UFRJ), PQ-2 do CNPq, CNE da FAPERJ e membro sênior do IEEE. Miguel é ainda editor do periódico Annals of Telecommunications e diretor técnico-científico do LARC. Ele foi por duas vezes consecutivas JCNE da FAPERJ, membro afiliado da ABC, coordenador da CE-ReSD da SBC e membro do CADM da RNP. Suas áreas de atuação são redes de computadores, ciência de dados, cibersegurança e aprendizado de máquina aplicado.

Rossana M. C. Andrade é professora da Universidade Federal do Ceará, Departamento de Computação, DT 1D, Cientista chefe da FUNCAP e membro sênior da SBC, IEEE e ACM. Foi da diretoria do LARC e, atualmente, é coordenadora da CE-ReSD da SBC. As suas pesquisas atuais estão voltadas para Internet das Coisas e Computação Ubíqua.

Referências

Carpenter, B. E. (1996). Architectural Principles of the Internet. RFC 1958. Acessado em `https://www.rfc-editor.org/info/rfc1958`.

Cetic.br (2023). Pesquisa sobre o Uso das Tecnologias de Informação e Comunicação nos Domicílios Brasileiros. TIC Domicílios. Acessado em https://www.cgi.br/media/docs/publicacoes/2/20240826111431/tic_domicilios_2023_livro_eletronico.pdf.

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., Postel, J., Roberts, L. G., and Wolff, S. (2009). A brief history of the Internet. *ACM SIGCOMM Computer Communication Review*, 39(5):22–31.

Weiser, M. (1999). The computer for the 21st century. *SIGMOBILE Mob. Comput. Commun.*, 3(3):3–11.

Mundo Desconectado e Invisível: Desafios e Oportunidades para Mitigar a Desigualdade Digital

Alberto Schaeffer-Filho[1], Antônio Abelém[2], Gabriel Nazar[1], Jéferson Nobre[1], Juliano Wickboldt[1], Lisandro Granville[1], Luciano Gaspary[1], Weverton Cordeiro[1]

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Porto Alegre – RS

[2]Instituto de Ciências Exatas e Naturais – Universidade Federal do Pará (UFPA)
Belém – PA

Resumo. *Apesar da popularização, um terço da população mundial continuou desconectada da Internet em 2023. Mesmo durante a pandemia, dois terços das crianças em idade escolar no mundo ficaram sem acesso à Internet para desenvolver suas atividades em seus lares. Embora superlativos e preocupantes, esses números infelizmente revelam muito pouco sobre a existência de um verdadeiro "mundo desconectado e invisível". Disparidades socioeconômicas e educacionais, além de questões culturais, de gênero e de raça, entre outras, também afetam uma parcela significativa da população mundial que, embora tenha acesso a tecnologias digitais e à Internet, enfrenta grandes dificuldades no dia a dia digital. Por exemplo, a falta de acessibilidade em websites, apps, etc., muitas vezes priva pessoas com necessidades particulares do direito a participar plenamente no mundo digital. Mesmo pessoas que têm um smartphone às vezes são incapazes de acessar a Internet por falta de eletricidade ou cobertura de rede celular. Esses exemplos, que afetam o acesso a serviços básicos de saúde, educação, cultura e cidadania e comprometem o desenvolvimento econômico e social, impõem grandes desafios multidisciplinares em áreas como Interface Humano-Computador, Projeto de Sistemas Eletrônicos e Computacionais, Engenharia de Software, Redes de Computadores, entre outros. Neste documento, elencamos desafios e oportunidades de pesquisa em Computação que a comunidade científica possui à frente para mitigar a exclusão e desigualdade digital, e promover a inclusão digital em regiões com necessidades específicas.*

Abstract. *Despite the widespread popularity of the Internet, one-third of the world's population remained disconnected in 2023. Even during the pandemic, two-thirds of school-age children worldwide did not have access to the Internet to carry out their activities at home. Although superlative and worrying, these numbers unfortunately reveal very little about the existence of a truly "disconnected and invisible world". Socioeconomic and educational disparities, as well as cultural, gender and racial issues, among others, also affect a significant portion of the world's population who, despite having access to digital technologies and the Internet, face great difficulties in their daily digital lives. For example, the lack of accessibility on websites, apps, etc., often deprives people with special needs of the right to fully participate in the digital world. Even people who have a smartphone are sometimes unable to access the Internet due to a lack of electricity or cellular network coverage. These examples, which affect*

access to basic health, education, culture and citizenship services and compromise economic and social development, pose major multidisciplinary challenges in areas such as Human-Computer Interface, Electronic and Computational Systems Design, Software Engineering, Computer Networks, among others. In this document, we list challenges and research opportunities in Computing that the scientific community has ahead to mitigate digital exclusion and inequality, and promote digital inclusion in regions with specific needs.

1. Introdução

De acordo com o Censo IBGE 2022 [IBGE 2022], 28,2 milhões de brasileiros não têm acesso à Internet[1], e quase um milhão de brasileiros não têm acesso à energia elétrica[2]. Esses números infelizmente refletem internamente um problema persistente no restante do mundo. Em plena era digital, cerca de 675 milhões de pessoas ainda vivem sem eletricidade [Chade 2023], 2,7 bilhões sem Internet [Andrade 2023] e dois terços das crianças em idade escolar não tem acesso à Internet em casa [UNICEF 2020].

Embora superlativos e preocupantes, esses números revelam muito pouco sobre a existência de um "mundo desconectado e invisível". Disparidades socioeconômicas e educacionais, além de questões culturais, de gênero e de raça, entre outras, também afetam uma parcela significativa da população mundial que, embora tenha acesso a tecnologias digitais e à Internet, enfrenta grandes dificuldades no dia a dia digital e acaba sendo privada de uma *conectividade significativa* [NIC.Br 2024]. Esse conceito, que refere-se à capacidade do indivíduo em alcançar benefícios pessoais e oportunidades decorrentes do acesso à Internet, abrange aspectos como qualidade do acesso, dispositivos disponíveis para uso e habilidades digitais. Por exemplo, a falta de acessibilidade em websites, apps, etc., muitas vezes priva pessoas com necessidades especiais do direito a participar plenamente no mundo digital [Bentley et al. 2024]. Mesmo pessoas que têm um smartphone às vezes são incapazes de acessar a Internet por falta de eletricidade ou cobertura de rede celular, ou porque a interface não atende às necessidades particulares dessas pessoas.

Esses exemplos, que afetam o acesso a serviços básicos de saúde, educação, cultura e cidadania e comprometem o desenvolvimento econômico e social, impõem um grande desafio multidisciplinar para desenvolver soluções computacionais – a partir de avanços em áreas como Interface Humano-Computador, Engenharia de Software, Projeto de Sistemas Eletrônicos e Computacionais, Redes de Computadores, entre outros – para promover a conectividade significativa da população. Neste documento, começamos pela definição do conceito de "mundo desconectado e invisível", passando pelas causas e consequências da existência desse mundo desconectado e invisível, e elencamos um conjunto de desafios e oportunidades de pesquisa em Computação que a comunidade científica possui à frente para mitigar a exclusão e desigualdade digital, e promover a inclusão digital efetiva para as populações necessitadas. A importância e relevância deste tema é inclusive destacada pela temática do Congresso da Sociedade Brasileira de Computação 2024 (*Deserto Digital: O Mundo Desconectado e Não Visto*) e a Palestra P3 proferida no SEMISH / CSBC 2024, também sobre o tema de mundo desconectado e invisível[3].

[1] InfoMoney: 28,2 milhões de brasileiros não têm acesso à Internet, diz IBGE - https://www.infomoney.com.br/consumo/282-milhoes-de-brasileiros-nao-tem-acesso-a-internet-diz-ibge/

[2] Jornal da USP: Falta de acesso à energia elétrica ainda é uma realidade no Brasil - https://jornal.usp.br/radio-usp/falta-de-acesso-a-energia-eletrica-ainda-e-uma-realidade-no-brasil/

[3] SEMISH / CSBC 2024: Palestra P3: Mundo Desconectado e Invisível: Desafios e Oportunidades de Pesquisa para Mitigar a Desigualdade Digital

2. Mundo Desconectado e Invisível: Conceitos e Definições

O fenômeno da exclusão digital, que leva à existência de um mundo desconectado e invisível, é referenciado e estudado pelo menos desde a década de 90, com os anos 2000 testemunhando o surgimento de uma pesquisa interdisciplinar focada nos aspectos tecnológicos, psicológicos, sociológicos, econômicos, e educacionais relacionados [Van Dijk 2017]. Em resumo, a exclusão digital descreve a lacuna existente entre pessoas que possuem níveis satisfatórios de acesso a tecnologias digitais (incluindo aqui as capacidades econômicas para adquiri-las) bem como as habilidades para usar essas tecnologias efetivamente, em contrate às pessoas que possuem acesso limitado e habilidades digitais reduzidas [Soomro et al. 2020, Lythreatis et al. 2022]. Para compreender esse fenômeno a partir da dimensão tecnológica [Wilson-Menzfeld et al. 2024], algumas perguntas importantes devem ser respondidas:

O que esse mundo engloba? Para além de classificar as pessoas entre as que possuem e não possuem acesso a tecnologias digitais e à Internet, esse mundo compreende pessoas que, embora tenham acesso a tecnologias, são excluídas de uma vida plenamente digital por fatores diversos. Por exemplo, Sin *et al.* [Sin et al. 2021] recentemente definiram o conceito de "marginalização pelo projeto digital" (*digital design marginalization*), que se refere a decisões de projeto de interfaces de software (incluindo aqui apps de smartphones) que ativamente criam barreiras para certos grupos de usuários em usufruírem de serviços online. Um exemplo ilustrativo, e bastante debatido recentemente, é o de cardápios exclusivamente digitais em restaurantes[4], que pode criar dificuldades no acesso ao serviço para alguns grupos de pessoas mais velhas ou pessoas com dificuldades visuais. Essas barreiras podem ter diversas consequências, como por exemplo perda de acesso de qualidade a serviços digitais, aumento de estigma social [Caldeira et al. 2022], menor auto-estima, entre outros [Robinson et al. 2015, Sin et al. 2021].

Por que esse mundo existe (e persiste)? De acordo com Myrdal [Myrdal 1968], as próprias desigualdades econômicas regionais podem representar obstáculos ao progresso econômico (a pobreza se torna sua própria causa). Da mesma forma, pode-se argumentar que o mundo desconectado e invisível se retroalimenta, impedindo que muitos indivíduos possam se integrar a uma sociedade digital devido ao contexto social, cultural, econômico, em um momento em que tanto serviços quanto oportunidades consistentemente migram para o mundo digital. Assim, aspectos como baixa renda familiar, lacunas educacionais, infraestrutura deficiente, limitações de acesso a dispositivos, esparsidade populacional e ausência de políticas públicas mais efetivas contribuem tanto para marginalizar digitalmente diversos grupos sociais, como também pode impedir indivíduos desses grupos possam alcançar individualmente a conectividade significativa.

Qual o impacto desse mundo para a sociedade? Há diversos estudos que mapeiam os efeitos prejudiciais da marginalização digital nas mais variadas dimensões. Em termos econômicos, um estudo do Google e da consultoria McKinsey revelou i) que o brasileiro domina apenas o básico na Internet, ii) que quanto maior a renda, melhor a qualidade do acesso e as oportunidades que a pessoa tem acesso e iii) que maior maturidade digital invariavelmente levaria a maior produtividade e menor desemprego[5]. Essa é a mesma

(Palestrante: Weverton Cordeiro) - https://csbc.sbc.org.br/2024/semish/

[4]Globo.com: Cardápio digital através de QR Code divide opiniões e gera debate até em casas legislativas do RJ e MG - https://g1.globo.com/jornal-hoje/noticia/2023/05/12/cardapio-digital-atraves-de-qr-code-divide-opinioes-e-gera-debate-ate-em-casas-legislativas-do-rj-e-mg.ghtml

[5]O Globo: Inclusão digital pode engordar PIB em US$ 70 bilhões - https://oglobo.globo.com/economia/inclusao-digital-pode-engordar-pib-em-us-70-bilhoes-23550013

conclusão de estudos independentes feitos pela PwC Brasil, e pelo Movimento Brasil Competitivo e a Fundação Getúlio Vargas. O primeiro estima em 7.7% o aumento potencial do PIB para a América Latina pelo aprimoramento de competências da força de trabalho até 2028[6]. O segundo, determina que acelerar a digitalização da economia brasileira pode gerar até 1,1 trilhão em ganhos para o PIB[7]. Além de prejudicar no acesso à informação e a serviços básicos, a falta de letramento digital da população representa uma amarra ao desenvolvimento econômico social, impedindo a geração coletiva de riqueza e de bem estar social.

3. Áreas Habilitadoras, Desafios e Oportunidades

A Computação tem muito a contribuir com soluções para mitigar, reduzir e mesmo prevenir a marginalização digital via esforço multidisciplinar, com desafios científicos e tecnológicos em áreas habilitadoras, discutidas de forma não exaustiva a seguir:

Interação Humano-Computador (IHC). *Desafios:* Sin *et al.* [Sin et al. 2021] argumentam que muitos serviços essenciais estão sendo migrados para o ambiente digital e que, como resultado, certas populações estão sendo excluídas. Esse processo de marginalização ocorre, entre outros, devido a decisões de projeto que não levam em consideração as situações socio-econômicas, culturais, físicas e mentais da diversidade da população alvo dos serviços. *Oportunidades:* O *design inclusivo* tem sido uma abordagem na indústria para o desenvolvimento de interfaces que levem em consideração a diversidade humana. A literatura de IHC é rica em estudos e pesquisas para guiar o projeto de interfaces para populações vulneráveis e menos favorecidas [Anuyah et al. 2023].

Engenharia de Software. *Desafios:* Assim como na área de IHC, uma ampla literatura tem reconhecido há tempos na disciplina de Engenharia de Software que o projeto de software deve abordar as diversas necessidades dos usuários [Paiva et al. 2021]. *Oportunidades:* Aqui, pode-se vislumbrar diversas preocupações, como i) desenvolvimento de sistemas energeticamente eficientes (para uso em comunidades com acesso dificultado à energia elétrica), ii) desenvolvimento de software que garanta a evolução dos serviços sem constantes atualizações de hardware (que acabe limitando o acesso de populações mais carentes com dispositivos mais antigos) e iii) processos de atualização de software que considerem o levantamento de requisitos e implementação cientes dos grupos populacionais alvos dos softwares.

Redes de Computadores. *Desafios:* Embora o acesso universal à Internet seja um direito humano fundamental [United Nations 2016], vários desafios nos mantêm longe de fornecer acesso facilitado, tanto para *populações isoladas tecnologicamente* como para regiões mais carentes. Várias iniciativas públicas e privadas visam mitigar esses desafios [Crowcroft et al. 2015, A4AI 2018, IETF GAIA 2024], apesar dos desafios ambientais, sociais e econômicos que implicam em altos custos de capital e custeio de infraestrutura de rede e, ao mesmo tempo, baixo retorno financeiro sobre investimento de soluções tradicionalmente usada em grandes centros urbanos. Nesse contexto, as redes comunitárias se destacam como uma solução viável para facilitar o acesso à internet em regiões afastadas dos grandes centros urbanos, ao promover modelos de compartilhamento de custos de aquisição, montagem, operação e manutenção da rede, permitindo que as comunidades

[6]PwC Brasil: O abismo digital no Brasil - *Como a desigualdade de acesso à Internet, a infraestrutura inadequada e a educação deficitária limitam nossas opções para o futuro* - https://www.pwc.com.br/pt/estudos/preocupacoes-ceos/mais-temas/2022/O_Abismo_Digital.pdf

[7]Valor Econômico: Hora de apressar o passo da digitalização - https://valor.globo.com/brasil/coluna/hora-de-apressar-o-passo-da-digitalizacao.ghtml

possam crescer de maneira sustentável e conectada. *Oportunidades:* A principal direção de pesquisa é como desenvolver soluções tecnológicas que possam permitir a conexão à Internet em locais mais distantes de centros urbanos, sem cobertura de rede celular e mesmo de energia elétrica, para um público sem condições de custear serviços via satélite (por ex. na região Amazônica, onde zerar a exclusão digital ainda é um desafio[8]). Dispositivos de baixa potência, baixo custo e longo alcance são soluções tecnológicas interessantes, ao mesmo tempo que pesquisa em transmissão de dados otimizada para links precários [Scheibe et al. 2021].

Projetos de Sistemas Eletrônicos e Computacionais. *Desafios:* A industria de semicondutores tem focado principalmente em soluções *high-end*, por ex. placas gráficas avançadas que vão além de sua funcionalidade original e permitem realizar atividades em inteligência artificial, criptoeconomia, etc. No entanto, as novas soluções em hardware baseiam-se essencialmente no aumento do custo e consumo energético para o aumento na vazão do processamento de dados [Carro and Nazar 2023]. *Oportunidades:* Com a perda de força da Lei de Moore em um momento em que também há grande preocupação com computação verde e energeticamente eficiente, vislumbra-se como oportunidade pesquisas em projetos de hardware que permitam reduzir o custo/consumo energético mantendo a vazão no processamento de dados. Soluções podem incluir, por ex., processamento de dados onde os mesmos são gerados, sem movimentos desnecessários entre processador e memória (que drenam energia), redução de precisão para atender requisitos não funcionais diversos (como uso de recursos, etc.).

Inteligência Artificial. *Desafios:* IA representa uma grande promessa de revolucionar o acesso à informação, aprendizado personalizado, criar oportunidades de trabalho, entre outras [Božić 2023]. Com a revolução em andamento da inteligência artificial e a emergência de plataformas como ChatGPT, no entanto, há uma grande preocupação sobre o surgimento de um "*AI-divide*" (exclusão digital promovida pela inteligência artificial) [Bentley et al. 2024]. O temor é que que alguns grupos possam ter maior acesso às vantagens proporcionadas pela IA, enquanto outros grupos estarão mais vulneráveis ou terão acesso a menos oportunidades. Exemplos incluem demanda por trabalhadores com maior expertise tecnológico e extinção de profissões que possam ser vistas como automatizáveis por IA[9]. *Oportunidades:* Além da necessidade de uma uma maior investigação sobre o impacto de IA em populações menos favorecidas, vislumbra-se a necessidade de desenvolvimento de educação e treinamento em IA, desenvolvimento de um framework para Ética em IA, acesso com equidade a oportunidades, recursos e ferramentas em IA, promoção de diversidade e inclusão em IA, entre outros [Farahani and Ghasemi 2024].

Cibersegurança. *Desafios:* Com a maior adoção de serviços online, incluindo serviços financeiros, cria-se um ambiente propício para o lançamento de crimes cibernéticos contra populações mais vulneráveis, as quais acabam (in)voluntariamente se excluindo do ambiente digital por receio de furtos, roubos, golpes, etc. [Holgersson et al. 2021]. *Oportunidades:* Os desafios em cibersegurança vão desde o letramento e alfabetização digital, a compreensão das necessidades específicas dos grupos mais vulneráveis, desenvolvimento de ferramentas de segurança com interfaces simplificadas e análise de riscos específicos para grupos marginalizados. Soluções podem envolver, inclusive, o uso de Inteligência

[8] O Liberal Amazônia: Mais que dobra o número de locais rurais com cobertura total de internet no Pará - https://www.oliberal.com/economia/mais-que-dobra-o-numero-de-locais-rurais-com-cobertura-total-de-internet-no-para-1.600970

[9] Para a pesquisa que fundamentou a escrita desse trecho do documento, utilizou-se a ferramenta de IA Google Gemini, meramente para ilustrar o impacto da IA já vivenciado no cotidiano científico, onde ferramentas apoiam a escrita de artigos científicos - https://gemini.google.com/app

Artificial, permitindo automatizar a análise de riscos para grupos mais vulneráveis e permitir o acesso seguro a serviços digitais.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 e Grant #88887.954253/2024-00. Este estudo foi financiado em parte pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil (CNPq) - Grants #444978/2024-0, #314506/2023-3, e #405940/2022-0. O trabalho também recebeu apoio da Fundação de Amparo a Pesquisa de São Paulo (FAPESP) Grant #20/05183-0 (Skynet), #21/00199-8 (SMARTNESS), #23/00673-7 (IDRCIC), e #23/00816-2 (Low End Networks).

Referências

A4AI (2018). A4AI Affordability Report 2018 [Online]. Available: `https://a4ai.org/affordability-report/report/2018/`.

Andrade, G. (2023). Crescimento da internet desacelera e mais de 2,7 bilhões estão sem acesso. https://gizmodo.uol.com.br/crescimento-da-internet-desacelera-e-mais-de-27-bilhoes-estao-sem-acesso/.

Anuyah, O., Badillo-Urquiola, K., and Metoyer, R. (2023). Characterizing the technology needs of vulnerable populations for participation in research and design by adopting maslow's hierarchy of needs. In *2023 CHI Conference on Human Factors in Computing Systems*, pages 1–20.

Bentley, S. V., Naughtin, C. K., McGrath, M. J., Irons, J. L., and Cooper, P. S. (2024). The digital divide in action: how experiences of digital technology shape future relationships with artificial intelligence. *AI and Ethics*, pages 1–15.

Božić, V. (2023). Artifical intelligence as the reason and the solution of digital divide. *Language Education and Technology*, 3(2).

Caldeira, C., Nurain, N., and Connelly, K. (2022). "i hope i never need one": Unpacking stigma in aging in place technology. In *2022 CHI Conference on Human Factors in Computing Systems*, pages 1–12.

Carro, L. and Nazar, G. L. (2023). Desafios para a computação energeticamente eficiente. *Sociedade Brasileira de Computação*.

Chade, J. (2023). Em plena era digital, 675 milhões de pessoas ainda vivem sem eletricidade. https://noticias.uol.com.br/colunas/jamil-chade/2023/06/06/em-plena-era-digital-675-milhoes-de-pessoas-ainda-vivem-sem-eletricidade.htm.

Crowcroft, J., Wolisz, A., and Sathiaseelan, A. (2015). Towards an Affordable Internet Access for Everyone: The Quest for Enabling Universal Service Commitment (Dagstuhl Seminar 14471). *Dagstuhl Reports*, 4(11):78–137.

Farahani, M. S. and Ghasemi, G. (2024). Artificial intelligence and inequality: challenges and opportunities. *Qeios*, 7:1–14.

Holgersson, J., Kävrestad, J., and Nohlberg, M. (2021). Cybersecurity and digital exclusion of seniors: What do they fear? In *International Symposium on Human Aspects of Information Security and Assurance*, pages 12–21. Springer.

IBGE (2022). Panorama do censo 2022. https://censo2022.ibge.gov.br/panorama/.

IETF GAIA (2024). IETF GAIA. Available: `https://datatracker.ietf.org/rg/gaia/about/`.

Lythreatis, S., Singh, S. K., and El-Kassar, A.-N. (2022). The digital divide: A review and future research agenda. *Technological Forecasting and Social Change*, 175:121359.

Myrdal, G. (1968). *Teoria econômica das regiões*. Saga. https://institutomyrdal.wordpress.com/wp-content/uploads/2015/03/teoria-econc3b4mica-das-regic3b5es-subdesenvolvidas.pdf.

NIC.Br (2024). Conectividade significativa: Propostas para medição e o retrato da população no brasil. https://cetic.br/media/docs/publicacoes/7/20240415183307/estudos_setoriais-conectividade_significativa.pdf.

Paiva, D. M. B., Freire, A. P., and de Mattos Fortes, R. P. (2021). Accessibility and software engineering processes: A systematic literature review. *Journal of Systems and Software*, 171:110819.

Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., Schulz, J., Hale, T. M., and Stern, M. J. (2015). Digital inequalities and why they matter. *Information, communication & society*, 18(5):569–582.

Scheibe, A., Reichert, W., Gaspary, L., and Cordeiro, W. (2021). Programmable low-end networks: Powering internet connectivity for the other three billion. In *2021 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, pages 187–195. IEEE.

Sin, J., L. Franz, R., Munteanu, C., and Barbosa Neves, B. (2021). Digital design marginalization: New perspectives on designing inclusive interfaces. In *2021 CHI Conference on Human Factors in Computing Systems*, pages 1–11.

Soomro, K. A., Kale, U., Curtis, R., Akcaoglu, M., and Bernstein, M. (2020). Digital divide among higher education faculty. *International Journal of Educational Technology in Higher Education*, 17:1–16.

UNICEF (2020). Dois terços das crianças em idade escolar no mundo não têm acesso à internet em casa, diz novo relatório do unicef-itu. https://www.unicef.org/brazil/comunicados-de-imprensa/dois-tercos-das-criancas-em-idade-escolar-no-mundo-nao-tem-acesso-a-internet.

United Nations (2016). The promotion, protection and enjoyment of human rights on the internet. *UN Digital Library*, A/HRC/32/L.20(32):1–4.

Van Dijk, J. (2017). Digital divide: Impact of access. *The international encyclopedia of media effects*, 1:1–11.

Wilson-Menzfeld, G., Erfani, G., Young-Murphy, L., Charlton, W., De Luca, H., Brittain, K., and Steven, A. (2024). Identifying and understanding digital exclusion: a mixed-methods study. *Behaviour & Information Technology*, pages 1–18.

Desafios Computacionais para Resiliência a Desastres Naturais

Alberto Schaeffer-Filho, Gabriel Nazar, Jéferson Nobre, Juliano Wickboldt, Lisandro Granville, Luciano Gaspary, Weverton Cordeiro

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

{alberto,glnazar,jcnobre,jwickboldt,granville,paschoal,wlccordeiro}
@inf.ufrgs.br

Resumo. *Com a frequente ocorrência de desastres naturais, as infraestruturas da sociedade são postas à prova. Do ponto de vista de sistemas computacionais e de comunicação, a operação dos mesmos é desafiada não só pela própria natureza dos desastres, como inundações e tempestades, mas também pela carga crítica que se impõe aos sistemas durante ações de atendimento às emergências. Sistemas de monitoramento e de alerta devem agir com celeridade para notificar a população e as autoridades. A infraestrutura de comunicação deve manter-se disponível para uso por populações em risco e por equipes de resgate. Neste documento, apresentamos o grande desafio de desenvolver sistemas de computação e de comunicação que incorporaram a resiliência a desastres de forma sistemática, capazes de atuar com eficácia na minimização de prejuízos humanos, sociais e econômicos causados por tais fenômenos.*

Abstract. *With the frequent occurrence of natural disasters, society's infrastructures are put to the test. From the perspective of computer and communication systems, their operation is challenged not only by the nature of disasters themselves, such as floods and storms, but also by the critical load imposed on systems during emergency response actions. Monitoring and warning systems must act quickly to notify the population and authorities. Communication infrastructure must remain available for use by populations at risk and by rescue teams. In this document, we present the major challenge of developing computing and communication systems that systematically incorporate disaster resilience, enabling effective action to minimize the human, social, and economic losses caused by such events.*

1. Introdução

Desastres naturais, incluindo eventos climáticos extremos como tempestades intensas, ciclones e inundações, têm tido um impacto crescente nas infraestruturas urbanas e rurais, desafiando a resiliência e a sustentabilidade de cidades inteiras [Marasco et al. 2022]. A intensidade e a frequência desses eventos, exacerbadas pelas mudanças climáticas, podem sobrecarregar os sistemas de drenagem, danificar redes de transporte, além de comprometer a integridade dos canais de comunicação e de infraestruturas de computação. Como um exemplo muito representativo, recentemente a enchente histórica no Rio Grande do Sul afetou a infraestrutura da região metropolitana de Porto Alegre e de diversas outras

partes do estado. Nos primeiros cinco dias de chuvas já havia mais de 400 mil pontos sem energia elétrica, pelo menos 186 municípios sem sinal de internet e telefone e mais de 1 milhão de residências sem abastecimento de água [BBC 2024].

Sistemas computacionais e de comunicação para monitoramento e de alerta rápido, como os utilizados para avisar populações locais sobre a ocorrência de *tsunamis*, furacões ou enchentes repentinas, devem operar com extrema rapidez e precisão. Falhas nesse tipo de sistema ou mesmo o atraso em minutos, ou em certos casos até mesmo segundos, pode significar a diferença entre vidas salvas ou perdidas. Esses sistemas precisam ser robustos o suficiente para funcionar mesmo quando parte da infraestrutura está comprometida pelo próprio desastre, o que implica na necessidade de soluções distribuídas, redundantes e com capacidades de autossuficiência energética e de capacidade processamento de dados.

Além dos sistemas específicos para monitoramento e alerta de desastres, a própria infraestrutura básica de comunicação, tanto para as populações em risco quanto para as equipes de resgate, deve manter-se disponível a todo momento. No entanto, durante um desastre, essa comunicação é frequentemente interrompida, seja pela destruição das torres de celular, pela falha no fornecimento de energia ou pela ruptura de cabos de transmissão. Ademais, durante um desastre, o volume de dados gerados e processados aumenta significativamente. Sensores ambientais, dispositivos de monitoramento e sistemas de alerta rápido são acionados simultaneamente, sobrecarregando redes e servidores. Adicionalmente, a conectividade da população com esses sistemas, seja para obter informações vitais ou para solicitar ajuda, cria um aumento abrupto na demanda de comunicação, que muitas vezes não foi dimensionada para lidar com picos tão elevados.

Neste contexto, surge o grande desafio de desenvolver sistemas de computação e comunicação que incorporem a resiliência a desastres de forma sistemática. Resiliência pode ser entendida como a capacidade de lidar com, e superar, situações adversas. Estratégias de resiliência podem tipicamente envolver etapas ativas (antecipação e preparação) e reativas (detecção e mitigação) e empregam princípios como redundância, diversidade, conectividade e associação [Sterbenz et al. 2010]. Nesse documento, defendemos que um dos Grande Desafios a serem superados é a capacidade de desenvolver e utilizar mecanismos e técnicas computacionais para promover resiliência de uma forma sistemática, especialmente em situações de desastres naturais. Soluções emergenciais, como redes móveis temporárias, satélites e drones equipados com estações rádio base, podem ser exploradas para garantir a continuidade dos serviços. O uso de tecnologias emergentes, como inteligência artificial e aprendizado de máquina, podem também potencialmente otimizar a alocação de recursos e a priorização de tarefas durante as fases críticas de um desastre, aumentando a eficácia das ações de resposta. Por fim, ressalta-se que esse desafio está alinhado ao manifesto "Tecnologias Digitais para o Meio Ambiente" aprovado na assembleia da SBC em 2022 [Sociedade Brasileira de Computação 2022], o que reforça uma demanda já expressa recentemente pela comunidade.

2. Resiliência: Um Modelo Conceitual de Referência

Resiliência é uma noção aplicável a diversas situações, e denota a capacidade de "se recuperar" em ou após situações adversas. Qualquer estratégia para resiliência terá um conjunto familiar de etapas que são basicamente ativas (antecipação e preparação) e re-

ativas (detecção e mitigação). O modelo conceitual [Dobson et al. 2019] descrito nesta seção fornece orientação sobre o projeto e a operação de um sistema resiliente, aplicando princípios como redundância, diversidade, conectividade e associação. Ele captura os principais componentes do processo de resiliência, dentro de um *loop* de controle *online* e de um *loop offline* denotado por D^2R^2+DR, conforme mostrado na Figura 1.

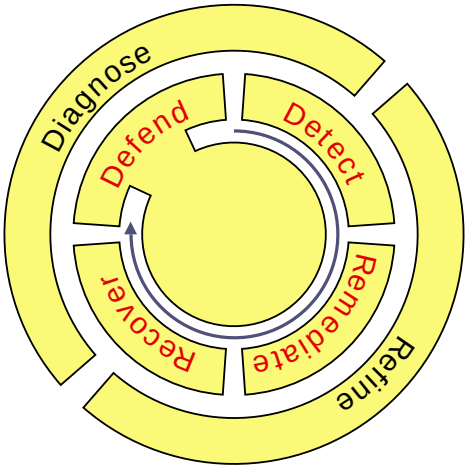

Figura 1. Modelo conceitual de resiliência D^2R^2+DR [Dobson et al. 2019]

A primeira parte interna da figura, representada por D^2R^2, especifica um *loop* de controle em tempo real: *defender*, *detectar*, *remediar*, *recuperar*. O centro do círculo denota defesas estruturais, como caminhos redundantes geodiversos. A primeira parte do *loop* de controle consiste em defesas operacionais ativas, por exemplo, mecanismos de proteção previamente instalados. Caso uma situação adversa ocorra e passe pelas defesas, a detecção precisa identificar e sinalizar a ameaça. Os mecanismos de remediação são, então, aplicados para garantir o melhor serviço possível durante um evento adverso em andamento (como uma sobrecarga operacional, uma falha ou um ataque) ou após um evento que destruiu partes da infraestrutura (como um desastre em larga escala). Por exemplo, o tráfego pode ser redirecionado em torno de áreas de infraestrutura com falha. Finalmente, a recuperação retorna a infraestrutura (no exemplo, de rede) às operações normais assim que o evento adverso cessar e os elementos de infraestrutura com falha forem reparados. O *loop* externo DR, *diagnose* e *refinement*, representa um controle não em tempo real por meio do qual os aspectos de infraestrutura relacionados à resiliência, bem como as estratégias operacionais, são analisados e aprimorados, refletindo as experiências e os desenvolvimentos de longo prazo dentro do ambiente.

3. Desastres e Adversidades: o Estudo de Caso da Grande Enchente no RS

Nos meses de abril e maio de 2024, fortes chuvas atingiram o estado do Rio Grande do Sul. A cheia dos cinco rios do interior do estado que deságuam no lago Guaíba[1] que banha a capital Porto Alegre levou a uma elevação histórica do nível das águas (Figura 2), que superou o nível observado na, até então, maior enchente de 1941. Os impactos foram sem precedentes: mais da metade dos bairros de Porto Alegre foram inundados; 70% da população da cidade ficou sem abastecimento de água devido à inundação de estações de

[1]Popularmente conhecido como rio Guaíba, tecnicamente trata-se de um lago e não de um rio.

(a) Antes **(b) Depois**

Figura 2. Imagens de satélite mostrando a elevação do nível das águas em maio de 2024 [MetSul 2024]

tratamento de água; 19 das 23 estações de bombeamento de água que deveriam eliminar a água da enchente deixaram de operar; muitos bairros ficaram sem luz; o aeroporto Salgado Filho interrompeu suas operações, e as principais rodovias de acesso à cidade foram alagadas [Ling and Inglês 2024]. A infraestrutura de comunicação de dados foi severamente afetada, uma vez que antenas 5G de alcance limitado ficaram inoperantes (sendo necessário, quando possível, o uso de tecnologias "legadas" de 3G ou inferiores). O maior *datacenter* de Porto Alegre, que está localizado em uma área que foi completamente alagada, teve sua operação muito dificultada. Além disso, o Centro de Tecnologia da Informação e Comunicação do Estado do Rio Grande do Sul (PROCERGS) e a Companhia de Processamento de Dados de Porto Alegre (PROCEMPA) tiveram sua operação paralisada ou em nível de grande contingência. Esses múltiplos eventos "em cascata" levaram ao colapso do funcionamento da cidade e da região por mais de trinta dias, acarretando a morte de centenas de pessoas e a prejuízos incalculáveis.

Posicionando lado-a-lado o estudo de caso recém referido e o modelo conceitual de resiliência D^2R^2+DR introduzido na Seção 2, podemos realizar, de maneira simplificada, o seguinte *enquadramento*. A cidade de Porto Alegre conta com um sistema contra inundação construído na década de 1970. O sistema de proteção consiste de um conjunto de mecanismos de defesa, incluindo diques, comportas e casas de bombeamento, além de um muro que atinge 6m acima do nível do mar. No entanto, os mecanismos que deveriam realizar a *defesa* quanto a ameaças previamente antecipadas falharam: a água do Guaíba vazou por comportas que se romperam e inundou as casas de bombas que deveriam eliminar a água. Mecanismos de *detecção* como sensores que medem o nível de elevação do Guaíba e sistemas de monitoramento igualmente falharam ou produziram resultados inconsistentes. As etapas de *mitigação* envolveram a evacuação de bairros inteiros e tentativas de contenção da água que transbordava através do uso de sacos de areia. Ainda, nesta etapa, podemos incluir o desligamento preventivo de um dos datacenters da PROCEMPA [Procempa 2024] e o uso de infraestruturas alternativas de sensoriamento do nível da água [CORREIO DO POVO 2024]. Mais de três meses após o incidente, ações de *recuperação* ainda persistem, ao passo que diversos bairros ainda passam por reconstrução e retomada. A médio prazo, ações de *diagnóstico* e *refinamento* precisarão ser tomadas, incluindo reforços nas barreiras de proteção (*e.g.*, comportas, di-

ques e estações de bombeamento) e o aprimoramento dos sistemas de monitoramento, alerta e protocolos de resposta a incidentes.

4. Desafios, Áreas Habilitadoras e Oportunidades

A área de Computação tem muito a contribuir na prevenção e no gerenciamento de desastres como o relatado anteriormente, seja de forma direta ou indireta. A seguir, são discutidos desafios e oportunidades de pesquisa em um conjunto amplo de áreas da Computação.

Tecnologias de comunicação. *Desafio:* Devido a seu baixo alcance e posicionamento, parte da infraestrutura de 5G foi destruída durante as fortes chuvas e enchente. Além disso, em certas regiões as operadoras de telefonia tiveram sua operação completamente comprometida. *Oportunidade:* Tecnologias legadas como 3G assumiram o papel, e empresas de telefonia estabeleceram acordos para permitir que clientes de empresas concorrentes compartilhassem a infraestrutura de comunicação. Em certas localidades que ficaram sem cobertura, Mobile Ad Hoc Networks (MANETs) poderiam ter sido utilizadas se as soluções estivessem implementadas nos celulares dos moradores [Safari et al. 2023]. Contudo, o cenário de desastres naturais demanda MANETs com características até então não contempladas, requerendo avanços de pesquisa substanciais. Aspectos abertos a investigação incluem: *trade-off* entre comunicação em padrões mínimos e eficiência energética; interfaceamento oportunístico com a Internet (por subconjunto de equipamentos 3G operacionais ou via poucos pontos-chave em que haja tal comunicação disponível); estratégias inovadoras de monitoramento e teste de infraestruturas que, em tese, só serão ativadas em meio a crises; aplicações projetadas para comunicações de emergência e de uso simplificado (requisito particularmente importante em países como os do Hemisfério Sul em que a população é menos "educada" para episódios de catástrofe).

Infraestruturas de TIC. *Desafio:* Com o alagamento de bairros inteiros, a estrutura física de *datacenters* e de outras infraestruturas críticas de computação foram comprometidas, causando interrupção de serviços. *Oportunidade:* Aspectos para pesquisa nos próximos dez anos incluem, por exemplo, modelagem dessas infraestruturas e verificação formal de requisitos de resiliência em situação de catástrofe; e projeto de estratégias "agressivas" de migração e provisionamento sob demanda de infraestruturas de rede e computação, com preservação de estado e usando, como ponto de partida, funcionalidades básicas oferecidas por tecnologias de virtualização.

Sensoriamento e ciência de dados. *Desafio:* Coleta em tempo real de dados de monitoramento da infraestrutura física de proteção (níveis das águas, integridade das comportas, operação dos diques) e o pré-processamento dessas informações. *Oportunidade:* Redes de sensores (ambientais, hidrológicos, etc) podem ter um papel crucial no aprimoramento dos sistemas de monitoramento e emissão de alertas, utilizando técnicas para acompanhamento de tendências e mudanças em variáveis críticas para antecipar possíveis desastres [Ray et al. 2017]. Apesar de décadas de pesquisa na área, desastres naturais elevam os requisitos observados até agora em ambientes como os de *smart cities*, demandando pesquisa básica e aplicada em: projeto de dispositivos robustos;

métodos computacionais inovadores capazes de "compensar" coletas incompletas e falhas; mecanismos para processamento de grandes volumes de dados em tempo "quase real"; modelos preditivos acurados de fenômenos bastante "caóticos"; projeto de ferramentas e mecanismos de visualização para apoiar a tomada rápida de decisão; e integração com atuadores visando à automação de determinadas ações, eliminando humanos "desesperados" e despreparados de parte do processo de remedição.

Planejamento, robótica e drones. *Desafio:* A dificuldade de acesso a áreas alagadas e a escassez de embarcações da defesa civil dificultaram as ações de resgate. *Oportunidade:* O uso de drones em ações de resgate oferece diversas oportunidades de pesquisa, como o desenvolvimento de algoritmos para mapeamento em tempo real de áreas afetadas combinado com algoritmos de processamento de imagens para localizar sobreviventes [Shaheen et al. 2023]. A pesquisa pode incluir o planejamento de rotas de entrega eficientes para suprimentos em locais de difícil acesso e o emprego de drones para estabelecer redes de comunicação temporárias em áreas impactadas [Flores 2024]. Além disso, a integração de sensores para monitorar condições ambientais e o avanço em técnicas de inteligência artificial para analisar dados e tomar decisões rápidas são áreas promissoras.

5. Métricas de Avaliação

Para avaliar o progresso de soluções computacionais voltadas à resiliência frente a desastres, é essencial definir métricas claras, contextualizadas e multidimensionais:

- Do ponto de vista *técnico-operacional*, o tempo médio de resposta a emergências, a cobertura de sensores e redes de comunicação, e a precisão de modelos preditivos são fundamentais. Também devem ser observados aspectos como dependabilidade, confiabilidade e tolerância a falhas [Macêdo et al. 2023], que são temas correlatos tratados em sistemas críticos.
- Na dimensão *social*, métricas como número de pessoas alcançadas por sistemas de alerta, tempo médio de evacuação em áreas de risco e o grau de inclusão digital da população usuária são cruciais para medir o impacto humano.
- A *sustentabilidade* das soluções pode ser avaliada por indicadores como o custo por área atendida, a capacidade de operar em ambientes com infraestrutura limitada e o tempo necessário para implantação.

Por fim, indicadores qualitativos, como a confiança da população nas tecnologias, o engajamento em treinamentos e a coleta de feedback por plataformas participativas, complementam a avaliação. Essas métricas, quando combinadas, permitem um acompanhamento estruturado do impacto e da efetividade das soluções.

6. Conclusões

Diante dos desafios enfrentados pelo Brasil frente a desastres naturais cada vez mais frequentes e severos, as diversas áreas da Computação oferecem oportunidades únicas para tornar nossa sociedade mais resiliente. Tecnologias de comunicação tolerantes a falhas, redes ad hoc emergenciais, sensores ambientais robustos, infraestruturas computacionais

adaptativas e o uso de inteligência artificial para previsão e tomada de decisão são apenas alguns exemplos de ferramentas que podem ser desenvolvidas ou adaptadas para o contexto brasileiro.

No entanto, é fundamental que essas tecnologias sejam pensadas de forma contextualizada, levando em conta as particularidades socioeconômicas, geográficas e infraestruturais do país. Isso inclui desde o desenvolvimento de soluções acessíveis e usáveis por populações vulneráveis até a adaptação de tecnologias internacionais – como drones de resgate, modelos de previsão com aprendizado de máquina, ou redes de sensores para cidades inteligentes – à nossa realidade. Investir em pesquisa aplicada, formação interdisciplinar e parcerias com órgãos públicos e defesa civil é essencial para transformar essas oportunidades em soluções concretas e eficazes para o Brasil.

Referências

BBC (2024). A cronologia da tragédia no rio grande do sul. BBC News Brasil: https://www.bbc.com/portuguese/articles/cd1qwpg3z77o. Acessado em setembro de 2024.

CORREIO DO POVO (2024). Sistema de monitoramento dá suporte nas enchentes de porto alegre. https://www.correiodopovo.com.br/not%C3%ADcias/chuvasnors/sistema-de-monitoramento-d%C3%A1-suporte-nas-enchentes-de-porto-alegre-1.1490745.

Dobson, S., Hutchison, D., Mauthe, A., Schaeffer-Filho, A., Smith, P., and Sterbenz, J. P. G. (2019). Self-organization and resilience for networked systems: Design principles and open research issues. *Proceedings of the IEEE*, 107(4):819–834.

Flores, H. (2024). Opportunistic multi-drone networks: Filling the spatiotemporal holes of collaborative and distributed applications. *IEEE Internet of Things Magazine*, 7(2):94–100.

Ling, A. and Inglês, R. (2024). Como porto alegre ficou debaixo d'água. Folha de São Paulo: https://www1.folha.uol.com.br/cotidiano/2024/05/como-porto-alegre-ficou-debaixo-dagua.shtml. Acessado em agosto de 2024.

Macêdo, R. J., de Sá, A. S., Freitas, A. E. S., Veríssimo, P. E., Gorender, S., and de Sá, M. O. S. (2023). *42a Jornada de Atualização em Informática*, chapter Confiabilidade e Segurança nos Sistemas Distribuídos Físico-Digitais. SBC. https://doi.org/10.5753/sbc.12853.0.2.

Marasco, S., Kammouh, O., and Cimellaro, G. P. (2022). Disaster resilience quantification of communities: A risk-based approach. *International Journal of Disaster Risk Reduction*, 70:102778.

MetSul (2024). Imagens de satélite mostram enchente arrasadora na grande porto alegre. https://metsul.com/imagens-de-satelite-mostram-enchente-arrasadora-na-grande-porto-alegre/.

Procempa (2024). Procempa mantém operações apesar de alagamento. https://prefeitura.poa.br/procempa/noticias/procempa-mantem-operacoes-apesar-de-alagamento.

Ray, P. P., Mukherjee, M., and Shu, L. (2017). Internet of things for disaster management: State-of-the-art and prospects. *IEEE Access*, 5:18818–18835.

Safari, F., Savić, I., Kunze, H., Ernst, J., and Gillis, D. (2023). A review of ai-based manet routing protocols. In *2023 19th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 43–50.

Shaheen, M. M. Z., Amer, H. H., and Ali, N. A. (2023). Robust air-to-air channel model for swarms of drones in search and rescue missions. *IEEE Access*, 11:68890–68896.

Sociedade Brasileira de Computação (2022). Tecnologias digitais para o meio ambiente: Manifesto SBC. coordenação de Marcelo Rita Pias e Raimundo José de Araújo Macêdo. 10.5753/sbc.rt.2022.07.01.

Sterbenz, J. P., Hutchison, D., Çetinkaya, E. K., Jabbar, A., Rohrer, J. P., Schöller, M., and Smith, P. (2010). Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines. *Computer Networks*, 54(8):1245 – 1265. Resilient and Survivable networks.

Kids Online: como contribuir para a proteção de crianças e adolescentes em ambientes de mídias sociais?

Humberto T. Marques-Neto[1], Jussara M. Almeida[2], Fabrício Benevenuto[2]

[1]Departamento de Ciência da Computação
Pontifícia Universidade Católica de Minas Gerais (PUC MINAS)
30.535-901 – Belo Horizonte – MG – Brazil

[2]Departamento de Ciência da Computação
Universidade Federal de Minas Gerais (UFMG)
31.270-010 – Belo Horizonte – MG – Brazil

humberto@pucmians.br, {jussara,fabricio}@dcc.ufmg.br

Abstract. *Recent studies and reports point to an increase in risk situations for children and adolescents in social media environments. These risk situations occur despite the existence of several access control tools. This proposal argues that a significant challenge for the Computer Science community is the development of more effective solutions to make the experience of children and adolescents on social media platforms safer and more promising..*

Resumo. *Estudos e relatos recentes apontam para um aumento de situações de risco para crianças e adolescentes em ambientes de mídias sociais. Isto ocorre a despeito da existência de várias ferramentas de controle de acesso. Esta proposta argumenta que um grande desafio para comunidade de Ciência da Computação é o desenvolvimento de soluções mais eficazes em tornar a experiência de crianças e adolescentes em plataformas de mídias sociais mais segura e promissora.*

1. Contexto

As plataformas de mídias sociais oferecem ambientes online importantes para realização de diversas atividades, uma vez que criam diferentes experiências para interação social e para disseminação e consumo de conteúdo digital. Essas plataformas formam um ecossistema muito complexo, com muitos elementos que contribuem para a dinâmica do ambiente. Neste contexto, destaca-se a atuação de influenciadores digitais, que, por meio da constante produção e postagem de conteúdo online, atraem uma grande quantidade de seguidores. A interação entre seguidores e influenciadores, por meio de curtidas e

comentários associados às postagens, pode criar um sentimento de pertencimento a um grupo ou a uma comunidade virtual. Por outro lado, influenciadores digitais buscam a monetização a partir da postagem de conteúdos que geram maior engajamento de seus seguidores, bem como a partir do patrocínio de marcas, que pagam os influenciadores para promover seus produtos ou serviços em seus canais online.

De acordo com o IBGE, em 2023 o uso de plataformas de mídias sociais esteve presente no cotidiano de mais de 80% dos brasileiros[1], inclusive de crianças e adolescentes. De fato, a grande presença de usuários infanto-juvenis nessas plataformas é inegável[2] Por exemplo, uma pesquisa recente[3] aponta para um aumento da presença de crianças entre 8 e 12 anos em plataformas como Instagram, Snapchat e Facebook, mesmo embora estas plataformas exigem que usuários cadastrados tenham no mínimo 13 anos. Embora entretenimento e interação com os pares sejam razões primárias do interesse deste público pelas mídias sociais, ressalta-se que o papel destas plataformas na vida de crianças e adolescentes não se restringe somente a isto. Por exemplo, é cada vez maior o uso de conteúdo online como apoio ao ensino tradicional adotado em salas de aula em escolas no Brasil e no mundo [Neumann et al. 2022]. Segundo a literatura, vários canais do YouTube com conteúdo educacional, como o Khan Academy, apresentam resultados educacionais muito positivos [Drew 2018, Nagumo et al. 2020].

Por outro lado, o uso de plataformas de mídias sociais por crianças e adolescentes também gera muitas preocupações. Diversos pediatras e neuropediatras alertam para os potenciais riscos associados ao uso destas plataformas, que incluem desde depressão e outros problemas psicológicos, como doenças associadas à falta de sono e o desenvolvimento de vícios, até cyberbullying e pedofilia [Bozzola et al. 2022]. Logo, as interações online deste público precisam ser monitoradas para garantir que os espaços utilizados sejam ambientes seguros e propícios para o desenvolvimento infanto-juvenil. Nesta direção, a forma como crianças e adolescentes são envolvidas no processo de consumo de informação em ambientes online vem sendo avaliada, dentro do

[1] https://www.abranet.org.br/Noticias/IBGE%3A-164%2C5-milhoes-de-brasileiros-acessaram-a-Internet-em-2023-5078.html

[2] https://cetic.br/pt/pesquisa/kids-online/

[3] https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-tweens-and-teens-2021

possível, por entidades internacionais como o UNICEF[4], regida por leis locais de cada país (LGPD e GDPR) e monitorada por ONGs, como o Instituto Alana[5].

Mais ainda, várias ferramentas de controle parental têm sido desenvolvidas (e.g., Assis et al. (2024), Kumar et al. (2021)), permitindo aos pais gerenciar as atividades de seus filhos em diferentes plataformas e ambientes online. Este tópico também é pauta de preocupação de grandes empresas fornecedoras de tecnologia. Como exemplos, a Meta anunciou mudanças no uso do Instagram, em resposta à pressão social sobre a proteção de crianças e adolescentes[6], a Apple passou a permitir configurar um iPhone e um iPad para uso exclusivo de crianças, enquanto o YouTube continua oferecendo uma versão alternativa voltada para o público infantil (i.e., YouTube Kids).

Entretanto, a despeito de todos os esforços já feitos, estudos e relatos apontam para um aumento de situações de risco para crianças em ambientes online, inclusive crianças menores de 13 anos [Murthy 2023]. *Por quê*?

Falta de adesão por parte de pais ou ainda uso indevido pelas próprias crianças que, conseguem contornar (e até mesmo burlar) os limites impostos, são possíveis razões para isto. Sejam quais forem as razões, está claro que a inserção cotidiana da população infanto-juvenil, os chamados nativos digitais, nos ambientes online, em especial em aplicações de mídias sociais, é um caminho sem volta que requer atenção e cuidado. Portanto, é essencial que a comunidade de Ciência da Computação ofereça alternativas mais eficazes para protegê-los nessas plataformas.

Os modelos de controle e segurança adotados pelas várias ferramentas disponíveis atualmente são fundamentalmente baseadas na restrição de funcionalidades e/ou proibição de acesso, assim como na denúncia de comportamento indevido para os pais. Como argumentado em Akter et al. (2022), este tipo de modelo de segurança pode gerar conflito devido ao desequilíbrio de poder, principalmente entre adolescentes, podendo não ser suficiente para proteger alguns jovens em formação e outros ainda vulneráveis a comportamentos maliciosos. É preciso buscar modelos alternativos e adaptativos que possam ser eficazes para uma maior parcela do público infanto-juvenil.

---

[4] https://www.unicef.org/parenting/child-care/keep-your-child-safe-online

[5] http://www.alana.org.br

[6] https://www.nytimes.com/2024/09/17/technology/instagram-teens-safety-privacy-changes.html

2. Desafio para Ciência da Computação

Neste contexto, um desafio para a Ciência da Computação é o **projeto de algoritmos, modelos, metodologias e soluções eficazes em tornar a experiência de crianças e adolescentes em mídias sociais mais segura e promissora**. Nós argumentamos que soluções eficazes devem ser pautadas nos objetivos de educar e conscientizar (pais, cuidadores, crianças e adolescentes) sobre o uso das plataformas de mídia social. Mais do que proibir, as soluções deveriam ajudar crianças e adolescentes a compreenderem os riscos e como se manterem seguras enquanto estão buscando informação ou divertimento em ambientes online. Este é um grande desafio por ser complexo e naturalmente multidisciplinar, uma vez que envolve a interação e a comunicação com o público infanto-juvenil sob diversas perspectivas.

Nesta direção, nós vislumbramos duas grandes frentes complementares a serem exploradas: (1) *desenvolvimento de soluções de mentoria e acompanhamento da experiência das crianças e adolescentes online* e (2) *monitoração dos espaços online para detecção de padrões anômalos e ambientes inadequados ou inseguros*. Como eixo comum, nós argumentamos pela necessidade de **soluções descentralizadas e independentes de plataforma.** Em outras palavras, é essencial promover a adoção e engajamento de crianças, adolescentes, pais e educadores, a consistência multiplataforma e multidispositivo, assim como o alinhamento com leis oficiais e diretrizes de instituições globais (como da Organização Mundial de Saúde). Para atingir esses objetivos, essas frentes devem poder executar independentemente de políticas específicas adotadas pelas plataformas. Logo, soluções propostas devem ser altamente configuráveis e capazes de se adaptar aos diferentes comportamentos dos envolvidos.

Quanto à *primeira frente*, sabe-se que os ambientes online de compartilhamento e disseminação de conteúdo para crianças e adolescentes, assim como o ambiente computacional que viabiliza o seu funcionamento, possuem muitos dados, não necessariamente estruturados, que podem subsidiar a construção de modelos de inteligência artificial (IA) que poderiam ser utilizados na mentoria e acompanhamento desses jovens usuários. Assim, podemos destacar algumas questões que podem orientar o desenvolvimento de trabalhos da área de Ciência da Computação:

- como coletar dados da dinâmica da interação de crianças e adolescentes para criar modelos de IA sem depender das plataformas e suas respectivas APIs?
- como as ferramentas baseadas em IA (e.g. LLMs e tecnologias estado-da-arte) poderiam contribuir com a produção e conscientização de acesso às

categorias de conteúdos digitais voltados ou consumidos por crianças e adolescentes?

- como projetar e construir soluções baseadas em IA capazes de comunicar adequada e eficazmente com crianças e adolescentes a fim de "conquistar" a sua confiança para que utilizem ferramentas que contribuam com a sua proteção, e, por que não dizer, sua formação, em ambientes online?
- como garantir a integridade, a transparência e explicabilidade de funcionamento dos sistemas baseados em IA utilizados na mentoria e acompanhamento de crianças e adolescentes em ambientes online?

Quanto à *segunda frente*, observa-se que apesar de haver legislação que direciona o uso de plataformas de mídias digitais e ambientes online, tais espaços não são imunes à atuação de usuários maliciosos que podem comprometer a segurança e a saúde de crianças e adolescentes. Em especial, a exposição a conteúdo inadequado e tóxico é uma preocupação não somente para o público em geral (e.g., discurso de ódio, *fake news*) mas também para o público infanto-juvenil. Neste último caso, a exposição a propagandas (proibido segundo algumas legislações), conteúdos que incentivem práticas nocivas ou perigosas (e.g., auto-mutilação, suicídio), com conotação sexual ou com violência excessiva são preocupações particulares para esse público. Assim, o monitoramento contínuo e inteligente de espaços online, sem depender diretamente das suas organizações gestoras, apresenta desafios de pesquisas importantes para a Ciência da Computação. A seguir, há uma lista de questões fundamentais:

- como identificar e classificar conteúdos digitais (em forma de texto, som, imagem e vídeo) inadequados para crianças e adolescentes que são disseminados em ambientes online com a "presença virtual" deste público?
- como utilizar os recursos da inteligência artificial para identificar e reportar tentativas de "mascaramento" de conteúdo digital (e.g. propaganda velada) impróprio para crianças e adolescentes?
- como utilizar recursos da inteligência artificial para monitorar, detectar e denunciar assédios, pedofilia e *cyberbulying* realizados contra crianças e adolescentes em espaços online?
- como monitorar o uso de espaços online de diferentes plataformas digitais acessadas em múltiplos dispositivos, tais como, smartphone, TV, computador, consoles de jogos, etc. de forma contínua, dinâmica e inteligente?

Quem receberia o resultado desta monitoração? Crianças, adolescentes, pais e cuidadores, com certeza. Mas, obviamente essa monitoração também pode ser utilizada para conscientizar e orientar influenciadores digitais, assim como as plataformas sobre violações observadas. Ela também pode prover insumos para organizações internacionais, ONGs e legisladores que atuam na proteção de crianças e adolescentes.

Ressalta-se que, embora possam ser exploradas de forma independente, as duas frentes propostas se complementam no sentido de que evidências reveladas pela monitoração poderiam alimentar os algoritmos de mentoria na tomada de decisão quanto a ambientes possivelmente inadequados ao público infanto-juvenil. Da mesma forma, observações feitas pelas ferramentas de mentoria, capturando as experiências dos usuários em tempo real, também poderiam alimentar e direcionar as decisões de monitoração. Este potencial de cooperação entre as várias soluções também adiciona desafios próprios ao projeto e desenvolvimento dos softwares baseados em IA a fim de garantir eficiência e escalabilidade, enquanto atende os requisitos de descentralização, proteção de privacidade, justiça, transparência e explicabilidade.

Em suma, um grande desafio para profissionais de Ciência da Computação para os próximos anos é contribuir com a proteção e com a saúde (em especial a saúde mental) de crianças e adolescentes a partir da criação de algoritmos e modelos de inteligência artificial construídos com dados da dinâmica de comportamento desses jovens usuários das tecnologias de informação e comunicação em ambientes online.

3. Sobre os autores

Os autores desta proposta possuem doutorado em Ciência da Computação e bastante experiência em realização de projetos de pesquisa sobre análise de comportamento de usuários em sistemas distribuídos em larga escala, tais como as plataformas de mídias sociais. O Prof. Humberto Marques-Neto da PUC Minas possui publicações recentes sobre as estratégias utilizadas por influenciadores digitais em plataformas como o Instagram e o TikTok, além de experiência em gestão de projetos de pesquisa e inovação envolvendo sistemas distribuídos em larga escala. A Profa. Jussara M. Almeida, pesquisadora nível 1C do CNPq, possui vários trabalhos, muitos em colaboração com pesquisadores internacionais, em temas relacionados à modelagem de comportamento de usuários e computação social. O Prof. Fabrício Benevenuto, pesquisador nível 1D do CNPq, possui vários artigos reconhecidos e citados pela comunidade acadêmica, incluindo o artigo vencedor do *test-of-time award* do ICWSM 2020.

Referências

Akter, M., Godfrey, A., Kropczynski, J., Lipford, H. and Wisniewski, P. (2022). From Parental Control to Joint Family Oversight: Can Parents and Teens Manage Mobile Online Safety and Privacy as Equals? *Proc. ACM Hum.-Comput. Interact*. 6, CSCW1, Article 57, 28 pages.

Assis, J. V. and Valença, G. (2024). Is My Child Safe Online? - On Requirements for Parental Control Tools in Apps used by Children. *Journal on Interactive Systems*, *15*(1), 823–838.

Bozzola, E., Spina, G., Agostiniani, R., Barni, S., Russo, R., Scarpato, E., Di Mauro, A., Di Stefano, A. V., Caruso, C., Corsello, G. and Staiano, A. (2022). The Use of Social Media in Children and Adolescents: Scoping Review on the Potential Risks. *Int'l Journal of Environmental Research and Public Health*, *19*(16), 9960.

Drew, Christopher. (2018). Four Questions to Ask When Using YouTube in the Classroom. *eLearn* 2018, 2, Article 3 (02-01-2018).

Kumar, M., Dwivedi, V., Sanyal, A., Bhatt, P. and Koshariya, R. (2021). Parental Security Control: A tool for monitoring and securing children's online activities. In *Proc. ACM 13th International Conference on Contemporary Computing*, 469–474.

Murthy, V. (2023). *Social Media and Youth Mental Health – The U.S. Surgeon General's Advisory*. Disponível em: https://www.hhs.gov/sites/default/files/sg-youth-mental-health-social-media-advisory.pdf

Nagumo, E., Teles, L. F., and Silva, L. de A. (2020). A utilização de vídeos do Youtube como suporte ao processo de aprendizagem. *Revista Eletrônica De Educação*, *14*, e3757008.

Neumann, M. M., Park, E., Soong, H., Nichols, S. and Selim, N. (2022). Exploring the social media networks of primary school children. *Education 3-13*, 1–15.

O Combate à Desinformação nas Plataformas Digitais: Desafios e Oportunidades

Julio C. S. Reis[1], Philipe Melo[2], Márcio Silva[3], Fabrício Benevenuto[4]

[1]Departamento de Informática – Universidade Federal de Viçosa (UFV)

[2]Instituto de Ciências Exatas e Tecnológicas - Universidade Federal de Viçosa (UFV) – Florestal

[3]Faculdade de Ciência da Computação - Universidade Federal de Mato Grosso do Sul (UFMS)

[4]Departamento de Ciência da Computação - Universidade Federal de Minas Gerais (UFMG)

jreis@ufv.br, philipe.freitas@ufv.br,
marcio.inacio@ufms.br,fabricio@dcc.ufmg.br

Abstract. *Digital platforms have evolved into powerful ecosystems for the production and dissemination of information in the online environment. They have drastically changed the way we communicate and receive information, opening up unprecedented opportunities and creating complex challenges, e.g. in the fight against misinformation. This is one of the most serious and complex challenges of our time, with profound implications such as political polarization, attacks on democracy and risks to public health. This complexity requires solutions that span multiple fields of knowledge and the collaboration of different sectors of society. This article explores the challenges and opportunities that arise in combating misinformation on digital platforms.*

Resumo. *Plataformas digitais se tornaram poderosos ecossistemas de produção e disseminação de informação no ambiente online, transformando drasticamente o modo como nos comunicamos e nos informamos, abrindo oportunidades imprevistas e criando desafios complexos, como os relacionados ao combate à desinformação. Este é um dos desafios mais graves e complexos da atualidade, com impactos profundos, como a polarização política, ataques à democracia e riscos à saúde pública. Essa complexidade exige soluções que envolvam múltiplas áreas do conhecimento e a colaboração de diversos setores da sociedade. Este artigo explora os desafios e oportunidades que surgem no combate à desinformação em plataformas digitais.*

1. Introdução

Plataformas digitais deixaram de ser simples ferramentas de interação entre pessoas (*e.g.,* amigos e familiares) para se tornarem poderosos ecossistemas de produção e distribuição de informação no ambiente online, transformando radicalmente o modo como nos comunicamos e nos informamos. A democratização da inclusão digital e acesso à informação, impulsionada com a proliferação de *smartphones*, trouxe diversos benefícios para a sociedade, mas também revelou vulnerabilidades deste formato de comunicação. O espaço digital não facilita somente a disseminação de informações verídicas e notícias oriundas de fontes confiáveis, mas também favorece a propagação de conteúdos enganosos, que

podem distorcer percepções e influenciar decisões críticas, como nas esferas da política e saúde pública [Reis et al., 2023]. Enquanto até pouco tempo atrás tínhamos um fluxo relativamente controlado de informações, mediado por fontes mais confiáveis, hoje temos cada vez mais um universo gigantesco de dados amplificados por algoritmos e inteligência artificial, onde a desinformação se espalha com uma eficiência, velocidade e escala, muitas vezes, até maior que o conhecimento verificado [Vosoughi et al., 2018].

Nesta direção, plataformas como Facebook, Instagram, Twitter/X, e aplicativos de mensagem instantânea como WhatsApp e Telegram, hoje acessados por bilhões de pessoas, desempenham um papel crucial, pois intensificam ainda mais esse problema ao priorizar o engajamento[1] sobre a precisão e confiança do conteúdo compartilhado, ou permitindo ainda conteúdo viral e anônimo circular por uma rede mais privada e particular protegida por camadas de criptografia [Benevenuto and Melo, 2024]. Parte do desafio é inerente à natureza das próprias plataformas: *(i)* muitas vezes é mais oportuno e menos dispendioso produzir e consumir informações (*i.e.,* notícias) nesses ambientes em comparação com meios noticiosos tradicionais, como jornais, rádio ou televisão; e *(ii)* é mais fácil compartilhar, comentar e discutir com amigos ou outros leitores em plataformas digitais, o que melhora e/ou impulsiona a comunicação e as interações entre os usuários nestes sistemas [Shu et al., 2017].

A desinformação é um dos desafios mais graves da atualidade[2][3], com impactos profundos, como a polarização política, ataques à democracia e riscos à saúde pública. Ela se propaga por qualquer plataforma digital com grande audiência e permeia todos os tipos de conteúdo como imagens, texto, áudio e vídeo. Essa desinformação acaba tendo um impacto devastador na ciência, em particular em áreas como saúde pública [Suarez-Lledo and Alvarez-Galvez, 2021] e crise climática [Treen et al., 2020], ao distorcer fatos e minar a confiança no conhecimento científico. Essa complexidade exige soluções que envolvam múltiplas áreas do conhecimento e a colaboração de diversos setores da sociedade. Do ponto de vista da computação, a produção de tecnologias que possam identificar, monitorar e conter a disseminação de campanhas de desinformação em plataformas digitais é essencial, mas insuficiente se não for acompanhada de uma coordenação entre pesquisadores, tecnólogos, legisladores, educadores e a sociedade civil para reforçar a resiliência cívica e proteger a integridade das informações e dos processos democráticos. Este artigo explora os desafios e oportunidades que surgem no combate à desinformação em plataformas digitais, analisando as estratégias tecnológicas emergentes e as abordagens regulatórias necessárias para restaurar a confiança pública e proteger a integridade dos processos informacionais e democráticos.

2. O Desafio do Combate à Desinformação

Combater a desinformação é uma típica luta adversária. A cada evento crítico, as campanhas de desinformação exploram novas estratégias de manipular a informação/opinião, enquanto novos mecanismos de defesa são criados visando mitigar tais campanhas de desinformação e conter seu avanço [Wu et al., 2024]. Em contextos como saúde, por exemplo, campanhas de desinformação podem ter motivações completamente diferentes [Swire-Thompson et al., 2020], o que torna o problema ainda mais desafiador.

[1] https://www.facebook.com/business/help/718033381901819
[2] https://tinyurl.com/4jn5tzvz
[3] https://tinyurl.com/2efvfvpa

Aqui, são elencadas algumas questões que consideramos críticas de serem abordadas e que comitantemente configuram oportunidades de pesquisa nesta temática. Uma visão geral dessas questões é apresentada na Figura 1.

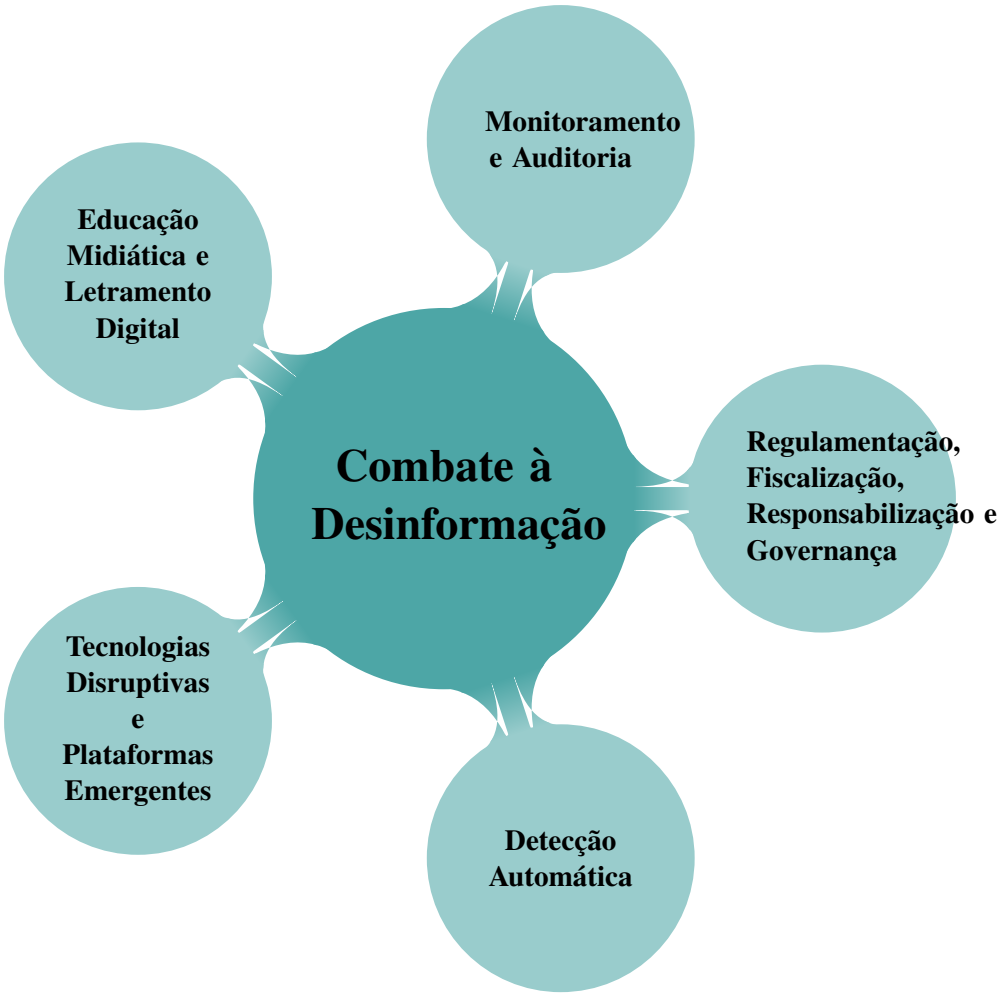

Figura 1. Dimensões que abarcam desafios e oportunidades de pesquisa no combate à desinformação em plataformas digitais.

Monitoramento e Auditoria de Plataformas Digitais

As plataformas digitais passam por constantes mudanças que impactam diretamente como as informações são disseminadas. O WhatsApp, por exemplo, implementou recentemente o recurso de comunidades, que potencializa o disparo massivo de mensagens, enquanto também colocou rótulos que auxiliam o usuário a identificar uma mensagem compartilhada várias vezes dentro da plataforma [Melo et al., 2024]. Além disso, em vários casos, as plataformas digitais são utilizadas apenas como veículos (ou vitrine) para disseminação de conteúdos produzidos por *websites* externos dedicados exclusivamente à produção e disseminação de desinformação [Couto et al., 2024]. Neste cenário, *viabilizar o acesso e a obtenção dos dados para monitoramento e/ou auditoria desses sistemas é algo essencial*. O monitoramento e auditoria de plataformas digitais consistem em medidas para trazer maior transparência ao espaço midiático. Ademais, elas viabilizam uma avaliação crítica do ecossistema online, com potencial para colocar em evidência a (des)informação

que tem circulado de forma viral nestes ambientes, além de revelar como os algoritmos operam, investigando quais critérios são utilizados na disseminação e potenciais estratégias de priorização de conteúdo. Estas estratégias contribuem para o aumento da confiança do público nas plataformas, e também estimulam uma cultura de responsabilidade entre as empresas de tecnologia, incentivando a colaboração entre diferentes atores, como governos, ONGs e pesquisadores, na criação de soluções eficazes para o combate à desinformação.

Regulamentação, Fiscalização, Responsabilização e Governança

Ter regulamentações efetivas é fundamental no combate à desinformação, pois estabelece diretrizes claras para o funcionamento das plataformas digitais e responsabiliza os devidos atores pela disseminação de informações errôneas de forma deliberada e intencional. Essas regulamentações podem incluir exigências para a transparência na publicidade política, a rotulação de conteúdo falso e a proibição de práticas fraudulentas, como a criação de perfis falsos. Com regras bem definidas, é possível limitar o alcance de campanhas de desinformação e garantir que os usuários tenham acesso a informações mais precisas e verificáveis. Além disso, regulamentações robustas podem incentivar as plataformas a adotarem tecnologias e práticas que ajudem a detectar e remover conteúdos enganosos de forma mais eficaz. Neste cenário, protocolos como *C2PA*[4] que se baseiam em criptografia podem ser utilizados para codificar detalhes sobre as origens de um conteúdo em uma plataforma digital. Tais medidas podem ser utilizadas no combate à desinformação, bem como para fiscalização e responsabilização das plataformas. Ainda, para a criação de marcos regulatórios eficazes, é necessária a colaboração entre governos, especialistas de diferentes áreas do conhecimento e a sociedade civil. Isso garante que as regras sejam atualizadas e adaptáveis às rápidas mudanças no ambiente digital. Finalmente, a implementação de mecanismos de fiscalização e sanções para aqueles que desrespeitam as normas é igualmente essencial ao assegurar que as plataformas levem a sério a responsabilidade de monitorar e gerenciar o conteúdo. Essas ações, juntamente com estratégias de governança [Almeida et al., 2024], são fundamentais para a criação de um ambiente mais responsável, confiável, monitorável, e por fim, propício ao combate à desinformação.

Detecção Automática de Desinformação

O desenvolvimento de abordagens de detecção automática de desinformação representa um desafio complexo e multifacetado. À medida que as estratégias para disseminação de desinformação se tornam mais sofisticadas e crescem em velocidade e escala, é essencial que as ferramentas de detecção automática evoluam simultaneamente para acompanhar esse ritmo. Neste contexto, técnicas de aprendizado de máquina e inteligência artificial têm sido empregadas para analisar características de linguagem, padrões de propagação, identificar fontes suspeitas e classificar conteúdos [Reis et al., 2019]. No entanto, a eficácia dessas abordagens é dependente da qualidade dos dados utilizados, bem como da capacidade de adaptação às nuances culturais e contextuais que influenciam a percepção da verdade. Além disso, a detecção automática de desinformação enfrenta desafios éticos e práticos. Questões como a privacidade dos usuários, a liberdade de expressão e a possibilidade de viés algorítmico devem ser cuidadosamente consideradas durante o projeto

[4] `https://c2pa.org/`

e desenvolvimento de tais ferramentas. Neste contexto, é necessário criar soluções eficazes que não apenas protejam a integridade da informação, mas também respeitem os direitos individuais das pessoas. Ainda neste sentido, também é crucial que os sistemas de detecção não sirvam apenas para identificar a desinformação, mas também ofereçam transparência sobre como essas determinações são feitas, o que pode ser útil para substanciar e/ou suportar vereditos fornecidos por agências de checagem de fatos, por exemplo. É fundamental ressaltar que a desinformação não aparece apenas por meio de textos. Atualmente, ela também se espalha nas plataformas digitais num formato multimídia, como áudio, imagem e vídeos (como *deepfakes*). Assim, é imprescindível que as abordagens automáticas de detecção sejam projetadas, adaptadas ou desenvolvidas para lidar com diferentes tipos de mídia.

Tecnologias Disruptivas e Plataformas Emergentes

O ambiente online muda constantemente, surgindo novas plataformas e tecnologias. Neste contexto, campanhas de desinformação se beneficiam do surgimento de "novidades", especialmente aquelas que apresentam menos controle e regulamentação. Logo, *estratégias de combate à desinformação precisam ser propostas, revistas e atualizadas constantemente para acompanhar e contemplar evoluções e/ou mudanças tecnológicas*. Em resposta a esses desafios, tecnologias emergentes, como modelos de linguagem avançados, estão sendo desenvolvidas para auxiliar no combate à disseminação de informações falsas [Liu et al., 2024]. No entanto, a própria sofisticação dessas ferramentas também introduz novos riscos. Modelos de linguagem, como o ChatGPT[5] e DeepSeek[6], têm o potencial de gerar conteúdo incorreto, inclusive muitas vezes de forma a parecer confiável [Kreps et al., 2022]. Além disso, a utilização maliciosa dessas ferramentas por atores com intenções de manipulação aumenta o risco de danos sociais e políticos. Modelos generativos que, devido ao acesso massivo e à facilidade no uso sem critérios nem regras bem definidos, têm apresentado potencial para serem explorados como recurso para a propagação de desinformação. Neste contexto, é fundamental uma análise constante de como essas tecnologias e plataformas emergentes estão/podem ser utilizadas para a disseminação da desinformação no ambiente online, provendo uma avaliação diagnóstica que compreenda uma investigação, inclusive, de questões éticas relacionadas, que forneça insumos para criação de novas e/ou atualização de medidas de combate.

Educação Midiática e o Letramento Digital

A educação midiática e o letramento digital são indispensáveis no combate à desinformação. Elas capacitam os indivíduos a analisar criticamente as informações que consomem e compartilham [Jones-Jang et al., 2021]. Ao aprender a identificar fontes confiáveis, entender a diferença entre fatos e opiniões e reconhecer técnicas de manipulação, os cidadãos se tornam mais resistentes à conteúdos enganosos. Essa formação melhora a capacidade de discernimento em relação às informações compartilhadas em plataformas digitais e também promove uma cultura de verificação e responsabilidade no consumo de informação, essencial em um mundo saturado de dados. Além disso, outro desafio neste contexto está relacionado a como promover uma educação midiática mais inclusiva e acessível. Programas educacionais que ensinam habilidades de

[5] `https://chatgpt.com/`
[6] `https://www.deepseek.com/`

navegação na Internet, análise crítica de conteúdos e compreensão das dinâmicas desses sistemas podem ter um impacto significativo na forma como as pessoas interagem com a informação.

3. Considerações Finais

A desinformação é uma das questões mais desafiadoras ultimamente, com consequências graves, incluindo polarização política, ataques à democracia e riscos para a saúde pública, bem como para diversas outras áreas. A desinformação manifesta-se em qualquer plataforma com uma grande base de usuários, incluindo redes sociais online, aplicações de mensagens e em última instância a Internet. Logo, medir o impacto causado pela propagação deste tipo de conteúdo nestes ambientes é algo extremamente desafiador. É difícil aferir o quanto a desinformação influencia decisões políticas ou votos, por exemplo. Neste cenário, propor e/ou explorar métricas mais palpáveis que podem ser acessadas por pesquisas de opiniões no sentido de identificar, por exemplo, se as pessoas acreditam em teorias conspiratórias, na ciência, nas vacinas, nas mudanças climáticas, etc., pode ser uma maneira de mensurar seu impacto e gerar insumos para a proposição de estratégias efetivas contra o problema.

Finalmente, a desinformação permeia todas as formas de mídia e conteúdo, incluindo imagens, texto, áudio e vídeo. Tudo isso torna a desinformação um problema complexo e desafiador, que requer o envolvimento de diferentes áreas do conhecimento e segmentos da sociedade. Nesse artigo, apresentamos uma série de questões que consideramos críticas de serem abordadas pela sociedade brasileira para mitigar o problema da desinformação nos próximos anos.

4. Currículo Resumido dos Autores

Os autores deste artigo possuem doutorado em Ciência da Computação, são professores e pesquisadores com diversas publicações científicas de impacto no tema desinformação. Todos os autores atuaram no desenvolvimento de sistemas reais para o combate à desinformação através do projeto "Eleições sem Fake"[7], vinculado ao Programa Permanente de Enfrentamento à Desinformação do Tribunal Superior Eleitoral (TSE). O Prof. Julio C. S. Reis fez sua tese de doutorado sobre detecção automática de desinformação em plataformas digitais (vencedora do CTD-SBSI e premiada no CTD-SBC). A tese de doutorado do Prof. Márcio Silva envolveu a construção de um sistema para trazer transparência para as propagandas realizadas em plataformas digitais, resultando em premiações relevantes internacionais, como o concorrido prêmio de proteção à privacidade dado pelo CNIL e INRIA de 2020. O Prof. Philipe Melo manteve no ar um sistema de monitoramento de grupos públicos de WhatsApp e Telegram por 4 anos, oferecendo material a dezenas de reportagens jornalísticas e para checagem de fatos, com destaque para duas reportagens do Fantástico. Finalmente, o Prof. Fabrício Benevenuto, é pesquisador nível 1D do CNPq e coordenador do projeto "Eleições sem Fake". Ele possui dezenas de artigos científicos relacionado ao tema, incluindo o artigo vencedor do *test-of-time award* do ICWSM 2020. Fabrício também é membro do conselho consultivo de segurança do TikTok, dentre outras atividades ligadas ao combate à desinformação.

[7] `http://www.eleicoesemfake.dcc.ufmg.br/`

Agradecimentos

Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).

Referências

Almeida, V., Almeida, J. M., and Meira, W. (2024). The role of computer science in responsible ai governance. *IEEE Internet Computing*, 28(3):55–58.

Benevenuto, F. and Melo, P. (2024). Misinformation campaigns through whatsapp and telegram in presidential elections in brazil. *Communications of the ACM*, 67(8):72–77.

Couto, J. M., Reis, J. C., and Benevenuto, F. (2024). Can computer network attributes be useful for identifying low-credibility websites? a case study in brazil. *Social Network Analysis and Mining (SNAM)*, 14(1):153.

Jones-Jang, S. M., Mortensen, T., and Liu, J. (2021). Does media literacy help identification of fake news? information literacy helps, but other literacies don't. *American Behavioral Scientist*, 65(2):371–388.

Kreps, S., McCain, R. M., and Brundage, M. (2022). All the news that's fit to fabricate: Ai-generated text as a tool of media misinformation. *Journal of experimental political science*, 9(1):104–117.

Liu, A., Sheng, Q., and Hu, X. (2024). Preventing and detecting misinformation generated by large language models. In *Proc. of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 3001–3004.

Melo, P. F., Hoseini, M., Zannettou, S., and Benevenuto, F. (2024). Don't break the chain: Measuring message forwarding on whatsapp. In *Proc. of the Int'l AAAI Conference on Weblogs and Social Media (ICWSM)*, pages 1054–1067.

Reis, J. C., Correia, A., Murai, F., Veloso, A., and Benevenuto, F. (2019). Supervised learning for fake news detection. *IEEE Intelligent Systems*, 34(2):76–81.

Reis, J. C., Melo, P., Silva, M., and Benevenuto, F. (2023). Desinformação em plataformas digitais: Conceitos, abordagens tecnológicas e desafios. *Congresso da Sociedade Brasileira de Computação (CSBC). Jornada de Atualização em Informática (JAI)*.

Shu, K., Sliva, A., Wang, S., Tang, J., and Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1):22–36.

Suarez-Lledo, V. and Alvarez-Galvez, J. (2021). Prevalence of health misinformation on social media: systematic review. *Journal of medical Internet research*, 23(1):e17187.

Swire-Thompson, B., Lazer, D., et al. (2020). Public health and online misinformation: challenges and recommendations. *Annu Rev Public Health*, 41(1):433–451.

Treen, K. M. d., Williams, H. T., and O'Neill, S. J. (2020). Online misinformation about climate change. *Wiley Interdisciplinary Reviews: Climate Change*, 11(5):e665.

Vosoughi, S., Roy, D., and Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380):1146–1151.

Wu, J., Guo, J., and Hooi, B. (2024). Fake news in sheep's clothing: Robust fake news detection against llm-empowered style attacks. In *Proc. of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pages 3367–3378.

A Urgente Necessidade da Literacia Digital

Glaucio de Sousa Santos[1,2], Claudia Lage Rebello da Motta[1]

[1]Programa de Pós-Graduação em Informática
Universidade Federal do Rio de Janeiro (UFRJ)
Rio de Janeiro – RJ - Brasil

[2]Diretoria de Tecnologia da Informação e Telecomunicação
Polícia Civil do Estado do Rio de Janeiro, RJ, Brasil.

{glaucio75@hotmail.com, claudiam@nce.ufrj.br}

Abstract. *The ubiquity of technology brings significant new benefits to modern society, but also presents new challenges to the ongoing safety and well-being of users who have not been properly educated to understand the risks they face in the digital world—or how to mitigate them. This essay offers a critical reflection on the absence of a national public policy for digital literacy aimed at protecting ordinary citizens within the cyber ecosystem. It ultimately argues that the scientific community—particularly in the field of Computing—has a duty to lead the development of broad educational strategies that promote critical, ethical, and secure digital citizenship.*

Resumo. *A ubiquidade da tecnologia traz novos grandes benefícios à sociedade moderna, mas também traz novos desafios no constante a segurança e bem estar dos usuários, que não foram educados adequadamente para ter consciência dos riscos a que estão expostos no mundo digital e nem como mitigá-los. Este ensaio propõe uma reflexão crítica sobre a ausência de uma política pública nacional de literacia digital voltada à proteção do cidadão comum no ecossistema cibernético. Argumenta-se ao final que a comunidade científica, especialmente no âmbito da Computação, tem o dever de liderar a construção de estratégias educacionais de largo alcance que promovam a cidadania digital crítica, ética e segura.*

1. Contextualização

Nos dias atuais, a tecnologia está constante e massivamente presente na vida das pessoas, ultrapassando barreiras como idade, classe social e nível educacional, devido a enorme utilidade que proporciona. O telefone celular é a maior interface para interação humano computador, com uma disponibilidade jamais vista na história da humanidade. A conectividade não está mais limitada pela disponibilidade de infraestrutura de redes de torres de telefonia. A conexão via satélite está disponível aos locais mais remotos. Com essa ubiquidade, difícil imaginar alguém vivendo um dia de vida sem qualquer contato com a tecnologia.

Todavia, há aspectos negativos na disseminação da tecnologia e sua democratização. Como toda evolução tecnológica – e aqui a palavra 'tecnológica' é

usada em uma acepção mais ampla -, desde a descoberta do fogo, passando pela escrita, pela energia hidráulica, energia elétrica, o motor a combustão, o mundo sempre precisou se educar em como tirar o melhor proveito delas e também como evitar o mau uso.

A tecnologia avança em uma velocidade sem precedente. Na mesma toada, as ameaças cibernéticas. Frise-se, o termo 'cibernético' aqui é utilizado com o alcance mais amplo possível, para além das ameaças e proteção a sistemas, redes e dados, mas também abarcar todo ecossistema cibernético, seus usuários e as comunicações entre eles. É nesse ponto que fazemos o recorte da nossa abordagem.

A sociedade brasileira está preparada para lidar com as tecnologias mais recentes de maneira minimamente segura? A quem compete estabelecer e implementar esta educação?

O objetivo da presente proposição é demonstrar que, apesar da certeza da sociedade estar exposta a inúmeros riscos e ameaças cibernéticas, ainda não se tem uma definição clara de um programa educacional com alcance suficiente para colocar a sociedade ciente de tais riscos e preparadas para minimizá-los e, desta forma, assegurar um ecossistema cibernético mais seguro nesta nova era tecnológica da humanidade.

A seguir, vamos elencar os principais riscos a que todos estamos expostos, dada a ubiquidade da tecnologia.

Por fim, nosso objetivo será alcançado se a presente proposição levar a uma reflexão da comunidade acadêmica representada pela Sociedade Brasileira da Computação que, salvo melhor juízo, detém a legitimidade para trazer a luz este assunto tão caro ao desenvolvimento da sociedade brasileira.

2. Relevância

A ciência e a sociedade empregam muito esforço no que diz respeito às ameaças e proteção de sistemas, redes e dados. O tópico é tratado por profissionais altamente especializados, que empregam técnicas e procedimentos que estão na fronteira do saber. Esse campo é impulsionado por um mercado competitivo e lucrativo, onde a lei da oferta e demanda fomenta uma corrida incessante pelo desenvolvimento de novas soluções de segurança.

A necessidade de inovação constante e de respostas rápidas às ameaças emergentes estimula tanto o aprimoramento das tecnologias existentes quanto a criação de novos métodos para proteger as infraestruturas digitais e os dados críticos. Mas é comum ouvir que o usuário é o elo mais fraco de qualquer sistema de proteção. As consequências disso estão muito além da proteção de tais sistemas. É urgente discutir a necessidade de proteção de todos os aspectos da vida do usuário de tecnologia.

O *Center for Humane Technology*, organização independente sem fins lucrativos, tem como foco expor os mecanismos por trás de todas as tecnologias aplicadas para capturar e influenciar nossos pensamentos, comportamentos, emoções e ações. Seus fundadores são os autores do documentário 'O dilema das Redes', assistido por mais de 100 milhões de pessoas e ganhador do Prêmio Emmy, exibido na plataforma Netflix e do vídeo *THE AI DILEMMA*, disponível na plataforma youtube,

exibido em 09 de março de 2023, com mais de 3,4 milhões de visualizações até 20 de setembro de 2024 [HARRIS & RASKIN, 2023].

Em um breve resumo, os autores relatam que tudo começa com engenheiros criando e aplicando algoritmos para aumentar o engajamento de usuários nas redes sociais através da predição de outputs personalizados que conseguissem capturar a intenção dos usuários por mais tempo. Todavia, isso gerou efeitos indesejados: *Doomscrolling* (‘doom’ - catástrofe, ruína; ‘scrooling’ – ato de rolar página) ou ‘rolagem da perdição’ ou ‘rolagem para a desgraça’, assim entendido como a prática de consumir compulsivamente notícias ruins; Diminuição da capacidade concentração infantil; Adição em pornografia; Pedofilia digital; Sobrecarga informacional; Polarização; *Fake news*; Ameaça a democracia; Colapso da realidade, da verdade e confiança; Criação de cyber weapons, guerra assimétrica; Relacionamentos superficiais.

Os autores alertam que estes problemas não foram devidamente enfrentados e que serão agravados exponencialmente com o surgimento dos *Gollem – Generative Large Language Multi-Modal Models* – Modelos Geração de Linguagem Multi-Modais de Larga Escala, com capacidade de gerar sons, imagens. Relatam preocupação com o fato de que em 2022 o número de pesquisas em segurança no uso da Inteligência Artificial foi 30 vezes menor do que as pesquisas em Sistemas de Processamento Neural, sugerindo um preocupante aumento na busca pela habilidade de modelar o pensamento humano em detrimento da segurança no uso da IA [HARRIS & RASKIN, 2023].

O usuário não representa apenas o elo mais fraco na proteção dos sistemas de informação. Ele é, em si, na qualidade de ser humano, o mais frágil e mais exposto a perigos. Será que as pessoas comuns percebem a gravidade dos riscos cibernéticos? Qual o nível de conscientização sobre boas práticas de segurança digital? Qual o impacto do uso massivo de tecnologia sem a consciência dos perigos que esta traz? Com que frequência nos deparamos com comportamentos de risco, como o uso de senhas fracas, clicar em links suspeitos ou compartilhar informações pessoais sem cuidados? Quais os impactos na saúde mental das pessoas com o uso inconsciente de tecnologia? As pessoas estão cientes de problemas como polarização, filtros de bolha, medo de ficar de fora (*Fear of Missing Out*)?

3. Impactos do Desafio na Sociedade

A sociedade está vivenciando as consequências da ausência de educação digital. Ameaças à verdade, à democracia, ao bem-estar e saúde mental das pessoas. Crianças, adolescentes, idosos, adultos, absolutamente todos, estão expostos a diferentes tipos e graus de vulnerabilidade. A tecnologia não respeita espaços sociais, privados, a separação de classes, os limites territoriais.

Em muitos casos, a segurança cibernética é vista como custo e não investimento. A regulamentação insipiente do espaço cibernético contribui para este cenário de risco. A conveniência oferecida por muitos dispositivos e serviços tecnológicos incentivam um comportamento menos cuidadoso. Por exemplo, o uso de autenticação automática e a confiança excessiva em dispositivos sem considerar as configurações de segurança

adequadas expõem a vulnerabilidade. As pessoas muitas vezes priorizam conveniência sobre segurança, o que as torna suscetíveis a fraudes.

Muitas fraudes e ataques cibernéticos não ocorrem por causa de falhas técnicas, mas sim porque as pessoas são manipuladas para fornecer informações, baixar arquivos ou clicar em links maliciosos. Ensinar as pessoas a reconhecerem e evitarem técnicas de engenharia social é um dos maiores desafios, pois esses ataques exploram a psicologia humana, e não apenas falhas tecnológicas. Vivemos uma verdadeira epidemia de fraudes no Brasil.

O crime de estelionato, assim definido como aquele previsto no art. 171 do Código Penal vem aumentando a incidência ano a ano por todo o Brasil, segundo o Anuário Brasileiro de Segurança Pública de 2024, publicado pelo Fórum Brasileiro de Segurança Pública.

No Estado do Rio de Janeiro, observamos o mesmo fenômeno do aumento da incidência dos crimes de estelionato ao longo dos últimos anos.

Uma análise do comportamento de outros indicadores de crimes no Estado do Rio de Janeiro mostra um comportamento semelhante entre os crimes de Estelionato e os crimes de Roubo e Furto de Celular:

Os aumentos nos anos de 2022 e 2023 sugerem uma correlação entre os crimes de estelionato com os crimes de furto e roubo de celular. A hipótese sugerida é que criminosos estão se aproveitando da falta de educação digital do cidadão para criar mecanismos de proteção de seus dispositivos, que estariam sendo utilizados pelos criminosos para prática de crimes de estelionato.

Essa hipótese é reforçada pelo comportamento completamente diferente dos outros indicadores de crimes que aferem o comportamento da criminalidade no Rio de Janeiro. A seguir, exibimos os indicadores de roubo de veículos, roubo de carga e ocorrências com morte violenta – letalidade violenta, que são historicamente usados pelo Governo do Estado do Rio de Janeiro.

Os crimes de estelionato digitais provocam prejuízos financeiros e emocionais nas suas vítimas. Com uma educação adequada, dando as pessoas conhecimento necessário para que se protejam da engenhosidade social perpetrada pelos criminosos, esse impacto poderia ser minimizado.

A disrupção do nosso meio de vida trazida pela tecnologia não foi percebida como uma ameaça. Todos os dias, milhares de pessoas entram no espaço cibernético como se estivessem sentados no banco do motorista de um veículo sem saber como guiá-lo. O desafio é descobrir como criar uma cultura de uso seguro da tecnologia, manter a educação sobre cibersegurança sempre atualizada para acompanhar as novas vulnerabilidades e técnicas de ataque que surgem constantemente. Hackers estão sempre desenvolvendo novos métodos de exploração, e manter o público informado sobre os perigos emergentes é uma tarefa contínua.

Os dispositivos que utilizam sistemas computacionais estão por toda parte da vida do cidadão comum. Difícil apontar um empreendimento que não faça uso de um sistema. Seja um smartphone, usado pelo microempreendedor, ou data centers enormes,

usados por grandes empresas, a economia deslocou-se do eixo industrial para uma base tecnológica.

A evolução tecnológica sem dúvidas traz incontáveis benefícios para a sociedade, ganhos de produtividade, acesso a informação, expansão do mercado consumidor, automação de atividades, eficiência operacional, controle, etc. Mas, assim como toda tecnologia ao longo da história da humanidade, traz também aspectos negativos, relacionados ao seu mal uso e consequências não pensadas.

Quando se fala em cibersegurança, muito se fala sobre a proteção contra ataques aos sistemas de informação, a proteção dos dados, privacidade, criação destes mecanismos. Todavia, não importa quão moderno, bem desenvolvido e arquitetado são as soluções de proteção, a maior vulnerabilidade está na falta de conhecimento dos usuários comuns. Assim como ele é o elo fraco desta corrente, também é, em si, o mais vulnerável e necessitado de proteção. Saber como se comportar, usar, o que evitar, é imperioso para evitar ataques cibernéticos

Tabela 1 – FÓRUM BRASILEIRO DE SEGURANÇA PÚBLICA. 18º Anuário Brasileiro de Segurança Pública. São Paulo: Fórum Brasileiro de Segurança Pública, 2024. Disponível em: https://publicacoes.forumseguranca.org.br/handle/1234 , consultado em 21/09/2014

TABELA 15

Estelionato e Estelionato por meio eletrônico [1]

Brasil e Unidades da Federação – 2022-2023

Brasil e Unidades da Federação	Estelionato					Estelionato por meio eletrônico				
	Ns. Absolutos		Taxas [2]		Variação (%)	Ns. Absolutos		Taxas [2]		Variação (%)
	2022 [3]	2023	2022	2023		2022 [3]	2023	2022	2023	
Brasil	**1.816.438**	**1.965.353**	**894,4**	**967,8**	**8,2**	**207.125**	**235.393**	**173,1**	**196,8**	**13,6**
Acre	5.649	5.538	680,6	667,2	-2,0	154	248	18,6	29,9	61,0
Alagoas	20.087	21.058	642,2	673,3	4,8	4.973	5.729	159,0	183,2	15,2
Amapá	8.587	7.146	1.170,3	973,9	-16,8	380	618	51,8	84,2	62,6
Amazonas	15.430	17.464	391,5	443,1	13,2	419	877	10,6	22,2	109,3
Bahia	91.223	98.934	645,1	699,6	8,5	4.183	7.515	29,6	53,1	79,7
Ceará	68.754	68.235	781,7	775,8	-0,8	...	...	...	...	...
Distrito Federal	52.995	50.387	1.881,0	1.788,4	-4,9	15.749	16.060	559,0	570,0	2,0
Espírito Santo	37.391	35.168	975,3	917,3	-5,9	15.277	14.578	398,5	380,3	-4,6
Goiás	74.163	74.906	1.051,0	1.061,5	1,0	1.488	2.167	21,1	30,7	45,6
Maranhão	14.799	14.782	218,4	218,1	-0,1	5.781	6.278	85,3	92,6	8,6
Mato Grosso	20.261	20.565	553,8	562,1	1,5	9.253	11.257	252,9	307,7	21,7
Mato Grosso do Sul	13.647	13.357	495,0	484,5	-2,1	5.027	5.414	182,3	196,4	7,7
Minas Gerais	132.120	141.649	643,2	689,6	7,2	35.878	40.906	174,7	199,2	14,0
Pará	35.038	36.845	431,5	453,7	5,2	13.099	16.884	161,3	207,9	28,9
Paraíba	5.669	6.443	142,6	162,1	13,7	406	1.030	10,2	25,9	153,7
Paraná	141.777	145.205	1.238,8	1.268,8	2,4	5.738	7.029	50,1	61,4	22,5
Pernambuco	59.499	58.460	656,8	645,3	-1,7	14.164	13.941	156,4	153,9	-1,6
Piauí	15.310	16.663	468,0	509,4	8,8	276	584	8,4	17,9	111,6
Rio de Janeiro	123.841	120.218	771,3	748,8	-2,9	...	...	...	...	...
Rio Grande do Norte	25.945	26.018	785,6	787,8	0,3	...	...	...	...	...
Rio Grande do Sul [4]	95.182	87.627	874,6	805,2	-7,9	...	6.577	...	...	...
Rondônia	15.376	18.393	972,4	1.163,2	19,6	5.932	6.532	375,2	413,1	10,1
Roraima	5.396	5.359	847,5	841,7	-0,7	778	1.893	122,2	297,3	143,3
Santa Catarina	95.706	94.944	1.257,6	1.247,6	-0,8	64.646	64.482	849,4	847,3	-0,3
São Paulo	611.572	750.430	1.377,1	1.689,7	22,7	...	...	...	...	...
Sergipe	19.734	17.763	892,9	803,8	-10,0	446	751	20,2	34,0	68,4
Tocantins	11.287	11.796	746,8	780,4	4,5	3.078	4.043	203,6	267,5	31,4

Fonte: Secretarias Estaduais de Segurança Pública e/ou Defesa Social; Instituto de Segurança Pública/RJ (ISP); Polícia Civil do Estado do Acre; Polícia Civil do Distrito Federal; Polícia Civil do Estado de Roraima; Instituto Brasileiro de Geografia e Estatística (IBGE) – Censo Demográfico 2022; Fórum Brasileiro de Segurança Pública.

(...) Informação não disponível.

(1) Em 2021, o crime de Estelionato - Fraude eletrônica passou a ser tipificado pelos parágrafos 2ºA, 2ºB e 3º do art. 171 do Código Penal.

(2) Taxas por 100 mil habitantes.

Nesse sentido, um dos grandes desafios tecnológicos é a ausência de literacia digital dos usuários de todos os níveis. A tecnologia avança rapidamente e, com ela, as ameaças cibernéticas. A comunidade acadêmica está diante de um desafio enorme. Estas ameaças e vulnerabilidades são bem conhecidas, definidas e estudadas no âmbito acadêmico. Mas todo esse conhecimento precisa ultrapassar os limites da academia. Precisamos pensar nas melhores e mais eficientes estratégias para enfrentar este desafio.

Referências

HARRIS, Tristan and RASKIN, Aza. (2023) “The AI Dilemma. [S.l.]: Center for Humane Technology”

https://www.youtube.com/watch?v=xoVJKj8lcNQ. Acesso em: 19 set. 2024.

Universalização da Cidadania Digital

Flavia Bernardini[1], Raissa Barcellos[2], Claudia Cappelli[2], José Viterbo[1], Cristiano Maciel[3]

[1] Universidade Federal Fluminense (UFF)
Niterói – RJ – Brazil

[2] Universidade do Estado do Rio de Janeiro (UERJ)
Rio de Janeiro – RJ – Brazil

[3] Universidade Federal de Mato Grosso (UFMT)
Cuiabá – MT – Brazil

{fcbernardini,viterbo}@ic.uff.br,
{raissa.barcellos,claudia.cappelli}@ime.uerj.br,
cristiano.maciel@ufmt.br

Abstract. *Neste trabalho discutimos os desafios para a universalização da cidadania digital no Brasil, destacando a importância do acesso às tecnologias, do letramento digital, da capacidade de interpretar dados e da participação cidadã em plataformas digitais. Propomos um modelo conceitual composto por quatro níveis: acesso à informação, letramento digital, análise e consumo de serviços digitais, e e-participação. Apresentamos os principais componentes, metas e indicadores necessários para sua implementação. Defendemos que a construção de um ecossistema digital inclusivo, acessível e eficaz exige a colaboração entre governo, academia, sociedade civil e setor privado.*

Resumo. *In this work, we discuss the challenges of universalizing digital citizenship in Brazil, emphasizing the importance of access to technology, digital literacy, data interpretation skills, and civic participation through digital platforms. We propose a conceptual model structured into four levels: access to information, digital literacy, data analysis and digital service consumption, and e-participation. We present key components, goals, and indicators required for implementation. We argue that building an inclusive, accessible, and effective digital ecosystem demands collaborative efforts among government, academia, civil society, and the private sector.*

1. Introdução

A cidadania está relacionada à noção de pertencimento a uma comunidade política. Como conceito, está fortemente conectada ao território onde a pessoa nasce ou escolhe viver. Ela conecta as pessoas umas às outras sob a jurisdição do estado. Também conecta as pessoas ao estado de formas específicas, podendo ser por meio de direitos (como o voto) ou deveres (como o pagamento de impostos). A existência e demanda por direitos cívicos,

sociais e políticos que garantam aos cidadãos igualdade e proteção são componentes chave da cidadania.

Nas últimas décadas, o avanço das tecnologias digitais tem transformado profundamente a sociedade, trazendo consigo novas oportunidades para a participação cívica e a inclusão digital (Maciel et al., 2016). Surge o conceito de cidadania digital, sendo a atuação do cidadão no meio digital para exercer seus direitos e deveres (Hintz et al, 2019). No entanto, à medida que o mundo digital se expande e evolui, surgem novos desafios que precisam ser enfrentados para garantir que essa cidadania seja acessível, inclusiva e eficaz (Viterbo e Bernardini, 2022). Um dos motivos é a profunda transformação em como a sociedade é organizada e como as decisões são tomadas. Todos esses mecanismos podem gerar muitas oportunidades para ações cidadãs, como o engajamento político por meio de debates em mídias sociais e o acesso a serviços com interações online. Por outro lado, as tecnologias podem criar e reforçar desigualdades, como se tem visto nos últimos anos. Por exemplo, durante a pandemia, diversos serviços de governo passaram a ser oferecidos de forma totalmente online, mas a falta de acesso ou inclusão digital, de letramento digital, de capacidade de entendimento e análise crítica dos dados, excluiu muitos cidadãos no Brasil.

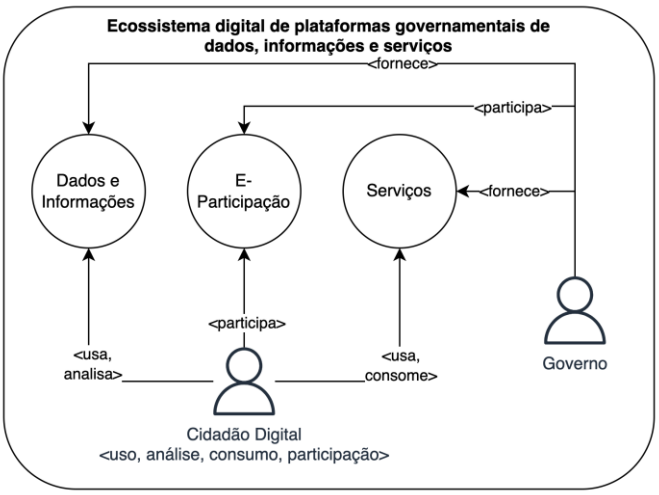

Figura 1. Ecossistema digital de plataformas governamentais de dados, informações e serviços e seus componentes

Na Figura 1, apresentamos um esquema do ecossistema digital de plataformas governamentais de dados, informações e serviços. O ator governo atua como fornecedor de dados, informações e serviços para os cidadãos. O ator cidadão encontra, entende, analisa e interpreta dados e informações e usa e consome serviços do governo. Nas plataformas de e-participação, que podem ser fornecidas pelo governo ou pela sociedade civil, tanto os cidadãos como os governos participam. Nesse contexto, o grande desafio de ter uma cidadania digital universal é amplo e multifacetado. Daí, deve envolver o acesso igualitário à tecnologia, o letramento digital, o conhecimento da proteção da privacidade e da segurança online, a habilidade de analisar, interpretar e entender as plataformas governamentais, e a necessidade de fortalecer a participação pública em plataformas digitais. Ao mesmo tempo, a crescente complexidade dos ambientes digitais, com a integração de inteligência artificial, big data e novas ferramentas de comunicação,

exigirá que governos, sociedade civil e setor privado colaborem para criar políticas e práticas que assegurem uma cidadania digital inclusiva e equitativa.

Este documento está organizado como segue: Na Seção 2, são apresentadas as principais barreiras observadas nas diversas sociedades e comunidades da atualidade para alcançar a cidadania digital universal. Na Seção 3, são apresentados os componentes necessários da cidadania digital universal, com o objetivo de dividir a cidadania digital universal em partes que possam ser tratadas por diferentes grupos, considerando suas diversas expertises, já que a cidadania digital universal é altamente multifacetada. Na Seção 4, são apresentados os desafios, agrupados por atividades e características da cidadania digital universal, que devem ser individualmente tratados, bem como a interação entre eles. Na Seção 5, é apresentado um conjunto nuclear de metas e indicadores que devem ser considerados para avaliação da evolução da cidadania digital universal. Finalmente, na Seção 6, são apresentadas as considerações finais deste trabalho.

2. Barreiras para a Cidadania Digital Universal

Uma das principais barreiras neste tema é a **desigualdade de acesso à tecnologia**, conhecida como a "exclusão digital". Milhões de pessoas, especialmente em áreas rurais ou de baixa renda, ainda não têm acesso à internet de qualidade ou a dispositivos tecnológicos como computadores e smartphones. Sem esse acesso, é impossível para essas populações desenvolverem o letramento digital necessário para participar plenamente da sociedade digital. Esse desafio é agravado em países em desenvolvimento, onde a infraestrutura tecnológica ainda é insuficiente.

Outra barreira importante é a **falta de formação adequada em letramento digital ainda no sistema educacional**. Muitas escolas, especialmente nas regiões menos favorecidas, ainda não integram o uso das tecnologias digitais de forma eficiente em seus currículos. Além disso, a formação de professores muitas vezes não inclui o uso de ferramentas digitais, o que dificulta a transmissão dessas habilidades para os alunos. Mesmo quando há acesso a dispositivos, muitas vezes faltam programas educativos que ensinam como usá-los de maneira crítica e produtiva.

A **barreira geracional** também é um fator relevante. Pessoas mais velhas, que não cresceram em um ambiente digital, frequentemente enfrentam dificuldades para se adaptar às novas tecnologias. Elas podem encontrar resistência ou ansiedade ao lidar com dispositivos e plataformas digitais, o que limita seu letramento digital. Programas de capacitação para essa faixa etária ainda são escassos, dificultando sua inclusão plena na sociedade digital.

Além disso, a **complexidade crescente do ambiente digital** cria constantemente novas barreiras para o letramento digital. Com a proliferação de diferentes plataformas, aplicativos e sistemas, torna-se difícil para o usuário médio acompanhar todas as inovações e compreender plenamente como navegar com segurança e eficácia no mundo digital. A desinformação, as ameaças à privacidade e os ataques cibernéticos também adicionam camadas de dificuldade, exigindo que os cidadãos tenham habilidades avançadas para identificar fontes confiáveis e proteger suas informações pessoais.

Por fim, a **exclusão digital de grupos vulneráveis**, como pessoas com deficiência, ainda é uma barreira considerável. Muitos sites e aplicativos não são projetados com acessibilidade em mente, dificultando o uso dessas tecnologias por pessoas com deficiências visuais, auditivas ou motoras. A falta de inclusão digital para esses grupos representa uma barreira significativa para o desenvolvimento do letramento digital.

É importante observar que, para diminuir as barreiras, é fundamental que equipes interdisciplinares sejam formadas. Portanto, além de profissionais com formação e atuante em computação são atores importantes nesse tema, já que muitas das barreiras envolvem a evolução no desenvolvimento de artefatos computacionais, é importante que designers, especialistas em educação, e outros profissionais de áreas correlatas sejam envolvidas.

3. Componentes Necessários à Universalização da Cidadania Digital e suas relações

São cinco os componentes necessários para a Universalização da Cidadania Digital:

1. **Acesso às TICs e à Informação:** Este componente é o primeiro e mais básico pilar da cidadania digital. Todos os cidadãos devem ter acesso às tecnologias digitais e à internet para poderem encontrar informações de maneira rápida, acessível e inclusiva.
2. **Letramento Digital:** O acesso e uso eficaz das plataformas digitais são cruciais para a inclusão digital e social. Sem habilidades de letramento digital, os cidadãos podem enfrentar exclusão ao não conseguir acessar informações e oportunidades online. O letramento digital é essencial para a participação ativa e cidadania, especialmente considerando o letramento digital crítico deles. A utilização de Linguagem Simples, que visa simplificar a comunicação ao eliminar jargões e informações desnecessárias, é fundamental para tornar o conteúdo digital mais acessível e inclusivo, especialmente para pessoas com baixo nível de escolaridade, idosos ou iniciantes em tecnologia. Isso contribui para a democratização do acesso à tecnologia e o desenvolvimento do letramento digital.
3. **Análise e Interpretação de Dados:** Dados e informações devem ser precisos, completos, consistentes, coerentes e organizados, de maneira que transmita significado ao usuário, estimulando a formação de conhecimento e o engajamento. A interpretabilidade é entendida como a capacidade de (i) analisar um conjunto de dados e informações e (ii) de interpretar e analisar estatísticas, gráficos, relatórios e indicadores. A interpretação adequada dos dados e informações é fundamental para o empoderamento dos cidadãos e para promover uma participação cidadã efetiva. Além disso, a interpretação é também medida pelas perspectivas e habilidades dos usuários. É, portanto, essencial estabelecer uma cooperação eficaz entre governo e cidadãos.
4. **Consumo de Serviços Digitais:** Refere-se à capacidade de os cidadãos acessarem e aplicarem as informações encontradas, bem como utilizar o serviço digital com êxito. Isso inclui a navegação em portais governamentais, a solicitação de serviços públicos online e o uso de ferramentas digitais para resolver problemas cotidianos.
5. **Participação Pública:** Cidadãos utilizam as tecnologias digitais para se engajar com o governo e com as questões de interesse público. Isso pode ocorrer por meio de consultas públicas online, votações eletrônicas, plataformas de deliberação colaborativa, fóruns e participação em discussões políticas digitais.

Para estes componentes apresentados, pontuamos que a implementação de uma identidade digital única, interoperável entre diferentes órgãos públicos, pode simplificar o acesso aos serviços, reduzir burocracias e permitir a personalização de ofertas digitais (Robles-Carrillo, 2024). É essencial que o modelo adotado garanta o controle do cidadão sobre seus próprios dados, promovendo a confiança e a adesão à solução.

A Figura 2 apresenta um diagrama dos componentes da cidadania digital e seus níveis relacionados, baseado em características dos Modelos de Transparência Organizacional

(Cappelli, 2009) e de Interpretabilidade de Dados Governamentais Abertos (Barcellos et al., 2021). O diagrama estrutura a cidadania digital em quatro níveis: o Nível 1 é o acesso à informação; o Nível 2 é o letramento digital, que envolve o entendimento e uso de dados e informações (Trevisan et al., 2022); o Nível 3 refere-se à análise e interpretação de dados, assim como o consumo de serviços digitais; e o Nível 4 é a e-participação, que permite a contribuição ativa na tomada de decisões políticas. Destaca-se que os cidadãos podem possuir habilidades parciais em todos os níveis, o que não o impediria de transitar por diversos níveis sem o atingimento da maturidade completa em cada um deles.

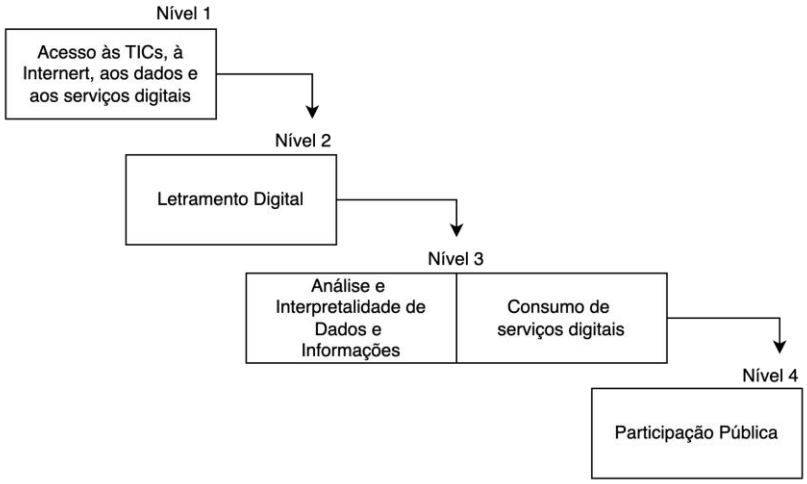

Figura 2. Níveis de cidadania digital universal e seus componentes

4. Desafios para a Cidadania Digital Universal

Os desafios foram agrupados em cinco tarefas ou aspectos da cidadania digital universal, e são apresentados a seguir:

1. Acesso às TICs e à Informação: Garantia que toda a população tenha acesso universal à internet e a dispositivos tecnológicos com acesso a Dados Abertos Governamentais. Criação de plataformas governamentais que sejam inclusivas e acessíveis para todos.

2. Letramento Digital: Criação de programas de capacitação digital desde o ensino fundamental até formações independentes, que ensinem desde habilidades básicas até o uso de sistemas de dados públicos complexos. Estabelecimento de padrões de interface que garantam a redução da complexidade e o aumento da transparência. Definição de padrões para construção de apresentações de dados acessíveis, visualizações interativas de dados e simplificação das formas de apresentação de informações ao cidadão. Além da preocupação com a exclusão digital, é necessário reconhecer um desafio inverso, mas igualmente crítico: o uso excessivo e acrítico de tecnologias digitais, especialmente entre crianças e adolescentes. O letramento digital, portanto, deve ser entendido não somente como a aquisição de habilidades técnicas, mas também como a formação de uma consciência crítica sobre o uso da tecnologia, prezando por um letramento digital crítico (Técnico-operacional em TIC, Informacional em TIC e Social no Uso das Mídias). Isso implica a necessidade de reformular práticas educacionais, integrando o uso responsável das tecnologias ao currículo escolar, incentivando a mediação ativa por educadores e

famílias, e promovendo a alfabetização midiática desde os primeiros anos da educação formal.

3. **Análise e Interpretação de Dados:** Criação de ferramentas interativas que permitam a análise de dados. Desenvolvimento de métodos e ferramentas que apoiem a interoperabilidade entre os portais de dados e informações governamentais. Implementação de técnicas computacionais de padronização de dados e estruturação de informações, análise semântica e de sentimentos, e visualização em portais governamentais de dados, informações e serviços. Criação de novos mecanismos de coleta e recuperação de dados e informações relevantes para o cidadão.
4. **Consumo de Serviços Digitais:** Criação de modelos de garantia da qualidade de interação de plataformas de serviços digitais. Simplificação de sistemas e processos digitais governamentais, tornando-os mais fáceis de usar. Garantia da transparência dos processos governamentais para consciência do cidadão quanto a seu papel em tais processos.
5. **E-Participação:** Criação de mecanismos de incentivo e engajamento de participação dos cidadãos nas decisões políticas e na governança. Desenvolvimento de plataformas de participação direta e de democracia líquida, como sistemas de consultas públicas, audiências online e fóruns de debate sobre políticas públicas, que garantam requisitos de privacidade e outras leis que se enquadrem nesse processo.

Além dos desafios técnicos e operacionais, é fundamental considerar os princípios de direitos fundamentais que devem sustentar qualquer ecossistema de cidadania digital. A garantia da privacidade, especialmente em plataformas que coletam dados sensíveis, deve ser assegurada por mecanismos transparentes de proteção e consentimento. O tratamento isonômico na oferta de serviços digitais deve evitar discriminações, intencionais ou não, causadas por algoritmos ou critérios de acesso. Por fim, é essencial promover a accountability dos agentes públicos e das tecnologias envolvidas, assegurando que cidadãos possam compreender, questionar e fiscalizar o funcionamento das plataformas e das decisões tomadas por sistemas automatizados. A garantia dos direitos fundamentais na cidadania digital requer uma abordagem centrada nos direitos, conforme destacado pela OECD (2022), que enfatiza a importância de políticas e práticas que assegurem a proteção da privacidade, a equidade no acesso e a responsabilidade dos atores envolvidos na era digital.

4. Metas iniciais e indicadores de medição

Foram definidas cinco metas a serem consideradas como as primeiras metas para tratar as barreiras e desafios, com respectivos indicadores a serem considerados. É importante observar que novas metas podem ser incluídas no futuro. As metas são:

- **Meta 1**: Acesso universal dos cidadãos à internet: Indicadores do CGI.br (Comitê Gestor da Internet no Brasil), que monitora a cobertura e acessibilidade da internet no país; Indicadores de letramento digital da população.
- **Meta 2**: Uso universal dos portais de dados, informações e serviços governamentais: Indicadores de engajamento dos cidadãos no uso dos portais governamentais, no sucesso

em atingir seus objetivos nas interações, e na redução na procura pelos serviços de forma presencial com o governo.

- **Meta 3**: Uso universal por parte do governo da Técnica de Linguagem Simples para oferta de dados, informações e serviços aos cidadãos: Indicadores relativos ao percentual de portais de dados, informações e serviços que usam Linguagem Simples.
- **Meta 4**: Uso universal por parte do governo de técnicas de coleta, recuperação e análise de dados e informações por parte do cidadão: Indicadores relativos ao percentual de portais governamentais que adotaram tais práticas.
- **Meta 5**: Adoção universal por parte do governo de mecanismos de interoperabilidade de portais ou de portais únicos para simplificar o acesso aos dados, informações e serviços de forma padronizada e simplificada: Indicadores relativos ao percentual de portais que oferecem mecanismos simplificados de coleta, recuperação e auxílio ao uso de dados, informações e serviços de forma simplificada.

Do ponto de vista computacional, o desenvolvimento de artefatos computacionais para atingir tais metas é fundamental. Isso, pois somente os órgãos de governo podem não conseguir resolver todos os desafios. Por exemplo, a falta de recursos humanos em TIC em muitas prefeituras de cidades de pequeno e médio porte inviabiliza várias das atividades apresentadas. É necessário, portanto, haver também a participação da academia e de movimentos de inovação na iniciativa pública e privada.

5. Considerações Finais

Contextualizamos neste trabalho as barreiras, componentes, desafios e metas para evoluirmos a cidadania digital universal em uma sociedade, em especial a sociedade brasileira. Como este grande desafio apresentado está contextualizado na Sociedade Brasileira de Computação, compondo os grandes desafios da computação para os próximos 10 anos a partir de 2025, focamos principalmente no desenvolvimento de artefatos computacionais que vão ao encontro da cidadania digital universal. No entanto, é importante ressaltar que equipes multidisciplinares de pesquisa são essenciais, já que muitos componentes sociais e educacionais estão presentes nas barreiras e desafios no tema. Além disso, as metas e indicadores apresentados servem de base para avaliações iniciais. No entanto, modelos de referência para a cidadania digital podem ser propostos e avaliados de forma iterativa. Por fim, mecanismos de avaliação do nível de cidadania digital universal em sociedades e comunidades com demandas específicas também devem ser propostos e aplicados para a evolução do entendimento do cenário de cidadania digital universal.

Referências

Barcellos, R., Bernardini, F. and Viterbo, J. (2022). “Towards defining data interpretability in open data portals: Challenges and research opportunities”. Information Systems, 106, 101961. https://doi.org/10.1016/j.is.2021.101961

Cappelli, C. (2009). “Uma Abordagem para Transparência em Processos Organizacionais Utilizando Aspectos”. Tese de Doutorado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brasil.

Hintz, A., Dencik, L. and Wahl-Jorgensen, K. (2019). “Digital Citizenship in a Datafied Society”. Polity Press, United Kingdom.

Maciel, C.; Slaviero, C. ; Cappelli, C. ; Garcia, A. C. B. (2016). Technologies for popular participation: a research agenda. In: 17th Annual International Conference on Digital Government Research, Shanghai-China. Internet Plus Government: New Opportunities to Solve Public Problems. New York: ACM, 2016. p. 202-211. https://doi.org/10.1145/2912160.291219

Robles-Carrillo, M. (2024). Digital identity: an approach to its nature, concept, and functionalities. International Journal of Law and Information Technology, 32(1), eaae019. https://doi.org/10.1093/ijlit/eaae019

Trevisan, D., Maciel, C., & Souza, T. F. M. de. (2022). O Lugar da Crítica na Mobilização de Letramentos Digitais. Revista de Educação do Vale do Arinos - RELVA, 9(2), 110-133. https://doi.org/10.30681/relva.v9i2.6023

OECD (2022). *Rights in the Digital Age: Challenges and Ways Forward*. OECD Digital Economy Papers, No. 347. Disponível em: https://read.oecd.org/10.1787/deb707a8-en.

Viterbo, J.; Bernardini, F. (2022). Empoderamento Digital: O Papel da Computação na Construção de uma Sociedade Inclusiva e Democrática. Computação Brasil, 48, p. 12–14. https://doi.org/10.5753/compbr.2022.48.2777

PATROCÍNIO