

## Capítulo

# 9

## Levantamento bibliográfico utilizando a ferramenta TheEnd

Martony Demes da Silva, Gleison de Andrade e Silva

### *Abstract*

*The bibliographic survey is one of the main stages in the construction of scientific research. This is due to the need to search, in databases, works related to the research topic. In this way, it is possible to catalog and reference relevant works that encourage research. However, this activity can be complex if a well-defined method and support tools are not used. This chapter aims to teach academics and researchers a method of conducting a bibliographic survey using Systematic Mapping of Literature using the resources of the tool TheEnd. It is hoped, therefore, that the researcher can build a bibliographic survey quickly and systematically .*

### *Resumo*

*O levantamento bibliográfico é uma das principais etapas da construção da pesquisa científica. Isso se deve pela necessidade de buscar, em bases de dados, trabalhos relacionados ao tema da pesquisa. Desta forma, é preciso catalogar e referenciar trabalhos relevantes que fomentem a pesquisa. Porém, essa atividade pode ser complexa caso não se utilize um método bem definido e ferramentas de apoio. Este capítulo pretende ensinar aos acadêmicos e pesquisadores um método de realizar um levantamento bibliográfico por meio do Mapeamento Sistemático da Literatura utilizando os recursos da ferramenta TheEnd. Espera-se, com isso, que o pesquisador possa construir um levantamento bibliográfico de maneira rápida e sistemática.*

### **9.1. Introdução**

Levantamento bibliográfico ou Pesquisa Bibliográfica (PB) é a primeira etapa no processo de investigação científica [Detroz et al. 2015]. PB é o procedimento de consultar bases de dados de trabalhos científicos com o objetivo de descobrir e reunir informações que servirão de base para a construção da investigação a partir de um determinado tema.

De certo, a PB é um levantamento de toda a bibliografia publicada em forma de livros, revistas, artigos, publicações avulsas e impensas escritas [Lakatos and Marconi 1991]. Com efeito, o pesquisador terá oportunidade de visualizar diversos aspectos oportunos em determinado assunto.

Conseqüentemente, com a evolução das pesquisas, ocorre um aumento do número de publicação e resultados [Petersen et al. 2008]. Nesse sentido, é importante realizar um levantamento do estado da arte desse tema. Esse levantamento possibilita que pesquisadores que obtenha uma visão geral do tema de interesse.

Porém, a condução de uma PB de forma parcial ou incompleta pode levar a resultados tendenciosos [Detroz et al. 2015]. Isso inviabiliza a replicação dos métodos utilizados e conseqüentemente conduzindo à conclusões incorretas. Diante desse quadro, o estudante de graduação que tem pouca maturidade em pesquisas bibliográficas pode ser levado a um caminho não linear de pesquisa científica.

Diante desse cenário, a proposta deste capítulo é apresentar uma prática de elaborar uma pesquisa bibliográfica por meio de um método denominado Mapeamento Sistemático da Literatura (MSL). A estratégia terá suporte de um ambiente digital Web e colaborativo *TheEnd* [Braga et al. 2016]. Ao final, o resultado será um levantamento bibliográfico de um determinado tema.

Nesse âmbito, sabe-se que há alguns tipos de PB os quais serão apresentados nesse trabalho. Entretanto, a escolha do Mapeamento Sistemático da Literatura justifica-se por ser um método científico capaz identificar, interpretar e sumarizar os trabalhos relevantes para determinada linha de pesquisa, área ou fenômeno de interesse de forma não tendenciosa e replicável [Kitchenham 2004a].

Como exposto, a proposta pretende utilizar ferramentas específicas de suporte à PB. São elas:

i) O *software TheEnd*: é um serviço Web capaz de auxiliar o planejamento, execução e sumarização dos resultados de um MSL, tornando a pesquisa mais ágil, colaborativa e replicável- *link de acesso à ferramenta*;

ii) A ferramenta *Google Sheets*: com este software, pode-se gerenciar os trabalhos selecionados e as informações extraídas do *Software TheEnd* de modo *on-line*.

O Diferencial desta proposta é a possibilidade de elaborar referencial teórico (e bibliográfico) de modo sistemático, replicável e colaborativo. Essas características facilitam o entendimento pelo o estudante iniciante em pesquisas científicas. Além disso, a ferramenta de MSL citada permite que vários pesquisadores trabalhem ao mesmo tempo em uma mesma pesquisa, centralizando e compartilhando dúvidas e conhecimento.

### 9.1.1. Objetivos

O objetivo geral do curso é explanar sobre levantamento bibliográfico com enfoque na PB sistemática. Os objetivos específicos são:

- Explicar sobre pesquisa bibliográfica e mapeamento sistemático;
- Apresentar a Ferramenta *Software TheEnd*;

- Apresentar uma prática de uma PB do tipo Mapeamento Sistemático com uso do *Software TheEnd*.

### 9.1.2. Plano do Curso

O plano do minicurso é ilustrado na Tabela 9.1. Vale destacar que, devido ao momento vigente (2020), a apresentação e prática será realizada na modalidade a distância.

**Tabela 9.1. Plano do Minicurso**

Item	Descrição
Público Alvo	Público em geral, pesquisadores, acadêmicos de graduação e pós-graduação
Requisitos para participar	Conhecimentos básicos de metodologia científica e de informática básica
Carga Horária	A definir
Método didático	Aulas expositivas onlines e prática com softwares
Período	10 de Setembro
Bibliografia	[Kitchenham 2004a, Keele et al. 2007, Braga et al. 2016, da Silva et al. 2020]

## 9.2. Pesquisa Bibliográfica

O passo inicial do processo de investigação é a Pesquisa Bibliográfica (PB). PB é o levantamento de toda a bibliografia já publicada, em forma de livros, revistas, publicações avulsas e impensas escritas [Lakatos and Marconi 1991]. A condução desta etapa produz os referenciais teóricos, que são os trabalhos relacionados ou o estado da arte sobre um determinado tema [da Silva et al. 2020].

A PB é fundamental para encontrar lacunas, as quais são oportunidades de novas investigações. Além disso, é possível avaliar o que já foi estudado exhaustivamente e que, talvez, o tema esteja saturado.

Diante disso, há a alternativa de realizar um levantamento de trabalhos relacionados de modo organizado, sistematizado e replicável. Esse mecanismo é denominado de Pesquisa Bibliográfica Sistemática (PBS). Ela contribui para aumento do domínio do conhecimento do pesquisador. Além disso, esse método favorece como ponto inicial de análise para possíveis novas contribuições. Na subseção 9.2.1 a seguir serão apresentados os tipos de pesquisas bibliográficas.

Pesquisas publicadas sobre PB são chamados de **estudos secundários**. Os **estudos primários** apresentam relatos de descobertas científicas como experimentos ou estudos de casos. Uma pesquisa ou estudo secundário, portanto, realiza a revisão de vários estudos primários relevantes com objetivo de sintetizar as evidências de um determinado tema. Há, ainda, o **estudo terciário**, que realiza uma revisão dos artigos secundário relacionados.

### 9.2.1. Tipos de Pesquisas Bibliográficas

As pesquisas bibliográficas são classificadas em:

- **Tradicional:** são pesquisas que não seguem um processo sistemático definido. A experiência do pesquisador guia o processo de pesquisa. Diante disso, não há garantia de resultados imparciais: o pesquisador pode enfatizar somente os dados que concordam com as hipóteses apoiadas por ele;
- **Bibliométrica:** são PBs que usam métodos estatísticos e matemáticos para a obtenção de informações quantitativas sobre livros e artigos. Identifica e quantifica aspectos da produção bibliográfica em uma determinada área. Nesse tipo de PB, a relevância de um artigo ou a influência de um pesquisador a partir do número de citações feitas a esse artigo contribuem para a coleta dos dados;
- **Revisão Sistemática da Literatura (RSL):** em inglês, *Systematic Literature Review* (SLR)) este tipo de PB sistemática que objetiva identificar e avaliar de forma confiável, imparcial e replicável os trabalhos relevantes sobre determinado assunto. Assim, a revisão tem o propósito de identificar e entregar todas as pesquisas relevantes à linha de pesquisa, e não somente aquelas que corroborem para a hipótese apoiada pelo pesquisador. RSL também pode ser denominada de *survey*, *review*, levantamento ou panorama;
- **Mapeamento Sistemático da Literatura (MSL):** em inglês, *Systematic Literature Mapping* (SLM) é uma PB que faz visão geral da área de estudo, quantificando os resultados. Assim, é possível identificar tendências e lacunas de pesquisas a serem estudadas [da Silva et al. 2020]. Esse tipo pode ser identificado por outros nomes, como mapeamento sistemático de estudos, *overview* ou meta-análise;
- **Revisão Terciária:** são revisões que produzem levantamento de revisões secundárias já publicadas sobre determinado assunto. Pode ser uma Revisão Sistemática ou Mapeamento Sistemático.

Vale destacar que a revisão sistemática possui maior nível de evidência científica, tanto no contexto de validade como de confiabilidade [Bork 2005]. Isso corrobora sobre a importância da construção de estudos de pesquisa bibliográfica sistematizada, como os RSL e MSL.

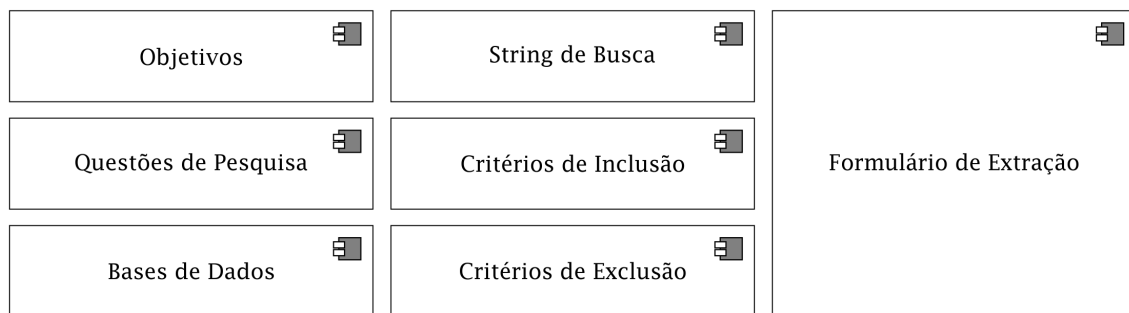
### 9.3. Mapeamento Sistemático da Literatura

A construção do levantamento bibliográfico proposto neste capítulo utilizou-se dos procedimentos de Mapeamento Sistemático da Literatura de [Kitchenham 2004b, Petersen et al. 2008]. Um MSL é definido como um conjunto de procedimentos com o objetivo de identificar, interpretar e sumarizar os trabalhos relevantes para determinada linha de pesquisa, área ou fenômeno de interesse de forma não tendenciosa e replicável [Kitchenham 2004b].

As justificativas para escolha do MSL (ou SLM) são:

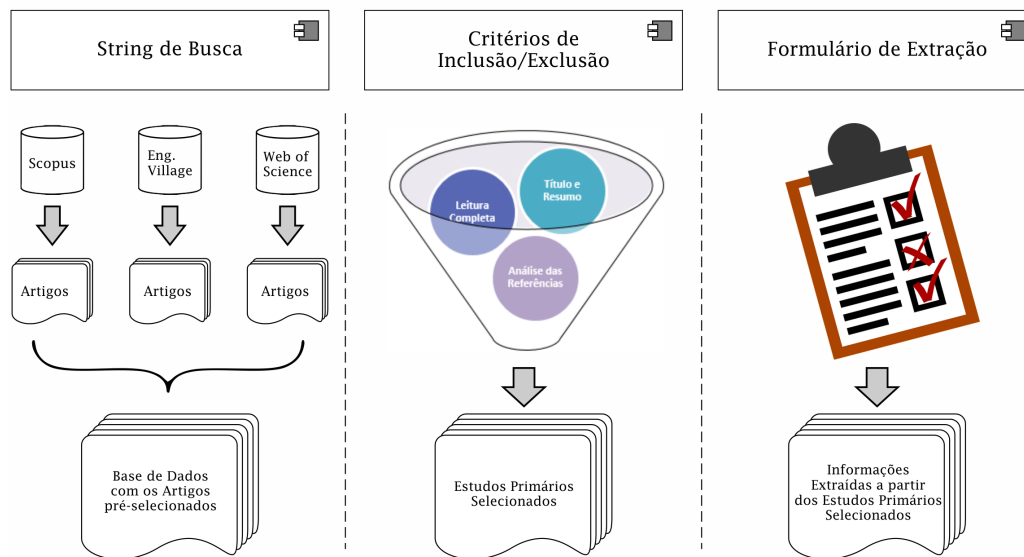
1. Organizar e sintetizar todas evidências empírica sobre determinado tratamento ou tecnologia;
2. Identificar lacunas na área;
3. Apresentar uma visão geral de determinado tema de pesquisa [Kitchenham 2004b].

Segundo [Kitchenham 2004b] as etapas de construção de um MSL são: **Planejamento, Execução e Sumarização (resultados)**. No planejamento, os pesquisadores elaboram o protocolo de execução do estudo. Nesse sentido, são definidos os i) objetivos da pesquisa; ii) as questões de pesquisa; iii) as bases de dados que serão utilizadas; iv) as strings de busca a serem aplicadas nas bases de dados; v) os critérios para inclusão e exclusão dos trabalhos obtidos e vi) os formulários para extração de informações relevantes, dentre outros aspectos. A Figura 9.1 apresenta os componentes necessários para o planejamento citados anteriormente.



**Figura 9.1. Planejamento do MSL**

A execução, representada na Figura 9.2, são aplicados os processos ou itens planejados anteriormente. Inicialmente, o pesquisador aplica as *strings* de busca nas bases de dados para obter o conjunto de estudos iniciais. Em seguida, é necessário aplicar os critérios de inclusão/exclusão. Isso é feito realizando a leitura do título e do resumo dos trabalhos obtidos na busca. O resultado são os artigos selecionados. O passo seguinte é leitura completa dos trabalhos selecionados para extração de dados relevantes.



**Figura 9.2. Execução do MSL**

Por último, na terceira etapa, realiza-se a sumarização dos resultados. As informações são organizadas em tabelas e constrói-se os gráficos para facilitar o entendimento do tema pesquisado. Nesta etapa, ocorre a discussão sobre os resultados obtidos.

Nas próximas subseções deste capítulo serão detalhadas as etapas de um MSL. Para efeito didático, será utilizado um exemplo prático para facilitar o entendimento. A proposta de um MSL deste capítulo é sobre **mineração de dados** e **autismo**. Ou seja, que pesquisas tratam sobre práticas de mineração de dados em prol do autismo.

### 9.3.1. Planejamento

O Planejamento é a etapa de elaboração do protocolo que guiará o Mapeamento Sistemático. Os elementos que compõe esse plano são fundamentais para desenvolver a pesquisa de modo sistemático. Como citado anteriormente os elementos são: os objetivos, *strings* de busca, bases de dados, questões de pesquisa, critérios de inclusão e exclusão e os formulários de extração.

O primeiro passo é elaborar os **objetivos** da pesquisa. A partir destes, as **questões de pesquisa** e as **strings de buscas** são definidas. Em seguida, as **bases de dados** relevantes ao tema. No planejamento também são construídos os **Critérios de seleção** dos trabalhos, que são objetivos, de inclusão e de exclusão. Por fim, criam-se os **formulários** de extração alinhados com os objetivos da pesquisa.

Vale destacar que esse protocolo deve ser elaborado de forma que a PB fique sistemática, objetiva e replicável. Para isso, é importante realizar testes preliminares à medida em que os elementos são elaborados. Por exemplo, ao pensar nas strings de buscas, é interessante fazer testes nas bases de dados e avaliar se os trabalhos encontrados estão condizentes com o propósito da pesquisa. Além disso, caso seja necessário alterar algum elemento deste protocolo pode tornar necessário realizar novamente toda a PB ou mesmo compromete-la.

### 9.3.1.1. Questões de pesquisa

A metodologia de [Kitchenham 2004b], ressalta sobre a necessidade da existência de uma Questão Principal(QP), a qual está relacionada com o objetivo do mapeamento. E para alcançar os objetivos da pesquisa, faz-se uso de Questões Secundárias (QSs). Essas QS's também podem categorizar os trabalhos conforme se deseja. Para a prática apresentada neste capítulo as questões estão descritas a seguir:

- **QP** – Como a *data mining* tem sido usada para apoiar causa autista?
- **QS1** – Qual técnica de *data mining* mais usadas?
- **QS2** – Qual finalidade a *data mining*?
- **QS3** – Quais as áreas tem mais *data mining* com autismo?

### 9.3.1.2. Strings de Busca

As *strings* são as próprias palavras-chaves do tema da pesquisa. Deve-se elaborar essas *Strings* com base nas técnicas ou métodos da pesquisa, no conceito ideia ou área de aplicação. A Figura 9.2 demonstra a composição de strings de busca para o exemplo deste curso.

**Tabela 9.2. Strings de busca**

<b>Grupo</b>	<b>Palavra-chave</b>
Técnica ou método	<i>data mining</i>
Conceito ou idéia	<i>autism, autistic</i>

### 9.3.1.3. Base de dados

Deve-se escolher bases de dados ou fontes de buscas de trabalhos científicos condizentes com o tema do MSL. Geralmente, a escolha por essas plataformas é justificada pelo objetivo de coletar o máximo possível de trabalhos com abrangência internacional.

Vale destacar que essas bases de buscas são indexadoras de outras bases de dados. Para este caso exemplo, deve-se escolher bases que possuem trabalhos da área de computação (Exemplo: *Scopus*, IEEE entre outras) e também pode-se utiliza bases da área da saúde (APA, PubMed). Para simplificar a demonstração será utilizado apenas *Scopus*.

### 9.3.1.4. Critérios de Seleção

Os critérios de seleção visam eliminar artigos que não tem relevância ao tema ou estão fora do escopo desejado. Esses critério são: i) **Objetivos** - para restringir com base em determinados parâmetros, como ano, disponibilidade do trabalho (gratuidade), duplicidade

entre outros; ii) **Exclusão**: visa retirar artigos que não condiz com um tema específico e iii) **Inclusão**: todos os artigos selecionados no mapeamento deve estar dentro desse critério. A Tabela 9.3 apresenta os critérios do exemplo abordado neste capítulo.

**Tabela 9.3. Critérios de Seleção**

<b>Critérios Objetivos</b>	
CO1	Publicação no período de 2019 a 2020
CO2	Artigos em inglês
CO3	Artigos de acesso gratuito e na íntegra
CO4	Artigos publicados em periódicos ou eventos
CO5	Artigos não duplicados
<b>Critérios de Exclusão</b>	
CE1	Artigos que não propõe uso de <i>data mining</i>
CE2	Artigos que não estão relacionados com autismo
CE3	Artigos incompletos (menor que 4 páginas)
CE4	Artigos que não são primários
<b>Critérios de Inclusão</b>	
CI1	Artigos que apresentem indícios do uso de <i>data mining</i> com autismo

### 9.3.1.5. Formulário de Extração

O formulário é, de fato, as questões da pesquisa, especificamente as questões secundárias, ilustradas anteriormente. Acrescenta-se, neste item, as opções de respostas para as questões. Vale ressaltar que é importante que as questões sejam objetivas, como apresentadas na Tabela 9.4 a seguir.

**Tabela 9.4. Formulários de extração**

<b>QS1. Qual técnica de <i>data mining</i> mais usadas?</b>	
Opções:	Estatística, Visualização, Árvore de decisão, Regras de associação, Redes neurais ou Classificação
<b>QS2. Qual finalidade a <i>data mining</i>?</b>	
Opções:	Planejamento, análises, tomada de decisão, outra
<b>QS3. Quais as áreas tem mais <i>data mining</i> com autismo?</b>	
Opções:	Saúde, educação, pesquisa, governo ou outra

A etapa de elaboração do formulário para extração dos dados requer uma análise previa de alguns artigos, um experimento preliminar. Isso evita de incluir respostas que não tenha nas opções ou excessivamente a opção outra.

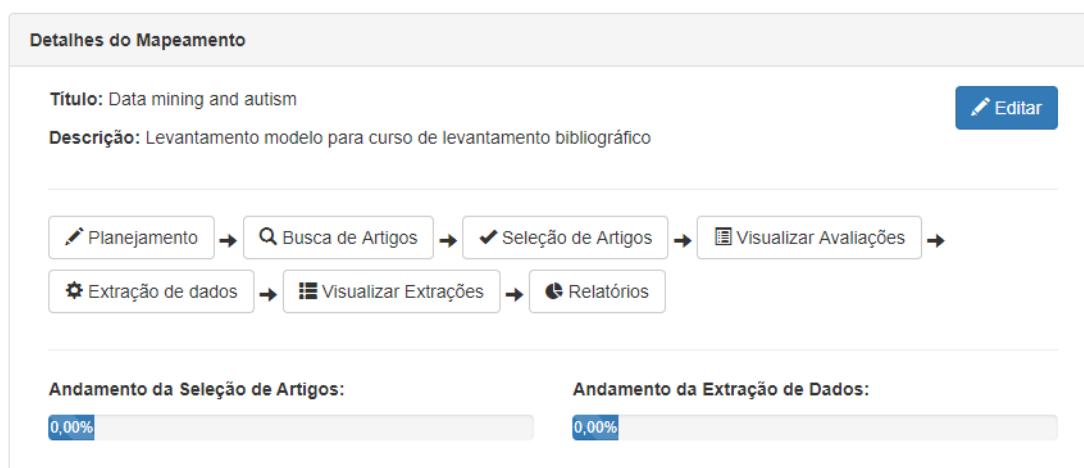
### 9.3.2. Execução

Nesta etapa é posto em prática o que foi planejado na etapa anterior. Como já citado, aplicaram-se as strings de busca nas quatro bases de dados. O resultado da busca foi um



conjunto inicial de artigos pré-selecionados. Esse número de trabalhos pode ser alto, pois não passou por nenhum critério de seleção.

O próximo passo é a seleção desses artigos por meio dos critérios predefinidos. Em seguida, faz-se a extração dos dados desses artigos selecionados. Essa extração é realizada por meio da resolução das questões do formulário de extração. Para otimizar toda esse processo, utilizou-se um *software* para MSL: A *TheEnd* [Braga et al. 2016]. Esse ambiente permite realizar todas as etapas e processos de MSL. O diferencial desta ferramenta é que o mapeamento pode ser feito de modo colaborativo com vários pesquisadores. A Figura 9.3 a seguir ilustra os botões acesso a todas as etapas do mapeamento da ferramenta.



**Figura 9.3. Tela inicial de acesso as etapas do MSL na *TheEnd* [Braga et al. 2016]**

Por fim, após a extração, os dados produzidos podem ser sumarizados e gerar resultados, que incluem: categorizar as informações, encontrar *gaps*(ou lacunas), tendências e outras. Isso é feito por meio de tabelas e gráficos gerados na etapa de sumarização.

Ressalta-se que os procedimentos do exemplo prático deste capítulo serão apresentados na Seção 9.4.

### 9.3.3. Sumarização

A sumarização ou o resultado é o momento em que constrói-se os gráficos e as tabelas para ilustrar o resultado do mapeamento. Demonstra-se quantitativamente uma visão geral das informações e permite visualizar as respostas para as questões da pesquisa. Nesse âmbito, os resultados apresentam as tendências, classificação de informações e lacunas dentro do tema da pesquisa.

Dado o embasamento teórico sobre revisão bibliográfica e MSL, a próxima etapa do curso é apresentar o detalhamento da prática, dentro do objetivo do curso. Na Seção 9.4 será apresentado a prática do MSL com uso da ferramenta *TheEnd*.

## 9.4. Software de MSL: *TheEnd*

A ferramenta *TheEnd* é um software *Web*, colaborativo e disponível gratuitamente no endereço: <https://easii.ufpi.br/theend/home/login>. Ela foi desenvolvida pelo Laboratório de Engenharia de Software e Informática Industrial - EaSII - do Programa de Pós Graduação em Ciências da Computação. Por meio dessa ferramenta, é possível construir e gerenciar Mapeamentos Sistemáticos.

Diante desse recurso, pretende-se construir o levantamento bibliográfico com as informações coletadas do mapeamento. Nos próximos passos a seguir, objetiva-se apresentar as principais telas e ferramentas.

### 9.4.1. Acesso e cadastro de um MSL

Como ponto de partida, ao acessar o software *TheEnd*, deve-se cadastrar um usuário com um endereço de e-mail. De posse das credenciais de acesso, o pesquisador pode cadastrar um mapeamento. A Figura 9.4 apresenta a tela de criação de um novo MSL.

The screenshot displays the 'Criar Mapeamento' (Create Mapping) interface. On the left is a sidebar with a 'Dashboard' menu and options for 'Home', 'Listar Mapeamentos', 'Novo Mapeamento', and 'Conta'. The main area has a breadcrumb trail 'Home / Lista de Mapeamentos / Criar Mapeamento' and a title 'Criar Mapeamento'. The form includes a 'Titulo' field containing 'Data mining and autism' and a 'Descrição' field containing 'Levantamento modelo para curso de levantamento bibliográfico'. A 'Criar Mapeamento' button is positioned at the bottom right of the form.

Figura 9.4. Tela que ilustra a criação de um novo do MSL na *TheEnd* [Braga et al. 2016]

Logo após, o pesquisador terá acesso a tela principal do seu MSL, apresentada na Figura 9.5. A imagem apresenta destacado 5 partes numeradas: A área 1 ilustra os botões de acesso as etapas do MSL; na parte 2 e 3 apresentam o andamento da seleção e da extração, respectivamente; na área destacada 4 mostra que é a opção de adicionar mais participantes (pesquisadores) no MSL; e na parte 5 destaca qual é o tipo de participante cadastrado.

Figura 9.5. Tela principal de uma mapeamento na *TheEnd* [Braga et al. 2016]

Dando um enfoque na Figura 9.5, detalha-se cada etapa do MSL no sistema *TheEnd*. Visando facilitar o entendimento, numerou-se cada botão da etapa do MSL na Figura 9.6 a seguir. Essa figura será referência para entendimento do sistema.

Figura 9.6. Botões para acessar as etapas do MSL na *TheEnd* [Braga et al. 2016]

#### 9.4.2. Planejamento do MSL na prática

Conforme já exposto neste capítulo, o planejamento é a primeira etapa do MSL. Isso é feito, claro, acessando o item 1, identificado na Figura 9.6. Da mesma forma que é proposto por [Kitchenham 2004b], o protocolo é elaborado nessa etapa. A Figura 9.7 apresenta por meio de guias as opções para realizar um cadastro simples dos itens planejados que compõem o planejamento: i) objetivos; ii) questões de pesquisa; ii) strings de buscas; iv) os critérios e o v) formulário. Ressalta-se que as informações desses itens já foram apresentadas na subseção 9.3.1.

**Figura 9.7. Tela de planejamento do MSL na *TheEnd* [Braga et al. 2016]**

De todos os itens, vale destacar e ilustrar o cadastro do formulário de extração. A Figura 9.8 apresenta um exemplo de questão de pesquisa cadastrada. Como já explicado na Subseção 9.3.1.5, o formulário é formado a partir das questões de pesquisa com as opções de respostas para tornar o resultado mais objetivo.

**Figura 9.8. Formulário de extração de dados do MSL na *TheEnd* [Braga et al. 2016]**

### 9.4.3. Execução do MSL na Prática

A execução do protocolo planejado inicia-se com a busca dos artigos utilizado as strings de buscas aplicando-as nas bases de dados escolhidas. Como já informado, a base de dados escolhida para essa demonstração é a *Scopus*. A Figura 9.9 apresenta uma busca preliminar na referida base de dados utilizado as strings: "*data mining*" AND "*autism*".

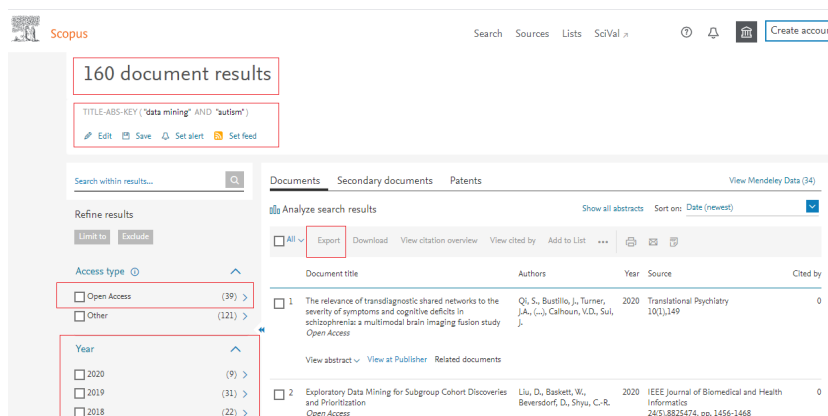


Figura 9.9. Busca de artigos no Scopus

A busca inicial encontrou 160 trabalhos. O próximo passo é aplicar os critérios de seleção, já expostos na Subseção 9.3.1.4. Primeiro aplica-se os critérios objetivos (CO1, CO2, CO3 e CO4). Esses critérios são feitos ainda na plataforma da Scopus. Depois disso, ficaram apenas 10 trabalhos.

Os dados desses artigos devem ser exportados no formato *.bib* para inseri-los na ferramenta *TheEnd*. Para tanto, deve-se utilizar a opção de *Export* do Scopus, ilustrado na Figura 9.10 a seguir.

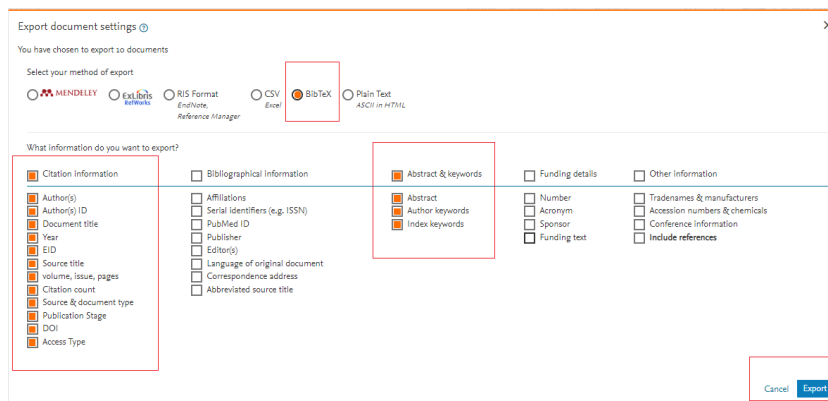


Figura 9.10. Exportação de artigos no formato *BibTex* no Scopus

O formato *.bib*, ou *BibTex* para o  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  foi criado para facilitar a separação da bibliografia e uso no texto. Esse formato segue o mesmo conceito da distinção do conteúdo com o estilo do texto utilizada no próprio LaTeX, XHTML, CSS e outros. A Figura 9.11 apresenta um exemplo de conteúdo de um arquivo *.bib*. Esse arquivo contém informações dos artigos selecionados (título, ano, resumo, autores etc).

```
scopus (2).bib x
Scopus
EXPORT DATE: 27 July 2020
@ARTICLE{Qi2020,
author={Qi, S. and Bustillo, J. and Turner, J.A. and Jiang, R. and
and Yang, X. and Stevens, M. and Zhuo, C. and Xu, Y. and Calhoun, V
title={The relevance of transdiagnostic shared networks to the severe
multimodal brain imaging fusion study},
journal={Translational Psychiatry},
year={2020},
volume={10},
number={1},
doi={10.1038/s41398-020-0834-6},
art_number={149},
note={cited By 0},
url={https://www.scopus.com/inward/
```

Figura 9.11. Recorte de um trecho do arquivo .bib utilizado neste MSL

Dado o arquivo *BibTex*, o passo seguinte é carregá-lo na ferramenta *TheEnd*. Para isso, deve-se acessar o botão **2, Busca de Artigos**, destacado na Figura 9.6. Nesta opção de busca de artigos, é possível inserir informações sobre os artigos de forma manual, ou seja, um de cada vez, ou em conjunto, por meio de arquivo BibTex. A Figura 9.12 destaca essas opções.

Figura 9.12. Tela de busca de artigos: i) por meio de arquivo *BibTex* ou ii) manualmente

Após carregar os dados dos artigos, na mesma tela é possível *Refinar artigos*. Essa propriedade possibilita remover artigos duplicados ou semelhantes, de mesmo autor, por exemplo.

O próximo passo da execução é selecionar os artigos. Nessa opção é possível por meio do botão *Seleção de artigos*. Ao acessar a tela de seleção, o pesquisador utilizará os critérios de exclusão e de inclusão escolher os artigos almejados, conforme os critérios. Conforme exposto, a escolha se baseia na leitura do título e resumo [Kitchenham 2004b]. A Figura 9.13 apresenta um exemplo com um dado artigo.

**Artigo - 193696**

**Título:** Consecrate recurrent neural network classifier for autism classification

**Abstract:** Most recent discoveries in Autism Spectrum Disorder (ASD) detection and classification studies reveal that there is a substantial relationship between Autism disorders and gene sequences. This work is indented to classify the autism spectrum disorder groups and sub-groups based on the gene sequences. The gene sequences are large data and perplexed for handling with conventional data mining or classification procedures. The Consecrate Recurrent Neural Network Classifier for Autism Classification (CRNNC-AC) work is introduced in this work to classify autism disorders using gene sequence data. A dedicated Elman [1] type Recurrent Neural Network (RNN) is introduced along with a legacy Long Short-Term Memory (LSTM) [2] in this classifier. The LSTM model is contrived to achieve memory optimization to eliminate memory overflows without affecting the classification accuracy. The classification quality metrics [3] such as Accuracy, Sensitivity, Specificity and F1-Score are concerned for optimization. The processing time of the proposed method is also measured to evaluate the pertinency. ? BEIESP.

**Palavras-chave:** Autism spectrum disorder classification; Elman Network; Gene sequence-based autism disorder detection; Long Short-Term Memory; Recurrent Neural Network

**Autor:** Padmapriya, S. and Murugan, S.

**Fonte:** Scopus

**Ano:** 2019

**Tipo de documento:** Article

---

**Critérios de Inclusão:**

C1: Artigos que apresentem indícios do uso de data mining com autismo

**Comentários:**

**Incluir**

**Critérios de Exclusão:**

CE1: Artigos que não propõe uso de data mining

CE2: Artigos que não estão relacionados com autismo

CE3: Artigos incompletos (menor que 4 páginas)

CE4: Artigos que não são primários

**Comentários:**

**Rejeitar**

**Figura 9.13. Tela ilustrando os critérios de inclusão e exclusão para um dado artigo**

Ao finalizar a seleção, o pesquisa passa para o item **4** (da Figura 9.6) - Visualizar Avaliações. Nesta opção, os pesquisadores podem comparar as decisões de seleção (no botão *Comparação das Avaliações*) entre si (caso haja mais de um pesquisador). Assim, se houve discordância, podem decidir sobre qual optar em reunião. Para esse MSL, ao final da seleção, restou apenas **8 artigos** selecionados para etapa de extração.

### Detalhes do Mapeamento

Título: Data mining and autism  
Descrição: Levantamento modelo para curso de levantamento bibliográfico  
Andamento da Avaliação: 100,00%

---

Comparação das Avaliações: [Comparação das Avaliações](#)

Exportar Resultado:

Meus Aceites  .xls

---

### Resultados das Avaliações

Quantidade de Artigos Avaliados: 10  
Quantidade de Artigos Aprovados: 8  
Quantidade de Artigos Rejeitados: 2  
Quantidade de Artigos a serem Avaliados: 0

---

Quantidade de Artigos Filtrados: 0  
Quantidade de Artigos Repetidos: 0  
Quantidade de Artigos sem Termos (Regex): 0  
Quantidade de Artigos sem Autores: 0  
Quantidade de Artigos sem Abstracts: 0

**Figura 9.14. Tela com os resultados das avaliações e a opção para comparar as avaliações**

O próximo passo é a extração dos dados. Para tal, deve-se responder as questões do formulário de extração. A Figura 9.15 apresenta como estão dispostas as questões. Observa-se que são questões objetivas com múltiplas escolhas.



have heuristic value and advocate specific approaches to refine available treatment strategies for comorbid conditions in schizophrenia.  
? 2020, The Author(s).

Palavras-chave:

Formulário de Extração

Realizar Extração de Dados

QS1. Qual técnica de data mining é aplicada?:

- Arvore de decisão
- Classificação
- Estatística
- Redes neurais
- Regras de associação
- Visualização

**Figura 9.15.** Tela ilustrando as questões do formulário de extração

Com mesmo propósito da seleção, na extração também é realizada a comparação de avaliações de extração, acessando o botão destacado com número 6, na Figura 9.6.

#### 9.4.4. Sumarização do MSL na prática

Por fim, tem-se a etapa de construir gráficos e tabelas para apresentar os resultados. A planilha geral com os resultados consolidados é extraído dentro da opção *Visualizar Extrações*, demonstrado na Figura 9.16, na opção *Exportar Extrações: CSV*. De posse dessa planilha, pretende-se utiliza-la no *software Google Sheets* para gerenciar mudança, trabalhar de modo colaborativo e construir gráficos adicionais que a *TheEnd* não gera.

Detalhes do Mapeamento

Título: Data mining and autism  
Descrição: Levantamento modelo para curso de levantamento bibliográfico  
Andamento da Avaliação: 100.0%

Comparação das Extrações: Comparação das Avaliações

Análise de concordância: Aplicar

Exportar Extrações: CSV (Minhas Extrações) CSV (Todas Extrações)

Extrações do usuário

QS2. Nesta pesquisa, data mining é para qual finalidade? ▼ ✕

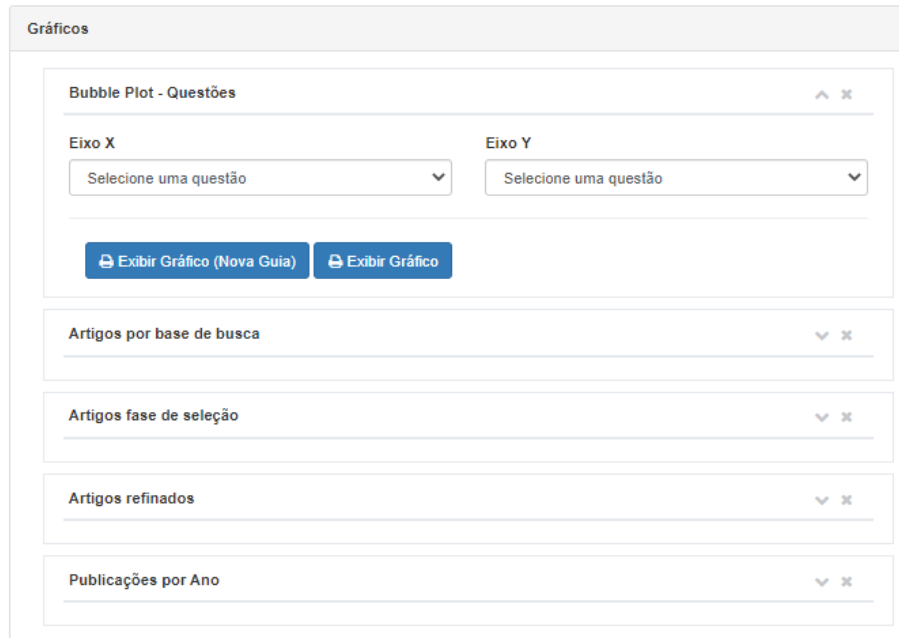
QS3. Qual a área é aplicado data mining com autismo, neste artigo? ▼ ✕

QS1. Qual técnica de data mining é aplicada? ▼ ✕

**Figura 9.16.** Tela apresenta a comparação das extrações, análise de concordância e exportar extrações

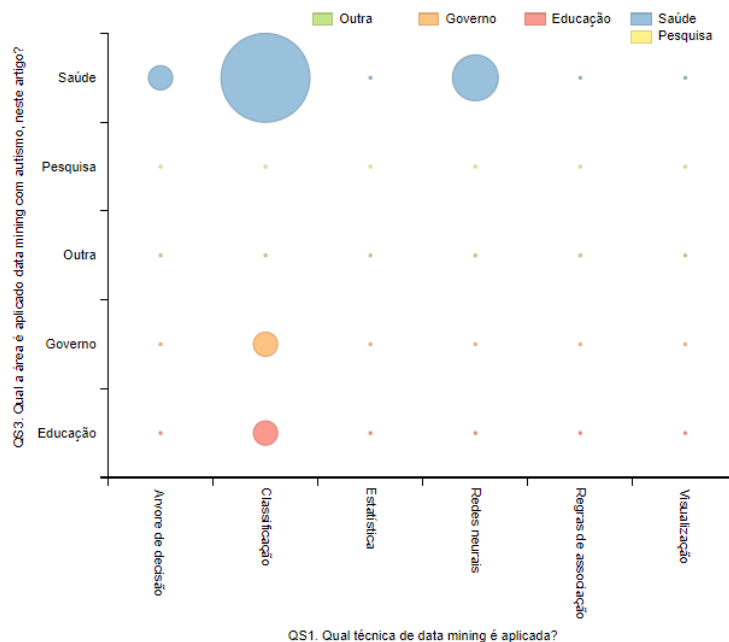
Já os gráficos podem são produzidos de maneira automática na própria ferramenta *TheEnd*, no botão *Relatórios*, destacado com número 7, da Figura 9.6. Observa-se, por

meio da Figura 9.17, que é possível construir diversos gráficos na *TheEnd*, de modo automático.



**Figura 9.17. Opções de gráficos da *TheEnd***

Dentre as opções de gráfico, o *Bubble Plot - Questões*, ou gráfico em bolha. Por meio desse, é possível avaliar a associação entre respostas das questões de pesquisa. Nesse cenário, é possível mapear os resultados do MSL. A Figura 9.18 apresenta um gráfico em bolha que cruza os dados da área com a técnica de *data mining*.



**Figura 9.18. Exemplo de gráfico em bolha extraído da *TheEnd***

Com base na Figura 9.18, pode-se visualizar, por exemplo, que não nenhuma pesquisa de educação com uso de redes neurais, regras de associação e visualização. Isso sugere prováveis lacunas e oportunidades de pesquisas. Por outro lado, há considerável quantidade de pesquisas que envolvem a área da saúde e técnicas de classificação. Infere-se que é uma opção que vem sendo bem explorada.

## 9.5. Considerações Finais

Pretende-se, com este capítulo, que o estudante construa um levantamento bibliográfico por meio do Mapeamento Sistemático da Literatura de maneira objetiva. Para tanto, ele contará com suporte da ferramenta *TheEnd* como alternativa que permite elaborar um trabalho sistemático, colaborativo e replicável.

Os resultados do levantamento produzido por meio do método deste curso, possibilita encontrar lacunas e tendências de pesquisa em um determinado nicho de pesquisa. Com isso há um potencial de publicação desses resultados em meios científicos (congressos e revistas).

Portanto, o participante do curso estará apto a elaborar referencial teórico consistente e organizado.

## Referências

- [Bork 2005] Bork, A. M. T. (2005). *Enfermagem baseada em evidências*. Guanabara Koogan.
- [Braga et al. 2016] Braga, R., Oliveira, P. A., Souza, M., Neto, P. S., Rabêlo, R., and Britto, R. (2016). Estudo prático sobre mapeamento sistemático da literatura utilizando

a ferramenta theend. *Anais Escola Regional de Informática do Piauí*, 1(1):77–91.

- [da Silva et al. 2020] da Silva, M. D., Soares, A. C. B., and Moura, I. C. (2020). Aplicação de ferramentas computacionais para o desenvolvimento do ensino de crianças com autismo: um mapeamento sistemático da literatura. *Revista Brasileira de Informática na Educação*, 27(03):351.
- [Detroz et al. 2015] Detroz, J. P., Hinz, M., and da Silva Hounsell, M. (2015). Uso de pesquisa bibliográfica em informática na educação: um mapeamento sistemático. *Revista Brasileira de Informática na Educação*, 23(01):28.
- [Keele et al. 2007] Keele, S. et al. (2007). Guidelines for performing systematic literature reviews in software engineering. Technical report, Technical report, Ver. 2.3 EBSE Technical Report. EBSE.
- [Kitchenham 2004a] Kitchenham, B. (2004a). Procedures for performing systematic reviews. Keele university. technical report tr/se-0401, Department of Computer Science, Keele University, UK, 2009-04-28T08:33:13.000+0200.
- [Kitchenham 2004b] Kitchenham, B. (2004b). Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33(2004):1–26.
- [Lakatos and Marconi 1991] Lakatos, E. M. and Marconi, M. d. A. (1991). Metodologia científica. 2ª edição. *São Paulo: Atlas*.
- [Petersen et al. 2008] Petersen, K., Feldt, R., Mujtaba, S., and Mattsson, M. (2008). Systematic mapping studies in software engineering. In *12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12*, pages 1–10.