

## Proposta de catalogação de modelos de Machine Learning usando BIBFRAME e RDA

Marcos Antonio Siqueira

[marcos.a.siqueira@unesp.br](mailto:marcos.a.siqueira@unesp.br)

### 1. INTRODUÇÃO

Os modelos de Machine Learning, ML, como a tradução indica, é uma subárea da Inteligência Artificial, IA, que permite que as máquinas aprendam, por exemplo, a fazer o diagnóstico de doenças sem a intervenção de um médico. Esse uso de ML já é uma realidade que destaca a capacidade de modelos de ML aprenderem e os impactos sociais que a IA já está causando na sociedade.

Para esclarecer a relevância das questões envolvidas com o desenvolvimento e uso de modelos de ML analise o cenário hospitalar e pense em um médico ou médica que está prestes a ser contratado por um hospital. Com certeza é esperado que a apresentação do diploma e histórico escolar deste profissional seja um requisito indispensável. Não é possível imaginar a possibilidade de um médico ou médica ser contratado sem a apresentação desses documentos básicos. É interessante pensar que um diploma e um histórico escolar detalham todo o ‘treinamento’ que esse médico ou médica recebeu ao longo da graduação. O nome da instituição de ensino vai dizer muito sobre a qualidade do ‘treinamento’ recebido, mas em especial o histórico escolar detalha como esse ‘treinamento’ foi realizado e até como aquele médico, enquanto aluno, se saiu visto que as notas estão disponíveis. E quando um modelo de ML é adquirido pelo hospital? É solicitado seu ‘diploma’ e ‘histórico escolar’? Existe um documento padrão que permita ter uma ideia de como o modelo de ML foi treinado?

Os modelos de ML são, por assim dizer, ‘treinados’ assim como um aluno de medicina é ‘treinado’ ao longo da graduação. A seguir será apresentado de forma resumida como é realizado o treinamento de um modelo de ML, porém, nesse momento, é possível perceber porque a tese de que é necessário a catalogação de modelos de ML usando BIBFRAME e RDA é justificável. A tese tem um interesse especial em catalogar os dados referentes ao processo de treinamento a que o modelo de ML foi submetido. É muito mais do que catalogar um software comum representando título, autor, versão, URI, palavras-chave, por exemplo. O

foco é catalogar, por assim dizer, o ‘histórico escolar’ do modelo do ML usando padrões internacionalmente aceitos e respeitados para que aqueles que venham a se interessar em usar o modelo de ML tenham todas as informações relevantes que determinam o quanto este é confiável.

Entender o processo de desenvolvimento de um modelo de ML é fundamental para compreender a importância da catalogação de modelos de ML usando BIBFRAME e RDA. É possível resumir as etapas de desenvolvimento de um modelo de ML em quatro: 1) Coleta e preparação dos dados usados no treinamento, 2) escolha do algoritmo de ML apropriado, 3) treinamento do modelo e finalmente 4) validação e talvez reinício do processo. Para detalhar essas 4 etapas, será usado neste momento um exemplo hipotético do desenvolvimento de um modelo de ML capaz de diagnosticar uma doença específica. É necessário esclarecer que o treinamento de um modelo de ML é muito específico, ou seja, o modelo de ML treinado para diagnosticar a doença X não será capaz de diagnosticar a doença Y. Em IA e ML os projetos são bem específicos.

No exemplo hipotético de desenvolvimento de um modelo de ML capaz de diagnosticar uma doença específica usado nesta introdução, a etapa 1) Coleta e preparação dos dados usados no treinamento, envolveria, coletar os dados de exames de, hipoteticamente, 5 milhões de pessoas que tiveram ou não esse tipo específico de doença. Na prática uma planilha com dezenas ou até centenas de colunas com 5 milhões de linhas seria criada com casos de pessoas reais. A criação desta planilha depende de um profissional da área da Ciência da Computação, CC, que terá que coletar os dados de diferentes fontes e reunir tudo em uma única planilha. Pensando na proposta desta tese de catalogar o modelo de ML desenvolvido, seria necessário catalogar cada uma dessas colunas indicando os tipos de cada dado e a importância deles, o que na CC, é denominado Dicionário de Dados. Enfatizando, cada projeto de ML terá quantidades de colunas e linhas diferentes com tipos e objetivos de dados diversos, mas que influenciam decisivamente o treinamento do modelo de ML, por isso, a necessidade de incluir na catalogação de modelos de ML os dados sobre essa etapa 1) Coleta e preparação dos dados usados no treinamento.

Na etapa 2) escolha do algoritmo de ML apropriado, diversos algoritmos já prontos estão disponíveis para os profissionais da área da CC. Na maioria dos casos, os programadores ou programadoras de computadores trabalham com uma linguagem de programação, por exemplo, Java ou Python, que já oferecem bibliotecas de algoritmos para

ML. Entre os profissionais de programação, essas bibliotecas são denominadas pacotes. Por exemplo, os programadores que usam a linguagem Python têm o pacote Scikit-Learn à disposição. Esse pacote contém diversos algoritmos já prontos. Portanto, nesta etapa 2) escolha do algoritmo de ML apropriado, o desafio é selecionar o algoritmo mais adequado para o treinamento que se pretende realizar com o modelo de ML em desenvolvimento. Note que algoritmo de ML é uma coisa e modelo de ML é outra, ou seja, um algoritmo de ML é selecionado para facilitar o desenvolvimento de um modelo de ML. Se o pacote Scikit-Learn, por exemplo, não existisse na linguagem de programação Python, isso significaria que os programadores e programadoras de computadores teriam que desenvolver desde o início todas as complexas fórmulas matemáticas necessárias para o projeto tornando o trabalho mais difícil. A tese proposta acredita que os detalhes dessa escolha de algoritmo de ML mais apropriado também devem ser catalogados porque influenciam decisivamente na capacidade final desse modelo de ML aprender o que se espera. Ainda nesta etapa, após escolher o algoritmo, será necessário configurar diversos parâmetros do mesmo. Esses detalhes também são incluídos na catalogação proposta.

Antes de detalhar a etapa 3) treinamento do modelo, é necessário esclarecer que o processo de treinamento de um modelo de ML pode levar muitas horas e até dias. No exemplo hipotético usado nesta introdução para apresentar as etapas de desenvolvimento de um modelo de ML, uma planilha com 5 milhões de linhas e talvez centenas de colunas está disponível para o treinamento e, por isso, milhares de cálculos usando fórmulas complexas da área de matemática, estatística e probabilidade serão necessários para identificar padrões matemáticos contidos nos dados desta planilha. É exatamente esses padrões matemáticos que um modelo de ML aprende, ou seja, nos casos positivos existe um padrão nos números apresentados nos exames dos pacientes que indicam o diagnóstico. O mesmo ocorre para casos negativos. Alguns modelos de ML podem levar dias para serem treinados. Isso na prática significa que o computador usado neste treinamento ficará ligado dia e noite durante alguns dias até esse treinamento terminar e o modelo de ML ser criado para uso posterior. É um bom momento para lembrar que um algoritmo treinado para identificar a doença X não identificará a doença Y. São modelos treinados para aprender coisas bem específicas, portanto, em projetos mais complexos, talvez vários modelos de ML diferentes terão que ser desenvolvidos e treinados. Embora as etapas estejam sendo apresentadas de forma resumida nesta introdução, é possível perceber quantos detalhes

podem influenciar na capacidade de aprendizado de um modelo de ML e, consequentemente, como é justificável questionar até que ponto o modelo de ML desenvolvido seria confiável. O tempo gasto no treinamento, a quantidade de treinamentos realizados e a infraestrutura de hardware usada também são incluídos nesta proposta de catalogação de modelos de ML.

Finalmente chegamos na etapa 4) validação e talvez reinício do processo. No exemplo hipotético usado nesta introdução, 5 milhões de linhas estão na planilha, correspondendo a 5 milhões de pessoas com diagnósticos positivos ou negativos para uma doença específica. Destes 5 milhões de linhas da planilha, 4 milhões foram usados para o treinamento na etapa anterior, por isso, existe 1 milhão de linhas, ou casos reais, que o modelo treinado não conhece. Agora é hora de apresentar esse 1 milhão de casos desconhecidos para o modelo de ML e verificar se este realmente aprendeu e é confiável. Nesta etapa é possível também mensurar com diversas métricas a eficiência do modelo de ML treinado. Por exemplo, é possível constatar que em 97% dos casos o algoritmo aprendeu e, portanto, foi capaz de fazer o diagnóstico correto. São diversas métricas geradas nesta etapa, portanto, a tese proposta inclui catalogá-las também. É comum essas métricas não serem satisfatórias, exigindo que todas as 4 etapas sejam repetidas. Portanto, se esse treinamento fosse repetido 5 vezes num modelo que precisou de 5 dias de treinamento, seriam 25 dias dedicados apenas a etapa 3) treinamento do modelo, sem contar o tempo necessário para coletar e preparar os dados necessários.



Fonte: Adaptado de <https://www.datascienceacademy.com.br/>

Depois de todas as 4 etapas acima serem realizadas um modelo de ML que estudou 5 milhões de casos reais estará treinado e pronto para ser usado no diagnóstico de uma doença específica. Uma consideração válida de ser citada é que aquela planilha com 5 milhões de linhas e talvez centenas de colunas não é mais necessária para o funcionamento do modelo de ML após a conclusão das 4 etapas descritas porque sua função era treinar o modelo. Na prática significa que outros softwares, aplicativos ou sites terão que ser desenvolvidos para que o modelo de ML seja funcional. Essa fase é conhecida como o deploy do modelo. Em outras palavras, por exemplo, um site restrito apenas aos médicos será criado para permitir que os dados dos exames médicos de um novo paciente específico sejam digitados para que o modelo de ML realize o diagnóstico. Outra possibilidade para evitar a digitação manual desses dados e minimizar erros, seria desenvolver um software que coleta os dados gerados após diversos exames realizados diretamente dos sistemas da clínica e automaticamente envia esses para o modelo de ML realizar o diagnóstico.

Depois dessa breve e resumida apresentação das etapas necessárias para o desenvolvimento de um modelo de ML, é possível perceber que este desenvolvimento é trabalhoso, repleto de detalhes significativos e complexos. Portanto, questionar a qualidade desses modelos é algo comum que os profissionais da área da CC devem estar prontos para

esclarecer. Existem cenários em que a catalogação de modelos de ML usando BIBFRAME e RDA poderiam aumentar a confiança nestes modelos. Por exemplo, na relação entre 1) empresas desenvolvedoras de modelos de ML e seus clientes, 2) empresas desenvolvedoras de modelos de ML e suas próprias filiais e equipes de desenvolvimento, 3) sites que disponibilizam modelos de ML já treinados seguindo o modelo open source e desenvolvedores que usam esses sites e 4) casos jurídicos e pessoas que se sentiram prejudicadas por modelos de ML.

Portanto, a tese destaca a necessidade de catalogar não apenas as informações básicas sobre título, autor, versão e URI de um software comum, mas também, todos os detalhes relacionados com as 4 etapas de desenvolvimento de um modelo de ML descritos anteriormente. A hipótese é que os estudos de Alves (2010) podem ser aplicados no desenvolvimento de uma proposta de catalogação de modelos de ML usando BIBFRAME e RDA. A hipótese também inclui usar o modelo conceitual proposto por Coneglian (2020) como prova de conceito para esta tese, pois será possível apresentar a catalogação dos modelos de ML necessários para a implantação da proposta deste autor. A proposição é pesquisar os detalhes técnicos envolvidos no desenvolvimento de um modelo de ML consultando as referências bibliográficas da área da CC contextualizando para a área da Ciência da Informação, CI. Além de permitir detectar todos os detalhes do desenvolvimento de um modelo de ML que, em tese, necessitam ser catalogados, esta etapa do estudo também permitirá trazer para a área da CI detalhes de como os modelos de ML são desenvolvidos para uma possível identificação de outras propostas de pesquisas futuras envolvendo ML. A seguir, a proposição inclui um estudo das disciplinas de Catalogação e Metadados nas referências bibliográficas da área da CI para determinar como o BIBFRAME e RDA podem ser usados para a catalogação desse tipo específico de software denominados modelos de ML.

O cenário desta tese é a grande quantidade de modelos de ML sendo desenvolvidos e pouca informação sobre estes sendo disponibilizados em um padrão que permita a interoperabilidade dos dados. Neste cenário a CI, embasada nas disciplinas de Catalogação e Metadados, surge como ferramenta que auxiliará na resolução do problema. A tese proposta também contribui para afirmar a importância da CI em um cenário em que a IA passa a ser cada vez mais comum.

## **2. OBJETIVOS**

O problema que esta pesquisa se propõe investigar é: Quais são os dados e metadados que precisam ser identificados e catalogados relacionados ao treinamento de modelos de Machine Learning, ML? É possível representar a catalogação referente ao desenvolvimento de modelos de ML usando BIBFRAME e RDA?

### **2.1. OBJETIVO GERAL**

Propor uma forma de catalogação de modelos de Machine Learning que incluam os detalhes relacionados ao desenvolvimento deste modelo e possibilitar a interoperabilidade do registro em BIBFRAME e RDA.

### **2.2. OBJETIVOS ESPECÍFICOS**

São os seguintes: 1) Identificar todos os dados e metadados relacionados ao desenvolvimento de modelos de Machine Learning, ML, e propor uma forma de catalogação desse tipo particular de software; 2) verificar se é possível representar todos os dados catalogados referente ao desenvolvimento de modelos de ML usando o BIBFRAME e RDA; caso não seja possível; 3) Identificar um padrão de metadados existente que possibilite a representação; e, por fim, caso ainda não seja possível; 4) Propor um modelo para um padrão de metadados que atenda a essa necessidade específica.

## **3. PROCEDIMENTOS METODOLÓGICOS**

Pesquisas na área das ciências sociais aplicadas são diversificadas, por isso, é possível usar tipos diversos de pesquisas, conforme Valentim (2005). A autora destaca ainda que o pesquisador pode escolher o tipo que mais se adequa às questões relacionadas ao problema de pesquisa porque existem diversos tipos e possibilidades de pesquisa científica. É também possível que a pesquisa seja esquematizada como pura e aplicada simultaneamente, pois uma abordagem contribui para a outra no desenvolvimento técnico e científico, como defende Gil (2002). Seguem detalhes de como a pesquisa está organizada para esta tese.

### **3.1. NATUREZA DA PESQUISA**

Quanto à sua natureza, a pesquisa caracteriza-se como qualitativa.

### **3.2. TIPO DA PESQUISA**

Quanto ao tipo de pesquisa, esta pode ser definida como descritiva e exploratória porque realiza-se uma revisão bibliográfica visando compreender melhor os temas envolvidos: 1) Metadados no contexto do desenvolvimento de modelos de Machine Learning, ML, e 2) Metadados no contexto da Catalogação; para deste modo, compreender o problema de pesquisa a ser investigado, ou seja, como catalogar modelos de ML usando BIBFRAME e RDA. O tipo de pesquisa descritiva e exploratória permite munir-se de subsídios para a reflexão e discussão do tema de pesquisa proposto visando a elaboração do referencial teórico.

Os autores Cervo e Bervian (2003) destacam que a pesquisa exploratória se caracteriza pela busca de informações relacionados com o assunto estudado, tendo em mente as possíveis facetas do problema, para encontrar uma solução. Pesquisas exploratórias, segundo Köcher (2002, p. 126), possibilitam ainda “[...] um processo de investigação que identifique a natureza do fenômeno e aponte as características essenciais das variáveis que se quer estudar”. Conforme apresentado na tese, hipótese, proposição e objetivos, detalhar os aspectos do objeto de estudo aqui em foco, modelos de ML e seus metadados, é característica da pesquisa descritiva (COSTA, COSTA, 2001).

### **3.3. MÉTODO DE PESQUISA**

Na introdução desta pesquisa foi possível perceber que uma proposta de catalogação de modelos de Machine Learning, ML, usando BIBFRAME e RDA inclui diversos detalhes que devem ser levados em conta. Por exemplo, existem situações condicionais, ou seja, a quantidade de metadados catalogados para um modelo de ML nem sempre serão os mesmos porque, como apresentado, cada modelo pode ter suas particularidades e necessidades. Por isso, o método de pesquisa aplicado para este estudo usa mapas conceituais para apresentar os caminhos e decisões escolhidos para justificar a proposta de catalogação sugerida.

Os mapas conceituais sintetizam, representam e organizam os conceitos relacionados ao conhecimento que está em estudo. Embora possa ser usado em qualquer parte ou momento da pesquisa para representar o conhecimento, optou-se nesta pesquisa, usá-lo para representar os detalhes envolvidos na catalogação de um modelo de ML usando BIBFRAME e RDA. Também é possível incluir parágrafos, antes ou depois de cada mapa



conceitual, detalhando os pontos que se deseja enfatizar. Legendas, cores, tipos de seta, formatos e outros recursos auxiliam na leitura do mapa conceitual.

O passo a passo para a construção dos mapas conceituais desta pesquisa incluem 1) identificar o tema ou a pergunta principal que o mapa representa, neste caso, a catalogação de um modelo de ML usando BIBFRAME e RDA; 2) verificar os conceitos principais: ML, BIBFRAME e RDA; 3) organizar todos conceitos relacionados em uma lista; 4) agrupar os conceitos secundários, palavras ou símbolos que representam todos os metadados identificados nas etapas de desenvolvimento de um modelo ML, que conseqüentemente serão incluídos na catalogação; 5) definir as ligações existentes entre os conceitos; e 6) revisar a estrutura dos mapas conceituais desenvolvidos.

Criado inicialmente para ajudar no aprendizado de crianças, o sucesso dos mapas conceituais neste contexto possibilitou inúmeras outras aplicações para esta ferramenta ao longo dos anos. (NOVAK e CAÑAS, 2006) No contexto da organização e representação do conhecimento Rodrigues e Cervantes (2014) afirmam que um mapa conceitual “tem se demonstrado como opção para diversas atribuições, a começar por sua definição de organizar e representar o conhecimento.” Confirmando a potencialidade dos mapas conceituais em contextos diversos, Carabetta Júnior (2013) aplicou esta ferramenta com alunos de um curso de medicina. O autor afirma que a utilização dos mapas conceituais foi “relevante por se alinhar a uma formação teórica adequada às necessárias intervenções na realidade estudada e por facilitar a apropriação de conceitos científicos pelos alunos”. Por todos os motivos apresentados e com base nos autores acima, esta pesquisa escolheu os mapas conceituais como método de pesquisa.

### **3.4. TÉCNICA DE COLETA DE DADOS**

A Revisão Sistemática da Literatura, RSL, foi usada nesta pesquisa porque todas as conclusões devem estar baseadas na literatura da área da Ciência da Computação, especificamente sobre o desenvolvimento de modelos de Machine Learning e também na literatura da área da Ciência da Informação, focada nas disciplinas de Catalogação e Metadados. Outro motivo para esta escolha, é a possibilidade desta técnica de coleta de dados permitir que o autor da pesquisa expresse sua opinião sobre a literatura consultada (GALVÃO; PEREIRA, 2014). Por fim, a RSL, permite definir os critérios ou limites da revisão.

### **3.5. INSTRUMENTOS DE COLETA DE DADOS**

Esta pesquisa segue a orientação de Okoli (2019) sobre a importância do planejamento do protocolo RSL ser feita antes do início da Revisão Sistemática da Literatura. O autor afirma que o pesquisador deve estar ciente dos procedimentos que seguirá, atento a todos os detalhes, o que exige o detalhamento do protocolo (OKOLI, 2019). Visto que existem categorias de modelos de Machine Learning, ML, por exemplo, baseado em aprendizados supervisionados, não supervisionado e outros, o protocolo RSL desta pesquisa considera este fato. Outro ponto importante que foi considerado no protocolo RSL usado nesta pesquisa são as 4 etapas que o desenvolvimento de um modelo de ML passa. Portanto o protocolo RSL deixa claro a qual categoria e fase de desenvolvimento os dados coletados se referem porque estes detalhes influenciarão na proposta de catalogação sugerida. Em conclusão, o uso do protocolo RSL permitirá que outros pesquisadores repitam os procedimentos que conduziram às conclusões e sugestões desta pesquisa.

### **3.6. TÉCNICAS DE ANÁLISE DE DADOS**

Optou-se por realizar uma triangulação de dados tendo em mente uma análise qualitativa, descritiva e indutiva (pelo pesquisador) usando como instrumento de análise os mapas conceituais. No contexto da gestão da informação e do conhecimento em organizações a triangulação metodológica pode ser utilizada por pesquisadores quando envolve diversas interpretações de dados. A proposta é utilizar perspectivas diversas ao analisar o mesmo conjunto de objetos, que nesta pesquisa envolve metadados e Machine Learning (MORAES e FADEL, 2008).

## **4. CONSIDERAÇÕES FINAIS**

O presente estudo aplica uma pesquisa qualitativa buscando gerar conhecimentos para aplicação prática e uso da sociedade ao considerar que a utilização de modelos de Machine Learning, ML, estão se tornando cada vez mais comuns. Incluindo um caráter exploratório, esta pesquisa apresenta detalhes sobre como modelos de ML são desenvolvidos tornando possível entendê-los melhor, com o objetivo de propor uma catalogação desses modelos usando padrões reconhecidos como o BIBFRAME e RDA.

Visto que os mapas conceituais permitem sintetizar, representar e organizar os conceitos relacionados ao conhecimento que está em estudo, nesta pesquisa focada em

modelos de ML, BIBFRAME e RDA, a escolha desta ferramenta como método, permite desenhar todo o percurso metodológico de pesquisa e estruturar os instrumentos de coleta e análise de dados. Uma proposta de catalogação de modelos ML usando BIBFRAME e RDA deve considerar diversos detalhes que podem incluir situações condicionais, por exemplo, a quantidade de metadados catalogados para um modelo de ML pode ser diferente de outro porque cada modelo pode ter suas particularidades. Sendo assim, a utilização dos mapas conceituais permite apresentar os caminhos e decisões escolhidos para justificar a proposta de catalogação sugerida, validando os resultados obtidos na pesquisa.

Um mapa conceitual pode ser utilizado como uma atividade meio ou uma atividade fim, ajuda organizar informações para facilitar a recuperação, facilita a revisão do conhecimento, libera a mente de um pensamento linear contribuindo para um maior fluxo de ideias e ajuda a organizar as ideias de forma coerente. Em conjunto com uma Revisão Sistemática da Literatura, RSL, que utiliza o protocolo RSL, a coleta e posterior análise dos dados auxilia no entendimento de como BIBFRAME e RDA podem ser aplicados na catalogação de modelos de ML. Por fim, com a triangulação de dados e os mapas conceituais é possível considerar de forma científica as diversas interpretações e perspectivas que envolvem uma proposta de catalogação de modelos de ML.

## 5. REFERÊNCIAS

ALVES, R. C. V. **Metadados como elementos do processo de catalogação**. 2010. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010. Disponível em: <<https://repositorio.unesp.br/handle/11449/103361>> Acesso 08/07/2022.

BERTI JUNIOR, D. W.; ANDRADE, I. A. de; CERVANTES, B. M. N. **Mapas conceituais: uma ferramenta tecnológica aplicada a organização da informação e do conhecimento**. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA 191 INFORMAÇÃO – CBBBD, 24., 2011. **Anais...** Maceió, 2011. p.1-9.

CARABETTA JÚNIOR, V. A utilização de mapas conceituais como recurso didático para a construção e interrelação de conceitos. **Revista Brasileira de Educação Médica**, v.37, n.3, p.441-447, 2013. Disponível em: <<https://www.scielo.br/pdf/rbem/v37n3/17.pdf>>

CONEGLIAN, Caio Saraiva, **RECUPERAÇÃO DA INFORMAÇÃO COM ABORDAGEM SEMÂNTICA UTILIZANDO LINGUAGEM NATURAL: A INTELIGÊNCIA ARTIFICIAL NA CIÊNCIA DA INFORMAÇÃO**, 195 f, Tese (Doutorado), Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, 2020. Disponível em <<https://repositorio.unesp.br/handle/11449/193051>> Acesso 08/07/2022.

CERVO, A. L.; BERVIAN, P. A. **Metodologia científica**. 5. ed. São Paulo: Prentice Hall, 2003.

COSTA, M. A. F. da; COSTA, M. de F. B. da. **Metodologia da pesquisa: conceitos e técnicas**. Rio de Janeiro: Interciência, 2001.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

GALVÃO, T. F.; PEREIRA, M. G. Revisões sistemáticas da literatura: passos para sua elaboração. **Epidemiologia e Serviços de Saúde**, Brasília, v.23, n.1, p.183-184, jan./mar. 2014. Disponível em: <<https://www.scielo.br/pdf/ress/v23n1/2237-9622-ress-23-01-00183.pdf>>

KÖCHE, J. C. **Fundamentos da metodologia Científica: teoria da ciência e iniciação à pesquisa**. 20. ed. Petrópolis: Vozes, 2002.

MORAES, C. R. B.; FADEL, B. **Triangulação metodológica para o estudo da gestão da informação e do conhecimento em organizações**. In: VALENTIM, M. L. P. (Org.). **Gestão da informação e do conhecimento no âmbito da Ciência da Informação**. São Paulo: Polis; Cultura Acadêmica, 2008. p.27-40.

NOVAK, J. D; CAÑAS, A. J. The origins of the concept mapping tool and the continuing evolution of the tool. **Information Visualization Journal**, n.5, p.175-184, Jan. 2006.

Disponível em:

<<http://cmap.ihmc.us/publications/researchpapers/originsofconceptmappingtool.pdf>>

OKOLI, C. Guia para realizar uma revisão sistemática da literatura. **EaD em Foco**, v.9, n.1, e748, 2019. Disponível em: <<https://eademfoco.cecierj.edu.br/index.php/Revista/article/view/748/359>>

RODRIGUES, M. R.; CERVANTES, B. M. N. Organização e representação do conhecimento por meio de mapas conceituais. **Ciência da Informação**, Brasília, v.43, n.1, p.154-169, jan./abr. 2014. Disponível em: <<http://revista.ibict.br/ciinf/article/view/1425/1603>>

VALENTIM, M. L. P. (Org.). **Métodos qualitativos de pesquisa em Ciência da Informação**. São Paulo: Polis, 2005. 176p. (Coleção Palavra-Chave, 16) Disponível em: <<http://www.abecin.org.br/e-books/colecao-palavra-chave/>>

YIN, R. K. **Estudo de caso: planejamento e métodos**. 3.ed. Porto Alegre, Bookman, 2005.