



UNIVERSIDADE
FEDERAL DO CEARÁ

UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE HUMANIDADES
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO
CURSO DE BIBLIOTECONOMIA

MARISTELA RODRIGUES DE CARVALHO

A CONTRIBUIÇÃO DA MINERAÇÃO DE DADOS NA RECUPERAÇÃO DA
INFORMAÇÃO E SUAS RELAÇÕES COM A BIBLIOTECONOMIA

FORTALEZA-CE
2019

MARISTELA RODRIGUES DE CARVALHO

A CONTRIBUIÇÃO DA MINERAÇÃO DE DADOS NA RECUPERAÇÃO DA
INFORMAÇÃO E SUAS RELAÇÕES COM A BIBLIOTECONOMIA

Monografia apresentada ao Curso de Biblioteconomia do Departamento de Ciências da Informação da Universidade Federal do Ceará, como requisito parcial para obtenção do título de Bacharel em Biblioteconomia.

Orientador: Prof. Dr. Osvaldo de Souza

FORTALEZA-CE
2019

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca de Ciências Humanas

C325c Carvalho, Maristela Rodrigues de.
A contribuição da mineração de dados na recuperação da informação e suas relações com a biblioteconomia / Maristela Rodrigues de Carvalho. – 2019.
47 f. : il.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Humanidades, Curso de Biblioteconomia, Fortaleza, 2019.
Orientação: Prof. Dr. Osvaldo de Souza.

1. Mineração de Dados. 2. Bibliotecário. 3. Recuperação da Informação. I. Título

CDD 020

MARISTELA RODRIGUES DE CARVALHO

A CONTRIBUIÇÃO DA MINERAÇÃO DE DADOS NA RECUPERAÇÃO DA
INFORMAÇÃO E SUAS RELAÇÕES COM A BIBLIOTECONOMIA

Monografia apresentada ao Curso de Biblioteconomia do Departamento de Ciências da Informação da Universidade Federal do Ceará, como requisito parcial para obtenção do título de Bacharel em Biblioteconomia.

Orientador: Prof. Dr. Osvaldo de Souza

Aprovada em ___/___/_____.

BANCA EXAMINADORA

Prof. Dr. Osvaldo de Souza (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Hamilton Rodrigues Tabosa
Universidade Federal do Ceará (UFC)

Prof. Francisco Lima
Faculdade de Inovação e Negócios EASE Brasil

A Deus.

Aos meus pais, Antonio e Auristela (in memoriam), por me darem a vida e pela dedicação ao ensino dos valores morais e à minha mulher, Clariana, pelo incentivo em todas as minhas realizações e grande amor.

AGRADECIMENTOS

A Deus, pela vida, saúde e cognição.

À minha mulher Clariana, pelo incansável apoio, paciência, confiança e amor.

Ao Prof. Dr. Osvaldo de Souza, pela excelente orientação.

Aos professores participantes da Banca examinadora, Hamilton Tabosa e Francisco Lima, pelo tempo dedicado e valiosas colaborações e sugestões.

Aos professores que durante todo o período de graduação se dedicaram em ensinar as disciplinas com esmero e amor.

Aos meus gestores no estágio, Julio Guido e Mayara, pelos ensinamentos, compreensão e exemplos.

Aos colegas da turma de graduação, pelo excelente convívio, pela amizade construída, pelas reflexões, críticas e contribuições.

“Viver cercada por regras pode ser muito restritivo. É a liberdade que nos define. ”
(ANA ROŠ – Documentário Chef’s Table, 2018)

RESUMO

A relevância do uso de ferramentas inteligentes para a ampliação dos meios e métodos voltados à recuperação da informação com qualidade é inquestionável. Este trabalho se propõe a investigar a contribuição e relevância da mineração de dados para a recuperação da informação e suas relações com a biblioteconomia, contribuindo com sugestões de possíveis aplicações desta técnica em bibliotecas, objetivando elaborar melhorias e inovações em procedimentos de recomendações de obras baseadas em grupos de usuários de mesmo interesse, preferências, palavras-chave, formação de coleções e classificação/indexação automáticas. Na elaboração foi utilizado o método de pesquisa exploratória na forma de revisão bibliográfica, de autores que tiveram contato com o tema com a pretensão de amplificar e inter-relacionar a compreensão obtida por eles. E é nesse cenário, invadido pela explosão de dados dia a dia, que se encontram usuários, pluralizados por suas exigências, idiomas, urgências, demandas informacionais distintas, ávidos por satisfazerem tais necessidades, aqui identificados como protagonistas neste processo. Nessa tríade, a última faceta é dedicada aos bibliotecários, com foco nas mudanças positivas das atribuições formais da profissão, nas alterações de conteúdos curriculares, objetivando a equivalência dos seus afazeres com a sua função social, motivadas pelas novas e diversificadas expectativas que emergem da sociedade, e portanto, pode-se estabelecer como conclusão, que o uso da mineração de dados em unidades de informação contando com a participação de um bibliotecário, será muito positiva para a profissão, para o profissional, para a unidade de informação e sobretudo e em especial para seus usuários e a comunidade a qual eles estão inseridos. Todas essas mudanças, somadas à educação continuada, pensamento crítico, forte engajamento e sentidos voltados para a curiosidade e o aperfeiçoamento tornam o bibliotecário capaz de conquistar novas ocupações no acirrado mercado de trabalho.

Palavras-chave: Mineração de dados. Bibliotecário. Recuperação da Informação.

ABSTRACT

The use of intelligent tools to expand the means and methods for information retrieval is unquestionable. This paper means to do a biographical summary on the contributions of data mining, which, when used with the rigor it requires, can help uncover valuable information, many of which weren't even thought of before, and will be used in the decision making process. It is in this scenario, overtaken by the explosion of day-to-day data, that we find the users: plural in their demands, languages, urgency, informational demands, eager to get their needs met, who are recognized as protagonists in this process. This relevance places them under a magnifying glass when we identify the amount of existing studies on communities, users, and their importance. In this triad, the last facet is dedicated to librarians, with emphasis on the positive changes on the current assignments of the profession, on curricular changes, aiming to reach the equivalence between their job and the social function of the librarian, motivated by the new and diversified expectations that emerge from society. All of the changes that have occurred, added to the continuing education and the engagement of the librarian, make them able to conquer new occupations in the competitive work space.

KeyWords: Data Mining. Librarian. Information Retrieval.

LISTA DE ILUSTRAÇÕES

Figura 1: O processo de KDD- Visão Geral.....	22
Figura 2: Possíveis aplicações em Bibliotecas	31
Figura 3: Processo de Recomendação.....	32
Figura 4: Matriz de Swot apresentada no workshop.....	35
Figura 5: Competências dos Profissionais da Informação do século XXI	37

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
CBO	Classificação Brasileira de Ocupações
FEBAB	Federação Brasileira de Associações de Bibliotecários
IBM	International Business Machines
IFLA	International Federation of Library Associations and Institutions
IHC	Interação Humano-Computador
ISO	International Organization for Standardization
KDD	Knowledge Discovery in Database
MTECBO	Ministério do Trabalho Classificação Brasileira de Ocupações
PUC	Pontifícia Universidade Católica
RBBD	Revista Brasileira de Biblioteconomia e Documentação
RS	Rio Grande do Sul
SAS	Statistical Analysis System
SWOT	Strengths Weaknesses Opportunities Threats
UFC	Universidade Federal do Ceará
UFPB	Universidade Federal da Paraíba
UFPE	Universidade Federal de Pernambuco
UFPR	Universidade Federal do Paraná

SUMÁRIO

1. INTRODUÇÃO	13
1.1 Problemática	14
1.2 Justificativa.....	15
1.3 Objetivos.....	15
1.3.1 Objetivo geral	16
1.3.2 Objetivos específicos.....	16
2. O USUÁRIO COMO PROTAGONISTA	17
3. A CIÊNCIA DE DADOS PARA AS BIBLIOTECAS E SERVIÇOS DE INFORMAÇÃO.....	20
3.1 Mineração de Dados suas ferramentas e técnicas.....	21
3.2 Aplicações estratégicas da Mineração de Dados.....	29
4. NOVOS BIBLIOTECÁRIOS: COMPETÊNCIAS, HABILIDADES E MERCADO DE TRABALHO.	34
5. CONCLUSÃO	41
REFERÊNCIAS.....	42

1. INTRODUÇÃO

A revolução científica do século XIX foi também responsável por dar subsídios ao advento da Ciência da Informação no século seguinte, trazendo com ela a ampliação de possibilidades nas atividades desenvolvidas em ambientes informacionais e a quebra de fronteiras, visando mitigar a insegurança e incerteza dos rumos que estávamos tomando. A partir do século XV, com a criação da imprensa as bibliotecas ganham independência no que se refere ao registro, disseminação e acesso a informação. Já nos anos 80, segundo autores como Figueiredo (1979), Vieira (1983), e Saracevic (1992), concluíram que à biblioteconomia, cabe a prática e construção de um conjunto de normas de maneira a sistematizar os serviços da biblioteca. Entretanto, com o desenvolvimento das tecnologias inicia-se um conjunto de sustentações teóricas que passam a discutir sobre os rumos da informação, difusão e compartilhamento, particularmente por via de uma rede.

A partir de então, com o uso crescente de computadores, a quantidade de informações produzidas aumenta abruptamente, tendo o seu ápice com a popularização da internet. Como consequência dessa ação, enormes repositórios de informação estruturada e não estruturada, são gerados em sistemas de bancos de dados, distribuídas em arquivos de textos, imagens, áudios, etc. Olhar para esses repositórios como fontes de informação, deu origem ao uso de técnicas e ferramentas capazes de trata-los tornando essa atividade indispensável e muito rentável.

Este trabalho foi elaborado utilizando o método de pesquisa exploratória na forma de revisão bibliográfica.

Metodologia é uma preocupação instrumental. Trata das formas de se fazer ciência. Cuida dos procedimentos, das ferramentas, dos caminhos. A finalidade da ciência é tratar a realidade teórica. Para atingirmos tal finalidade, colocam-se vários caminhos. Disto trata a metodologia. (DEMO, 1985, p. 19).

A pesquisa de autores, artigos, textos e leis, foi feita utilizando exclusivamente a *internet*. Foram consultados: o portal de periódicos da Capes, Scielo, portais específicos de universidades como PUC-Rio, PUC-RS, UFPB, UFPE, UFPR, plataformas corporativas como *Computerworld*, Oficina da NET, IBM Brasil, SAS. *Sites* de categorias de classes e governamentais como o IFLA, FEBAB, RBBB,

MTECBO, Biblioteca Nacional, Casa Civil e ABNT. Para a realização da pesquisa foram utilizadas as seguintes expressões de busca: mineração de dados, mineração de dados na recuperação da informação, mineração de dados na biblioteconomia, bibliomineração, bibliometria, ciência de dados na biblioteconomia, competências do bibliotecário, bibliotecário no mercado de trabalho, entre outros.

A partir dos resultados úteis encontrados foram feitas análises as quais permitiram construir os relacionamentos demonstrados na Figura 2 – “A” e “B”. No conjunto das investigações, percebemos a oportunidade de vislumbrar as aplicações propostas na parte “B”, decorrentes das características tecnicistas da parte “A”, operando em concomitância às competências ilustradas na Figura 5. Embora essas ações, aplicações e competências (Figura 2 e Figura 5), possam existir independente uma da outra, acreditamos que quando elas acontecem de maneira estruturada, aumenta a relevância da mineração de dados na recuperação da informação com qualidade, feita com a colaboração de um bibliotecário.

1.1 Problemática

A obrigação de olhar para o futuro imaginando o acirramento das exigências sobre as competências individuais de todos os profissionais, faz com que não seja exagero dizer, que a tomada de decisões baseadas no uso de tecnologias que viabilizem a extração de informações e conhecimentos nas esferas públicas e privadas, pode mudar governos, empresas e hábitos sociais. Também não é exagero dizer, que todas essas transformações deram origem a um novo perfil do bibliotecário, exigindo dele competências técnicas mais aprofundadas, como também, sua participação ampliada como agente de transformação social.

O questionamento que se estabelece é: qual a contribuição que a mineração de dados traz para a recuperação da informação com qualidade? Quais dessas influências contribuem na formação de um novo perfil do Bibliotecário?

Neste contexto, a Ciência de Dados é uma área da ciência muito disciplinada que objetiva estudar as informações, seu processo de captura, transformação, geração e posteriormente a análise de dados, sempre pensando em transformar dados em informação ou produtos de informação. Embutida na Ciência de Dados encontramos o componente que trata dos métodos do descobrimento do

conhecimento, a Mineração de Dados, uma ferramenta de análise de dados capaz de traçar perfis e relacionamentos não triviais, motivada pela inteligência computacional. A ideia de minerar dados não é nova, desde os anos sessenta já se trabalhava com o propósito de encontrar correlações sem uma hipótese em bases de dados com ruído. Nos anos oitenta, avanços na área consolidaram o termo mineração de dados que ao longo do tempo e até o presente momento, evolui na utilização de ferramentas e métodos.

1.2 Justificativa

O meu interesse no tema de recuperação da informação utilizando técnicas como por exemplo, a mineração de dados, se dá pelo fato de querer aprofundar meus conhecimentos na previsão de tendências futuras. Segundo Elder (2017) “A aprendizagem a partir dos dados é extremamente poderosa e seu uso está transformando a tomada de decisões de empresas dos mais diversos setores em ritmo acelerado [...], para encontrarmos excelência nesse trabalho, precisamos ouvir atentamente com o objetivo de transformar um desafio do mundo real em um problema parecido, mas solucionável”. O resultado desse aprofundamento pode ser um ponto de partida ou motivação para que outros colegas possam dar continuidade ao tema, contribuindo para o engrandecimento da profissão, ao torna-la imprescindível em qualquer setor, segmento e mercado. Não seria justo omitir minhas pretensões pessoais associadas a sonhos antigos ainda não realizados, ao perceber o potencial de poder fazer acontecer. A adrenalina que me motiva está atrelada a mão na massa do mercado de trabalho, muito embora esse perfil não garanta sucesso, nem facilidades. Há que se ter foco, saúde, fé e sobretudo, compromisso com os propósitos estabelecidos.

1.3 Objetivos

Este trabalho se propõe a apresentar dois tipos de objetivos, divididos em geral e específico, conforme a seguir:

1.3.1 Objetivo geral

O objetivo geral deste trabalho é investigar a contribuição e relevância da mineração de dados na recuperação da informação com qualidade e suas relações com a Biblioteconomia.

1.3.2 Objetivos específicos

- a) Apresentar as características da mineração de dados;
- b) Procurar possíveis relacionamentos entre as características da mineração de dados e as atividades no campo da biblioteconomia;
- c) Identificar possíveis relacionamentos entre as características da mineração de dados e as competências do bibliotecário, inerentes ou requeridas.

2. O USUÁRIO COMO PROTAGONISTA

Não é característica exclusiva da última década, a importância sobre atender com qualidade às necessidades informacionais de usuários cada vez mais globais, numerosos, exigentes e diversos. Mas, a cada dia aumenta a relevância dos estudos sobre o comportamento informacional natural, emocional, cognitivo, contextos de vida, experiências e necessidades, voltados para aspectos individuais dos usuários no processo de pesquisa, incluindo estudos das táticas de busca, compreensão da relevância pela ótica do usuário e como encontram informações dentro e fora da *web*. São estudos centrados e concentrados nos seres humanos, suas necessidades e comportamentos informacionais. “O conhecimento do usuário é importante por ele ser o responsável pela existência do serviço e, conseqüentemente, por avaliá-lo” (DIAS; PIRES, 2004, p.9).

São evidentes o comprometimento do estudo de usuários para a biblioteconomia, “mais um exemplo da integração útil e benéfica das técnicas da ciência da informação ao aperfeiçoamento das técnicas bibliotecárias” (FIGUEIREDO, 1994, p.10).

Simultaneamente aos esforços para conhecer seu usuário e não menos importante, é a pesquisa que visa conhecer a comunidade em que ele está inserido, considerando aspectos políticos, econômicos, demográficos, social e cultural, englobando usuários reais e potenciais. De acordo com Bohmerwald (2005), as pesquisas sobre sistemas de informação e seus usuários apresentam uma análise mais completa quando utilizam características de estudos de usuários, comunidades e estudos de usabilidade. Todo esse aparato visa não somente ao aprimoramento das técnicas de pesquisa e de ferramentas para a recuperação da informação com qualidade, mas, sobretudo, atender interesses distintos de organizações nas tomadas de decisão.

Informação é algo que necessitamos quando deparamos com uma escolha. Qualquer que seja seu conteúdo a quantidade de informação necessária depende da complexidade da escolha. Se nos deparamos com um grande espectro de escolhas igualmente prováveis, se qualquer coisa pode acontecer, precisamos de mais informação do que se encarássemos uma simples escolha entre alternativas. (MILLER, 1966, p. 8)

Não há como falar de usuários e suas necessidades informacionais sem mencionar a interação entre pessoas e sistemas computacionais ou como também é conhecida, IHC – Interação Humano-Computador, relação que permite o desenvolvimento de interfaces de qualidade. Segundo Barbosa e Silva (2010) a melhoria das interfaces traz diversos benefícios: redução de erros, redução de custos de treinamento e suporte técnico, aumento da produtividade e da fidelidade de usuários, entre outros.

Da mesma forma, Kafure *et. al.* (2013) apontam que nos estudos de usabilidade, é necessário identificar o perfil do usuário desde o ponto de vista emocional e cognitivo. Apesar do conceito de usabilidade ser facilmente compreendido, na prática, ainda estamos caminhando em busca de melhorias nesse aspecto. Entende-se por usabilidade a facilidade funcional, a interação quase que automática do usuário com interfaces de programas e aplicações.

A norma ISO 9241-11 (ABNT, 2011, p. 3), dispõe orientações de interface para trabalho com computadores, define a usabilidade como “o grau em que um produto é usado por usuários específicos para atingir objetivos específicos com eficácia, eficiência e satisfação em um contexto de uso específico”.

Se o usuário é o ponto focal de todos os sistemas de recuperação da informação, se o incremento de estudos sobre o comportamento informacional individual colaboram para enriquecer e melhorar os métodos de busca e ainda, se é inegável que surgem a cada dia ferramentas e estudos ergonômicos para o desenvolvimento de interfaces que permitam a interação e operação do computador objetivando garantir que sistemas e dispositivos estejam adaptados ao usuário, então, por que ainda encontramos dificuldade em recuperar com qualidade e precisão a informação demandada? Infelizmente, todos esses estudos são feitos separadamente por diversas áreas que não trabalham em conjunto. Sistemas para recuperação de informação são complexos e exigem a soma de conhecimentos de profissionais com domínio de conhecimento em diversas áreas.

O profissional de ciência da informação pode contribuir com conhecimentos específicos para a criação de bancos de dados e oferecer uma perspectiva técnica para indexação e recuperação de informação (DZEKANIYAK, 2004). Para que o usuário localize a informação mais adequada às suas necessidades, é importante que seja realizado um trabalho de estruturação da informação, entendida como “sua organização lógica para posterior recuperação e linguagem de busca, como os

comandos que permitem a recuperação da informação através de palavras contidas nos títulos, resumos ou outros campos de dados” (BRANSKI, 2004). Ferreira *et. al.* (2003) apontam que diversos estudos revelam problemas básicos relativos à falta de consistência entre a organização da informação e sua expressão no sistema. Outro ponto relevante é o conceito subjetivo da necessidade informacional, ocorrida apenas na mente de cada indivíduo e, portanto, não acessível ao observador. Neste caso, a necessidade só pode ser descoberta por dedução, observando o comportamento ou por declaração do próprio indivíduo, que nem sempre traduz em sua totalidade todas as suas necessidades.

Cooper (1971, p. 22) afirma que uma necessidade informacional é algo não observável diretamente. Não podemos, por exemplo, ver suas estruturas, no entanto, a necessidade informacional existe, pelo menos na mente do usuário.

A descoberta de determinados padrões particulares no contexto de grupos de usuários ou de uma grande base de dados, cria possibilidades para métodos preditivos e descritivos importantes na recuperação da informação muitas vezes, sequer imaginada, mas que pode ser potencialmente útil. Uma das técnicas utilizadas nesse sentido é a Mineração de Dados, parte integrante do processo da Descoberta de Conhecimento nas Bases de Dados ou KDD (*Knowledge Discovery in Database*). Segundo Quoniam, *et. al.* (2001), o KDD pode ser visto como o processo da descoberta de novas correlações, padrões e tendências significativas por meio da análise minuciosa de grandes conjuntos de dados armazenados.

Considerando a união de todos esses esforços pode-se acreditar que continuamos progredindo na construção de meios, processos, ferramentas e estudos, visando atender com qualidade as necessidades informacionais de cada indivíduo.

3. A CIÊNCIA DE DADOS PARA AS BIBLIOTECAS E SERVIÇOS DE INFORMAÇÃO

A Biblioteconomia é um marco no contexto da organização do conhecimento por ser considerada uma atividade milenar, sobretudo pelas questões que dizem respeito à organização e registro das informações. No início, a organização, registro e classificação eram realizados de maneira primitiva, mas cumpria o seu papel de assegurar a memória da humanidade pelos procedimentos voltados para o acesso à informação, mesmo que por muitos séculos, esse acesso estivesse restrito a divisões sociais específicas.

Segundo Silva e Freire (2012), a partir do século XVI, devido à grande quantidade de indivíduos que foram alfabetizados na Europa, a biblioteca ganha a função de instituição socializadora e passa a ter uma autonomia social. Mas foi somente no “século XIX que a Biblioteconomia aperfeiçoa seus estudos técnicos e epistemológicos para a organização do conhecimento, sobre como tratar a informação” (SILVA; FREIRE, 2012). Muito mais recente é a Ciência da Informação, que no século XX estimula novas reflexões em torno da organização do conhecimento, tanto pelo olhar técnico, quanto cognitivo.

Informação é o termo que designa o conteúdo daquilo que permutamos com o mundo exterior ao ajustar-nos a ele, e que faz com que nosso ajustamento seja nele percebido. Viver de fato é viver com informação. (NORBERT, 1960, p. 25).

Quebrar paradigmas nesta área do conhecimento, não é novidade. Conforme afirma Figueiredo (1996, p. 16), o visionário Belga Paul Otlet, previu no século passado, o desenvolvimento das tecnologias com o uso de novos tipos de sistemas mecânicos integrados para o uso da informação, que transformariam o meio ambiente e as práticas do pesquisador. E, portanto, a relevância da ciência de dados não deve ser vista como algo surpreendente em qualquer seara do conhecimento ou atividade profissional.

Tendo como foco a representação, a organização e o armazenamento da informação, para que seja possível a recuperação, o acesso, o uso e reuso e a disseminação, a Ciência da Informação constatou que as tecnologias da informação e comunicação possibilitariam a concepção de sistemas e serviços que permitissem a comunicação, o uso e o armazenamento da informação. (CONEGLIAN; GONÇALVES; SEGUNDO, 2017, p. 3).

A avalanche de dados produzidos a cada hora, considerando a velocidade em que devem ser tratados e apresentados e ainda numa numerosa variedade de fontes e suportes em que se encontram, torna o profissional sem recursos, estéril no cumprimento de suas atividades e crescimento pessoal/profissional. Segundo Mayer-Schönberger e Cukier (2013) o *Big Data* representa “[...] uma nova fonte de valor econômico e informação”, visto que os dados tem se tornado uma das principais fontes de renda e informação das organizações. Essa conjuntura tem provocado uma revolução no modo como as análises de dados estão ocorrendo, tanto no âmbito organizacional, quanto no acadêmico.

Segundo Shera (1976), a meta do Bibliotecário é maximizar a utilidade social dos registros gráficos para benefício da humanidade. Neste aspecto, a investigação quantitativa da literatura, pode dar celeridade no atingimento desse alvo, encontrando a chave para conhecer melhor comunidades de usuários objetivando adequar os serviços oferecidos a essas comunidades e aos diferentes grupos de usuários. Na prática, isso significa dizer que Bibliotecários podem fazer mais e muito melhor, sem perder sua essência, utilizando novos recursos, quebrando paradigmas, alçando novos voos, exercendo atividades ainda não imaginadas para a sua classe.

3.1 Mineração de Dados suas ferramentas e técnicas

A Ciência de Dados é uma área interdisciplinar voltada para o estudo e a análise de dados estruturados ou não, que segundo Silveira (2016), visa a extração de conhecimento que podem ser usados para tomar decisões e fazer previsões e não simplesmente, interpretar números. Essa ciência é um campo que existe há mais de trinta anos e vem ganhando destaque devido ao desenvolvimento de técnicas e ferramentas, assim como também à concorrência acirrada por conhecimento. Parecida com a computação, pode ser aplicada à diferentes áreas do conhecimento, o domínio dessas técnicas, ajuda na identificação de problemas específicos.

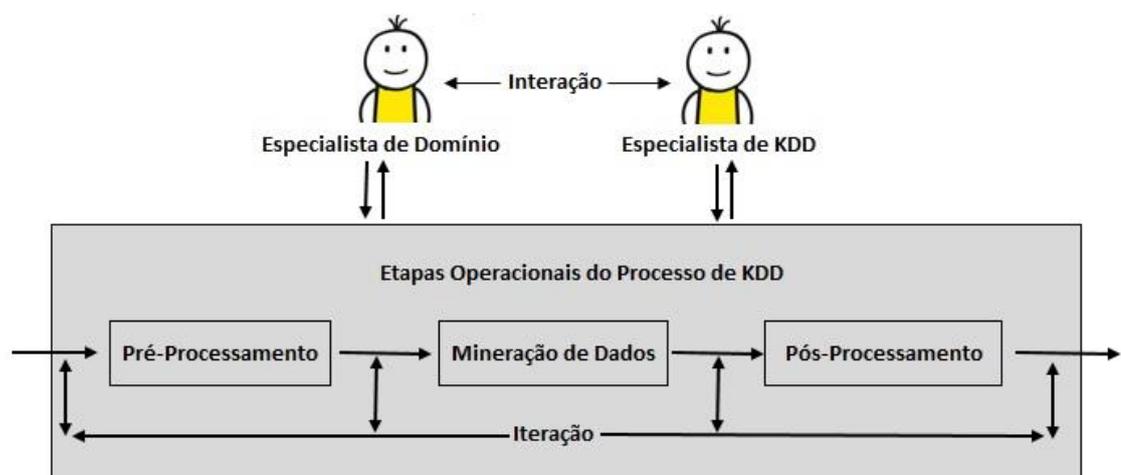
Contida na ciência de dados está a Mineração de Dados que atua a partir de uma quantidade grande e pesada de massa de dados. Segundo Cios *et. al.* (2007) e Fayyad (1996), a Mineração de Dados é uma peça na engrenagem maior, chamada Descoberta de Conhecimento nas Bases de Dados – cuja sigla em inglês é KDD.

Diversas ciências vêm utilizando a mineração de dados para o aprimoramento do saber em seus campos, entre elas: a administração, o marketing,

a astronomia, a medicina, a física e a ciência da informação. A mineração de dados envolve enquadrar modelos a (ou encontrar padrões em) dados observados. A maioria dos métodos de mineração é baseada em técnicas testadas de aprendizado de máquina e reconhecimento de padrões e estatística, como classificação, associação, regressão, *clustering*, sumarização, modelagem de dependência, entre outros (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1997, p. 10). A mineração de dados é uma técnica, que se bem utilizada, torna-se um grande diferencial na tomada de decisão. Porém, a extração de *insights* úteis das correlações e padrões identificados durante a coleta, somente será possível com o profundo conhecimento do negócio em questão, só assim, a atribuição de valor de cada descoberta pode ser bem executada. Quanto mais dados forem coletados, haverá mais evidências para identificação de público e mercado. A escolha apropriada da técnica de mineração a ser utilizada é essencial para que sejam extraídos os conhecimentos pretendidos, assim como, é fundamental a qualidade dos dados selecionados, caso contrário não haverá qualidade no resultado final.

Visão Pragmática do processo de KDD segundo Goldschmidt *et. al.* (2002) conforme figura 1:

Figura 1: O processo de KDD- Visão Geral



Fonte: Adaptada de Goldschmidt *et. al.* (2002)

São muitas as definições de Mineração de Dados encontradas na literatura. Como por exemplo:

“A mineração de dados é um processo altamente cooperativo entre homens e máquinas, que visa a exploração de grandes bancos de dados, com o objetivo de extrair conhecimentos através do reconhecimento de padrões e relacionamento entre variáveis. ” (CÔRTEZ; PORCARO; LIFSCHITZ, 2002).

“Mineração de dados é a extração de informações potencialmente úteis e previamente desconhecidas de grandes bancos de dados, serve para descobrir perfis de consumidores e outros comportamentos que não seriam identificados nem por especialistas. ” (GUIZZO, 2000).

“Mineração de Dados é uma técnica para determinar padrões de comportamento, em grandes bases de dados, auxiliando na tomada de decisão. ” (SILVA, 2000).

“Extração de conhecimento de base de dados (Mineração de Dados) é o processo de identificação de padrões válidos, novos, potencialmente úteis e compreensíveis embutidos nos dados. “ (FAYYAD *et. al.* 1996).

É fato que apesar dos avanços, não é consenso os benefícios práticos da Mineração de Dados. Em seus estudos, Han (2006), afirma que ela é “rica em dados, pobre em informação”, já Witten *et. al.* (2005), Bramer (2007) e Olson *et. al.* (2008), acreditam que a Mineração de Dados pode ser aplicada satisfatoriamente em áreas da saúde, financeira, eleitoral, varejo, segurança, recursos humanos e muitos outros. No entanto, todos concordam que a realização da Mineração de Dados deve ser dividida em fases interativas.

Sua aplicação pode ser um processo de verificação onde o usuário tenta provar sua hipótese acerca da relação entre dados, como um processo de descoberta de padrões, fazendo uso de técnicas como redes neurais, algoritmos genéticos, regras de associação, árvores de decisão, regressão, entre outros. (FRACALANZA, 2009, p. 15).

Em geral, são técnicas úteis para a tomada de decisão ou avaliação de resultados, de inteligência artificial em busca de afinidades ou discrepâncias entre dados, a fim de localizar modelos, padrões e anormalidades com a finalidade em transformar dados aparentemente ocultos, em informações.

Considerando todas essas explicações, somos levados a crer que o processo de extração de conhecimento ocorre de maneira totalmente automática, isso não é verdade. Na prática, as ferramentas auxiliam na execução dos algoritmos de mineração, entretanto, os resultados precisam da análise humana, sem essa análise crítica, não há como identificar a relevância sobre os padrões encontrados.

Como demonstra a Figura 1, uma das etapas do processo de KDD é a de pré-processamento, a qual engloba o tratamento e a preparação dos dados. É nessa etapa do processo que podem atuar os bibliotecários, realizando a curadoria de dados, atividade que envolve o recebimento, validação, transferência, conversão, identificação de metadados e descritores, seleção de níveis de acesso, análise, manutenção e atualização dos dados. Segundo Rezende *et. al.* (2003) e Witten; Frank (2009), para que sejam descobertos padrões de qualidade é importante que essa etapa seja cuidadosamente executada. Ainda, de acordo com Facelli *et. al.* (2011), o desempenho dos algoritmos de aprendizado de máquina geralmente é afetado pelo estado em que os dados se encontram, ou seja, pela qualidade dos dados disponíveis. O objetivo desta etapa é tornar os dados estatisticamente de melhor qualidade para a extração de padrões.

Abaixo, segundo Fayyad *et. al.* (1996), a divisão do pré-processamento de KDD em seis fases:

1. Preparação dos Dados: inclusão do conhecimento relevante para a aplicação, definindo também as metas que o processo precisará atingir;
2. Limpeza dos Dados: subtração dos dados que possam distorcer a análise, objetivando a remoção de ruídos, tratando atributos perdidos e até mesmo os transformando para reduzir o número de variáveis envolvidas no processo, a fim de aprimorar o desempenho do algoritmo de análise;
3. Seleção de Dados: seleção sobre qual conjunto ou subconjunto de dados em que o processo será aplicado;
4. Mineração de Dados: decisão sobre qual tarefa de mineração de dados será aplicada para atingir os objetivos do processo e qual a melhor técnica a ser utilizada;

5. Incorporação do Conhecimento Anterior: interpretação do modelo descoberto a fim de verificar com cuidado possíveis melhorias, possibilitando o retorno para qualquer fase anterior do processo, retirando padrões redundantes ou irrelevantes;
6. Interpretação dos Resultados: incorporação no sistema possibilitando a tomada de ações baseadas no conhecimento ou documentando-os e relatando-os às partes interessadas.

Após as seis fases de pré-processamento, passamos para a etapa de extração de modelos e padrões. Nesta etapa pode-se empregar diferentes métodos e técnicas de aprendizado de máquina, basicamente duas abordagens: aprendizado preditivo e aprendizado descritivo (FACELLI *et. al* 2011). O aprendizado preditivo analisa os dados que representam eventos pesados buscando relações entre estes que permitam prever situações em novos dados futuros. O aprendizado descritivo analisa os dados e identifica similaridades (agrupamentos) ou associações (regras de associação). Para identificar com facilidade a tarefa de mineração mais positiva no processo de busca de uma solução para um problema, é preciso saber qual resultado se deseja encontrar.

Segundo Carvalho e Dallagassa (2014), em geral, a natureza do problema em questão define a natureza do aprendizado a ser adotado na experimentação de mineração de dados, ou seja, geralmente os problemas são preditivos ou descritivos, entretanto, algumas vezes as duas estratégias podem ser adotadas de forma complementar. A parte mais importante de qualquer projeto de mineração de dados é definir claramente o problema a ser solucionado. Nenhum modelo conta sozinha a história completa. Não existe uma regra que diz quando você já esgotou os dados. (ABBOTT, 2017).

Segundo Fracalanza (2009), segue abaixo os tipos de predição e descrição.

1. Classificação: identifica a classe de um objeto através de um modelo classificador construído com informações de um conjunto de amostras. Categorizar clientes de acordo com o seu perfil de compras é um exemplo de tarefa de classificação. Um modelo de classificação é criado e os

atributos dos registros são analisados de acordo com as condições e características das classes pré-determinadas pelo modelo. Caso esses atributos coincidam, o registro é então classificado na classe em que melhor se encaixa.

2. Segmentação: conhecida como *clustering* se diferencia da classificação porque não pressupõe a existência de classes pré-definidas. Os registros são agrupados de acordo com a semelhança nos seus atributos, segmentando-os em *cluster* (subgrupos homogêneos) baseando-se no princípio de que os registros classificados em um grupo devem ser homogêneos entre si, e que os grupos devem ser heterogêneos entre si. Esse tipo de modelo é usado normalmente para otimizar campanhas de *marketing* direto, reter e expandir o relacionamento com clientes.
3. Regressão: avalia um acontecimento ou um padrão. Prevê padrões para uma variável de valor contínuo, com base nos outros atributos disponíveis e nas outras ocorrências disponíveis para análise. Ou seja, ela prevê um valor numérico desconhecido a partir de alguns atributos conhecidos. Geralmente, os problemas que a regressão auxilia estão relacionados à previsão da quantidade de itens em um determinado momento ou à previsão populacional. Ela é também a técnica mais fácil de usar, mas provavelmente é a menos poderosa. Um exemplo de utilização da tarefa regressão é prever o valor do aluguel de um imóvel a partir de um banco de dados de imóveis.
4. Associação: Devido a sua grande aplicabilidade as regras de associação encontram-se entre um dos mais importantes tipos de conhecimento que podem ser minerados em base de dados. Estas regras representam padrões de relacionamento entre itens de uma base de dados capaz de identificar transações que sempre ocorrem em conjunto. Na literatura também é conhecida como grupos de afinidade ou análise de cesta. Uma de suas típicas aplicações é a análise de transações de compras (*Market Basket Analysis*), um processo que examina padrões de compras de consumidores para determinar produtos que costumam ser adquiridos em conjunto. Esse estudo auxilia as lojas a encontrar a melhor maneira de

organizar seus produtos, de modo que a disposição deles nas prateleiras estimule as compras dos clientes.

A próxima etapa é a escolha das técnicas de Mineração de Dados que na verdade estão diretamente relacionadas à concepção de algoritmos para a procura de padrões nos dados. A preparação dos dados varia de acordo com as diversas técnicas utilizadas por meio de seus algoritmos, e suas particularidades exigem que o seu manipulador possua conhecimentos específicos sobre o seu funcionamento para, portanto, escolhê-la adequadamente de acordo com os objetivos a ela propostos.

Existem alguns *softwares* no mercado que possibilitam o uso das técnicas de Mineração de Dados, sendo que os mais robustos e amigáveis, são também os mais caros. Já a ferramenta WEKA (WITTEN; FRANK, 1999) é uma alternativa de *software* livre, que segundo Nicholson e Stanton (2004), contém muitas opções idênticas às de sistemas robustos, mas a custo menor, porém, de uso não tão amigável.

Segundo Fracalanza (2009), são apresentadas as técnicas de Mineração de Dados, conforme a seguir:

1. Algoritmo Genético: Nos anos 50 e 60 muitos biólogos começaram a desenvolver simulações computacionais de sistemas genéticos. Em seguida imitaram algumas etapas do processo de evolução natural das espécies incorporando-as a um algoritmo computacional. Essa ideia foi refinada ao longo do tempo e desde então, segundo Bala *et. al.* (1995), estes algoritmos são aplicados com sucesso nos mais diversos problemas de otimização, descoberta de padrões e aprendizado de máquina. Diferentemente dos métodos convencionais de mesmo objetivo, eles trabalham simultaneamente em conjuntos de soluções diferentes, realizando pesquisas adaptativas nos dados, modelando uma solução para um problema específico em estruturas de dados que são semelhantes a um cromossomo. Operadores são aplicados recombinando essas estruturas, gerando assim, novas combinações de regras de associação. Essa técnica é utilizada na classificação e na segmentação de dados, formulando hipóteses sobre a dependência dos atributos dos dados. Ao longo do tempo,

o algoritmo tende a aprender e a se aperfeiçoar, de maneira que somente as soluções com maior poder de sucesso na previsão são aceitas.

2. **Redes Neurais:** bastante utilizada em tarefas de classificação, regressão e segmentação. Conforme Haykin (2001), os dados são trabalhados com base no funcionamento do cérebro humano, aprendendo a tomar decisões baseadas nas experiências anteriores, ou seja, nas instâncias anteriores dos dados. Os neurônios do cérebro são representados por nós que estão conectados em outros nós por sinapses, formando uma rede de processamento. Os valores das entradas são multiplicados nos neurônios pelos pesos de suas sinapses, conforme vão caminhando na rede tem-se a classificação ou a previsão da entrada.
3. **Árvores de Decisão:** objetivam principalmente dividir as instâncias em classes. De acordo com Bramer (2007), cada nó da árvore testa o domínio de uma variável da entrada e o redireciona para o nó seguinte. Cada sub-árvore representa o resultado de um teste e a folha é a classificação que aquele registro recebeu. Ao final, cada nó terminal terá os registros da entrada adaptados às regras regidas por esse nó, representando assim, uma classe.
4. **Regras de Associação:** são definidas por uma correlação estatística entre alguns atributos da entrada com o objetivo de descobrir relações que ocorrem em comum dentro de um conjunto de dados. Cada registro é visto como uma transação e cada variável como um item dessa transação, deixando subentendido que a presença de um item implica necessariamente na presença de outro na mesma transação.
5. **Análise de Vizinhança:** determina a distância entre duas instâncias, ou seja, de uma função para identificar um conjunto de registros que estão próximos por determinada característica, essa técnica é empregada em análise de prognósticos e não para descoberta de conhecimento. Não é muito explorada na literatura.

3.2 Aplicações estratégicas da Mineração de Dados

Existe um vasto universo de possibilidades para a aplicação das técnicas de Mineração de Dados. Segundo Goldschmidt e Bezerra (2016), são factíveis a sua utilização nas áreas de telecomunicações, comércio, finanças, medicina, ação social, educação, energia, seguros, indústria, arrecadação de impostos e comportamento em redes sociais online. Mas foi Nicholson e Stanton (2004), que aproximaram ainda mais a Biblioteconomia à Mineração de Dados quando utilizaram o termo “Bibliomineração” para representar a Mineração de Dados aplicada à biblioteca.

O processo da Bibliomineração consiste em determinar áreas temáticas de interesse; identificar as fontes de dados internas e externas, coletar e depurar e tornar anônimos os dados destinados a um repositório, visando principalmente a descoberta de padrões subjacentes. Os resultados podem ser combinados com diferentes variáveis, antes de sua utilização na tomada de decisão pela administração da biblioteca. (NICHOLSON; STANTON, 2004, p. 2, tradução BARANOW).

Na prática, segundo Lima (1986), a aplicação de métodos quantitativos nas atividades de gestão de uma biblioteca pode facilitar decisões ao estabelecer soluções alternativas para questões administrativas de aquisição, manutenção ou descarte de documentos. Entretanto, análises descritivas ou preditivas só podem existir a partir da existência de uma base de dados. “A criação de um repositório de dados pode absorver até oitenta por cento do tempo necessário ao processo de bibliomineração” (BERRY; LINOFF, 1997). Uma vez implementada uma grande base de dados, “a biblioteca poderá ter ao seu alcance uma grande quantidade de relatórios e ferramentas analíticas, valiosos subsídios para conhecer profundamente suas comunidades de usuários e poder melhor adequar a utilização de seus recursos” (NICHOLSON; STANTON, 2004, p. 9, tradução BARANOW).

Considerando todo o potencial e características da Mineração de Dados, é possível sugerir a sua utilização, visando elaborar melhorias e inovações em procedimentos para unidades de informação. Na apresentação da Figura 2 - Parte “A”, as características da Mineração de Dados apresentadas, de acordo com Hoppen (2015), representam o poder da ferramenta que torna possível a sugestão de um pequeno elenco de possibilidades citados como exemplos (conforme apresenta a Figura 2 – Parte “B”) podendo, contudo, haver muitos outros.

São ao todo 5 sugestões de possíveis aplicações em bibliotecas na figura 2 “B”.

1. Por implementação de serviços de recomendações de obras baseadas em grupos de usuários de mesmo interesse e área, entendemos que é um serviço essencial para conhecer melhor usuários ativos, utilizando regras de associação em aprendizado preditivo. “O mais importante é prever para onde vão os clientes e se postar bem na frente deles.” (KOTLER, 2017).

2. Palavras-chave são o principal instrumento de uma pesquisa. Normalmente é a maneira utilizada pelo usuário de escrever a sua dúvida nos buscadores objetivando obter respostas e solucionar suas necessidades informacionais. A segunda sugestão de aplicações possíveis em uma biblioteca, considera a possibilidade de identificar entre uma massa de usuários, quais deles utilizaram palavras-chave idênticas, semelhantes, ou de mesmo sentido e, portanto, com demandas análogas, e fazer dessa atividade uma oportunidade de recomendação de conteúdo relevante.

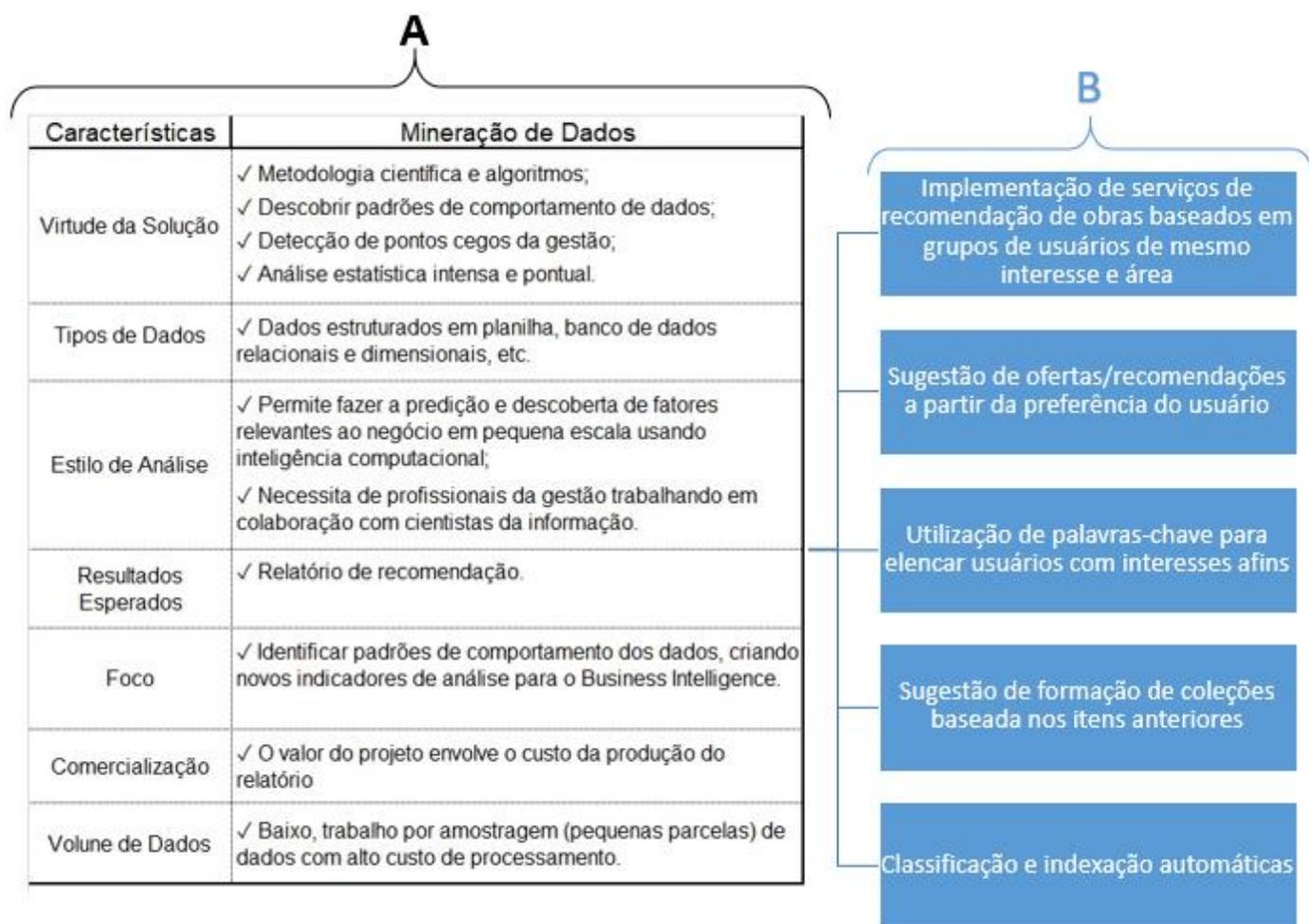
3. A terceira possibilidade apresentada de aplicação da mineração de dados em bibliotecas, faz uso do processo de recomendação a partir da preferência do próprio usuário. São algoritmos de filtragem que coletam dados e informações, a fim de realizar recomendações inteligentes baseadas no perfil de cada usuário, considerando seus gostos e preferências. Esse tipo de aplicação amplia o engajamento do usuário com o conteúdo. Os resultados da recomendação podem ser originados nos cliques, adições de reserva e comportamentos pautados nas decisões de escolha do item. Segundo a empresa SmartHint (2018), são algoritmos que coletam muitos dados que podem ser explícitos ou seja, imputados pelo próprio usuário, como comentários e avaliações de itens ou implícitos, que incluem o histórico de reservas, retiradas, número de visitas, cliques e registros de buscas.

4. As sugestões anteriores serão de extrema importância para a implementação da formação de coleções, uma vez que podem ser realizadas

diversas análises sobre usuários, que vão desde preferências, uso, reuso, desuso, entre outros.

5. A quinta sugestão refere-se à automação da classificação e indexação, que segundo Lapa e Correa (2014), afirmam poder ocorrer por meio da atribuição de sintagmas nominais considerando seus aspectos semânticos.

Figura 2: Possíveis aplicações em Bibliotecas



Fonte Figura 2 "A": Adaptada de Hoppen (2015)

Fonte Figura 2 "B": Elaboração da autora (2019)

A figura 2 demonstra uma relação de interdependência entre as características da Mineração de Dados e seus efeitos explicitados na Parte "A" e as sugestões de seu uso em Bibliotecas, explicitados na Parte "B". Considerando os elementos apresentados nas duas partes da Figura 2, compreendemos que atingimos os objetivos específicos:

- a) Apresentar as características da mineração de dados;
- b) Procurar possíveis relacionamentos entre as características da mineração de dados e as atividades no campo da biblioteconomia;

Para que seja possível criar recomendações é preciso que todo o sistema de recomendação siga processos bem definidos.

Um sistema de recomendação consiste de: dados contextuais, a informação que o sistema possui antes de começar o processo de recomendação, dados de entrada, a informação que o usuário precisa passar para o sistema conseguir gerar as recomendações e um algoritmo capaz de combinar os dados do contexto com os dados de entrada de forma a modelar adequadamente as recomendações. (SOUZA, 2011, p. 2).

Podem-se reconhecer na Figura 3, duas fontes de informação necessárias para a entrada no processo: dados dos usuários e as informações a respeito dos itens. Segundo Souza (2011), idealmente, estes dados relacionados ao perfil dos usuários deveriam ser fornecidos explicitamente pelos próprios usuários. Entretanto, há recursos via outras fontes, como a navegação pelas páginas, itens consumidos, forma como foram consumidos e período em que foram consumidos, capazes também de obtê-los.

Figura 3: Processo de Recomendação



Fonte: Adaptada de Souza (2011)

Os exemplos informados de serviços, atividades e projetos que podem ser desenvolvidos na biblioteca utilizando a Mineração de Dados, todos eles apresentam um diferencial ao oferecer o item mais apropriado a um determinado usuário, agregando valor não apenas ao bibliotecário, e nem somente à unidade de informação ao qual ele está inserido, mas sobretudo, esse tipo de ação, aumenta o nível de satisfação de quem o utiliza (satisfazendo assim, com qualidade, a sua necessidade informacional) e que pode garantir fidelidade à unidade de informação que o provê.

4. NOVOS BIBLIOTECÁRIOS: COMPETÊNCIAS, HABILIDADES E MERCADO DE TRABALHO.

Desde o surgimento das primeiras bibliotecas até os dias atuais, o bibliotecário vem evoluindo sendo acompanhado pela tecnologia. Pereira e Cunha (2007) defendem que a informação e conhecimento sempre estiveram relacionados ao processo de desenvolvimento humano, ao longo da história. Em decorrência desse processo, “o avanço das tecnologias de informação, encontra-se hoje no centro da reformulação de uma nova ordem mundial” (AMARO, 2018). Não se imagina mais o trabalho do bibliotecário sem o uso de tecnologias.

E é neste contexto, que se há alguma certeza para Bibliotecários, que almejam permanecer com o seu grau de empregabilidade crescente no mercado de trabalho, cada dia mais global, essa certeza pode ser chamada de muitos nomes: mudança, educação continuada, inovação, criatividade, adaptação, empatia, ação, resolução de problemas, construção de relacionamentos, entre outros.

Ribeiro e Ferreira (2018) afirmam sobre a necessidade das bibliotecas ressignificarem o seu papel, em razão das mudanças tecnológicas e sociais que impactam rotineiramente o comportamento, as demandas e o perfil de seus usuários. É inegável que existe uma inquietação quanto ao profissional do futuro e conseqüentemente, em como estão sendo formados hoje, esses profissionais. Conforme afirmam Nascimento; Figueiredo e Freitas (2003), independente às suas áreas de atuação, agora, mais do que nunca, será necessário recuperar informações, revisar conceitos, mantendo-se alertas e com os sentidos voltados para a curiosidade e o aperfeiçoamento.

No que toca ao aluno, é imprescindível que possua competências cognitivas necessárias para transcender do pensamento elementar e alcançar o pensamento crítico, que envolve a reorganização dinâmica do conhecimento de formas significativas e utilizáveis através de três competências gerais: avaliar, analisar e relacionar. (JONASSEN, 2007, p. 40).

O termo competência teve origem no ambiente empresarial na década de setenta e de acordo com a afirmação de Zabala e Arnai (2010), caracteriza uma pessoa capaz de realizar determinada tarefa real de forma eficiente. Entretanto, o que se entende hoje por eficiência do bibliotecário, possui uma abrangência que vai muito

além das competências compreendidas nos conteúdos curriculares, que se esforçam por acompanhar, adequando-se ao aumento das demandas informacionais atuais, que “há muito ultrapassaram os limites geográficos da biblioteca ou sistemas de informação tradicionais” (SOUTO, 2005). Por mais adequada que seja a formação acadêmica de um indivíduo, ela não conseguirá prepara-lo para todas as exigências do mercado de trabalho, visto que as transformações da sociedade ocorrem muito rapidamente, dando origem a novas carências e necessidades.

Há três anos, a Federação Internacional das Associações e Instituições Bibliotecárias - International Federation of Library Associations and Institutions - (Ifla), publicou um documento chamado *Ifla Trend Report 2016 Update*, que sintetiza o resultado de três anos de debates da comunidade internacional de bibliotecas. Esse documento questiona também, quais devem ser as habilidades necessárias aos bibliotecários para que possam auxiliar seus usuários a lidar com as informações digitais. Segundo Amaro (2018), esse também é um tema de reflexão da classe que analisa e discute quais são os novos papéis a serem desempenhados e qual será o perfil exigido dos profissionais na atualidade. Conforme Fabela (2005), a “cultura aprendente” se estabelece num ambiente de diversos atores o que contribui para a construção do conhecimento de forma partilhada.

No quadro da Figura 4, é apresentada a matriz de *swot* resultante da discussão ocorrida por um grupo de profissionais que se reuniu em janeiro de 2017, para discutir questões relacionadas aos desafios do bibliotecário nos dias atuais, sob o tema “O Bibliotecário do Século XXI”.

Figura 4: Matriz de Swot apresentada no workshop

Análise Swot	
<p>Forças</p> <p>Ressignificação do papel do bibliotecário na sociedade atual; Proatividade dos profissionais; Identidade do bibliotecário como profissional da informação; Atuação desvinculada aos ambientes tradicionais de trabalho; Regulamentação da profissão e competência profissional estabelecida por lei.</p>	<p>Oportunidades</p> <p>Crescimento do mercado de informação; Avanço das TICs; Necessidade crescente de informação na sociedade atual.</p>
<p>Fraquezas</p> <p>Estagnação justificada pela regulamentação da profissão; Identidade e atuação relacionadas exclusivamente a bibliotecas físicas; Falta de proatividade da classe profissional; Defasagem do currículo acadêmico; Cultura do papel do bibliotecário.</p>	<p>Ameaças</p> <p>Instabilidades do mercado de informação; Versatilidade no perfil dos concorrentes; Diminuição dos cargos público; Mudanças no perfil do usuário.</p>

Fonte: Adaptada de Côrte *et. al.* (2017)

Pode-se perceber que o mercado aparece ao mesmo tempo nos quadrantes de oportunidades e ameaças. Esse é um aspecto relevante, considerando que os bibliotecários que o veem como oportunidade, de acordo com Amaro (2018), tendem a ser profissionais competentes e proativos. Essa afirmação é corroborada por Takahashi (2000), quando afirma que os profissionais da informação precisam da contínua atualização e desenvolvimento de habilidades e competências, para atender aos novos requisitos técnico-econômicos e aumentar a empregabilidade.

Não há fórmulas nem manual na formação do profissional do futuro, então para que haja engajamento com o mercado e com a profissão, é preciso conhecer bem as funções inerentes à categoria e sobretudo quais habilidades desenvolver em favor de atitudes responsáveis e qualificadoras.

O perfil é delineado pelas habilidades, competências e atitudes necessárias para o desempenho da função profissional. A discussão dos problemas ligados ao perfil profissional é, na verdade, a discussão da função social da profissão, a qual, sujeita às influências do contexto, exige que a prática profissional se modifique, para atender expectativas novas e diversificadas que emergem da sociedade. (MUELLER, 1989, p. 63).

Muitos defendem a formação de um perfil especialista com foco no aprofundamento das especificidades advindas de cada área de atuação, contudo, não se pode esquecer que embora bibliotecários possam atuar em áreas diversas do conhecimento, todos trabalham com o mesmo objeto que é a informação, mantendo sempre o foco no usuário. Contudo, a relevância de sua participação em novas áreas contribuirá para a personificação, ou seja, a atribuição de qualidades pessoais, humanas, em atividades analíticas que por vezes, são executadas friamente por profissionais destinados e preparados a agir dessa maneira.

Na Figura 5, apresentamos um quadro contendo competências e habilidades pessoais necessárias aos profissionais da informação, para o desenvolvimento de qualquer que seja a sua área de atuação, elaboradas por Schumaker *et. al.* (2016). Após a sua análise cuidadosa, fica evidente que o bibliotecário do século XXI tem em mãos um importante papel de profissional da informação, não mais apenas, como o profissional dos livros.

Figura 5 : Competências dos Profissionais da Informação do século XXI

Competências Profissionais	
Área	Elementos
Serviços de informação e conhecimento	<ul style="list-style-type: none"> ✓ Reconhecer e articular necessidades de informação e conhecimento. ✓ Permitir o compartilhamento de conhecimento pelo contato e relações interpessoais, bem como por meio de sistemas e processos digitais ou eletrônicos. ✓ Priorizar o atendimento das necessidades operacionais ou estratégicas mais críticas da organização. ✓ Usar habilidades de gerenciamento de informação para aprender sobre um domínio, disciplina ou indústria. ✓ Compreender os diferentes aspectos do comportamento informacional humano. ✓ Analisar fluxos de informação e conhecimento relevantes para o contexto e as características da comunidade e os objetivos organizacionais. ✓ Defender a gestão e o uso efetivo de sistemas e processos de informação ✓ Ensinar, treinar e desenvolver competências informacionais e habilidades associadas para as partes interessadas. ✓ Aplicar os principais conhecimentos relacionados ao ambiente de trabalho para apoiar a missão organizacional.
Sistemas e tecnologias de informação e conhecimento	<ul style="list-style-type: none"> ✓ Envolver várias partes interessadas na recomendação de uma arquitetura da informação necessária para toda a organização ✓ Selecionar e usar ferramentas de gerenciamento de informações, como sistemas de gerenciamento de bibliotecas, sistemas de gerenciamento de conteúdo, plataformas de mídia social e ferramentas de recuperação e análise de informações. ✓ Desenhar interfaces para uma experiência de usuário intuitiva. ✓ Curar, publicar ou guardar informação em formatos utilizáveis. ✓ Selecionar e implementar sistemas de organização e conhecimento. ✓ Identificar sistemas e ferramentas para atender às necessidades de comunidades específicas. ✓ Fazer codificação usando <i>scripts</i> apropriados e outras ferramentas. ✓ Avaliar continuamente sistemas e tecnologias de informação e conhecimento.
Recursos de informação e conhecimento	<ul style="list-style-type: none"> ✓ Estabelecer orçamento para recursos e defender a alocação de fundos. ✓ Avaliar sistematicamente recursos novos ou desconhecidos mediante aplicação de estruturas e métodos analíticos. ✓ Gerenciar e entregar recursos relevantes de todos os tipos, mídia, e formatos, inclusive publicados e não publicados, internos e externos à organização, digitais, textuais, numéricos e visuais. ✓ Analisar continuamente a eficácia do portfólio de recursos, fazendo ajustes conforme necessário para garantir a relevância e fornecer aos usuários o melhor conteúdo de suporte à decisão. ✓ Realizar auditoria e mapeamento de recursos de informação e conhecimento disponíveis na organização, a fim de informar os usuários sobre os recursos relevantes para seus negócios. ✓ Alinhar a estratégia para gerenciar recursos de informação para apoiar os objetivos estratégicos das necessidades da organização e da comunidade. ✓ Fornecer recursos de informações autorizadas para atender às necessidades de um público em particular, cobrir um determinado tópico, campo ou disciplina, ou servir um propósito específico.

(Continua)

(Continuação)	
Competências Profissionais	
Área	Elementos
Recursos de informação e conhecimento	<ul style="list-style-type: none"> ✓ Negociar preços adequados e termos e condições para o licenciamento ou aquisição de recursos de informação. ✓ Identificar especialistas e fontes de facilitar o compartilhamento de conhecimento dentro das organizações. ✓ Ensinar a avaliar criticamente informações e fontes de informação.
Recuperação e análise de dados e informação	<ul style="list-style-type: none"> ✓ Entrevistar e consultar os membros da comunidade para identificar e esclarecer as necessidades de informação e conhecimento. ✓ Compreender os mecanismos de busca e os sistemas de recuperação de informação, incluindo as funcionalidades únicas fornecidas por diferentes sistemas e aplicar esse entendimento aos projetos de pesquisa e recuperação de informação ✓ Usar análise de dados, análise de texto, visualização e ferramentas similares apropriadas para analisar informações a fim de facilitar realização de inferências a extração de significados. ✓ Ensinar todas as competências em contextos formais e informais. ✓ Desenvolver sofisticadas estratégias de busca e recuperação para descobrir e recuperar informações de diferentes sistemas e repositórios. ✓ Avaliar a veracidade ou qualidade da informação e de suas fontes subjacentes nos motores de busca e sistemas de recuperação de informação. ✓ Comunicar os resultados dos projetos de recuperação e análise de informações de uma forma que seja útil ao público-alvo.
Organização de ativos de dados, informação e conhecimento	<ul style="list-style-type: none"> ✓ Usar padrões de metadados descritivos e temáticos reconhecidos para descrever ativos de informação. ✓ Desenvolver taxonomias e ontologias personalizadas, conforme as circunstâncias locais exigirem. ✓ Ministrando treinamentos sobre práticas efetivas para a organização e gerenciamento de informações ✓ Coordenar o desenvolvimento e a implementação de sistemas e processos de arquivamento personalizados para dar suporte às necessidades organizacionais. ✓ Desenvolver esquemas de metadados personalizados. ✓ Desenvolver políticas e procedimentos de retenção e destruição com base em requisitos legais e necessidades operacionais da organização. ✓ Aplicar práticas de controle de qualidade para assegurar a aplicação adequada de políticas e práticas para organização e gerenciamento de informações.
Ética da informação	<ul style="list-style-type: none"> ✓ Reconhecer questões éticas relativas à manipulação de informações, incluindo, entre outras, a privacidade, confidencialidade e segurança da informação, propriedade intelectual, direitos autorais e liberdade intelectual. ✓ Ensinar, influenciar e treinar pessoas sobre esse assunto. ✓ Avaliar e auditar a implementação organizacional de ética da informação. ✓ Modelar o comportamento ético relacionado à informação. ✓ Contribuir com as políticas, procedimentos e outras iniciativas organizacionais.

Fonte: Adaptada de Schumaker (2016).

Considerando as competências apresentadas na Figura 5, quando relacionadas à Figura 2 “A”, compreendemos que os atributos listados na página 37 para sistemas e tecnologias de informação e conhecimento e na página 38 para recuperação e análise de dados e informação, como também na mesma página, organização de ativos de dados, informação e conhecimento; todos relevantes, necessários e diretamente relacionados para que o bibliotecário possa executar atividades características da mineração de dados com maior propriedade e competência, como por exemplo, estruturar dados em planilha ou banco de dados relacionais, identificar padrões de comportamento dos dados, detectar pontos cegos da gestão, entre outros. E portanto, dito isso, considera-se atingido o objetivo específico abaixo:

c) Identificar possíveis relacionamentos entre as características da mineração de dados e as competências do bibliotecário, inerentes ou requeridas.

O mercado de trabalho mudou e não há mais motivação nem razão para querer ou pensar em construir uma carreira numa mesma organização por toda a vida.

[...] a forma de inserção no mercado de trabalho está mudando, e a quantidade de profissionais disponíveis não para de crescer. E, apesar da crise em curso, existem vagas principalmente em setores estratégicos e promissores. (SOUZA, 2018, p. 85).

As atribuições dos bacharéis em biblioteconomia são ratificadas pela Lei Federal nº 4.084 de 1962, que estabelece afazeres bastante restritivos aos profissionais, entretanto, a Classificação Brasileira de Ocupações – (CBO), do Ministério do Trabalho, estabelecida em 2002, apresenta uma ampliação das atividades, sobretudo quando menciona a prestação de serviços de consultoria e assessoria, o desenvolvimento de recursos informacionais, difusão cultural, desenvolvimento de estudos e pesquisas, entre outros. Essa visão ampliada, alarga também as possibilidades de áreas de atuação e conseqüentemente de abrangência de opções no mercado de trabalho, desde que, obviamente, haja qualificações do profissional, que estejam em conformidade com as possibilidades ofertadas. Não podemos deixar de mencionar que nesse modelo atual de inserção no mercado de trabalho, foram abaixo as barreiras de comunicação e geográficas, nos proporcionando ganhos no aumento de eficiência, eficácia, produtividade e criatividade.

O efeito dessas mudanças, segundo Souza (2018) aliada à diversidade de suportes, papéis, funções, usos e expectativas dos usuários na relação com o tratamento e uso da informação, ratifica a afirmação de Cunha (2013), quando diz que em um ambiente no qual o campo de atuação está passando por competições e apropriação de limites, impactada por tecnologias, é possível observar o crescimento de novos espaços de trabalho, incluindo a diversificação do próprio trabalho dos bibliotecários. Com efeito, seria omissivo não mencionar que para que esse cenário de ampliação de possibilidades e ambientes ocorra de fato, há um longo caminho a ser percorrido, que exigirá, de acordo com Souza (2018), diálogos com profissionais de outras áreas, adaptação às demandas atuais, e atenção às novas competências requeridas, como o conhecimento de técnicas de comunicação, interação, informática e gestão.

5. CONCLUSÃO

O trabalho aqui realizado estabelece relações diversas do tripé formado por bibliotecários, usuários e a mineração de dados. Essa investigação norteada a partir de uma revisão bibliográfica composta por alguns autores da área, possibilitou traçar conclusões, todas elas ratificadas no conteúdo do próprio trabalho, e sobretudo nas Figuras 2 e 5, que correspondem a uma avaliação positiva do uso da mineração de dados na recuperação da informação, assim como também é positivo, o vislumbre de que o bibliotecário pode e deve ocupar no mercado de trabalho e na sociedade, o seu devido lugar de destaque, como foi um dia, embora, com um viés bastante diferente, considerando que a maioria das empresas possui bancos de dados gerados a partir de seus sistemas legados, muitas vezes, sem nenhuma utilização para a tomada de decisão.

É fato, que há trabalho de casa a fazer, barreiras a quebrar, usuários a estudar, relacionamentos com outras áreas a ampliar e educação continuar. Não há tempo a perder e nem do que se queixar, é hora de agir. Assim como os bibliotecários, a mineração de dados também tem desafios a enfrentar, como a heterogeneidade dos dados em banco de dados (texto, áudio, imagem, multimídia) e a escalabilidade dos algoritmos com relação ao aumento do tempo de execução conforme se aumenta a base de dados.

Tudo isso é claro, muito bem relacionado com o entendimento e acompanhamento dos usuários e suas comunidades, no que diz respeito a compreensão de suas necessidades informacionais e todas as mudanças e exigências as quais elas estão inseridas. É importante ratificar que deve partir, sobretudo do próprio bibliotecário, a incorporação de competências como as apresentadas na Figura 5, pois a partir delas, pode ser possível ressignificar o seu papel, em razão das mudanças tecnológicas e sociais.

REFERÊNCIAS

AMORIM, M. C; et al. **Improving Financial Time Series Prediction Using Exogenous Series and Neural Networks Committees**. In: International Joint Conference Onn.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS - ABNT. **ISO 9241-11**. Requisitos ergonômicos para o trabalho com dispositivos de interação visual. Parte 11: orientações sobre usabilidade. Rio de Janeiro, 2011.

Aprenda a extrair mais valor dos seus dados. Disponível em: https://www.sas.com/pt_br/offers/sem/hbr-from-data-to-action.html?utm_source=Google&utm_medium=Search&utm_campaign=Big_Data&qclid=Cj0KCQjw_r3nBRDxARIsAJjleF0bs0asppMGOV5ohM7XrkIBcUDcPIfVBMQUpq_iNVoxnY49Ghpp8vAaAl-nEALw_wcB. Acesso em: 16 de maio 2019 as 20:00h

BERRY, Michael J. A, LINOFF, Gordon. **Data Mining techniques : for marketing, sales, and customer support**. New York : J. Wiley E Sons, 1997.

BAXENDALE, P. B. **Machine-made index for technical literature: an experiment**. IBM, Journal of Research and Development, 1958.

BORGES, G. S. Bruzinga. **Indexação automática de documentos textuais: critérios essenciais**. 111 f. Dissertação (Mestrado em Ciência da Informação)- Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2009.

BEZERRA. Eduardo. GOLDSCHMIDT. Ronaldo. **Exemplos de aplicações de data mining no mercado brasileiro**. Disponível em: <https://computerworld.com.br/2016/06/27/exemplos-de-aplicacoes-de-data-mining-no-mercado-brasileiro/> acesso em: 17 de maio 2019 as 22:00h

BRAMER, M. **Undergraduate Topics in Computer Science - Principles of Data Mining**. Springer, 2007.

BRANSKI, R.M. **Recuperações de informações na web. Perspectivas em Ciência da Informação**. 2004. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/File/351/160>>. Acesso em: 2 de maio de 2019.

CARVALHO. Deborah Ribeiro. DALLAGASSA. Marcelo Rosano. **Mineração de dados: aplicações, ferramentas, tipos de aprendizado e outros subtemas**. Disponível em: <https://revistas.ufpr.br/atoz/article/view/41340/25332>. Acesso em: 17 de maio de 2019 as 22:00h

Ciência de dados. Disponível em: https://pt.wikipedia.org/wiki/Ci%C3%Aancia_de_dados#cite_note-1. Acesso em: 18 de maio 2019.

COOPER, W. S. A definition of relevance for information retrieval. Information Storage and Retrieval, 1971.

CÔRTEZ, Sérgio da Costa; PORCARO, Rosa Maria; LIFSCHITZ, Sérgio. **Mineração de Dados – Funcionalidades, Técnicas e Abordagens**. PUC-Rio/InfMCC10/02, 2002. Disponível em: < ftp://obaluae.inf.puc-rio.br/pub/docs/techreports/02_10_cortes.pdf>. Acesso em: 30 de maio 2019.

CROFT, W. B; RUGGLES, L. The implementation of a document retrieval system. In: ANNUAL ACM CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL, 1982, West Berlin, Germany. **Proceedings...** New York, NY: Springer-Verlag New York, 1982. p. 28-37 . EDMUNDSON, H. P. New methods in automatic extracting. J. ACM, 1969.

DEMO, Pedro. **Introdução ao Ensino da Metodologia da Ciência**. In: _____. Introdução à Metodologia da Ciência. São Paulo: Atlas, 1985.

DIAS, M.M.; PIRES, D. **Usos e usuários da informação**. São Carlos: Edufscar, 2004.

DIAS, L.C. **Redes: emergência e organização**. In: CASTRO, I.E.; GOMES, P.C.C.; CORRÊA, R.L. (Org.). Geografia: conceitos e temas. 8. ed. Rio de Janeiro: Bertrand Brasil, 2006.

DUDZIAK, Elisabeth Adriana. **Curadoria de dados de pesquisa: o que é isso e como começar?** São Paulo: SIBiUSP, Jan. 2019. Disponível em: < <https://www.sibi.usp.br/noticias/curadoria-de-dados-de-pesquisa-o-que-e-isso-e-como-comecar/>> Acesso em: 04 julho 2019.

JONASSEN, David H. **Computadores, Ferramentas cognitivas: Desenvolver o pensamento crítico nas escolas**. Porto: Porto Editora. 2007.

FACELLI, K, Lorena et.al. **Inteligência Artificial: Uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC. 2011.

FAYYAD, U et.al. **From Data Mining to Knowledge Discovery in Databases**. American Association for Artificial Intelligence, 1996.

FAYYAD, U; STOLORZ, P. **Data mining and kdd: Promise and challenges**. Future generation computer systems, 1997.

FIGUEIREDO, N.M. **Estudos de usuários**. In: FIGUEIREDO, N.M. Estudos de uso e usuários da informação. Brasília: IBICT, 1994.

First Steps Towards Strategies for Stronger Libraries in Latin America and the Caribbean. Disponível em: <https://www.ifla.org/>. Acesso em: 19 de maio 2019 as 20:00h

FREIRE, I.M. et al. **Estudos de usuários: o padrão que une três abordagens**. Ciência da Informação, 2002. Disponível em: <http://www.scielo.br/pdf/ci/v31n3/a11v31n3.pdf>. Acesso em: 16 de maio 2019.

FREIRE, Paulo. **Pedagogia do Oprimido**. Rio de Janeiro: Editora Paz e Terra. 1996.

FREDDY, Angel et al. **Uma revisão dos algoritmos de radicalização em língua portuguesa. Information Research**, 2007.

HAN, J; KAMBER, M. *Data Mining: Concepts and Techniques*. Elsevier, 2006.

HAYKIN, Simon. “**Redes neurais- Princípios e Prática**”. tradução da segunda edição, editora Bookman, 1999.

GARVIN, P. L. et al. **Some opinions concerning linguistics and reformation processing**. Washington, D. C.: Center for Applied Linguistics, May . Available from National Technical Information Service, 1969.

GOLDSCHMIDT, Ronaldo; PASSO, Emmanuel. *Data Mining. Um guia Prático: Conceitos, Técnicas, Ferramentas, Orientações e Aplicações*. São Paulo: Editora Campus, 2005.

GOFFMANN & MORRIS, T.G. **Bradford's law and library acquisitions. Nature**, 2016.

GOTTSCHALG-DUQUE, C. *Bibliotecas e mídias sociais. In: RIBEIRO, A.C.M.L.; 2016*

FERREIRA, A. C. G. (Orgs.). **Biblioteca do século XXI: desafios e perspectivas**. Brasília: Ipea, 2016.

GOTTSCHALG-DUQUE, C. (Org.). **Ciência da informação: estudos e práticas**. Brasília: Thesaurus, 2016.

KAFURE, Ivette. **Usabilidade da imagem na recuperação da informação no catálogo público de acesso em linha**. 2004. 311 f. Tese (Doutorado em Ciência da Informação) Universidade de Brasília, Brasília, 2004.

HONORATO, Daniel de F. et al. *Utilização da indexação automática para auxílio à construção de uma base de dados para a extração de conhecimento aplicada à doenças pépticas. In: I WORKSHOP DE COMPUTAÇÃO, 2004.*

K. KLEVERTON. **O que é UX – User Experience**. Disponível em: <https://www.hostinger.com.br/tutoriais/ux-o-que-e-user-experience/acesso> dia: 20 de maio 2019 as 22:00h

IBM - Brasil. Disponível em: <https://www.ibm.com/br-pt>. Acesso em: 20 de maio de 2019 as 22:00h

LAKATOS, Eva Maria. MARCONI, M. de A. **Fundamentos de metodologia científica**. 3. ed. rev. e aum. São Paulo: Atlas, 1991.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. Brasília: Briquet de Lemos, 2004.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. Brasília: Briquet de Lemos, 1993.

LAPA, Remi; CORREA, Renato. **Indexação automática no âmbito da Ciência da Informação no Brasil. Informação & Tecnologia (ITEC)**, Marília/João Pessoa, 2014.

LEI Nº 4.084, DE 30 DE JUNHO DE 1962. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/1950-1969/L4084.htm. Acesso em: 20 de maio 2019 as 21:00h

LIMA, Regina Célia Montenegro de. **Estudo bibliométrico: análise de citações no periódico "Scientometrics"**. 1984.

MAMFRIM, Flávia P. B. **Representação de conteúdo via indexação automática em textos integrais em língua portuguesa**. Ci. Inf., Brasília, v. 20, n. 2, p. 191-203, jul./dez. 1991.

MARCONI, M. D. A.; LAKATOS, E. M. **Técnicas de pesquisa: planejamento e execução de pesquisas, amostragens e técnicas de pesquisas, elaboração, análise e interpretação de dados**. 3. ed. São Paulo: Atlas, 1996.

MARON, M. E.; KUHNS, J. L.; RAY, L. C. **Probabilistic indexing: a statistical approach to the library problem**. In: **NATIONAL MEETING OF THE ASSOCIATION FOR COMPUTING MACHINERY**, 14., ACM, 1959,

MATTAR, F. N. **Pesquisa de marketing**. São Paulo: Atlas, 1996.

Ministério do trabalho. CBO. Classificação brasileira de ocupações disponível em: <http://www.mtecbo.gov.br/cbsite/pages/pesquisas/buscaportitulo.jsf>. Acesso em: 22 de maio de 2019 as 23:00h

MOENS, Marie-Francine; DUMORTIER, Jos. **Automatic abstracting of magazine articles: the creation of 'highlight' abstracts**. In: **ANNUAL ACM CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL**, 21, ACM SIGIR, 1998, Melbourne, Australia. **Proceedings...** New York, NY: ACM, 1998.

MOREIRO GONZÁLEZ, José Antonio. **El contenido de los documentos textuales: su análisis y representación mediante el lenguaje natural**. Gijón: Ediciones Trea, 2004.

MORIN, Edgar. **Introdução ao pensamento complexo**. 2. ed. Lisboa: Instituto Piaget, 1990.

MUELLER, S. M. P. **Reflexões sobre a formação profissional para biblioteconomia e sua relação com as demais profissões da informação**. *Transinformação*, Campinas, 1989.

Norma Técnica. ABNT NBR ISO 9241-11:2011. Disponível em: <https://www.abntcatalogo.com.br/norma.aspx?ID=86090>. Acesso em: 24 de maio de 2019 as 22:00h

QUONIAM, Luc. et al. **Inteligência obtida pela aplicação de data mining em base de teses francesas sobre o Brasil. Ciência da Informação**, Brasília, v. 30, n. 2, p. 20-28, maio/ago. 2001.

OBERHOFER, Cecília Alves. **Conceitos e princípios para avaliação de sistemas de informação. Ciência da Informação**.1983.

O'CONNOR, J. **Automatic subject recognition in scientific papers: an empirical study. Journal of the Association for Computing Machinery**, 1965.

OLIVEIRA, Elias et al. **Um modelo algébrico para representação, indexação e classificação automática de documentos digitais. Rev. Bras. Biblio. Doc., Nova Série, São Paulo, 2007.**

ROBREDO, Jaime. **A indexação automática de textos: o presente já entrou no futuro.** In: Machado, U. D. (Org.). Estudos avançados em ciência da informação, Brasília, DF: Associação dos Bibliotecários do Distrito Federal, 1982. v. 1, p. 235-274.

ROBREDO, Jaime; CUNHA, Murilo Bastos da. **Aplicação de técnicas infométricas para identificar a abrangência do léxico básico que caracteriza os processos de indexação e recuperação da informação. Ci. Inf., Brasília, v. 27, n. 1, p. 11-27, jan./abr. 1998.**

SACCONI, L. A. **Nossa gramática: teoria.** São Paulo, Brasil: Atual. 1991.

SALTON, Gerard. Automatic text analysis. **Science**, Apr. 1970.

SALTON, Gerard. Introduction to moder information retrieval. McGraw-Hill. 1983.

SALTON, Gerard. Recent studies in automatic text analysis and document retrieval. **Journal of the Association for Computing Machinery**, v.20, n.2, p.258-27, Apr. 1973.

SALTON, Gerard; SMITH, Maria. On the application of syntactic methodologies in automatic text analysis. In: BELKIN, N. J.; RIJSBERGEN, C.,J. Van (Eds.). Annual International Acmsigir Conference On Research and Development in Information Retrieval, 12, 1989.

SARACEVIC, Tefko. **Ciência da informação: origem, evolução e relações. Perspec. Ci. Inf., Belo Horizonte, 1996.**

Sete características importantes para diferenciar BI, Data Mining e Big Data. Disponível em: <https://www.aquare.la/7-caracteristicas-importantes-para-diferenciar-bi-data-mining-e-big-data/> acesso em: 30 de maio de 2019 as 19:00h

SHERA, J. H. – **Introduction to library Science; basic elements of library servisse.** Littleton, Colo, **Libraries Unlimited**, 1976.

SILVA, Jonathas Luiz Carvalho. **Uma análise sobre a identidade da biblioteconomia: perspectivas históricas e objeto de estudo.** 2. ed. Recife: Ed. Do Autor, 2012.

SILVEIRA. Debora Pricila. **O que é Data Science? Disponível em:** <https://www.oficinadanet.com.br/post/16919-o-que-e-data-science>. Acesso em: 24 de maio 2019 as 23:00h

Significado de Pesquisa exploratória. Disponível em: <https://www.significados.com.br/pesquisa-exploratoria/> acesso em: 24 de maio 2019 as 23:00h

SOUZA, F. das C. de. **A construção escolar do bibliotecário brasileiro: ontem, hoje, amanhã. Ciência da Informação.** Brasília, 1991.

WIVES, Leandro K. **Indexação de documentos textuais.** 1997. 19 f. Trabalho Monográfico - Disciplina de Sistemas de Banco de Dados (Programa de Pós-Graduação em Ciência da Computação)- Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 1997.

ZABALA, A.; ARNAI, L. **Como aprender e ensinar competências.** Porto Alegre: Artmed, 2010. 197p.