

UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO
CENTRO DE CIÊNCIAS HUMANAS E SOCIAIS
ESCOLA DE BIBLIOTECONOMIA

LUIZ EDUARDO BARBOZA DA SILVA

BIG DATA:
Explorando um novo horizonte para a Biblioteconomia

Rio de Janeiro

2015

LUIZ EDUARDO BARBOZA DA SILVA

BIG DATA:
explorando um novo horizonte para a Biblioteconomia

Trabalho de Conclusão de Curso
apresentado como requisito para a
obtenção do título de Bacharel em
Biblioteconomia, pela Escola de
Biblioteconomia, da Universidade
Federal do Estado do Rio de Janeiro.

Orientador: Prof. Dr. Cláudio José Silva Ribeiro

Rio de Janeiro

2015

S581b

SILVA, Luiz Eduardo B. da

Big data: explorando um novo horizonte para a Biblioteconomia / Luiz Eduardo B. da Silva; orientação de Cláudio José Silva Ribeiro. – Rio de Janeiro, 2015.

58 f.: il. color; 30 cm.

Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2015.

1. Big Data. 2. Biblioteconomia. 3. Bibliotecário.
4. Profissional da Informação. I. Título

CDD 027.004.6

LUIZ EDUARDO BARBOZA DA SILVA

BIG DATA:

explorando um novo horizonte para a Biblioteconomia

Trabalho de Conclusão de Curso
apresentado como requisito para a
obtenção do título de Bacharel em
Biblioteconomia, pela Escola de
Biblioteconomia, da Universidade
Federal do Estado do Rio de Janeiro.

Aprovado em: _____ de _____ de _____.

BANCA EXAMINADORA

Orientador

Prof. Cláudio José Silva Ribeiro

UNIRIO

Membro avaliador

Prof. Maria Simone de Menezes Alencar

UNIRIO

Membro avaliador

Prof. Tatiana de Almeida

UNIRIO

Dedico este trabalho à Deus,
como retorno de tudo que tem feito por mim,
e a minha mãe, Jussara Barboza,
pelo carinho, cuidado, atenção e educação.

AGRADECIMENTOS

A realização deste trabalho não seria viável sem o auxílio e a colaboração de certas pessoas, que de suas maneiras, contribuíram para a minha formação e a conclusão deste trabalho. Então, para elas manifesto um especial agradecimento.

Primeiramente a Deus, por todo o cuidado e amor para comigo e minha família. Aos meus pais por todo encorajamento, carinho, educação, motivação, repreensão e exortação. Aos meus familiares e irmãos, pois família é à base de tudo.

Ao meu orientador Cláudio Ribeiro, por ter aceitado o meu convite de orientação, e o ter realizado com excelência. Aos mestres, que pertenceram à banca Simone Alencar e Tatiana de Almeida, por terem aceitado o meu convite.

Agradeço aos meus irmãos de faculdade João Paulo e Mateus Locatelli, porque sem eles não consigo nem imaginar a conclusão desse curso. Adriane também tem um papel especial nesta história.

As minha amigas da Academia Brasileira de Letras, a bibliotecária Kátia Marquet, e as conservadoras Maria Luiza e Marly Castro, por todo ensino, aprendizado para o mercado de trabalho e para vida.

A minha célula por ter me apoiado nesses quatros anos de faculdade, aos meus líderes e amigos, Alexandre, Giácomo e Fábio Jr. Lembro-me também de Rafael e Matheus Marques, por terem me ajudado com este trabalho e por serem amigos mais chegados do que irmão. E minha amiga Liliane por me ajudar a controlar toda a ansiedade de final de faculdade e por ser meu ponto de equilíbrio, me aconselhando e me auxiliando.

RESUMO

Este estudo aborda a importância do Big Data no novo mundo da informação, seus conceitos e suas aplicações, também aborda o profissional de Biblioteconomia, seu histórico e suas atividades. Essa pesquisa mostra através de literatura, todas as atribuições de um bibliotecário e suas relações com a sociedade. É um estudo descritivo com caráter exploratório, mapeando bibliografias que envolvem os dois objetos de estudo. Com essa pesquisa é possível verificar que o Bibliotecário possui condições para atuar nas aplicações do Big Data e nas pesquisas científicas relacionadas com a mesma. Big Data é uma proposta real para algumas empresas e instituições, e o Bibliotecário pode participar desse processo.

Palavras-chave: Big Data. Biblioteconomia. Bibliotecário. Profissional da Informação.

ABSTRACT

This study addresses the importance of Big Data in the new world of information, its concepts and its applications, also addresses the professional librarianship, its history and its activities. This research shows through literature, all assignments of a librarian, and their relationship with society. It is a descriptive study with exploratory character mapping bibliographies involving the two objects of study. With this search you can check that the Librarian has conditions to act Big Data applications and in scientific research related to the same. The Big data is a real proposal for some companies and institutions, and the librarian can participate in this process.

Keywords: Big Data. Library Science. Librarian. Professional information

LISTA DE QUADROS

Quadro 1 – Comparação entre os estilos de profissional de Biblioteconomia no decorrer do século XX.....	20
Quadro 2 – Funções que o bibliotecário pode exercer.....	21
Quadro 3 – Etapas da Descrição.....	24
Quadro 4 – Etapas da classificação.....	26
Quadro 5 – Etapas da Indexação.....	28
Quadro 6 – Etapas da Disseminação Seletiva da Informação.....	29
Quadro 7 – Facetas do Big Data.....	44
Quadro 8 – Atuações do Bibliotecário no Big Data.....	47
Quadro 9 – Análise de resultados.....	49

LISTA DE ILUSTRAÇÕES

Figura 1 – Representação da Web Invisível.....	34
Figura 2 – Processo de análise do Big Data.....	40

SUMÁRIO

1 INTRODUÇÃO.....	12
1.1 Objetivos gerais.....	14
1.2 Objetivos específicos.....	14
2 JUSTIFICATIVA.....	15
3 PROCEDIMENTOS METODOLÓGICOS.....	16
3.1 Caracterização da pesquisa.....	16
4 PERSPECTIVAS TEÓRICAS	18
4.1 Biblioteconomia.....	18
4.2 Bibliotecário.....	19
4.2.1 ATRIBUIÇÕES DO BIBLIOTECÁRIO.....	21
4.3 Tratamento documental.....	23
4.3.1 DESCRIÇÃO.....	24
4.3.2 CLASSIFICAÇÃO.....	25
4.3.3 INDEXAÇÃO.....	28
4.3.4 DISSEMINAÇÃO SELETIVA DA INFORMAÇÃO.....	29
4.4 Big Data.....	30
4.4.1 O 5 V'S DO BIG DATA.....	34
4.5 Big Data Analytics.....	39
4.5.1 DISCOVERY.....	40
4.5.2 MODEL PLANNING.....	43
5 COMO O BIBLIOTECÁRIO PODE CONTRIBUIR COM O BIG DATA?.....	45
6 ANÁLISES E RESULTADOS.....	48
7 CONSIDERAÇÕES FINAIS.....	51
REFERÊNCIAS.....	53

1. INTRODUÇÃO

Desde antiguidade, o ser humano percebeu a necessidade de observar, aprender e registrar o conhecimento adquirido de forma tácita, sendo essa percepção, o processo natural de criação de dados ou informação. Com o decorrer do tempo, desde a invenção da escrita, um pouco depois a invenção da imprensa, o volume de informação cresceu de forma explosiva, no século 18 já existia a preocupação de controlar essa informação, sendo essa preocupação mais atual do que imaginamos (SILVA; PINHO NETO; DIAS, 2013).

Segundo Cavalcanti (2015), o mercado atual é controlado por dois tipos de produtos: intangíveis e tangíveis. Com a modernização tecnológica algumas empresas estão convertendo suas informações de tangíveis para intangíveis, ou seja, um simples processo de digitalização de um documento representa essa conversão. Hoje, a maior empresa do mundo é a Alphabet Inc.¹ que vale 513,60 bilhões de dólares², uma empresa cujo 90% dos seus produtos são intangíveis.

Atualmente são gerados milhões de dados por segundo no mundo inteiro. Em 15 minutos a humanidade gera o triplo de informação pertencente a maior biblioteca do mundo, a Library of Congress nos Estados Unidos, sendo que nem 40% da humanidade tem acesso a celular e computadores (PETRY, 2013).

Petry (2013) relata que toda essa informação, pode ser dividida em dois tipos: a informação estruturada e não estruturada. A informação estruturada é a que tem valor para o usuário, porém a informação não estruturada era considerada inútil e excluída sem pensar, pois acarretava em uma perda de tempo e investimento. Esse panorama em relação à informação não estruturada se modifica por causa da inserção do Big data³, a informação considerada inútil acaba ganhando valor, na mão de empresários e usuários que saibam usa-la de forma estratégica.

¹ Alphabet Inc. é uma *holding* e um conglomerado que possui diretamente várias empresas que foram pertencentes ou vinculadas ao Google, incluindo o próprio Google.

² Valor de cotação visualizado no dia 06/11/15

³ Big data é um termo utilizado para descrever grandes volumes de dados e que ganha cada vez mais relevância à medida que a sociedade se depara com um aumento sem precedentes no número de informações geradas a cada dia.

Big data tem sido uma ferramenta de auxílio para grandes empresas que têm percebido essa necessidade de gerenciar e analisar uma grande quantidade de informação gerada a cada minuto.

O interessante é que Big data não está sendo uma utopia e nem um sonho, ao contrário, é realidade, alguns exemplos atuais do uso do Big data em âmbito mundial: segundo Kalena (2014) na Universidade de Harvard o professor Luke Bornn, elaborou uma disciplina optativa para o curso de Estatística chamada *Learning from Big data*, onde visa juntar os princípios da estatística, análise de dados a procura de relação, com o Big Data, os alunos desenvolvem projetos utilizando Big data, para resolver problemas de empresas com dificuldades em controlar o volume de dados, essa disciplina está sofrendo uma reformulação e só será ofertada na primavera de 2017.

Um segundo exemplo é o CityPulse, Marvão (2015) relata que o governo da Holanda junto com a empresa Atos, atuante no Brasil, utiliza o Big Data como instrumento de segurança. O sistema capta todas as informações de câmeras e microfones de vigilância em tempo real, e através de informações e comportamentos já captados anteriormente, foram criados indicadores em que os cientistas de dados, podem premeditar alguma ação e mover as autoridade competentes para que nessa ação não ocorra algo trágico. Esse cenário pode-se ver na ficção, o seriado americano *Person of Interest* possui esse enredo, onde é desenvolvido um software que pode premeditar ações terroristas e violentas.

Este trabalho pretende gerar uma visão mais detalhada do Big data, mostrar suas vertentes Discovery e Model Planning. Discutir sobre as suas atividades, como é inserida nas empresas, sua contribuição, os tipos de profissionais qualificados para gerencia-lo, as etapas da informação, desde a entrada, o processamento e a saída e outros assuntos pertinentes.

O Bibliotecário como um profissional de caráter interdisciplinar, pode trabalhar desde bibliotecas, empresas, cultura, cinematografia, tele novelas, pesquisa, e-commerce e entre outros. Big data é um horizonte exponencial onde o Bibliotecário pode explorar e marcar sua presença, como um profissional apto para trabalhar com Big data.

Nesse âmbito, este trabalho pretende segregar as características da informação no Big data, dentre as quais podemos destacar os 5 “v’s”, velocidade, volume, variedade, veracidade e valor, além de comparar com as competências do bibliotecário, desde o trabalho de pesquisa, referência e processamento técnico (catalogar, classificar e indexar).

Este trabalho está estruturado de maneira com que haja uma fácil compreensão e de forma que algumas seções mesmo que isoladamente se conectam nas seções finais. Além desta seção, a seção 2 apresenta argumentos para a justificação do trabalho, a seção 3 discute sobre os processos metodológicos aplicados, as dificuldades e as tomadas de decisões referentes ao trabalho. A seção 4 apresenta as perspectivas teóricas de Biblioteconomia e Big Data de forma isolada, e acompanhando a linha de perspectiva teórica, a seção 5 conecta os dois objetos de estudo, a seção 6 apresenta os resultados da pesquisa relatada nos processos metodológicos. E o trabalho é finalizado com a seção 7 que apresenta as considerações finais.

1.1 Objetivos gerais

Investigar a ascensão do Big data, e discutir suas particularidades paralelas com a Biblioteconomia.

1.2 Objetivos específicos

- Definir o conceito de Big data e sua aplicação no Brasil
- Desmembrar o Big data e relacionar com as atividades biblioteconômicas
- Investigar um novo campo de atuação para bibliotecários
- Mapear e avaliar o uso atual do tema Big Data em discussões publicadas na internet.

2 JUSTIFICATIVA

Big data é uma abordagem vista como perspectiva para o gerenciamento de informação e dados de empresas e/ou instituições. Esta abordagem não surge apenas com o objetivo de gerenciar, mas também ser um caminho para analisar a informação. As empresas serão cada vez mais adeptas ao uso do Big data, para que haja um crescimento harmônico.

A partir de uma análise da literatura pesquisada, percebeu-se que as instituições e empresas fora do Brasil, têm usado o Big Data como abordagem para identificar através de estudos e análises de pontos de interesse na população. Outra vertente também utilizada é o linked data, que parafraseando o estudo de Marcondes (2012) é a informação conectada de forma congruente, onde o usuário pode encontrar mais conteúdo para agregar na sua possível pesquisa ou interesse.

Esse trabalho pretende ajudar a desmitificar as dúvidas em relação à abordagem de Big data, entende-la de um jeito funcional e conhecer todas as suas funções. Essa pesquisa também discute se a Biblioteconomia pode atuar como uma área que poderá agregar para a abordagem de Big data, além de explorar o lado profissional do bibliotecário, que possui competências, que impulsionarão as etapas de trabalho nesta abordagem, sempre elencados os 5 v's do Big data que percorrem paralelamente com o processamento técnico da Biblioteconomia.

O Big data foi escolhido como o tema central porque é uma temática que está em ascensão, tanto no mercado competitivo, na área educacional, na área de segurança, e entre outros, mas Big Data pode ser implantado em qualquer área, usado de forma inteligente e adequado.

No Brasil existe pouca literatura acadêmica que abrange essa temática, uma grande parte das referências bibliográficas desse estudo está escrito na língua inglesa, e provavelmente esse é o primeiro estudo acadêmico em âmbito nacional, que possui como objeto de estudo Big Data e a Biblioteconomia, com cruzamento detalhado de informação resultando em um estudo que possa desmitificar o primeiro objeto.

Serão apresentados conceitos e definições das duas áreas, com comentários que irão contribuir com a captação do conhecimento tácito do leitor.

Este trabalho tem o objetivo de auxiliar a todos àqueles que se interessam pelo tema Big data, e querem entender um pouco dessa nova área que está surgindo. Adicionalmente, esta pesquisa busca incrementar a área de atuação dos bibliotecários, para conhecer novas oportunidades e áreas de investigação. A pesquisa pretende sanar todas as dúvidas do leitor, apresentar ambientes que utilizam a abordagem Big data, e fornecer informações que possam agregar conhecimento de forma cognitiva.

3 PROCEDIMENTOS METODOLÓGICOS

Neste capítulo serão expostos os procedimentos metodológicos utilizados para a realização deste estudo, baseados nos conceitos de metodologia de Gil (2008). Este capítulo está subdividido em dois tópicos: caracterização da pesquisa e universo da pesquisa.

3.1 Caracterização da pesquisa

A metodologia utilizada foi à bibliográfica com caráter descritivo e exploratório. A pesquisa bibliográfica consiste, segundo Gil (2008), em “proporcionar maior familiaridade com o problema (explicitá-lo). Pode envolver levantamento bibliográfico, entrevistas com pessoas experientes no problema pesquisado. Geralmente, assume a forma de pesquisa bibliográfica e estudo de caso.” E a pesquisa documental Gil (2008) afirma que:

É muito parecida com a bibliográfica. A diferença está na natureza das fontes, pois esta forma vale-se de materiais que não receberam ainda um tratamento analítico, ou que ainda podem ser reelaborados de acordo com os objetos da pesquisa. Além de analisar os documentos de “primeira mão” (documentos de arquivos, igrejas, sindicatos, instituições etc.), existem também aqueles que já foram processados, mas podem receber outras interpretações, como relatórios de empresas, tabelas etc.

Desse modo o trabalho se baseia no primeiro momento, numa pesquisa bibliográfica em relação à Biblioteconomia, para mapear a literatura publicada nas bases BRAPCI, Capes, IBICT e Worldcat. E uma pesquisa documental no segundo momento, para mapear as informações disponibilizadas sobre o assunto Big Data em Sites e Blogs.

Foi utilizada a ferramenta Google Alerta, para levantar bibliografias novas, e verificar se tem tido publicações ou postagens relacionando Big Data com Biblioteconomia. Foram realizadas algumas estratégias de buscas através de termos específicos, as respostas foram programadas para serem enviadas por e-mail nos meses de setembro e outubro, mas o processo não obteve resultados, então foi realizada outra busca com as mesmas estratégias, mas com os termos na língua inglesa, que obteve um melhor retorno de referências, que foram classificadas e tabeladas com observações agregando o valor da análise.

Este trabalho foi desenvolvido em dois campos de investigação, a Biblioteconomia e Big Data. Com a finalidade de demonstrar suas respectivas áreas e elencar as interseções entre os dois universos, buscou-se registros sobre debates e discussões.

4 PERSPECTIVAS TEÓRICAS

4.1 Biblioteconomia

A Biblioteconomia é uma ciência antiga, que possui seus princípios em costumes milenares, conforme Chassot (2002 apud RUSSO, 2010), o ano 288 a.C, com o surgimento da Biblioteca de Alexandria instituída por Ptolomeu I, foi o período que surge a Biblioteconomia, pois já existia a preocupação e a necessidade de organizar e classificar todo o conhecimento descrito, em suportes documentais da época.

Targino (2006) define a biblioteconomia como:

a área do conhecimento que se ocupa com a organização e a administração das bibliotecas e outras unidades de informação, além da seleção, aquisição, organização e disseminação de publicações sob diferentes suportes físicos.

Luisa Orera Orera (2006) conceitua a Biblioteconomia de uma maneira mais aprofundada:

uma ciência documental que tem por objeto de estudo as bibliotecas, entendidas como sistemas de transmissão de informação para seus usuários. Tal transferência ocorre por meio das coleções, que são formadas por um conjunto organizado de documentos. Assim o documento é compreendido a partir da ótica da coleção que integra o acervo de uma biblioteca. A coleção pode ser vista como composta fundamentalmente por livros, pois estes foram durante anos o principal suporte de informação. Entretanto, essa perspectiva alterou-se, dando espaço para outros tipos de documentos, como as publicações periódicas, folhetos, manuscritos, músicas impressas. Acrescenta-se ainda que, devido às mudanças no mundo da informação e na editoração eletrônica, as coleções de documentos convencionais passaram a conviver com os documentos eletrônicos, configurando a biblioteca como uma intermediária, entre os usuários e os documentos de qualquer tipo, que trazem consigo as informações.

No Brasil, a Biblioteconomia começou a ser estruturada com os surgimentos de bibliotecas de ordens religiosas, como os de Beneditinos, Franciscanos e Jesuítas. Russo (2010) relata que a Biblioteconomia tem como o marco no país, a fundação da Biblioteca Nacional e “esta biblioteca tem sua origem na Biblioteca Real d’Ajuda, que foi trazida pela Corte Real de Portugal, ao se

refugiar na sua colônia mais próspera, em 1808, mas a sua fundação, oficial, só ocorre em 1810.”.

A Biblioteca Nacional é a maior biblioteca da América Latina, hoje, é uma das dez maiores bibliotecas do mundo, contendo mais de dez milhões de itens no seu acervo⁴, Russo (2010), continua com seu relato expressando que a “sua administração fica a cargo de religiosos – como na maioria das bibliotecas brasileiras do período colonial – e só em 1846 é nomeado um Doutor em Medicina, o primeiro não religioso para dirigi-la.” No dia 10 de Abril de 1915, a Biblioteca Nacional inaugurou o primeiro curso de Biblioteconomia no Brasil, e em 1969 foi incorporado à grade de formação da Universidade Federal do Estado do Rio de Janeiro.

4.2 Bibliotecário

O profissional de Biblioteconomia tem sofrido fenômenos externos conforme o passar o tempo, e tem se adaptado as evoluções tecnológicas e culturais, o que torna um profissional adaptável, além de possuir um caráter interdisciplinar. No dicionário Aurélio (2010) o bibliotecário tem a seguinte definição: “relativo à biblioteca; bibliotecal; aquele que superintende uma biblioteca.” e no dicionário de Antônio Houaiss (2009), “aquele que administra uma biblioteca; qualquer membro do corpo profissional da biblioteca”.

Russo (2010) delimitou alguns passos na trajetória do profissional de Biblioteconomia no século XX no Brasil, e os períodos históricos que influenciaram e modificaram a atuação do bibliotecário, apresentados no quadro 1.

⁴ Disponível em <http://www.bn.br/biblioteca-nacional/historico?nu_pagina=11> Acesso em Set. 2015

Quadro 1 – Comparação entre os estilos de profissional de Biblioteconomia no decorrer do século XX

Ano	Tipo	Atribuição
1911	Bibliotecário erudito	De formação fortemente humanista, ligado à cultura e às artes – aspecto que norteou a criação do primeiro curso de Biblioteconomia do país – o da Biblioteca Nacional.
1930	Bibliotecário de formação técnica	Foca as atividades de tratamento e organização de documentos – que inspirou os primeiros cursos de São Paulo.
1960	Bibliotecário ligado às entidades profissionais	O influenciado pelo reconhecimento oficial da profissão como de nível superior e a criação dos órgãos de classe.
1970	Bibliotecário pesquisador	Atuante nos cursos de pós-graduação e acompanhando o surgimento dos primeiros periódicos científicos na área.
1980	Bibliotecário como agente cultural	Um novo profissional diante da reformulação curricular dos cursos de Biblioteconomia.
1990	Profissional da informação	Uma nova terminologia, que concebia um profissional de formação mais abrangente, envolvendo o trabalho com documentos e/ou informação.
2000	Bibliotecário autônomo	Formado para atuar como consultor (analista, arquiteto e gestor da informação) no ambiente flexível das organizações.

Fonte: Russo (2000)

4.2.1 ATRIBUIÇÕES DO BIBLIOTECÁRIO

O Conselho Regional de Biblioteconomia - 7ª região - descreve as funções desenvolvidas pelo profissional de Biblioteconomia:

O profissional de Biblioteconomia desenvolve atividades de organização, tratamento, análise e recuperação de informações em diversos níveis e suportes físicos, por meios manuais e automatizados, com vistas ao atendimento das necessidades informacionais de todos os segmentos da sociedade, ao avanço científico-tecnológicos e ao desenvolvimento social do país.

Tarapanoff (2000) elenca algumas habilidades que o profissional de Biblioteconomia, precisa possuir para atuar na nova sociedade que está, totalmente inserida em vários contextos informacionais, o quadro 2 apresenta essas habilidades.

Quadro 2 – Funções que o bibliotecário pode exercer

Função	Descrição
Preservar a informação	A partir do pressuposto que os conteúdos informacionais estarão sempre sendo produzidos, caberá aos bibliotecários à responsabilidade de localização, normalização, indexação, utilização de padrões, protocolos, metadados para garantir seu acesso futuro.
Organizar a informação para uso	Em função do paradigma da biblioteca como instituição social, esta deverá se constituir em um organismo mediador entre os indivíduos e os conhecimentos que eles precisam; com isso, os bibliotecários deverão viabilizar o acesso às fontes de informação, em qualquer tipo de suporte, assim como para qualquer tipo de indivíduo.
Conectar-se em redes e participar de consórcios	Visando o compartilhamento de recursos e a ampliação do acesso à informação para seus usuários, os bibliotecários deverão se organizar em redes, assim como participar de projetos cooperativos, a fim de racionalizar recursos e otimizar resultados nas suas tarefas regulares.

Ser empreendedor	A nova sociedade irá cada vez mais propor desafios aos profissionais de todas as áreas e os bibliotecários deverão desenvolver habilidades de filtrar na rede de informações extensa que se descortina à sua frente as informações relevantes, inovando na criação de serviços e produtos de informação mais sofisticados.
Trabalhar a informação, agregar valor	Significa imprimir aos produtos e serviços uma diferenciação que os torne indispensáveis aos usuários, como análise de informações, divulgação de notícias recentes, preparação de gráficos, de percentuais, de indicadores, a fim de transformar as informações em conhecimento para seus usuários.
Socializar a informação	Como as bibliotecas se constituem em organizações sociais sem fins lucrativos, os bibliotecários deverão proporcionar a prestação de serviços para os indivíduos e para a sociedade, quer sejam de forma tangível (impressos) ou intangível (eletrônicos).
Educar para a utilização da informação	A sensibilização para o uso da informação deve se constituir em um objetivo constante da biblioteca e os bibliotecários deverão procurar se educar e educar seus usuários a utilizar as fontes de informação, por meio das redes disponíveis, desenvolvendo um papel de educador nesse ambiente.
Valorizar o conceito econômico da informação	A escassez de recursos é uma realidade em todas as unidades de informação e os bibliotecários deverão saber quando e quanto cobrar por serviços de informação.
Criar, pesquisar e consumir informação	Com objetivo de melhor tecer a sua rede de informações, o bibliotecário deve sair da posição de apoio e de intermediador de informações para a de criador e consumidor das mesmas, produzindo conhecimentos e assumindo papéis mais pró-ativos.

Fonte: Tarapanoff (2000)

O Conselho Regional de Biblioteconomia – 7ª Região - reforça esse quadro de Tarapanooff:

O profissional de Biblioteconomia, que tradicionalmente atua em bibliotecas, encontra novas frentes de trabalho em sistemas e redes de informação de setores públicos, empresariais e industriais, escritórios de assessoria e consultoria, organização de arquivos e de documentação particulares, ensino e pesquisa, podendo atuar como analista da informação, como gestor de serviços de informação e também na área de normalização.

4.3 Tratamento documental

Para alguns profissionais, o processamento técnico é a medula espinhal do bibliotecário, porque é quando o mesmo trata da informação para disponibilizá-la para a sociedade. Mey (1995) relata que o processamento técnico “é como uma forma de representar e descrever um documento, seja sob os aspectos de conteúdo temático ou descritivo.”. O processamento técnico, para Mata (2004), possui suas particularidades, dependendo da unidade de informação:

Neste contexto, deve-se levar em conta o perfil e as necessidades dos usuários, bem como, os tipos de instrumentos e/ou ferramentas adotados na normalização para facilitar a busca e recuperação de documentos. Os instrumentos, conforme aponta a literatura, englobam: tabelas de classificação e de notação de autor, códigos de catalogação, tesouros e cabeçalhos de assunto, programas de catalogação cooperativa (...), softwares e bases de dados.

Este trabalho abordará apenas algumas etapas do processamento técnico como: classificação, descrição (catalogação) e indexação, e de dentro da referência a disseminação seletiva da informação.

4.3.1 DESCRIÇÃO

Para Mata (2004) “a catalogação identifica qualquer tipo de documento, com o objetivo de organizar catálogos de bibliotecas e bibliografias, independentemente da forma física como se apresentam.” Mey (1995) também relata:

Catalogação é o estudo, preparação e organização de mensagens codificadas com base em itens existentes ou passíveis de inclusão em um ou vários acervos, de forma a permitir interseção entre as mensagens contidas nos itens e as mensagens internas dos usuários.

A catalogação utiliza-se de catálogos como meio de ferramenta para o usuário, atualmente, o tipo de catálogo mais usado é o catálogo online, onde bibliotecas que possuem software, disponibilizam seu catálogo na rede e os usuários acessam. Mey (1995) enfoca que:

Catálogo é um canal de comunicação estruturado, que veicula mensagens contidas nos itens, e sobre os itens, de um ou vários acervos, apresentando-se sob forma codificada e organizada, agrupadas por semelhanças, aos usuários desse(s) acervo(s).

Para o processo de catalogação, Bruna e Alves (2011), ordenam em quatro etapas, os passos que o bibliotecário deve seguir para realizar uma descrição minuciosa do item:

Quadro 3 – Etapas da Descrição

Nº	Etapa	Descrição da etapa
1ª	Leitura técnica	Análise para levantar os dados informacionais necessárias para representar o mesmo. Esses dados são retirados da Fonte principal de informação: página de rosto e outras páginas que a antecedem, capa, colofão, encartes, apêndices, anexos, glossários, bibliografias, contêiner.
2ª	Descrição bibliográfica	Processo de caracterização do item no qual será descrita a ficha catalográfica que consiste nas

		seguintes áreas: título, indicação de responsabilidade, edição, local de publicação, editora, data de publicação, descrição física, séries, área de notas.
3 ^a	Pontos de acesso	Dados na quais os usuários consultarão a representação do item, em que se dividem em principal e secundários, no qual o primeiro é a primeira informação registrada e a segunda são dados mais os pontos de acesso além do principal, denominadas pistas.
4 ^a	Dados de localização	É o meio em que permite o usuário localizar o item no acervo, em que em bibliotecas são denominadas números de chamada, que geralmente compreendem o código da biblioteca e o número de chamada.

Fonte: Bruna e Alves (2011)

A ferramenta de catalogação mais adotada no Brasil é a AACR2, “um código de catalogação usado de forma internacional e atualmente em uso conjunto com o formato MARC o que favorece o intercâmbio de dados bibliográficos e catalográficos de forma internacional.” (BRUNA e ALVES, 2011).

O formato MARC é um conjunto de códigos e designações de conteúdos definido para codificar registros que serão interpretados por máquina. Sua principal finalidade é possibilitar o intercâmbio de dados, ou seja, importar dados de diferentes instituições ou exportar dados de sua instituição para outros sistemas ou redes de bibliotecas através de programas de computador desenvolvidos especificamente para isto (BRUNA e ALVES, 2011).

4.3.2 CLASSIFICAÇÃO

Depois da descrição, é necessário classificar a informação, para que possa ser recuperada quando for preciso, Guinchat e Menou (1994) explicitam que a classificação:

É uma operação da descrição de conteúdo de documentos, pela qual determina-se o assunto principal e, eventualmente, um ou dois assuntos secundários que são traduzidos para o termo mais apropriado da linguagem documental utilizada. Os termos da classificação são expressos, em geral, por notações. Mas podem ser utilizado também em termos genéricos de uma linguagem combinatória, como os nomes dos campos ou os grupos de uma lista de descritores ou de um tesouro.

Guinchat e Menou (1994) elencam algumas etapas para classificar uma informação, de maneira com que seja recuperada de forma rápida. Vale lembrar que esse esquema foi estruturado na década de 90, porém é muito válido no âmbito de biblioteca:

Quadro 4 – Etapas da classificação

1 ^a	Determinação do assunto principal do documento.
2 ^a	Determinação da classe a que pertence o assunto principal.
3 ^a	Determinação dos aspectos formais secundários do documento (lugar, tempo, forma do documento e língua) nos casos em que a classificação traz estas indicações.
4 ^a	Pesquisa dos números que correspondem à classe escolhida
5 ^a	Construção ou seleção do número correspondente, de acordo com as regras da classificação utilizada.
6 ^a	Colocação do número de chamada no documento.
7 ^a	Indicação do número de classificação na ficha catalográfica.

Fonte: Guinchat e Menou (1994)

Para Campos (2001) no Universo de Trabalho da Classificação, Ranganathan demarca três planos de trabalho: Plano Ideacional, Plano Verbal, Plano Notacional.

- Plano Ideacional

Plano onde é feito toda a análise de conteúdo, e dos conceitos, independente dos termos e da notação. (CAMPOS, 2001)

- Plano Verbal

Plano onde existe a conversão da análise em termos, em que a linguagem representa a ideias ou conceito, ela deve ser livre de homonímia e sinonímia, particularmente em se tratando de uma linguagem classificatória que não é uma linguagem natural. (CAMPOS, 2001)

- Plano Notacional

Plano onde é feito a estrutura notacional, do plano ideacional representado pela linguagem pelo plano verbal. (CAMPOS, 2001)

Atualmente, os manuais mais utilizados nas bibliotecas brasileiras são as CDD⁵ e CDU⁶, pois se utilizam de notações que otimizam o trabalho de organizar os livros nas estantes. Além dessas duas, Campos (2011) relata que em 1933, Ranganathan elaborou uma nova forma de classificar, se utilizando de facetas. A *Colon Classification*. O Esquema Rigidamente Facetado se caracteriza porque "as facetas e suas sequências são predeterminadas para todos os assuntos e acompanham uma classe básica" (RANGANATHAN, 1967 apud CAMPOS, 2001), ou seja, cada classe básica tem uma fórmula facetada e todos os elementos da fórmula devem estar presentes no assunto, o que por vezes, não ocorre. É isso que torna o esquema rígido (KUMAR, 1981 apud CAMPOS, 2001).

As facetas são manifestações de categorias do objeto de estudo, é um termo genérico usado para denotar algum componente, pode ser um assunto básico ou um isolado de um assunto composto, tendo ainda, a função de formar renques, termos e números. No contexto das classificações especializadas, é definida como uma manifestação das cinco categorias fundamentais. As facetas são de dois tipos: facetas básicas (que agrupa as áreas do conhecimento) e facetas isoladas (conceitos). (CAMPOS, 2001)

⁵ A Classificação de Dewey (DC) é de autoria do bibliotecário Melvin Dewey. A primeira edição data de 1876 e contava com 1000 entradas.

⁶ A Classificação Decimal Universal (CDU) foi elaborada pelos advogados belgas Pierre Otlet e Henry Lafontaine a partir de 1895. Tem como base a Classificação de Dewey e é atualizada pela Federação Internacional da Documentação (FID).

4.3.3 INDEXAÇÃO

Depois de descrever a informação, e de classificar, é preciso transformar toda análise realizada em descritores para auxiliar no processo de recuperação. Guinchat e Menou (1994) ponderam que:

A indexação é uma das formas de descrição de conteúdo. É a operação pela qual escolhe-se os termos mais apropriados para descrever o conteúdo de um documento. Este conteúdo é expresso pelo vocabulário da linguagem documental escolhida pelo sistema e os termos são ordenados para constituir índices que servirão à pesquisa.

Atualmente a indexação tem sido utilizada em várias empresas para automatizar o processo, e torna a informação, recuperável com alto grau de precisão, pode-se perceber a atuação da indexação nas *hashtags* em redes sociais, sites de compartilhamentos de vídeos possuem recursos para indexar o vídeo, em todos os sites de e-commerce os produtos estão organizados e são recuperáveis através de descritores.

De acordo com Guinchat e Menou (1994) a indexação pode possuir vários níveis, conforme a necessidade informacional do sistema, e deve-se atentar apenas para os assuntos principais, a indexação genérica identifica vários assuntos que a informação pode possuir. Na indexação média o bibliotecário elenca uma dezena de descritores para apresentar um trabalho preciso e detalhado, e finalmente a indexação exaustiva, que é resultado de uma análise minuciosa e não deixa despercebido nenhum detalhe contido na informação.

No quadro 5, Guinchat e Menou (1994) estruturam a etapa de indexação da seguinte maneira:

Quadro 5 – Etapas da Indexação

1º	Tomar conhecimento do documento;
2º	Determinar o assunto principal do documento;
3º	Identificar os elementos do conteúdo que devem ser descritos e extrair os termos correspondentes;
4º	Verificar a pertinência dos termos escolhidos
5º	Traduzir os termos da linguagem natural nos termos correspondentes da

	linguagem documental se for o caso
6º	Verificar a pertinência da descrição
7º	Formalizar a descrição se o sistema prevê regras particulares de apresentação ou de escrita.

Fonte: Guinchat e Menou (1994)

4.3.4 DISSEMINAÇÃO SELETIVA DA INFORMAÇÃO

A Disseminação seletiva da informação (DSI) não é necessariamente algo novo em âmbito de biblioteca, Lunh (1961 apud EIRÃO, 2009) entende a DSI como “aquele serviço dentro de uma organização que se refere à canalização de novos itens de informação, vindos de quaisquer fontes, para aqueles pontos onde a probabilidade de utilização, em conexão com o interesse corrente do usuário, seja alta”.

Conforme Eirão (2009), DSI é a tradução de *Selective Dissemination Information*, onde estão empregadas as três palavras: Disseminação (difundir e propagar), Seletiva (relação à seleção e escolhas) e Informação (conhecimento ou notícia trazida ao conhecimento de uma pessoa ou do público). A DSI possui também encorpada na sua funcionalidade “as duas leis de Rangantham, “Poupe o tempo do usuário” e “Para cada leitor seu livro”, ajudam a entender a Disseminação Seletiva da Informação, que pode ser um serviço personalizado, de valor agregado e direcionado para a necessidade particular de cada usuário.”. (EIRÃO, 2009).

A DSI possui algumas etapas que contribui de forma funcional para bibliotecas e empresas, conforme o quadro 6.

Quadro 6 – Etapas da Disseminação Seletiva da Informação

1º	Levantamento do perfil de interesse dos usuários	Descrição detalhada da qualificação, especialidade, necessidades e interesses dos usuários.
2º	Análise e tradução dos perfis	Atribuição de descritores, palavras-chave e códigos legíveis pelo sistema, que representem os temas a

		serem recuperados.
3º	Arquivamento dos perfis	Armazenamento no sistema dos perfis dos usuários, para processamento automatizado.
4º	Recuperação da informação	Realizada por computador, pelo confronto dos perfis dos usuários com a base de dados.
5º	Controle de qualidade	Verificação realizada para teste dos resultados, a fim de identificar possíveis erros de estratégia e de linguagem.
6º	Expedição aos usuários	Envio das listagens e ficha de avaliação, após os controles de expedição.

Fonte: Eirão (2000)

4.4 Big data

Segundo Moraes (2013), nos Estados Unidos em 2012, houve a eleição para a presidência da República, apenas dois candidatos se destacavam na corrida eleitoral: Barack Obama, que buscava a reeleição, entrou na história no ano de 2008 como o primeiro presidente negro dos Estados Unidos, e do outro lado Mitt Romney do partido Republicano.

No planejamento para elaborar a plataforma de campanha, o representante republicano não inovou, e optou pelo padrão: propagandas na televisão, preencher as caixas de correios com folhetos ou mandar e-mails automáticos para a população americana. Barack Obama teve um posicionamento diferente, transformou mais uma vez as eleições de forma radical.

O candidato Barack Obama utilizou-se das redes sociais e do Big Data para tornar a sua campanha mais humanística possível. Ele contratou analistas, e montou um grande banco de dados, já tendo incluso todas as informações da campanha de 2008, onde cidadãos americanos se cadastravam no site do candidato democrata, e contratou também Rayd Ghani como cientista-chefe, um pesquisador com um grande currículo na área de análise de dados.

Na aplicação, os detalhes são no mínimo surpreendentes, um exemplo dado pelo consultor Andrew Rasiej, mostra o poder do Big Data e como este ajudou

Barack Obama na sua reeleição, como já dito, tornando a campanha mais humanística.

Imagine uma mãe de duas crianças que vive em uma cidadezinha no estado de Ohio, no centro-oeste dos Estados Unidos. Ela votou na última eleição, cadastrou-se e navegou algumas vezes no site de Obama [...] Seus filhos estudam em escola pública. Ela costuma tuitar sobre o meio ambiente e ainda mantém uma página no Facebook sobre comida orgânica. Tudo isso ficaria registrado no banco de dados da equipe de Obama, “A campanha mandaria para ela e-mails de Michelle Obama sobre as políticas ambientais planejadas pelo presidente e sobre educação pública”. [...] Apenas mulheres com perfil semelhante receberiam esse tipo de mensagem. (MORAES, 2013)

Big data também auxiliou na arrecadação de fundos para a campanha de Obama, segundo a revista Time, Barack Obama realizou um jantar na Costa Oeste com George Clonney para angariar recursos financeiros, querendo repetir o sucesso do jantar, Obama solicitou para a sua equipe de analistas, identificarem a celebridade que possui mais prestígio na Costa Leste, o resultado esperado foi à atriz Sarah Jessica Parker, e se repetiu o sucesso. No final de tudo houve a arrecadação de quase um bilhão de dólares. (MORAES, 2013)

Esse é um ótimo exemplo para demonstrar a amplitude do Big Data, um presidente mais uma vez inovando, conseguindo chegar ao seu objetivo, agora trazendo para o ambiente de negócio, muitas pessoas já perceberam o efeito do filtro bolha. Quando se faz uma pesquisa sobre algum livro em um e-commerce e logo depois utilizar uma rede social ou até um site, percebe-se que existem propagandas do mesmo livro que foi pesquisado antes, onde é navegado existe uma perseguição do produto pesquisado, o que acaba influenciando na compra.

Mas o que é Big Data? Existem variados tipos de conceitos e definições de Big Data, alguns autores são mais sucintos, outros mais utópicos e sonhadores, mas todos com o mesmo foco, que é a grande natureza de dados. Segundo Akerkar (2014, apud NESSELO e FACHINELLI, 2014):

Big Data refere-se a conjuntos de dados, cujo tamanho está além das capacidades da tecnologia de banco de dados atual. É um campo emergente onde a tecnologia inovadora oferece alternativas para resolver os problemas inerentes que aparecem quando se trabalha com dados massivos, oferecendo novas maneiras de reutilizar e extrair valor a partir de informações.

Dumbill (2013, apud SHIRI, 2014, tradução nossa) deixa clara a limitação dos bancos de dados convencionais, o interessante é que existe a percepção da necessidade de Hardware potente para o uso do Big Data:

Big data são dados que excedem a capacidade de processamento dos sistemas de banco de dados convencionais. Os dados são muito grandes, se movem muito rápido, ou não se encaixam nas arquiteturas de banco de dados. Para ganhar valor a partir destes dados, você deve escolher um caminho alternativo para processá-lo.

O conceito de Big data ainda é pauta de muita discussão, por que ainda existe um pré-conceito em relação às empresas. Alguns dados para muitas empresas são desnecessários e facilmente descartados, mas “para ganhar o valor a partir desses dados, você deve escolher um caminho alternativo para processá-lo” (DUMBILL, 2013 apud SHIRI, 2014).

A Internacional Business Machines (IBM) é a maior empresa da área de tecnologia da informação, pioneira do Big Data no Brasil e em outros países. Conforme o ranking BrandZ de 2015 a IBM é quarta marca mais valiosa do mundo, perdendo apenas para a Apple, Facebook e Google. Por ser pioneira em vários países latinos e europeus, a IBM tem se tornado referência no estudo de Big Data. Devido a grande demanda, foi criada a Big Data University, uma iniciativa que disponibiliza cursos online relacionados ao Big Data, alguns gratuitos e outros pagos.

A IBM (2015) conceitua Big Data da seguinte maneira:

Big data é um termo utilizado para descrever grandes volumes de dados e que ganha cada vez mais relevância à medida que a sociedade se depara com um aumento sem precedentes no número de informações geradas a cada dia. As dificuldades em armazenar, analisar e utilizar grandes conjuntos de dados tem sido um considerável gargalo para as companhias.

Um estudo realizado pela IBM estima que são produzidos mais de 2,5 quintilhões de bytes todo dia. Se cada byte fosse um litro de água, provavelmente, teria como renovar toda a água da Baía de Guanabara no Rio de Janeiro em um dia, em dois meses toda a água do planeta. De acordo com Letouzé (2012 apud NESSELO e FACHINELLI, 2014, p.2) os dados digitais disponíveis em 2010 a nível

mundial são de 1200 exabytes (1.288.490.188.800 Gigabytes), esse numero deve aumentar em 40% no decorrer dos próximos anos, mas os números de servidores dobrarão a cada 20 meses.

Parisier (2009) relata que a produção de dados da humanidade desde o ano zero até o ano de 2005, é mensurada em cinco bilhões de gigabytes no decorrer deste tempo. Toda essa produção milenar, hoje, é desbancada em dois dias, devido à produção diária de 900 mil posts em blogs, 50 milhões tweets, 60 milhões atualizações de status no Facebook e 210 bilhões de e-mails, isso apenas abordando o contexto social, não estão incluídos as informações produzidas cientificamente, informações empresarias, e não se pode esquecer da Deep Web. Especialistas relatam que temos acesso apenas a 0,2% da surface da internet, os 99,8% restantes pertencem a deep web ou web invisível.

De acordo com Alencar (2014) web invisível refere-se a todo conteúdo que não pode ser indexado pelos sites de busca e, dessa forma, não está disponível diretamente para quem navega na internet. Não se deve atribuir a Web invisível apenas pelo lado negativo, como pornografia, pirataria ou crimes, mas todas as bases especializadas com ferramentas de busca dedicadas, formatos de arquivos não convencionais, informação de acesso restrita, informação paga e páginas dinâmicas. Alguns exemplos de Deep Web são as informações de bases de dados governamentais ou de empresas, que em geral respondem a uma busca no site da instituição, artigos indexados em bases de dados pagas, informações dinâmicas, em geral com captcha e informações em FTP.

Figura 1 – Representação da Web Invisível



Fonte Brand Powder⁷

4.4.1 OS 5 V'S DO BIG DATA

De acordo com Ribeiro (2014), Big data tem como aspectos estruturais os quatro V's: volume, variedade, velocidade e variedade, mas esse trabalho irá incluir mais um V importante, que é o valor da informação.

4.4.1.1 Volume

O aspecto de volume é o que mais caracteriza o Big Data, essa enorme quantidade de dados criados por dia, foi o início de todos os questionamentos. Esses dados são gerados por máquinas, redes de informação empresariais, produção científica e redes sociais. A seguir serão dados alguns exemplos sobre essa enorme criação, tanto no aspecto temporal presente, mas como algumas novas tecnologias que serão constantes.

De acordo com o site Dataflok (2013), os novos carros com o uso de sensores e de câmeras traseiras e frontais irão produzir 2 petabytes (2.097.151.999.999.992 de gigabyte) de informações por ano. Além disso, o setor

⁷ Disponível em < <http://litterisconsulting.com.br/internet-profunda-segredos-riscos-e-ameacas>>
Acesso em Set. 2015

agrícola tem gerado quantidades maciças de dados com sensores instalados em tratores. Os aviões geram aproximadamente a cada ano 2,5 bilhões de terabytes de informação, a partir de sensores instalados nos motores e turbinas. A Shell tem usado sensores supersensíveis para encontrar petróleo em alguns poços, e se instalarem estes sensores em todos os seus 10.000 poços, irá coletar aproximadamente 10 exabytes (10.737.418.240 gigabytes) de dados anualmente. Isso não se compara com o Square Kilometre Array, o maior telescópio do mundo, que está em construção, e a partir de 2017, irá produzir um exabyte (1.073.741.824 gigabytes) de dados por dia.

Se essa produção maciça de dados fosse nas décadas passadas, estaríamos em sérios problemas, mas atualmente, com investimento em grandes servidores e armazenamento em hadoop⁸, este é problema é amenizado.

Mas em relação ao volume, também se deve atentar para a questão dos servidores, segundo a Target Host (2015) o melhor servidor do mundo é o Tianhe-2 da Universidade Nacional de Tecnologia de Defesa da China, possui mais de três milhões de núcleos de processamento, mais de um milhão de gibabytes de memória e mais de 33 mil TFLOP (operações de ponto flutuante por segundo). O Segundo é o Titan do Laboratório Nacional de Oak Ridge, nos EUA, possui cerca de 500 mil núcleos de processamento, e aproximadamente 700 mil gigabytes de memória.

De acordo com Higa (2015), no Brasil, o maior servidor é o Grifo04 da Petrobras, que ocupa o 122º no ranking mundial, possui 17.408 núcleos de processamento, 12.288 Gigabytes de memória e processamento de 251,5 TFLOP.

4.4.1.2 Variedade

O segundo aspecto do Big Data é a variedade dos dados, antigamente toda a informação criada era denominada dado estruturado, atualmente houve uma mudança drástica, de toda informação gerada, 90% é de caráter não estruturado

⁸ Hadoop é uma plataforma de software em Java de computação distribuída voltada para clusters e processamento de grandes massas de dados.

(DATAFLOQ, 2015). Com essa variedade de formatos de informação, podemos elencar os dados estruturados, dados não estruturados e os dados semiestruturados. O maior desafio é como elaborar uma abordagem de análise para cada tipo de dado, e como armazená-lo também.

Claro (2015) pondera algumas características para dados estruturados, mas vale ressaltar, que dados estruturados são todas informações textuais e que podem ser utilizadas em tabela ou em banco de dados:

- Dados organizados em blocos semânticos (relações)
- Dados de um mesmo grupo possuem as mesmas descrições (atributos)
- Descrições para todas as classes de um grupo possuem o mesmo formato (esquema)

Dados não estruturados são as informações que não são textuais, como fotografias, vídeos e atualização em redes sociais. Claro (2015) relaciona para os dados não estruturados os seguintes aspectos:

- São os dados que não possuem uma estrutura definida.
- Normalmente caracterizados por documentos textos, imagens, vídeos, etc.
- Nem as estruturas são descritas implicitamente
- A Grande maioria dos dados atuais na Web e nas empresas segue este formato
- Sistemas tradicionais estão em colapso para processar estes dados

Os dados semiestruturados são os que apresentam a maior dificuldade para serem definidos, Claro (2015) enumera os seguintes atributos:

- Os dados semiestruturados são dados onde o esquema de representação está presente (de forma explícita ou implícita)
- Auto descritivo
- Uma análise do dado deve ser feita para que a sua estrutura possa ser identificada e extraída.

4.4.1.3 Velocidade

O terceiro aspecto do Big Data é a velocidade. Hoje em dia, as informações são criadas em tempo real ou quase real, esse aspecto lida com a velocidade que a informação é criada, armazenada, disponibilizada e analisada. Mas a questão é como lidar com esse aumento de velocidade? Como analisar essa massa de dados?

Antigamente, a informação era enviada por lotes, os bancos de dados eram atualizados pelo período da noite (DATAFLOQ, 2015), os livros demoravam para serem publicados e até os periódicos também possuíam um tempo prolongado, não era uma informação instantânea. Atualmente, esse cenário mudou com o uso da internet e dos computadores de mão ou portáteis, não é mais preciso esperar a segunda edição de um jornal, que é considerado o mais atualizado do dia, todas as informações estão disponibilizadas online, basta acessá-las.

Novos conceitos estão surgindo em relação ao aspecto de velocidade. Um conceito que vale a pena ser citado é a informação viral, informação que se espalha de forma incontrolável pelas páginas e redes sociais. Um exemplo recente é a fotografia tirada pela Agence France-Presse, onde uma criança no dia 02/09/2015 foi encontrada morta na praia de Brodum na Turquia, logo depois de duas embarcações com imigrantes naufragaram, fugindo da crise dos seus países. A Globo.com (2015) relata que “A foto virou um dos assuntos mais comentados no Twitter e diversos veículos da imprensa internacional o destacaram como emblemática da gravidade da situação, até mesmo com potencial para ser um divisor de águas na política europeia para os imigrantes”.

Datafloq (2015) enfoca que lidar com a velocidade tem sido desafiador para as grandes empresas, em um minuto, são realizadas 2,5 milhões de pesquisa no Google, 300 mil tweets são publicadas no Twitter, 20 milhões de fotografias são carregadas no Flickr, e 30 mil são visualizadas, 200 milhões de e-mails são enviados, e no Youtube mais de 100 horas de vídeos são carregados.

4.4.1.4 Veracidade

O quarto aspecto do Big Data é a veracidade, o volume de informação pode ser grande, pode possuir a velocidade em tempo real, nos variados formatos, mas se não for verídico, não possuirá uma finalidade relevante. De acordo com a Dataflop (2015) as empresas precisam tomar cuidado, pois uma informação incorreta pode acarretar em muitos prejuízos para as mesmas. É preciso elaborar um sistema de análise de informação, que possa verificar a veracidade, porque numa tomada de decisão automatizada, existe a necessidade do menor número de erro possível.

No âmbito de rede social, a confiabilidade e a confusão de dados é algo que vale a pena ser observado, onde qualquer indivíduo pode falar o que pensa, e mascarar algumas informações.

Exemplificando temos uma recente boa ação do ator e músico Jonny Deep que veio ao Brasil com a sua banda, para tocar em um evento musical. De acordo com a Globo.com (2015) a banda disponibilizou mais de 200 aparelhos auditivos para a população carioca, e zerou a lista de espera do Sistema único de Saúde (SUS), mas na rede social Facebook, uma imagem do ator e músico com a seguinte frase “Os governantes desse país deviam ter vergonha por não cuidar do seu povo.”, está em circulação com mais de meio milhão de compartilhamentos, porém o problema é que não foi encontrada nenhuma fonte que comprovasse essa frase, então é uma informação que deve ser desconsiderada por falta de veracidade.

4.4.1.5 Valor

Conforme a Dataflop (2015), todos os dados arrecadados criam certo valor para a empresa, sociedades e consumidores, seja ele um valor proveitoso ou ruim. Big data significa um grande negócio e cada indústria vai colher os benefícios dessa massa de dados.

O valor não está incluído na porcentagem e nem no dado em si, o valor está nas análises feitas sobre os dados e como os dados são transformados em informações e, eventualmente, em conhecimento. O valor da informação tem uma importância para as organizações, pois usarão esses dados para transformar sua estrutura organizacional que se baseia em conhecimentos derivados de análises de dados para a sua tomada de decisão.

4.5 Big Data Analytics

A análise de Big Data, segundo Ribeiro (2014) tem como objetivo “executar a análise preditiva dos dados por meio da execução de mining⁹ (minerações)”, essa análise possui alguns passos: a fase Discovery é o primeiro, logo depois vem “a preparação de modelos (fase de data preparation e model planning) que serão úteis na construção do grande conjunto de dados, chamado de lago de dados (data lake¹⁰).”. Segundo Rouse (2012, tradução nossa) Big Data Analytics é:

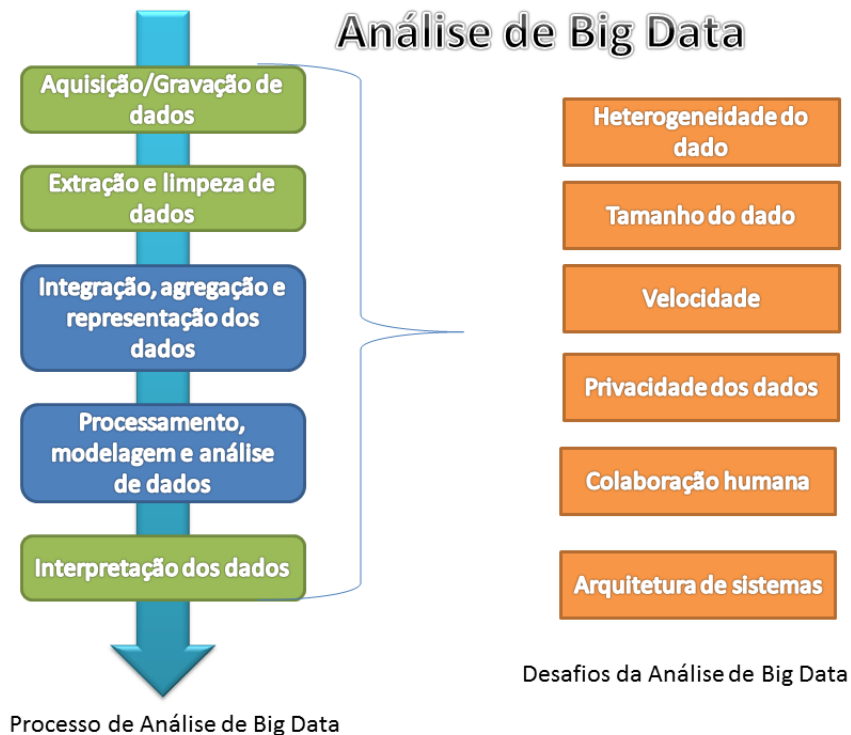
o processo de examinar grandes conjuntos de dados contendo uma variedade de tipos de dados - ou seja, Big Data - para descobrir padrões escondidos, correlações desconhecidos, tendências de mercado, preferências dos clientes e outras informações úteis sobre negócios.

Agrawal et. al.(2015), elaboraram um processo para a análise de Big Data, esse processo foi criado em conjunto com vários especialistas de empresas renomadas como Microsoft, IBM, Google HP e Yahoo! Na imagem a seguir, os grifos em verdes, são relacionados à fase de Discovery, e os grifos azuis: Data Preparation e Model Planning, e os grifos laranjas apresentam os desafios do Big Data.

⁹ A noção de mining de dados passa pela extração e análise de grandes volumes de informação em busca de padrões e comportamentos. (RIBEIRO, 2014)

¹⁰ Um lago de dados é um repositório de armazenamento que contém uma grande quantidade de dados bruto no seu formato original até que seja necessário.

Figura 2 – Processo de Análise de Big Data



Fonte: Agrawal et. al.(2015, tradução nossa)

4.5.1 Discovery

De acordo com Ribeiro (2014) “em suma, os projetos de Big Data são desenvolvidos com os objetivos de criar novos produtos, compreender novas necessidades dos clientes e seus comportamentos, bem como perceber novos mercados.” Oliveira (2013 apud RIBEIRO, 2014) levanta algumas questões que podem auxiliar na revisão e no entendimento da fase Discovery:

- Eu possuo o conhecimento suficiente do ambiente de dados e informação?
- Eu tenho informação suficiente para esboçar um plano analítico e compartilhar com meus pares?
- Eu consigo desenvolver trabalhos para organização para tipos de problemas? Categorizações e classificações de dados? Projeto de conjuntos (clusters) de dados?

- Eu consigo esboçar e realizar entrevistas para conhecer o contexto e domínio que será trabalhado?
- Eu posso identificar as diferentes fontes de dados?

Dumbill (2015, tradução nossa) apresenta uma ideia de percepção da fase do Discovery, que na tradução literal, é a descoberta do Big Data, quando existe a percepção e a necessidade de usar tal serviço:

muitas vezes a grandeza do volume e da velocidade é o que recebe o foco principal. A crescente variedade de entradas de dados, no entanto, é na verdade o cerne de produzir muitos benefícios de negócios para as organizações de hoje. E a adição de conjuntos de dados relacionados pode revelar insights poderosos.

Continuando no raciocínio, Dumbill (2015) apresentada três passos que cruzam com as indagações de Oliveira (2013 apud RIBEIRO, 2014): localizando dados, compreendendo a utilidade de uma série de dados, e dados multisourcing.

4.5.1.1 Localização de dados

Não existem mistérios nessa etapa, dados são produzidos em todo tempo, mas Dumbill (2015, tradução nossa) pondera o seguinte:

No mundo da ciência de dados, a obtenção de dados e a limpeza são frequentemente reconhecidas como 80 % do trabalho de qualquer análise, por causa dos bancos de dados dentro das organizações e um ecossistema complexo e opaco de fontes de dados externos. Os desenvolvedores de projetos de dados devem ser advertidos de que a descoberta de dados requer uma consideração significativa. E a descoberta de dados não é exclusivamente um problema técnico; uma estrutura organizacional é um fator chave no sucesso, usando os dados através de um negócio.

4.5.1.2 Compreendendo a utilidade de uma série de dados

Dumbill (2015) relata que depois de arrecadar um aglomerado de dados, deve-se ter a necessidade de compreender esses dados, que podem obter segmentos diferentes e meios diferenciados. O autor elenca quatro tipos de eixos

que a informação pode seguir: o primeiro eixo é a largura, onde é característico o dado possuir uma exaustiva ligação com outros tipos de dados, é muito semelhante com a ideia de linked data; o segundo eixo é a profundidade, tendo como característica, o dado possuir uma história, ou uma cronologia, um exemplo é o ano novo de um ano bissexto cair numa quinta-feira, existe uma explicação científica, essa explicação é a história; o terceiro eixo é a frequência, essa característica está relacionada à periodicidade da informação; e o quarto e último eixo é a latência, onde a informação tem o caráter de reação mais rápido, ou seja, que ajuda na tomada de decisão.

4.5.1.3 Dados Multisourcing

De acordo com Dumbill (2015), essa fase do Discovery, é logo depois da compreensão dos dados, e quando os requisitos não são preenchidos, podem-se consultar outras fontes, que precisam ser confiáveis e referências. O autor continua dizendo que:

Descobrir e selecionar as fontes de dados é um problema digno de consideração cuidadosa. As organizações não devem subestimar o esforço necessário para assegurar que os dados disponíveis são suficientes para resolver os problemas colocados pelas metas organizacionais. Através do multisourcing e a criação de serviços de dados, as organizações podem criar conjuntos de dados consistentes que sobrevivem à natureza evolutiva de qualquer negócio. (DUMBILL, 2015, tradução nossa).

É importante ressaltar que para ser feito à fase de Discovery, empresas especializadas em Big Data, realizam consultorias e oferecem esse serviço para as organizações que desejam aderir o uso do Big Data, mas cada empresa prestadora deste serviço possui seu próprio conceito em relação ao Discovery. Nesse trabalho foram abordados os conceitos de Dumbill (2015) que são utilizados pela IBM. Mas outras empresas como a Oracle¹¹, possuem seus próprios conceitos.

¹¹ Disponível em < <http://www.oracle.com/us/products/oracle-big-data-executive-brief-2415577.pdf> > Acesso em Out. de 2015

4.5.2 Model Planning

Todd (2012, tradução nossa) relata que o planejamento de modelo, é a “última etapa dos preparativos antes de executar os modelos analíticos e, como tal, exige que você complete o planejamento do trabalho analítico.”, mas para isso acontecer, é preciso levantar alguns questionamentos, conforme Oliveira (2013 apud RIBEIRO, 2014):

- Eu tenho um conjunto de dados que seja suficiente e de boa qualidade para iniciar a construção de um modelo?
- Eu tenho uma boa ideia sobre o tipo de modelo que vou testar?
- Eu posso refinar o modelo analítico?

A EMC (2011) declara que “O Big data requer uma nova plataforma de estudos analíticos a partir da qual tantos os negócios quanto à tecnologia possa obter vantagem competitiva”. O Big Data tem se tornado uma ferramenta poderosa em relação à tomada de decisão, por isso existe essa percepção de que é preciso novos estudos analíticos em relação ao Big Data. Logo depois a EMC (2011) continua relatando:

A possibilidade de apresentar novas escalas de capacidade de processamento para dar suporte a grandes conjuntos de dados permite a identificação contínua de percepções adicionáveis inerentes ao big data e proporciona a integração perfeita dessas percepções acionáveis ao ambiente de trabalho do usuário, onde quer que ele esteja.

Novo e Neves (2013), elencam uma visão importante de que “o big data abre a possibilidade de um modelo de decisão diferente, pois permite às empresas criação de experimentos controlados para testar hipóteses que guiarão a tomada de decisão”. Concluindo essa visão, a EMC (2011) versa que “essa nova plataforma de estudos analíticos pode liberar as organizações [...] estudos analíticos modernos e preditivos para as massas e aprimorando o processo de tomada de decisões em todos os níveis de organização.” e Brow (2011 apud NOVO E NEVES, 2013), finaliza dizendo que “possibilitando centenas ou milhares de experimentações é possível

distinguir entre simples correlação de eventos daqueles que realmente possuem uma ligação de causa e efeito”.

Ali Shiri (2014), elabora uma proposta biblioteconômica para auxiliar o Big Data na questão de Model Planning e Analytics, utilizando-se da Classificação Facetada elaborada por Ranganathan (1967) e aprimorada por Aitchison et al. (2002), que fornecem um conjunto mais específico e descritivo das categorias fundamentais que são úteis como uma base prática para análise facetada.

Quadro 7 – Facetas do Big Data

Facetas do Big Data	
Facetas propostas por Shiri	Categorias fundamentais propostos por Aitchison et al. (2002)
Atividade e operações	Energia
Pessoas	Agentes
Ambiente	Espaço
Data Type	Propriedade
Metadado	Entidades
Analytics	Tipo de Sistema

Fonte: Elaborado pelo autor

5 COMO O BIBLIOTECÁRIO PODE CONTRIBUIR COM O BIG DATA?

O Bibliotecário é um profissional de caráter interdisciplinar, podendo se adequar a qualquer campo semântico. Tálamo e Smit (2007) observam que “a área se relaciona tanto à organização de práticas científicas e profissionais, quanto ao acompanhamento das mudanças nas práticas culturais e nas modalidades de difusão e aquisições de conhecimentos.” Ribeiro (2014) afirma que também são observadas as variadas teorias e disciplinas que:

auxiliam na formalização da investigação em Ciência da Informação, dentre elas: teoria da informação, teoria das decisões, ciência cognitiva, lógica, análises em bibliometria, além da quantificação das estruturas do conhecimento e de seus efeitos. Assim, é possível propor que o profissional da informação acompanhe as pesquisas em Big Data, especialmente nos temas iniciais desenvolvidos na fase de Discovery, Data Preparation e Modeling Planning.

Gordon-Murnane (2012) relata que a manipulação e análise de dados são de suma importância para as tomadas de decisão de uma empresa, mas algumas empresas não estão sabendo explorar esses dados coletados de forma plena, e vários artigos e publicações na área de inteligência empresarial mostram que essa falha é resultado de outras falhas como: extrair valor de dados considerados irrelevantes, a má indexação e organização de informação na base de dados, estrutura de pesquisa mal formulada, a falta de conhecimento e de fontes que possam contribuir. O bibliotecário deve se envolver com Big Data, pois possui competência para organizar esse caos. Gordon-Murnane (2012, tradução nossa) elenca algumas habilidades do bibliotecário que podem ser aplicadas junto com o Big Data:

- Facilidade para a descoberta e recuperação de dados
- Manutenção da qualidade dos dados
- Através de catalogação, indexação e metadados, agregar valor aos dados
- Autenticação, arquivamento, gestão, preservação e representação de dados

Depois de elencar as habilidades, a bibliotecária Gordon-Murnane (2012, tradução nossa) cita que:

Os bibliotecários vêm fornecendo esses serviços para fontes de informação há muito tempo, acho que nós somos bem posicionados para mostrar o que sabemos e aplicá-lo aos desafios atuais e os problemas que enfrentam as empresas na era Big Data.

Desde 2014, os grandes centros de capacitação têm contratados especialistas e ofertados curso de especialização em Big Data, os cursos começaram como extensão, avançaram para pós-graduação, e hoje, existe graduação tecnológica em Big Data. Mas esses cursos possuem a sua matriz curricular voltada para o hardware, ou seja, para a estruturação de sistemas de dados, já existe esse curso, mas agora é da mesma escala do Big Data, ou seja, um grande sistema de dados estruturados.

Essa grande estrutura sistemática de dados não terá proveito, se não tiver uma boa programação de algoritmos para análise. A Biblioteconomia está engajada nesse processo, porque para criar um ótimo fluxo de informação é preciso delimitar algumas áreas, para que exista um trabalho congruente. As informações precisam ser indexadas de forma com que o sistema consiga lê-las, mas para isso, é preciso criar um vocabulário controlado baseado no campo semântico dos termos. Os estatísticos também estão se especializando em análise de dados, mas assim como os especialistas em Tecnologia da Informação, não possuem o lado humanístico para que a análise seja feita com o maior índice de acertos.

É importante ressaltar que o Bibliotecário, hoje, não pode trabalhar na estruturação do Big Data, mas sim nas suas aplicações.

Quadro 8 – Atuações do Bibliotecário no Big Data

Caso, situação ou processo de Big Data	Possível atuação do Bibliotecário
Eleição ganha por Barack Obama em 2012	Com o uso da Disseminação Seletiva da Informação, o Bibliotecário pode analisar e disseminar a informação adequada.
Em relação ao quarto “V” do Big Data, a Veracidade.	O bibliotecário sempre teve a necessidade de confirmar a veracidade da informação.
Em relação ao quinto “V” do Big Data, o Valor.	O profissional realiza o trabalho de mensuração da informação de forma estratégica, competências atribuídas por causa da Formação e Desenvolvimento de Coleções.
Big Data Analytics	O Bibliotecário possui domínio no fluxo de informação, e o fluxo do Big Data Analytics não se difere muito, então o Bibliotecário tem a condição de se adaptar, conhecer e auxiliar o processo de fluxo.
Localização de Dados e Multisourcing no Discovery	O Bibliotecário por ser especialista em fontes de informação consegue realizar essa atividade tranquilamente.
Compreendendo a utilidade de uma série de dados no Discovery	Alguns passos da classificação podem ser utilizados nessa etapa, por causa da análise da informação, sabendo descartar e utilizar os dados.

Model Planning	Esta etapa é onde o Bibliotecário pode trabalhar em peso. A modelagem de informação já é um exercício antigo, além disso, o Bibliotecário pode apontar e utilizar para a melhor forma de análise dos dados.
Pesquisa	O profissional de Biblioteconomia pode se interessar pelo Big Data, e estudá-lo de maneira científica, não apenas na sua aplicação, mas nos seus conceitos e seus desafios.

Fonte: Elaborado pelo autor

6 ANÁLISE E RESULTADOS

Por ser um trabalho de caráter exploratório e descritivo, foi utilizada a ferramenta Google Alerta da Alphabet, para mapear as publicações referentes à Biblioteconomia e ao Big Data. A pesquisa teve a duração de dois meses, tendo o começo no dia 01/09/15 e o término no dia 31/10/15. Foram programadas algumas pesquisas com combinações de descritores:

- Biblioteconomia and Big Data
- Ciência da Informação and Big Data
- Bibliotecário and Big Data
- Biblioteca and Big Data

O retorno obtido foi de nenhuma bibliografia encontrada, então houve a preocupação de ser perceber o que o mundo estava escrevendo sobre Biblioteconomia e Big Data de forma congruente. A segunda pesquisa referente à língua inglesa foi realizada do dia 25/09/15 ao dia 04/11/15, foram encontrados 21 referências, sendo 19 referências úteis para análise de dados, vale ressaltar que o Google alerta só trabalha com open data, as informações na web invisível não foram

recuperadas, porém provavelmente existem mais literaturas em base de dados. As pesquisas programadas foram com os seguintes termos:

- Library Science and Big Data
- Information Science and Big Data
- Librarian and Big Data
- Library and Big Data

Quadro 9 – Análise de resultados

Termos	Nº de alertas	Observação
Library and Big Data (Biblioteca e Big Data)	10 alertas	As referências deste tópico mostram algumas experiências de bibliotecas com Big Data, como a The British Library na Inglaterra; outra referência apresenta passos para inserir big data em bibliotecas de direito, e suas vantagens já que é um tipo de biblioteca com grande fluxo de informação; uma referência apresenta uma crítica sobre atuação do Big Data em uma determinada biblioteca; e outras referências são bem introdutórias e teorizam de forma sucinta o Big Data na biblioteca.
Librarian and Big Data (Bibliotecário e Big Data)	3 alertas	Essas referências comprovam que o bibliotecário pode trabalhar com Big Data, e possui totais condições para contribuir. Essas referências além de confirmar a proposta deste trabalho, mostram que existe essa preocupação de fazer com que o bibliotecário garanta seu lugar nas atividades com Big data, pois este tem contribuído com um novo mundo em que a informação é a matéria-prima mais valiosa.

Termos isolados	6 alertas	Algumas referências recuperadas apresentavam termos isolados, umas referências falavam apenas de Big Data, e foram apresentas termos em foram de contexto como Data and library, que não aborda Big Data, mas falava de biblioteca e os dados na biblioteca.
Referências não relevantes	2 alertas	Diferentes dos termos isolados, foram recuperados alguns termos, mas o artigo não abordava sobre biblioteca e nem Big Data, apenas foram citadas os termos dentro de outro campo semântico, a primeira psicologia e a outra nutrição.

Fonte: Elaborado pelo autor

Pode-se concluir que o Brasil está ainda iniciando na pesquisa de Big Data relacionada com Biblioteconomia em comparação aos outros países, na revisão bibliográfica foi possível identificar os pesquisadores brasileiros Cláudio José S. Ribeiro da Unirio e Marcos C. B. Cavalcanti da COPPE/UFRJ, que vêm desenvolvendo pesquisas sobre esta temática. Além das fontes utilizadas neste trabalho, não foram encontradas na surface publicações relacionadas à Ciência da Informação, e como já foi citado, as referências recuperadas confirmam a proposta desse trabalho, o bibliotecário tem espaço nas aplicações do Big Data.

7 CONSIDERAÇÕES FINAIS

O Big Data possui uma analogia interessante atribuída pela IBM (2013) “Big Data é um tsunami ainda em alto mar. Sabemos que vem mais e que ainda estamos brincando na praia. Mas, o assunto Big Data começa a chamar atenção.” Muitos estudos estão sendo realizados, muitas teses, alguns cursos estão sendo ofertados, ou seja, o Big Data tem tomado o seu espaço de maneira agressiva.

Através deste trabalho existe uma melhor compreensão em relação ao Big Data, foram encontradas algumas dificuldades, como pouca literatura na língua portuguesa, muitas referências apenas disponíveis em base de dados pagos, muitas teses relacionadas em apenas um objeto de estudo resultando em ambiguidade e incerteza na qual escolher para síntese. Dentro desse cenário podemos afirmar que o Big Data está se conhecendo, na fase de Discovery.

O Bibliotecário como profissional interdisciplinar, possui atribuições e competências para auxiliar e agregar valor ao Big Data, a questão escolhida nesse trabalho, foi respondida depois de uma análise exploratória e descritiva trazendo referências de especialista em suas respectivas áreas. Mas logo após essa conclusão outras perguntas são levantadas: será que o bibliotecário lutará por seu espaço? O Bibliotecário como profissional da informação irá se adequar as tecnologias?

A falta de literatura no Brasil é um fator preocupante, não apenas para a Biblioteconomia, mas também para os programadores e profissionais da Tecnologia da Informação, é preciso que haja incentivo a pesquisa sobre Big Data, uma área em aberto que pode ser preenchida por bibliotecários.

A escolha do tema deste estudo foi justamente para apresentar essa nova realidade para os bibliotecário e Bibliotecas, e ajudar a entender os conceitos de Big Data e suas aplicações, algumas Bibliotecas como Library of Congress e British Library estão utilizando Big Data para realizar alguns serviços.

É esperado que o Big Data cresça cada vez mais e comece tomar conta do cotidiano dos brasileiros, é esperado que o profissional bibliotecário alcance seu

espaço, e modifique o seu paradigma em relação à sociedade. Sabe-se que é um caminho longo a ser percorrido para que exista esse reconhecimento, portanto é fundamental que as instituições de ensino estejam preparadas para mostrar essas mudanças e encorajar aos futuros bibliotecários que existe Biblioteconomia fora da Biblioteca, a fim de coloca-los no mercado de trabalho como profissionais dispostos às novidades e mudanças na área.

“Cada vez mais se tornam indispensáveis os profissionais de mente aberta, atentos e flexíveis, capazes de enfrentar os desafios impostos pelas mudanças”. (MADUREIRA; VILARINHO, 2010).”

REFERÊNCIAS

AGRAWAL, Divyakant et al. **Challenges and Opportunities with Big Data**. Disponível em: <http://www.purdue.edu/discoverypark/cyber/assets/pdfs/BigDataWhitePaper.pdf>>. Acesso em: 01 out. 2015.

Aitchison, Jean, Gilchrist, Alan and Bawden, David. **Thesaurus construction and use: a practical manual**. London: Fitzroy Dearborn, 2002.

ALENCAR, Simone. **Técnicas de recuperação e disseminação da informação**. Rio de Janeiro: Video, 2014. 29 slides, color.

Biblioteca Nacional. **Primeiro curso de biblioteconomia no Brasil completa 100 anos**. Disponível em: <https://www.bn.br/noticia/2015/04/primeiro-curso-biblioteconomia-brasil-completa-100-anos>>. Acesso em: 10 ago. 2015.

BRUNA, Dayane; ALVES, Emanuele. Catalogação: análise e parâmetros gerais da representação da informação. **Xiv Encontro Regional de Estudantes de Biblioteconomia, Documentação, Ciência da Informação e Gestão da Informação**. Maranhão, jan. 2012. Disponível em: [http://rabci.org/rabci/sites/default/files/Catalogação análise e parâmetros gerais da representação da informação.pdf](http://rabci.org/rabci/sites/default/files/Catalogação%20análise%20e%20parâmetros%20gerais%20da%20representação%20da%20informação.pdf)>. Acesso em: 06 nov. 2015.

CAMPOS, Maria Luiza de Almeida. **Linguagem documentária: teorias que fundamentam sua elaboração**. Niterói, RJ: EdUFF, 2001.

CAVALCANTI, M. C. B. **Big data e pensamento complexo**. Rio de Janeiro: Video, 2015. 39 slides, color.

CLARO, Daniela Barreiro. **Dados extruturados x dados semiestruturados x dados não estruturados**. Bahia, 2015. Color. Disponível em: <http://homes.dcc.ufba.br/~dclaro/download/mate04/DadosEstruturadosxSemiEstruturadosxNaoEstruturados.pdf>>. Acesso em: 10 set. 2015.

CONSELHO REGIONAL DE BIBLIOTECONOMIA – 7ª. Região. **O profissional**. Disponível em:

<http://www.crb7.org.br/index.php?option=com_content&task=view&id=31&Itemid=1

>. Acesso em: SET. 2014.

DATAFLOQ. **Why The 3V's Are Not Sufficient To Describe Big Data**. Disponível em: <<https://datafloq.com/read/3vs-sufficient-describe-big-data/166>>. Acesso em: 10 out. 2015.

Dumbill, Edd. **Making sense of big data**. Big data 1: 1-2, 2013.

DUMBILL, Edd. **Data Discovery in the Data Value Chain**. Disponível em: <<http://www.ibmbigdatahub.com/blog/data-discovery-data-value-chain>>. Acesso em: 25 out. 2015.

EIRÃO, T. G. Disseminação seletiva da informação: uma abordagem. **Revista Digital de Biblioteconomia e Ciência da Informação**, v. 7, p. 20-29, 2009.

EMC. **Estudos analíticos do Big Data**. Disponível em: <<http://brazil.emc.com/collateral/emc-perspective/h8668-ep-cloud-big-data-analytics.pdf>>. Acesso em: 01 out. 2015.

FGV. **O Basômetro**. Disponível em: <<http://editora.fgv.br/blog/atualidades/o-basometro/>>. Acesso em: 18 set. 2015.

G1. **Foto chocante de menino morto revela crueldade de crise migratória**. Disponível em: <<http://g1.globo.com/mundo/noticia/2015/09/foto-chocante-de-menino-morto-vira-simbolo-da-crise-migratoria-europeia.html>>. Acesso em: 01 out. 2015.

G1. **Johnny Depp ajuda crianças com problema de audição no Rio**. Disponível em: <<http://g1.globo.com/musica/rock-in-rio/2015/noticia/2015/09/johnny-depp-ajuda-criancas-com-problema-de-audicao-no-rio.html>>. Acesso em: 21 out. 2015.

GORDON-MURNANE, Laura. **Big data: a big opportunity for librarians**. Disponível em: <<https://drive.google.com/file/d/0B9AibUksRnAYWTVrSU0tRmp0d28/view>>. Acesso em: 05 nov. 2015.

GUINCHAT, Claire, MENOU, Michael. **Introdução geral às ciências e técnicas da informação e documentação**. 2. ed. corr. e aum. Brasília: IBICT, 1994.

HENRIQUES, Marcos Santos Borges et al. BIG DATA. **Revista Pensar Tecnologia**, Belo Horizonte, v. 2, n. 3. jun. 2013. Disponível em: <http://revistapensar.com.br/tecnologia/pasta_upload/artigos/a55.pdf>. Acesso em: 18 maio 2015.

HIGA, Paulo. **Tianhe-2**: novo supercomputador mais rápido do mundo tem 1 petabyte de RAM. Disponível em: <<https://tecnoblog.net/133161/tianhe-2-supercomputador-mais-rapido/>>. Acesso em: 02 out. 2015.

HOLANDA, Aurélio Buarque de. **Dicionário Aurélio da língua portuguesa**. Coordenação de Marina Baird Ferreira, Margarida dos Anjos. 5 ed. Curitiba: Editora Positivo, 2010. 2222 p.

HOUAISS, Antônio. **Dicionário Houaiss da língua portuguesa**. Co-autoria de Mauro Villar, Francisco Manoel de Mello Franco; Instituto Antônio Houaiss de Lexicografia e Banco de Dados da Língua Portuguesa. 1. ed. Rio de Janeiro (RJ): Objetiva: Instituto Antônio Houaiss de Lexicografia, 2009. lix, 1986p.

IBM. **. Saiba o que é o Big Data e os desafios que as empresas enfrentam.** Disponível em: <http://www.ibm.com/midmarket/br/pt/infografico_bigdata.html>. Acesso em: 18 maio 2015.

IBM. **Cientista de Dados**: o profissional de Big Data. Disponível em: <https://www.ibm.com/developerworks/community/blogs/ctaurion/entry/cientista_de_dados_o_profissional_de_big_data?lang=en>. Acesso em: 06 nov. 2015.

KALENA, Fernanda. **Harvard coloca Big Data na sala de aula**. Disponível em: <<http://porvir.org/harvard-coloca-big-data-na-sala-de-aula/>>. Acesso em: 18 ago. 2015.

MADUREIRA, Helania Oliveira; VILARINHO, Lúcia Regina Goulart. A formação do bibliotecário para atuar em bibliotecas digitais: uma questão a aprofundar. **Perspectivas em Ciência da Informação**, v.15, n.3, p.87-106, set./dez. 2010. Disponível em: Acesso em: 21 ago. 2014.

Marcondes, Carlos Henrique . Linked data dados interligados - e interoperabilidade entre arquivos, bibliotecas e museus na web. **Encontros Bibli**, v. 17, p. 171-192-192, 2012.

MARVÃO, Susana. **Atos aproveita Big Data para tornar ruas mais seguras**. Disponível em: <<http://www.bitmaq.com.br/2015/08/atos-aproveita-big-data-para-tornar-ruas-mais-seguras/>>. Acesso em: 18 ago. 2015.

MATA, Maria Margarete Sell da. Biblioteconomia aplicada: experiência docente. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 9, n. 17, p.59-68, jan. 2004. 1º Semestre.

MEY, Eliane Serrão Alves. **Introdução à catalogação**. Brasília: Briquet de Lemos, 1995.

MORAES, M. **Big Data ajudou Obama a ganhar eleições**. Info, 15 janeiro 2013. Disponível em: <<http://www.cin.ufpe.br/~cjqf/%5Bpasta%5D%20TECNOLOGIA%20-%20material%20NAOCLASSIFICADO/BIG%20DATA%20revista%20Veja%202013.pdf>>. Acesso em: 18 maio 2015

NESELLO, P. ; FACHINELLI, Ana Cristina . Big Data: O novo desafio para a gestão. **Revista Inteligência Competitiva**, v. 4, p. 18-38, 2014.

NOVO, Rafael; NEVES, José Manoel Souza das. **Inovação na inteligência analítica por meio do Big Data**: características de diferenciação da abordagem tradicional. Disponível em: <http://www.cps.sp.gov.br/pos-graduacao/workshop-de-pos-graduacao-e-pesquisa/008-workshop-2013/trabalhos/desenvolvimento_de_tecnologia_e_sistemas/121191_32_44_FINAL.pdf>. Acesso em: 05 out. 2015.

ORACLE. **Big Data Discovery**. Disponível em: <<http://www.oracle.com/us/products/oracle-big-data-executive-brief-2415577.pdf>>. Acesso em: 22 out. 2015.

ORERA ORERA, Luisa. **Biblioteconomía**. In: LÓPEZ YEPES, José (Coord.). Manual de ciencias de la documentación. 2. ed. Madrid: Pirámide, 2006. p. 93-114.

PARISER, Eli. **O filtro invisível: o que a internet está escondendo de você**. Zahar, 2012.

PETRY, André. **O berço do Big Data**. Disponível em: <http://www.cin.ufpe.br/~cjgf/TECNOLOGIA_-_material_NAO-CLASSIFICADO/BIG_DATA_revista_Veja_2013.pdf>. Acesso em: 06 nov. 15.

RIBEIRO, C. J. S. . Big Data: os novos desafios para o profissional da informação. **Informação & Tecnologia**, v. 1, p. 96-105, 2014.

ROCHA, Cátia. **Pernambuco recorre a big data para previsão de imposto**. Disponível em: <<http://www.bitmag.com.br/2015/02/pernambuco-recorre-big-data-para-previsao-de-imposto/>>. Acesso em: 15 set. 2015.

ROUSE, Margaret. **Big data analytics definition**. Disponível em: <<http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics>>. Acesso em: 15 out. 2015.

RUSSO, Mariza. **Fundamentos de Biblioteconomia e Ciência da Informação**. Rio de Janeiro: E-papers, 2010. 177 p.

Schneider, R. D. **Hadoop For Dummies**, Special Edition. Mississauga, CAN: John Wiley & Sons Canada, 2012. 41 p.

SILVA, Maria Amélia Teixeira da; PINHO NETO, Júlio Afonso Sá de; DIAS, Guilherme Ataíde. Arquitetura da Informação para quê e para quem?: : uma reflexão a partir da prática em ambientes informacionais digitais. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Santa Catrina, v. 18, n. 37, p.283-302, maio 2013. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/viewFile/1518-2924.2013v18n37p283/25329>>. Acesso em: 06 nov. 15.

TÁLAMO, M. de F. G. M.; SMIT, J. W. Ciência da informação: pensamento informacional e integração disciplinar. **Brazilian Journal of Information Science**, v.1, n.1, jan./jun. 2007. Disponível em: <<http://www2.marilia.unesp.br/revistas/index.php/bjis/article/download/30/29>> Acesso em: 15 out. 2015

TARAPANOFF, K. O bibliotecário na sociedade pós-industrial. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 11. Florianópolis. Anais... Florianópolis: UFSC, 2000

Targethost. **Os maiores servidores do mundo**. Disponível em: <<https://www.targethost.com.br/maiores-servidores-do-mundo/>>. Acesso em: 02 out. 2015.

TARGINO, M. das G. A interdisciplinaridade da Ciência da Informação como área de pesquisa. In:_____.**Olhares e fragmentos**: cotidiano da Biblioteconomia e da Ciência da Informação. Teresina: Ed. da UFPI, 2006.,

TODD, Steve. **Phase 3 Innovation Analytics**: Model Planning. 2012. Disponível em: <http://stevetodd.typepad.com/my_weblog/2012/05/phase-3-innovation-analytics-.html>. Acesso em: 06 out. 2015.