



UNIVERSIDADE FEDERAL DA PARAÍBA - UFPB
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
MESTRADO EM CIÊNCIA DA INFORMAÇÃO

FABRÍCIO RODRIGUES DOS SANTOS GARRIDO

**BIBLIOTECONOMIA NO CONTEXTO DA *E-SCIENCE*: a formação do(a)
bibliotecário(a) de dados em instituições de ensino superior brasileiras**

João Pessoa, PB
2021



UNIVERSIDADE FEDERAL DA PARAÍBA - UFPB
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
MESTRADO EM CIÊNCIA DA INFORMAÇÃO

BIBLIOTECONOMIA NO CONTEXTO DA *E-SCIENCE*: a formação do(a)
bibliotecário(a) de dados em instituições de ensino superior brasileiras

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação do Centro de Ciências Sociais Aplicadas da Universidade Federal da Paraíba, na linha de pesquisa em Organização, Acesso e Uso da Informação, como critério parcial para aprovação no Mestrado em Ciência da Informação.

Orientadora: Profa. Dra. Gracy Kelli Martins
Gonçalves.

João Pessoa, PB
2021

Catálogo na publicação
Seção de Catalogação e Classificação

G241b Garrido, Fabrício Rodrigues dos Santos.

Biblioteconomia no contexto da e-Science : a formação do(a) bibliotecário(a) de dados em instituições de ensino superior brasileiras / Fabrício Rodrigues dos Santos Garrido. - João Pessoa, 2021.
184 f. : il.

Orientação: Gracy Kelli Martins Gonçalves.
Dissertação (Mestrado) - UFPB/CCSA.

1. Biblioteconomia - Formação profissional. 2. E-Science. 3. Bibliotecário(a) de dados. 4. Biblioteconomia - Ensino. 5. Ciência orientada a dados. I. Gonçalves, Gracy Kelli Martins. II. Título.

UFPB/BC

CDU 02 (043)



UNIVERSIDADE FEDERAL DA PARAÍBA CENTRO DE CIÊNCIAS
SOCIAIS APLICADAS PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
INFORMAÇÃO

ATA DE DEFESA DE DISSERTAÇÃO

Defesa nº 261

Ata da Sessão Pública de Defesa de
Dissertação do Mestrando **FABRÍCIO
RODRIGUES DOS SANTOS
GARRIDO** como requisito para obtenção
do grau de Mestre em Ciência da
Informação, Área de Concentração em
Informação, Conhecimento e Sociedade e
com Linha de Pesquisa em Organização,
Acesso e Uso da Informação.

Aos vinte e nove dias do mês de junho de dois mil e vinte e um (29/06/2021), das quatorze horas às dezoito horas e trinta minutos na sala virtual do Google Meet, conectaram-se via videoconferência a banca examinadora designada pelo Colegiado do Programa de PósGraduação em Ciência da Informação para avaliar o candidato ao Grau de Mestre em Ciência da Informação na Área de Concentração Informação, Conhecimento e Sociedade, o mestrando

FABRÍCIO RODRIGUES DOS SANTOS GARRIDO. Devido à pandemia do novo Coronavírus (COVID-19), considerando as estratégias de distanciamento social para contenção pandêmica e a Portaria Nº 323/GR/REITORIA/UFPB, de 16 de outubro de 2020, a videoconferência da defesa ocorreu com acesso por meio do link: <https://meet.google.com/wdo-zjtv-apo> . A banca examinadora foi composta pelos(as) professores(as): Dra. Gracy Kelli Martins Gonçalves - PPGCI/UFPB (Presidenta/Orientadora); Dr. Guilherme Ataíde Dias - PPGCI/UFPB (Examinador

Interno); Dra. Luana Farias Sales Marques - IBICT/UFRJ (Examinadora Externa); Dr. Henry Poncio Cruz de Oliveira - PPGCI/UFPB (Suplente Interno) e Dr. Denysson Axel Ribeiro Mota - UFCA (Suplente Externo). Dando início aos trabalhos, a Professora Dra. Gracy Kelli Martins Gonçalves, presidenta da banca examinadora, explicou aos presentes a finalidade da sessão e passou a palavra ao candidato para que fizesse oralmente a apresentação do trabalho de dissertação intitulado: **BIBLIOTECONOMIA NO CONTEXTO DA E-SCIENCE: a formação do bibliotecário de dados em instituições de ensino superior brasileiras**. Após a apresentação, o candidato foi arguido na forma regimental pelos examinadores. Respondidas todas as arguições, a Professora Dra. Gracy Kelli Martins Gonçalves, presidenta da banca examinadora, acatou todas as observações da banca e procedeu para o julgamento do trabalho, concluindo por atribuir-lhe o conceito:

☒Aprovado ☐Indeterminado ☐Reprovado.

Proclamados os resultados e encerrados os trabalhos, eu, Profa. Dra. Gracy Kelli Martins Gonçalves, presidenta da Banca Examinadora, lavrei a presente ata que segue assinada por mim, como representante dos participantes da banca, juntamente com os pareceres de avaliação da DISSERTAÇÃO e da defesa de dissertação do mestrando, devidamente assinados por seus respectivos avaliadores e em formato digital.

João Pessoa, 29 de junho de 2021.



Prof. Dra. Gracy Kelli Martins Gonçalves
Presidenta da Banca/Orientadora – PPGCI/UFPB

FABRÍCIO RODRIGUES DOS SANTOS GARRIDO

**BIBLIOTECONOMIA NO CONTEXTO DA *E-SCIENCE*: a formação do(a)
bibliotecário(a) de dados em instituições de ensino superior brasileiras**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação do Centro de Ciências Sociais Aplicadas da Universidade Federal da Paraíba, na linha de pesquisa em Organização, Acesso e Uso da Informação, como critério para conclusão do Mestrado em Ciência da Informação.

Aprovado em: ____/____/____

BANCA EXAMINADORA

Profa. Dra. Gracy Kelli Martins Gonçalves (Orientadora)
Universidade Federal da Paraíba (UFPB)

Prof. Dr. Guilherme Ataíde Dias (Examinador Interno)
Universidade Federal da Paraíba (UFPB)

Profa. Dra. Luana Farias Sales Marques (Examinadora Externa)
Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)

Em memória de vó e madrinha Maria de
Conceição Santos.
Fabrício Garrido

AGRADECIMENTOS

Agradeço, principalmente, aos meus pais Rose e Ademir que contribuíram em minha educação e formação inicial e as demais pessoas que de alguma forma também contribuíram Cecy Melo e Maria Lucy dos Santos, e ao maior incentivador (quase impositor) a fazer minha graduação, meu padrinho Arnaldo.

Ao corpo docente do PPGCI/UFPB, por todo conhecimento fornecido e paciência para com minhas perguntas e brincadeiras em sala de aula.

Em especial à minha orientadora, Profa. Dra. Gracy Kelli Martins Gonçalves, que iniciou sua orientação para comigo já na matrícula das primeiras disciplinas do mestrado, sempre com paciência e, o que ela tem de sobra, empatia sabendo dosar de forma impecável a liberdade que eu precisava para a pesquisa, e os simétricos puxões de orelhas igualmente necessários.

Aos Professore(a)s. Dra. Luana Farias Sales Marques, Prof. Dr. Guilherme Ataíde Dias, Prof. Dr. Denyson Axel Ribeiro Mota e Prof. Dr. Henry Poncio Cruz de Oliveira membros da banca que dispenderam, atenciosamente, seu tempo e disposição em meio a uma pandemia, para analisar e contribuir com este trabalho, nesse momento destaco tanto a Profa. Dra. Luana Marques como Prof. Dr. Guilherme Ataíde que foram além dos proformes habituais de uma banca, e forneceram um enorme contributo para pesquisa.

Aos profissionais extraclasse, e amigos pessoais, que contribuíram em minha formação, como Fernanda Mirelle, Zayr Cláudio e, em especial, à Rogério Marques e Rejane Freitas, sem cuja ajuda e paciência comigo dificilmente eu estaria aqui, agora, ou, pelo menos, o caminho percorrido até aqui teria sido bem mais árduo.

Aos meus antigos “chefes” Sheila Maluf, Nilton Generine e Laura Carvalho que me ensinaram tudo que sei sobre trabalho!

As minhas irmãs de consideração Ana Beck e Michele Beck, a esta devo eternamente pois graças a ela escolhi Biblioteconomia, e este curso não só foi a melhor escolha acadêmica de minha vida como foi quem me levou a conhecer minha esposa e, conseqüentemente, meu filho.

Aos meus irmãos Cláudia, Daniel e, em especial ao Kauê, que foi criado junto comigo e atormentou muito minha vida, mas também proporcionou algumas das melhores “farras” da minha vida.

Aos amigos que surgiram na minha vida de estudante de Biblioteconomia e se fixaram nela, e também fora da academia: os meus compadres Adriele (com bochechas destacadas) e Diógenes (Didica), Lidiane Herculano e em especial à Guilherme Martins (ex-Dumbo) e minha outra irmã de consideração Scheila Conrado (Tip-top).

E por fim, mas com certeza não menos importante, mas para estar em destaque, a mulher que me proporcionou ser e ter uma nova família, que me presenteou com nosso filho e o melhor: uma família com muito amor e cumplicidade, que só pode resultar em amor eterno, Taciany Kariny.

Sem ser dotado de força de corpo, pela
inteligência o homem se transformou no
mais forte dos animais, o mais terrível, e
mais criador, o mais destruidor.

O homem tem poder para transformar o
mundo num paraíso ou num deserto.

(Rubens Alves)

RESUMO

A profissão do(a) bibliotecário(a) surge de atividades tais como organizar, representar e disseminar informações, e a Ciência da Informação surge dessa mesma necessidade de forma expandida, após o evidenciado aumento exponencial de informação em meados do século XX. Os dados científicos existem desde o surgimento da ciência e a organização desses dados para uso por outros pesquisadores já foi verificado no século XVI. Porém, essa preocupação em organizá-los e disseminá-los se intensifica com o surgimento da *e-Science*, e está intimamente ligado com a profissão do(a) bibliotecário(a). Partindo desses pressupostos, este trabalho tem o objetivo analisar a formação do profissional da Biblioteconomia com enfoque em uma formação direcionada ao bibliotecário de dados, no contexto da *e-Science*, no Brasil. Para o alcance deste objetivo, delineou-se os seguintes objetivos específicos: discutir o surgimento e consolidação da *e-Science*; identificar o perfil profissional da Biblioteconomia de dados no contexto da *e-Science*; levantar os dados sobre os cursos de graduação em Biblioteconomia do Brasil; realizar um levantamento dos Projetos Pedagógicos de Curso; levantar e identificar nas ementas e programas de cada componente curricular quais se relacionam com a Biblioteconomia de Dados; e, por fim, consequentemente, revelar o panorama de formação dos cursos de Biblioteconomia do Brasil frente às demandas informacionais no contexto da *e-Science*. Trata-se, portanto, de uma pesquisa documental, de natureza qualitativa, que analisa os projetos políticos pedagógicos dos cursos de Biblioteconomia e utiliza a técnica de Análise de Conteúdo para identificar, nos seus respectivos conteúdos de formação, elementos que caracterizem as competências do profissional no contexto da *e-Science*, com base nos estudos desenvolvidos por Sales, Sayão, Maranhão, Drumond e Silva, no ano de 2019. Constatou-se que em todas as regiões brasileiras há ao menos um curso que aborde, direta ou indiretamente, alguma habilidade ou competência do(a) bibliotecário(a) de dados. No que tange a atualização dos PPCs, pouco mais de 50% foram atualizados nos últimos cinco anos, assim detectou-se a existência de um processo de atualização nos cursos, embora de forma não tão veloz. Há a expectativa que este estudo com seus dados possa contribuir para com os cursos que desejam abordar esse tema em sua matriz curricular.

Palavras-chave: *e-Science*; ciência orientada a dados; bibliotecário(a) de dados; ensino de Biblioteconomia.

ABSTRACT

The librarian's profession appears from activities such as organizing, representing and disseminating information, and Information Science appears from this same expanded need, after the evident exponential increase in information in the mid-20th century. Scientific data has existed since the emergence of science and the organization of this data for use by other researchers was already verified in the 16th century. However, this concern with organizing and disseminating them intensifies with the emergence of e-Science, and is closely *linked* to the profession of the Librarian. This work has the objective to analyze the training of the professional of Librarianship focusing on training directed to the data librarian, in the context of e-Science, in Brazil. To achieve this objective, the following specific objectives were outlined: to discuss the emergence and consolidation of e-Science; identify the profile of the data librarian in the context of e-Science; survey undergraduate courses in Library Science in Brazil; to raise the Pedagogical Course Projects; survey and identify in the menus and programs of each curricular component which are related to Data Librarianship; to point out the formation landscape of the Library Science courses in Brazil in the face of informational demands in the context of e-Science. This is a qualitative documentary research, which analyzes the pedagogical political projects of the Library Science courses, and uses the technique of Content Analysis to identify in their respective formation content, elements that characterize the professional's skills in the context of e-Science, based on the studies developed by Sales, Sayão, Maranhão, Drumond and Silva in 2019. It was found that in all Brazilian regions there is at least one course that directly or indirectly addresses some skill or competence of the data librarian. Regarding the updating of the PPCs, just over 50% were updated in the last five years, so it was detected the existence of an updating process in the courses, although not so fast. There is an expectation that this study with its data can contribute to courses that wish to address this topic in their curriculum.

Keywords: e-Science; data-driven Science; data librarian. librarian teaching.

LISTA DE FIGURAS

Figura 1 – Paradigmas da ciência	29
Figura 2 – Diagrama de Venn	37
Figura 3 – Base da <i>e-Science</i>	38
Figura 4 – Cauda longa da ciência	48
Figura 5 – Ciclo de vida dos dados	51
Figura 6 – Ciclo de vida dos dados para Ciência da Informação	54

LISTA DE QUADROS

Quadro 1 – Quadro de Definições sobre Dados	44
Quadro 2 – Princípios FAIR	56
Quadro 3 – Competências dos bibliotecários na gestão de dados de pesquisa	82
Quadro 4 – Termos relacionados a Bibliotecário(a) de dados na UFPB	95
Quadro 5 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFPB	96
Quadro 6 – Termos relacionados a Bibliotecário(a) de dados na UFCA	97
Quadro 7 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFCA	97
Quadro 8 – Termos relacionados a Bibliotecário(a) de dados na UFC	98
Quadro 9 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFC	99
Quadro 10 – Termos relacionados a Bibliotecário(a) de dados na UFPE	100
Quadro 11 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFPE	101
Quadro 12 – Termos relacionados a Bibliotecário(a) de dados na UFAL	102
Quadro 13 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFAL	102
Quadro 14 – Termos relacionados a Bibliotecário(a) de dados na UFS	103
Quadro 15 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFS	104
Quadro 16 – Termos relacionados a Bibliotecário(a) de dados na UFBA	105
Quadro 17 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFBA	106
Quadro 18 – Termos relacionados a Bibliotecário(a) de dados na UFRN	107
Quadro 19 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFRN	108
Quadro 20 – Termos relacionados a Bibliotecário(a) de dados na UNIR	110
Quadro 21 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UNIR	111
Quadro 22 – Termos relacionados a Bibliotecário(a) de dados na UFPA	112
Quadro 23 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFPA	113
Quadro 24 – Termos relacionados a bibliotecário(a) de dados na UFMT	115
Quadro 25 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFMT	116
Quadro 26 – Termos relacionados a Bibliotecário(a) de dados na UFG	116
Quadro 27 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFG	117
Quadro 28 – Termos relacionados a Bibliotecário(a) de dados na UnB	117
Quadro 29 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UnB	118
Quadro 30 – Termos relacionados a Bibliotecário(a) de dados na FESPSP	120
Quadro 31 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na FESPSP	121
Quadro 32 – Termos relacionados a Bibliotecário(a) de dados na UNIRIO (bacharelado)	123

Quadro 33 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UNIRIO	123
Quadro 34 – Termos relacionados a bibliotecário(a) de dados na UFRJ	124
Quadro 35 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFRJ	125
Quadro 36 – Termos relacionados a Bibliotecário(a) de dados na UNIVERSO	126
Quadro 37 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UNIVERSO	127
Quadro 38 – Termos relacionados a bibliotecário(a) de dados na PUC-Campinas	129
Quadro 39 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na PUC-Campinas	129
Quadro 40 – Termos relacionados a bibliotecário(a) de dados na USP	130
Quadro 41 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na USP	131
Quadro 42 – Termos relacionados a Bibliotecário(a) de dados na USP - Ribeirão Preto	133
Quadro 43 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na USP - Ribeirão Preto	134
Quadro 44 – Termos relacionados a Bibliotecário(a) de dados na UNESP	135
Quadro 45 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UNESP	136
Quadro 46 – Termos relacionados a Bibliotecário(a) de dados na UFSCar	137
Quadro 47 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFSCar	138
Quadro 48 – Termos relacionados a Bibliotecário(a) de dados na UFMG	139
Quadro 49 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFMG	140
Quadro 50 – Termos relacionados a Bibliotecário(a) de dados na UNIFORMG	140
Quadro 51 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UNIFORMG	141
Quadro 52 – Termos relacionados a bibliotecário(a) de dados na UFES	142
Quadro 53 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFES	143
Quadro 54 – Termos relacionados a bibliotecário(a) de dados na CEUCLAR	143
Quadro 55 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na CEUCLAR	144
Quadro 56 – Termos relacionados a bibliotecário(a) de dados na UFRGS	147
Quadro 57 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFRGS	148
Quadro 58 – Termos relacionados a bibliotecário(a) de dados na UEL	148
Quadro 59 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UEL	149
Quadro 60 – Termos relacionados a bibliotecário(a) de dados na UFSC	150
Quadro 61 – Ementas dos conteúdos curriculares relacionados com bibliotecário de dados na UFSC	151
Quadro 62 – Termos relacionados a bibliotecário(a) de dados na FURG	151
Quadro 63 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na FURG	152
Quadro 64 – Termos relacionados a bibliotecário(a) de dados na UDESC	153
Quadro 65 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UDESC	154
Quadro 66 – Termos relacionados a Bibliotecário(a) de dados na UNOCHAPECÓ	155

LISTA DE GRÁFICOS

Gráfico 1 – Panorama dos cursos ativos e não iniciados.	91
Gráfico 2 – Habilitação e modalidade de ofertas dos Cursos de Biblioteconomia.	91
Gráfico 3 – Tipo de categoria administrativa a qual pertence os Cursos de Biblioteconomia	92
Gráfico 4 – Quantidade de cursos de Biblioteconomia por Região, divididos entre públicos e privados	92
Gráfico 5 – Cursos de Biblioteconomia, segundo a modalidade	93
Gráfico 6 – Ano da última atualização dos PPCs das IES	157
Gráfico 7 – Disciplinas que abordam as temáticas relativas à Biblioteconomia de Dados/ <i>e-Science</i> .	158
Gráfico 8 – Cursos que abordam Biblioteconomia de Dados/ <i>e-Science</i> por região brasileira	159
Gráfico 9 – Segundo IES a quais abordam Biblioteconomia de Dados/ <i>e-Science</i> .	160
Gráfico 10 – Frequência de ocorrências dos termos	161

LISTA DE SIGLAS

ABEBD	Associação Brasileira de Escolas de Biblioteconomia e Documentação
ALA	American Library Association
ANCIB	Associação Nacional de Ciência da Informação e Biblioteconomia
BN	Biblioteca Nacional
CBBB	Congresso Brasileiro de Biblioteconomia e Documentação
CEPBB	Código de Ética Profissional bibliotecário(a) Brasileiro
CES	Câmara de Educação Superior
CEUCLAR	Claretiano Centro Universitário
CFB	Conselho Federal de Biblioteconomia
CFE	Conselho Federal de Educação
CGACGIES	Coordenação Geral de Avaliação de Cursos de Graduação e Instituições de Educação Superior
CNE	Conselho Nacional de Educação
CVD	Ciclo de Vida dos Dados
DASP	Departamento Administrativo do Serviço Público
DataONE	<i>Data Observation Network for Earth</i>
DCC	<i>Digital Curation Centre</i>
DCN	Diretrizes Curriculares Nacionais
DECIN	Departamento de Ciência da Informação
DTL	<i>Dutch Techcentre for the Life Sciences</i>
ECA	Escola de Comunicações e Artes
ELSP	Escola Livre e Sociologia Política
ENANCIB	Encontro Nacional de Pesquisa em Ciência da Informação
FaBCI	Faculdade de Biblioteconomia e Ciência da Informação
FEBAB	Federação Brasileira de Associações de Bibliotecários
FEFIEG	Federação das Escolas Isoladas do Estado da Guanabara
FESPSP	Fundação Escola de Sociologia e Política de São Paulo
FFCLRP	Faculdade de Filosofia Ciências e Letras de Ribeirão Preto
FGV	Fundação Getúlio Vargas
FURG	Universidade Federal do Rio Grande
IASSIST	<i>International Association for Social Science Information Service and Technology</i>
IBBD	Instituto Brasileiro de Bibliografia Documentação
IBICT	Instituto Brasileiro de Informação em Ciência e Tecnologia
IES	<i>Instituições de Ensino Superior</i>
LIBER	<i>Ligue des Bibliothèques Européennes de Recherche</i>
OCDE	Organização para Cooperação e Desenvolvimento Econômico
OST	<i>Office of Science and Technology</i>
PGD	Plano de Gestão de Dados
PPC	Projeto Pedagógico do Curso
PPP	Projeto Político Pedagógico
PUC	Campinas. Pontifícia Universidade Católica de Campinas
RCUK	<i>Research Councils United Kingdom, Research Councils United Kingdom</i>
SESu/MEC	Secretaria de Educação Superior do Ministério da Educação e do Desporto
STFC	<i>Science and Technology Facilities Council</i>
TDIC	Tecnologias Digitais de Informação e Comunicação

UDESC	Universidade Estadual de Santa Catarina
UEL	Universidade Estadual de Londrina
UFAL	Universidade Federal de Alagoas
UFBA	Universidade Federal da Bahia
UFC	Universidade Federal do Ceará
UFCA	Universidade Federal do Cariri
UFES	Universidade Federal do Espírito Santo
UFG	Universidade Federal de Goiás
UFMT	Universidade Federal do Mato Grosso
UFPA	Universidade Federal do Pará
UFPB	Universidade Federal da Paraíba, Universidade Federal da Paraíba
UFPE	Universidade Federal de Pernambuco
UFRGS	Universidade Federal do Rio Grande do Sul
UFRJ	Universidade Federal do Rio de Janeiro
UFRN	Universidade Federal do Rio Grande do Norte
UFS	Universidade Federal de Sergipe
UFSC	Universidade Federal de Santa Catarina
UFSCar	Universidade Federal de São Carlos
UnB	Universidade de Brasília
UNIFORMG	Centro Universitário de Formiga
UNIR	Universidade Federal de Rondônia
UNIRIO	Universidade Federal do Estado do Rio de Janeiro
UNIVERSO	Universidade Salgado de Oliveira
UNOCHAPECÓ	Universidade Comunitária da Região de Chapecó
USP	Universidade de São Paulo

SUMÁRIO

1 INTRODUÇÃO	21
1.1 JUSTIFICATIVA	23
1.2 QUESTÃO DA PESQUISA.....	26
1.3 OBJETIVOS.....	26
1.3.1 <i>Objetivo Geral</i>	27
1.3.2 <i>Objetivos Específicos</i>	27
2 QUARTO PARADIGMA CIENTÍFICO DE GRAY	28
3 DOS DADOS AOS METADADOS.....	39
3.1 DADOS.....	41
3.2 DADOS DE PESQUISA: OS CUIDADOS NECESSÁRIOS.....	45
3.3 METADADOS PARA DADOS DE PESQUISA	60
4 A ATUAÇÃO DO(A) BIBLIOTECÁRIO(A) E OS DADOS NA CIÊNCIA DA INFORMAÇÃO	64
4.1 BIBLIOTECONOMIA DE DADOS	71
4.2 BIBLIOTECÁRIO(A) DE DADOS.....	76
4.2.1 <i>Habilidades e competências</i>	79
5 PROCEDIMENTOS METODOLÓGICOS.....	84
6 ANÁLISE DOS CURSOS DE BIBLIOTECONOMIA.....	90
6.1 PANORAMAS GERAIS	90
6.2 ANÁLISE CURRICULAR	93
6.2.1 <i>Nordeste</i>	94
6.2.2 <i>Norte</i>	109
6.2.3 <i>Centro-Oeste</i>	115
6.2.4 <i>Sudeste</i>	119
6.2.5 <i>Sul</i>	146
6.3 PANORAMA GERAL DOS CURRÍCULOS	156
7 DISPOSIÇÕES FINAIS.....	163
REFERÊNCIAS.....	167
APÊNDICE A	179
APÊNDICE B.....	182
APÊNDICE C	184

1 INTRODUÇÃO

Na visão de Thomas Khun, a ciência vive em constante evolução marcada por revoluções científicas que levam às rupturas e às mudanças conhecidas como paradigmas. Jim Gray (2007) identificou quatro paradigmas científicos quando apresenta nitidamente as diferenças entre eles, porém, diferentemente dos paradigmas de Khun, eles não sofreram rupturas, mas se somam, tendo no quarto paradigma o três anteriores: o primeiro empírico, baseado em descrições de fenômenos naturais; o segundo teórico, baseado em modelos; o terceiro computacional, baseado em simulação de fenômenos complexos; e, atualmente, o quarto, com a exploração de dados baseada na associação da tecnologia computacional com o uso intensivo de dados, também conhecido como *e-Science*.

A *e-Science*, além de ser uma nova forma de se fazer pesquisa, é também um campo de estudo tendo como objeto a exploração de dados, por possuir definições diferenciadas em cada área do conhecimento, tanto no campo científico como no campo comercial, de natureza interdisciplinar, e seu estudo, enquanto campo científico, está fortemente atrelado à Ciência da Informação. Ferreira (2018) analisa um estudo realizado por Costa e Cunha (2015) em que "evidencia a pertinência de pesquisas relacionadas a *e-Science* entrelaçada à área da Ciência da Informação, devido a relevância e atualidade do tema e produção na área ainda pequena" (2018 p. 15-16).

Em 2001, foi realizado um painel internacional com 16 especialistas convocados pelo Comitê Diretor do *Research Councils United Kingdom* (RCUK), em nome de todos os Conselhos de Pesquisa do Reino Unido, com objetivo de revisar o Programa de Ciência Eletrônica do RCUK, com foco na *e-Science* e com atribuições de todos os conselhos de pesquisa, sendo essa uma das primeiras iniciativas relacionadas ao tema. Como entendimento geral, o painel aponta que "O antecedente do "e" na palavra *e-Science* raramente é explicitado, isto sem dúvida, começou como "eletrônico", mas agora significa melhor, 'melhorada' [*enhanced*] ou 'habilitada' [*enabled*]"¹ (RESEARCH COUNCILS UK, 2009, p. 13, tradução nossa).

Para este estudo se pondera o emprego da partícula “e” com o sentido do termo inglês *enhanced* (melhorada), por indicar uma melhor elucidação da Ciência contemporânea. Dessa forma, esta pesquisa adota o termo *e-Science* no intuito de ressaltar a partícula "e" a qual Ferreira (2018, p. 16) indica que "simboliza a grande

¹ Texto original: The antecedent of the “e” in *e-Science* is rarely made explicit. It no doubt began as “electronic” but now better stands for “enhanced” or “enabled”.

diferença que o eletrônico, melhorado e ampliado tem causado à ciência nos tempos atuais."

Para Gray (2007), do quarto paradigma emerge a Ciência de Dados, reconhecendo que a *e-Science* é o encontro da computação com a ciência, possibilitando pesquisas com uso e reúso intensivo de dados. Baseada nessa noção, percebe-se a busca por melhores definições de conceitos de dados, diferenciando-o do conceito de informação, bem como a necessidade de estudos de uma melhor gerência e curadoria desses dados, para que se libertem dos *hard disks* e *pen-drives* pessoais dos pesquisadores e alcancem repositórios digitais de dados, consolidando a pesquisa colaborativa advinda da *e-Science*.

Dessa forma, compreende-se que a Ciência de Dados emerge do *Big data*, proveniente principalmente do dilúvio de dados, originado pelas tecnologias digitais de comunicação e informação que, com sua popularização, e surgimento do Movimento Acesso Aberto, marcado pela Declaração de Budapeste de 2002, bem como a consolidação da Ciência Aberta (*Open Science*), culminaram numa maior facilidade de acesso, criação e compartilhamento de dados.

Para entender a Ciência de Dados, este trabalho pondera diferenciar dados e dados de pesquisa científica. Uma definição para o primeiro, baseada em Setzer (2015) e Borgman (2015), aponta como fatos, números, letras e símbolos que podem ser quantificados ou quantificáveis; o segundo são dados oriundos de pesquisas científicas e demandam um "tratamento especial", sendo essencial o seu uso para a Ciência de Dados. No que tange o tratamento desses dados, é preciso gerir todo seu ciclo de vida tomando cuidado com a coleta, o armazenamento, a recuperação e descarte, uma vez que eles serão usados, compartilhados e reusados, tanto pelo seu criador como por outros pesquisadores.

Dessa necessidade, surgem os princípios FAIR, acrônimo de: Encontráveis (*Findable*), Acessíveis (*Accessible*), Interoperáveis (*Interoperable*) e Reutilizáveis (*Reusable*). Para que esse uso e reúso sejam possíveis é essencial aplicação de metadados, que vai além da simples definição de "dados sobre dados", pois os metadados devem descrever os dados, tendo em vista que se trata de "[...] um objeto potencialmente informativo que descreve outro objeto potencialmente informativo²" (POMERANTZ, 2015, p. 26, tradução nossa).

² Texto original: [...] as potentially informative object that describes another potentially informative object.

Verifica-se assim que a *e-Science* está intimamente ligada à organização e à representação de dados, funções desenvolvidas tradicionalmente pelos bibliotecários(as) que, nesse contexto, como será visto, podem assumir a função de bibliotecário(a) de dados.

Tradicionalmente a Biblioteconomia fornece suporte à recuperação e ao acesso de coleções de materiais informacionais, por meio das técnicas de catalogação, indexação, classificação, serviços de referência, treinamentos, entre outros. No contexto da *e-Science* esses atributos se mantêm em relação ao tratamento do uso e reúso intensivo de dados dando suporte também à criação de metadados, além de promover treinamento aos pesquisadores para organização dessa nova forma de se fazer pesquisa, dando suporte às práticas recomendadas para tratamento de dados de pesquisa em todo o seu ciclo de vida, incluindo a estruturação do Plano de Gestão de Dados (PGD).

Sobre esse percurso de construção profissional estima-se que o primeiro bibliotecário que chegou ao Brasil foi Antônio Gonçalves, jesuíta, em 1604 (FONSECA, 1979 apud ALMEIDA; BAPTISTA, 2013), mas o primeiro curso de Biblioteconomia se iniciou mais de trezentos anos depois, quando a Biblioteca Nacional (BN), em 1911, o institui com quatro disciplinas de caráter humanista e, posteriormente, com disciplinas mais tecnicistas.

Em 1929, o colégio Mackenzie cria outro curso de caráter mais pragmático, sob influência estadunidense e poucos anos depois Adelpha Silva Rodrigues Figueiredo e Rubens Borba de Moraes criam a primeira Escola de Biblioteconomia de São Paulo seguindo uma terceira vertente, dando ênfase maior ao livro e erudição do(a) bibliotecário(a) (ALMEIDA; BAPTISTA, 2013).

Assim começa a caminhada bibliotecária no Brasil, através de interesses institucionais, com o objetivo de formar recursos humanos para trabalhar na BN. Somente a partir da criação do Curso da Escola de Sociologia de São Paulo é que a formação é direcionada para um perfil que atendesse às demandas informacionais da época, no que diz respeito ao mercado de trabalho.

1.1 Justificativa

Esta pesquisa se justificava por motivos pessoais, profissionais e científicos. Inicialmente, de maneira pessoal, minha atuação profissional, antes da formação em Biblioteconomia foi conduzida por atividades que me aproximavam dos estudos da tecnologia e da biblioteconomia, por trabalhar em uma editora de livros de informática,

onde meu trabalho era no desenvolvimento do sítio de *e-commerce* da empresa. Dessa experiência, ingressei no curso de Biblioteconomia e desenvolvi uma aptidão para os estudos referentes ao eixo tecnológico.

-

Na minha formação não foram abordados conteúdos que habilitassem seus egressos para atuação junto à *e-Science*. Só obtive conhecimento desse campo de estudos tempos depois, por uma colega de trabalho. Nas leituras e aprofundamento da temática, constatei a necessidade de buscar aperfeiçoamento para atuação neste contexto e, dessa forma, ampliar não só meus conhecimentos, como também minha atuação profissional.

Essa busca me inquietou a saber se, atualmente, dado o avanço das pesquisas sobre *e-Science* e gestão de dados, os cursos de Biblioteconomia já ofereciam uma formação que habilitasse o formando para atuar nesse nicho profissional. Desta inquietação, busquei por meio da pós-graduação ampliar meus conhecimentos e investigar como se encontrava o campo da formação profissional no Brasil quanto às competências desenvolvidas para que o bacharel em biblioteconomia possa atuar como bibliotecário de dados. Neste percurso, fundamentado pela literatura e por uma extensa busca nos currículos de Biblioteconomia no país, debrucei-me em uma investigação que atendesse aos meus anseios pessoais, profissionais e científicos, contribuindo para minha formação e para a discussão desse espaço de atuação que urge por profissionais capacitados para lidar com esse promissor contexto científico.

Dentre as justificativas possíveis, está o reconhecimento de que a organização, o acesso e o uso da informação, no escopo da Organização e Representação da Informação, seja ela física, eletrônica e/ou digital, visa à recuperação e ao acesso efetivo dos usuários (uso da informação), definida como espaço nuclear da Biblioteconomia (GUIMARÃES; FERREIR; FREITAS, 2013). Dessa forma, mediante a sua natureza mediadora entre a produção e o uso da informação, a Biblioteconomia está ligada à volumosa produção e o crescente compartilhamento de dados, quando grande parte da informação é gerada de forma individual ou colaborativa.

Nesse interim, as tecnologias digitais de informação e comunicação e a explosão de dados na internet implicam diretamente na atuação de diferentes profissionais. Destaca-se, nesse sentido, o(a) bibliotecário(a), como profissional da informação, visto que seu desempenho técnico e social se depara com a necessidade de novas formas de armazenamento, organização, acesso e uso da informação devido ao universo que o *Big*

data oferece, além dos diferentes conteúdos que precisam se readequar às condições e às técnicas de tratamento, visando torná-los acessíveis ao usuários no espaço digital.

De igual modo, os estudos sobre a formação do(a) bibliotecário(a) relacionado ao contexto do *Big data*, e especificamente à *e-Science*, no que tange aos processos de gestão/curadoria de dados, encontram-se em um nicho investigativo que busca discutir a importância desse profissional, tendo em vista sua capacitação técnica e de pesquisa, em uma aproximação às propostas de tratamento de dados.

Sobre essa questão, Anjos (2019) dedicou-se a investigar como os componentes inerentes a cada etapa do Ciclo de Vida dos Dados para a Ciência da Informação (SANT'ANA, 2013; 2019) estão contemplados nas disciplinas ministradas nos cursos de graduação em Biblioteconomia do Brasil. A autora utilizou como recorte a Nota de Conceito de Curso igual ou maior que 4, atribuída como conceito avaliativo pelo Ministério da Educação aos cursos de nível superior, sendo analisados nesse critério 21 cursos de Biblioteconomia no Brasil, obtendo como resultado que dentre os cursos analisados, uma parcela maior destes apresentam competências que podem ser aplicadas ao universo dos dados e ao Ciclo de Vida dos Dados em Ciência da informação. A pesquisadora aponta ainda que “não foram encontrados componentes curriculares obrigatórios que focassem diretamente na abordagem do quarto paradigma científico, no processo de gestão de dados ou no ciclo de vida dos mesmos” (2019, p. 235).

Diferente do que se confere em Anjos (2019), destaca-se nesta pesquisa o direcionamento da investigação para o perfil profissional do(a) bibliotecário(a), tendo como objetivo investigar os cursos oferecidos no país, incluindo os cursos à distância, por considerar como hipótese que os cursos EaD, por serem recém-criados, poderiam potencialmente contemplar um perfil mais atualizado de formação. A análise também se debruçou para além das disciplinas obrigatórias, buscando possíveis disciplinas optativas/complementares que tenham sido inseridas para suprir lacunas de formação nos Projeto Pedagógico de Curso (PPC), que não tenham passado por atualizações nos últimos anos.

O contexto que emerge do *Big data* implica diretamente nas formas de armazenamento, organização, recuperação e uso da informação, seja no contexto técnico-profissional ou científico e, para isso, o profissional da Biblioteconomia precisa desenvolver novos conteúdos formativos atendendo a uma necessidade científica, mercadológica e técnica, e é nessa perspectiva que justificamos a inclusão na investigação os cursos mais jovens, na expectativa de que pudessem potencialmente apresentar estruturas curriculares mais atualizadas para que os egressos de

Biblioteconomia possam atuar nesse campo consolidado, observando o uso e reuso intensivo de dados científicos.

Nessa perspectiva do uso intensivo de dados de pesquisa através da consolidação da *e-Science*, formam-se dois tipos de grupos de pesquisa: um grupo menor de pesquisadores pertencentes às grandes instituições, com grande capital de investimento e naturalmente que integram equipes bem estruturadas, inter e multidisciplinar capazes de fazer ótima curadoria e gestão dos dados de pesquisa ali gerados e/ou reusados; e um segundo grupo, ocupando a maior parte do universo da pesquisa científica, formado por pequenos pesquisadores, com menos recursos e equipes reduzidas, muitas vezes atuando sozinhos, e que precisam de mais apoio no que tange a curadoria e a gestão de seus dados de pesquisa.

Ademais, com a consolidação na pesquisa científica da *e-Science*, a necessidade emergente dos pesquisadores no que se refere à gestão, ao uso e ao reuso de dados de pesquisa científica, um profissional capacitado a atuar nesse seguimento se faz necessário, concebendo à Biblioteconomia um novo campo que o mercado de trabalho requer.

Sendo assim, no desenvolvimento desse novo perfil, consideramos que a graduação é o início da vida acadêmica dos pesquisadores científicos. Então, além de elencar os paradigmas do campo estudado, a graduação deve apontar as possibilidades de atuação que seus graduandos poderão seguir. Logo, busca-se evidenciar se os cursos de Biblioteconomia, dado o contexto atual, têm adequado seus núcleos de formação, a fim de atualizar seus currículos suprimindo a premência de formação para atuação no universo dos dados, do profissional e/ou pesquisador da área de Biblioteconomia e Ciência da Informação.

1.2 Questão da pesquisa

Mediante a discussão inicial, emerge-se a seguinte questão: *De que modo os Cursos de Biblioteconomia no Brasil contribuem para que os profissionais da Biblioteconomia atuem como bibliotecário(a)s de dados-no contexto da e-Science?*

1.3 Objetivos

Baseada na contextualização do objeto exposto e na questão ordenadora, os seguintes objetivos nortearam esta pesquisa:

1.3.1 Objetivo Geral

Analisar a formação do profissional responsável pelo elo entre Biblioteconomia e o contexto da *e-Science*, a partir dos currículos dos cursos de graduação em Biblioteconomia das instituições de ensino superior no Brasil.

1.3.2 Objetivos Específicos

- a) discutir o surgimento e a consolidação da *e-Science*;
- b) delinear o perfil do(a) bibliotecário(a) de dados no contexto da *e-Science*;
- c) identificar nos Projetos Pedagógicos de Curso as ementas que se relacionam com a formação voltada para o bibliotecário de dados;
- d) revelar o panorama de formação dos cursos de Biblioteconomia do Brasil frente às demandas informacionais no contexto da *e-Science*.

2 QUARTO PARADIGMA CIENTÍFICO DE GRAY

Antes de debater sobre o quarto paradigma, indicado por Gray em 2007, abordou-se alguns conceitos de *paradigma*, retomando uma discussão que se acentua principalmente com a publicação da famosa obra “Estrutura das Revoluções Científicas”, de Thoman Khun, com sua primeira publicação no ano de 1962. A publicação apresentou-se tão controversa à época, que o próprio Thomas Kuhn, quase sete anos depois, redigiu um posfácio rebatendo algumas críticas e trazendo novas interpretações sobre o conceito de *paradigma científico*. No entanto, nem mesmo a famosa obra apresenta um conceito geral de *paradigma*, tendo Masterman (1979), por exemplo, evidenciado vinte e dois atributos ao significado do termo *paradigma*, num compêndio constituído de quatro volumes reunidos nas “Atas do Seminário Internacional sobre Filosofia da Ciência de 1965”, realizado no Bedford College, Regent’s Park, Londres, de 11 a 17 de julho de 1965.

Porém, esse aglomerado de conceitos não diminui a importância da obra de Kuhn e traz diversas contribuições para a ciência ao empreender a busca por uma definição consensual ao termo *paradigma*, até então sendo usado de forma indiscriminada. A obra de Kuhn trata-se de uma realização sem precedentes ao se dedicar à unificação das diversas formas de uso do termo *paradigma* em torno de uma mesma definição, suficientemente aberta para permitir que toda espécie de problema possa ser resolvido por um grupo redefinido de praticantes de ciência nas suas diversas disciplinas. Nesse sentido, verifica-se que:

[...] o cientista deve preocupar-se em compreender o mundo e ampliar a precisão e o alcance da ordem que lhe foi imposta. Esse compromisso, por sua vez, deve levá-lo a perscrutar com grande minúcia empírica (por si mesmo ou através de colegas) algum aspecto da natureza. Se esse escrutínio revela bolsões de aparente desordem, esses devem desafiá-los a um novo refinamento de suas técnicas de observação ou a uma maior articulação de suas teorias. Sem dúvida alguma existem ainda outras regras desse gênero, aceitas pelos cientistas em todas as épocas." (KUHN, 2000, p. 65).

Thomas Kuhn pontua ainda que o desenvolvimento científico tem base em suas mudanças paradigmáticas, através das revoluções científicas que "[...] são aqueles episódios de desenvolvimento não cumulativo, nos quais um paradigma mais antigo é total ou parcialmente substituído por um novo, incompatível com o anterior " (2000, p. 125). Dessa forma, entende-se que um paradigma se constitui de teoria, métodos e

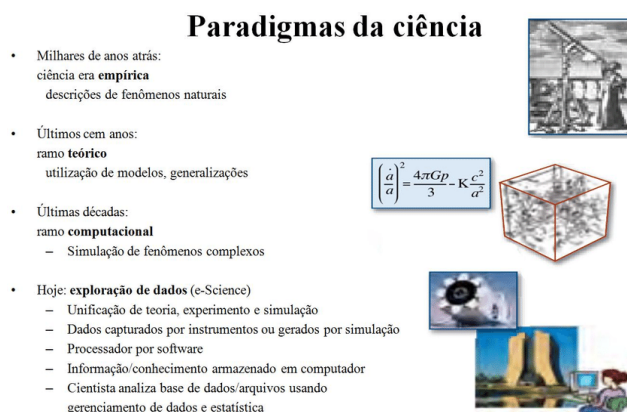
padrões que visam responder às investigações científicas e, enquanto o paradigma tem a capacidade de responder a estes problemas, o autor o define como “ciência normal”, porém quando o paradigma perde essa capacidade (de responder aos questionamentos científicos) é preciso modificá-lo ou criar um novo, momento que é chamado de “ciência extraordinária”, quando ocorre a mencionada revolução científica culminando em um novo paradigma. Sobre essa transformação dos paradigmas, temos que:

[...] Ao aprender um paradigma, o cientista adquire ao mesmo tempo uma teoria, métodos e padrões científicos, que usualmente compõem uma mistura inexplicável. Por isso, quando os paradigmas mudam, ocorrem alterações significativas nos critérios que determinam a legitimidade tanto dos problemas como das soluções propostas. (KUHN, 2000, p. 144).

Gordon Bell (2011, p. 11), no prefácio do livro organizado por Hey, Tansley e Tolle (2011), traz o surgimento de um dos aspectos do quarto paradigma através de Johannes Kepler, assistente do astrônomo Tycho Brahe, já no século XVI, quando pegou seu catálogo de observações astronômicas sistemáticas e formulou as leis do movimento planetário e, com isso, por um lado, estabeleceu uma divisão entre a mineração e a análise de dados experimentais coletados e cuidadosamente arquivados, por outro lado, tornou a criação de teorias possíveis, gerenciando os dados para outros pesquisadores.

Jim Gray ao ler um artigo sobre astronomia evidencia uma nova revolução científica e divide a ciência em quatro paradigmas: há mil anos baseado numa ciência empírica, há poucos séculos baseado num ramo teórico, nas últimas décadas baseado na computação e, atualmente, baseado na exploração de dados, como pode ser observado na Figura 1:

Figura 1 – Paradigmas da ciência



Fonte: Hey, Tansley, Tolle, (2011, p. 18) adaptado de Gray (2007)

Dessa forma, Gray (2007) expõe a característica do primeiro paradigma, quando a ciência era apenas experimental, há cerca de mil anos, circunstância em que o(a)s cientistas somente observavam e descreviam os fenômenos naturais até surgir a necessidade de se teorizar sobre os fenômenos.

O segundo paradigma nasce no momento em que somente observações não respondem mais aos questionamentos dos pesquisadores, a ciência se torna mais teórica, os pesquisadores começam a utilizar de modelos e generalizações como as Leis de Kepler, as Leis de Newton, as equações de Maxwell, entre outros, porém esses modelos transfiguram-se numa complexidade impeditiva de suas análises.

No terceiro paradigma, depois dos modelos teóricos se tornarem muito complicados para serem resolvidos analiticamente, concomitante com o surgimento dos computadores, os pesquisadores começam a fazer simulações de fenômenos complexos, essas nos acompanharam durante a maior parte do último milênio.

Já o quarto paradigma surge no ponto em que essas simulações, através de instrumentos complexos e de larga escala, produzem e acumulam uma quantidade acentuada de dados, a ciência passa a ser mais experimental com a exploração desses dados, conjuntura em que o pesquisador não precisa mais observar os fenômenos, os instrumentos captam dados de forma intensiva e, a partir desses dados, o pesquisador unifica teoria, experimento e simulação.

Gray (2007) aponta que a ciência utiliza a Exploração de Dados, ou *e-Science*, com dados capturados por instrumento, gerados por simulação e/ou processados por *software*, cuja informação/conhecimento fica armazenada em computadores e o(a) cientista analisa a base de dados e arquivos por meio do gerenciamento e das análises estatísticas desses dados.

A discussão sobre este quarto paradigma, conhecido como *e-Science*, é recente, o que pode gerar certa confusão terminológica. Apontando uma associação com ciberinfraestrutura, Costa (2017, p. 28) propõe que:

Em sentido *latu*, *e-science* e *cyberinfrastructure* são termos de grande abrangência que se referem à uma nova forma de se fazer ciência, bem como à infraestrutura tecnológica necessária para apoiar a pesquisa científica do Século XXI, como, por exemplo, a computação em *grid* e bancos de dados que suportem petabytes de dados não estruturados, com fluxo constante. No sentido *strictu*, *e-science* refere-se a uma nova forma de se fazer ciência, enquanto *cyberinfrastructure* refere-se a infraestrutura tecnológica que viabiliza a *e-science*. O termo *e-science* é predominantemente utilizado pelo Reino Unido, enquanto o

termo *cyberinfrastructure* é utilizado nos Estados Unidos, conforme será apresentado e esclarecido durante a revisão de literatura.

Com o desenvolvimento das Tecnologias Digitais de Informação e Comunicação (TDICs), a geração e o compartilhamento de dados foram facilitados, principalmente com o surgimento dos computadores pessoais e a popularização da internet, o que gerou uma simplificação nesses processos e intensificou o volume de dados que se tornou gigantesco.

Antes do advento dos computadores, os métodos e procedimentos adotados para a criação e compartilhamento de dados eram limitados a menos formas, como papel e filme. Atualmente, os mesmos dados podem ser convertidos para meios mais convenientes, como uma mensagem de correio eletrônico, um livro eletrônico, uma imagem em forma de bitmaps ou um filme digital (SOMASUNDARAM; SHRIVASTAVA, 2011, p. 27).

Esse grande volume de dados ocorre principalmente pela soma das necessidades dos indivíduos em compartilhar e das instituições/organizações em criarem e compartilharem esses dados para disseminarem e obterem informações, assim como afirmam Somasundaram e Shrivastava (2011, p. 28):

Formas baratas e fáceis de criar, coletar e armazenar todos os tipos de dados, junto a necessidades individuais e empresariais cada vez maiores, levaram ao crescimento acelerado do volume de dados, o que foi popularmente chamado de explosão de dados. Dados têm diferentes objetivos e importância, de modo que tanto os indivíduos quanto as empresas contribuíram em diversas proporções para essa explosão.

Nesse contexto, a preservação dos dados de pesquisa bem como seu compartilhamento e gerenciamento são de extrema importância, tendo em vista que esses dados poderão ser utilizados de formas diferentes e por pesquisas diferentes, inclusive contemplando dados gerados e não utilizados em uma pesquisa e que, em outra, se demonstre primordial.

Em termos práticos, no entanto, o que se entende por dado de pesquisa varia de acordo com a disciplina, a área do conhecimento, o contexto e até mesmo com sua destinação ou finalidade. A fotografia de uma construção municipal depositada em um arquivo histórico, por exemplo, pode não representar muito para um agrônomo; enquanto, para um historiador, aquela fotografia torna-se um dado de pesquisa. Vale ressaltar que nem todos os dados derivados de um projeto de pesquisa são publicados em artigos científicos; e ainda, que dados

produzidos no âmbito de um determinado projeto podem ser utilizados em uma agenda de pesquisa completamente diferente, por outro grupo de pesquisadores. Ou seja, o ‘ruído’ observado em uma determinada pesquisa pode consistir em ‘sinal’ para outro pesquisador, fato que reforça a importância do gerenciamento, compartilhamento e preservação dos dados de pesquisa (BERTIN; VISOLI; DRUCKER, 2017, p.38).

Na seção 4, 3 *DOS DADOS AOS METADADOS*, o conceito de *dado* será mais bem explorado, trazendo discussões importantes a formulação desse conceito tão fundamental a este trabalho. Para compreensão neste momento, pondera-se que *dado* é um elemento num sentido bruto, sem significado que pode ser armazenado e quantificado, sendo ele, por si só, não suficiente para compreensão do indivíduo, porém, a partir dele pode-se atribuir significado e chegar a alguma compreensão.

No cenário em que se edifica a popularização da internet, juntamente com a evolução tecnológica que permite um constante aumento na capacidade de armazenamento de dados, bem como seu fluxo cada vez mais veloz, a explosão informacional é inevitável, gerando o que se conhece como *Big data* (CONEGLIAN; SANTAREM SEGUNDO; SANT’ANA, 2016). Para estes teóricos,

Além das informações geradas por meios tradicionais, como bibliotecas, formulários, arquivos e banco de dados de empresas, existe uma geração de dado que extrapola tais meios, como as informações produzidos em redes sociais, as transações online, o GPS, os dados produzidos por dispositivos conhecidos como Internet das Coisas (Internet of Things - IoT), os smartphones, entre outros (ibidem, p. 5).

Assim, *Big data* não necessariamente se define como dados grandes, grande conjunto de dados e/ou ferramentas e procedimentos para manipular e analisar dados, mas, pode-se afirmar que parte de uma mudança no pensar e pesquisar cientificamente na qual se imputa as relações computacionais. Verifica-se essa definição no exemplo dado por Boyd e Crawford (2012, p. 663), ao indicar que uma coleção de dados do *Twitter*, considerados como enquadrados no *Big data*, não é maior do que os dados gerados por sistemas anteriores, como os dados dos censos, que não se enquadram no *Big data* e formam um conjunto maior de dados. Para os autores, o “*Big data* se refere menos ao grande volume de dados e mais a uma capacidade de pesquisar, agregar e

fazer referência cruzada de grandes conjuntos de dados”³ (BOYD; CRAWFORD, 2012, p. 663, tradução nossa).

Já os autores Silva Segundo e Araújo (2019, p. 113) chamam o fenômeno *Big data* de megadados, pois “[...] grande, na língua portuguesa, remete basicamente à quantidade, e mega denota tanto quantidade quanto qualidade [...]”. Os autores ressaltam ainda que a realidade atual do mercado de dados pode ser definida como um espaço onde nossos dados são coletados 24h por dia e vendidos para as grandes corporações traçarem o perfil do consumidor e fornecer propagandas personalizadas. Este universo gera uma quantidade diária imensa de dados, que apontam que 90% dos dados existentes do mundo foram gerados nos últimos anos.

Nesse momento, em que a ciência se utiliza progressivamente de uma perspectiva computacional, com grande apelo à simulação de fenômenos complexos, com base em Gray, (2007), Bell, Hey e Szalay (2009) trazem que do *Big data*, “[...] derivou-se a exploração de grandes volumes e uso intensivo de dados (*data-intensive*), em busca da interlocução de teorias, experimentos e simulações, dando lugar à *e-Science*.” (ALBAGLI; APPEL; MACIEL, 2013, p.7).

Jankowski (2007) afirma que o termo *cyberinfrastructure* surge em 2003 através do Relatório Atkins, que o define como uma infraestrutura de computação distribuída⁴ de informação e de tecnologias de comunicação que gera grandes volumes de dados e exige processamento de computador de alta velocidade, envolvida com as ciências naturais e biológicas como: a física de partículas, astronomia, meteorologia e pesquisa de DNA. Estas iniciativas tipicamente contam com a colaboração de profissionais da supercomputação e dos centros de pesquisa.

John Taylor, então diretor da *General of Research Councils* do escritório *Office of Science and Technology* (OST), utiliza o termo *e-Science* pela primeira vez ao anunciar o programa de fomento em pesquisa no Reino Unido, em 2001, para desenvolver as ferramentas, tecnologias e infraestrutura para apoiar essa ciência multidisciplinar e colaborativa (HEY; TREFETHEN, 2005). Assim, a “[...] *e-Science* trata da colaboração global em áreas-chave da ciência e da próxima geração de

³ Texto original: Big data is less about data that is big than it is about a capacity to search, aggregate, and cross-reference large data sets.

⁴ A computação distribuída visa resolver um determinado problema computacional através da utilização de diferentes recursos distribuídos geograficamente. Esse termo evolui para computação em grade quando passa a existir além de uma infraestrutura física uma infraestrutura lógica (*software*), aumentando a eficiência e eficácia na resolução dos problemas.

infraestrutura que possibilitará isso "⁵ (HEY; TREFETHEN, 2003, p. 1809, tradução nossa).

De forma resumida, Jim Gray, numa palestra ministrada em *Mountain View*, na Califórnia, em 11 de janeiro de 2007, para o NRC-CSTB (Conselho Nacional de Pesquisa e Conselho Nacional de Telecomunicações e Ciência da Computação [ambos dos Estados Unidos]), pontua a *e-Science* como "o ponto onde a 'TI [Tecnologia da Informação] encontra os cientistas'" (GRAY, 2007, p. 17), porém, Gray também a identifica como um novo paradigma científico complexo que envolve a geração, uso, reúso e compartilhamento de dados intensivos gerados por pesquisas, e seu principal desafio será gerenciar essa quantidade massiva de dados para que seja plausível de ser usada, reusada e compartilhada, e todo esse processo só será possível com a pesquisa se tornando cada vez mais computacional (2007).

Corroborando com o desafio do quarto paradigma e ilustrando sua potencialidade no enorme avanço que a ciência colaborativa pode significar para a própria ciência aliada às TDICs, Delaney e Barga (2011, p. 56-57) afirmam que:

[...] Se os dados arquivados se tornam imediatamente disponíveis via internet, o potencial para descobertas cresce substancialmente, por causa do número crescente de pesquisadores potenciais que podem explorar um espectro de 'espaço parâmetro' em rápida expansão. Para cientistas que operam nesse ambiente intensivo de dados, haverá necessidade do desenvolvimento de uma nova série de produtos de fluxo de trabalho científico que possa facilitar o arquivamento, a assimilação, a visualização, a modelagem e a interpretação da informação sobre todos os sistemas científicos de interesse. [...]

Com o quarto paradigma pontuado por Hey, Tansley e Tolle em 2009, a ciência se torna mais colaborativa, mais computacional e mais intensiva, evidenciando o aumento da dependência na tecnologia para a pesquisa científica com a infraestrutura cibernética e o estabelecimento da gestão de dados e compartilhamento de dados, contexto de pesquisa que passa a ser exigido por muitos organismos de financiamento à pesquisa (TENOPIR; et al, 2014, p. 84).

Medeiros e Caregnato (2012, p. 315), por sua vez, contextualizam outra faceta da *e-Science*, a partir da produção de grandes quantidades de conteúdos de pesquisa, gerando grandes quantidades de dados. Nesse contexto, associado com a pesquisa colaborativa mencionada, o pesquisador não precisa realizar os mesmos experimentos já

⁵Texto original: [...] *e-Science* is about global collaboration in key areas of science and the next generation of infrastructure that will enable it.

realizados anteriormente em outras pesquisas, basta utilizar os dados da pesquisa antecedente, reduzindo assim custos financeiros, de mão de obra e de tempo.

No que tange ao empirismo o qual resultará na pesquisa colaborativa e produção dessa grande quantidade de dados, Gray (2007) coloca a *e-Science* quanto a captura, a curadoria e a análise como ações-chaves para o uso de dados, cuja captura ocorre através de instrumentos. Cabendo ao pesquisador, nesse processo, codificar suas informações para que possam ser trocadas com outro(a)s cientistas, logo as informações precisam estar representadas de forma algorítmica, por meio de padrões de objetos em análise.

A partir dessa discussão, Sayão e Sales (2012) apontam necessidade da curadoria digital de dados de pesquisa como uma gerência necessária para preservar os dados de pesquisas, vistos como recursos digitais durante seu ciclo de vida, com o objetivo desses recursos se tornarem atemporais e atenderem às gerações atuais e às gerações futuras de usuários.

Complementando a fala de Sayão e Sales (2012), Albagli, Appel e Maciel (2013, p.4) afirmam que “A análise envolve o uso de tecnologias computacionais distribuídas e compartilhadas (os *grids*), as quais permitem o uso de recursos tecnológicos ou habilidades analíticas de cientistas posicionados em diferentes localidades”.

Logo, compreende-se a importância para pesquisador, em seu estudo, de todo processo de coleta dos dados de seu experimento, bem como gerenciamento, a guarda e disponibilização dos dados para que ele mesmo e demais pesquisadores, no presente e/ou no futuro, possam usufruir desses mesmos dados; finalizando com uma análise que, com o uso de tecnologias de comunicação, a *e-Science* pode envolver pesquisadores de diversas localidades, atribuindo um lastro maior às suas conclusões.

Dessa forma, o pesquisador ou grupo de pesquisa, ao realizar suas pesquisas devem prestar atenção especial aos dados (ou contar com um profissional que tenha esta competência), não somente no sentido qualitativo/quantitativo, mas no que se refere à gestão e/ou curadoria, na qualidade de organização, classificação e armazenamento para que esses dados possam ser compartilhados e reusados. É necessário frisar que a *e-Science* é uma ciência existente dentro de uma rede colaborativa, sendo possível, através de ferramentas e tecnologias que dão suporte a nova forma, fazer pesquisa em rede dando provimento e apoio às atividades de investigação científica. Nesse interim, Medeiros e Caregnato (2012) concluem que:

O advento da *e-Science* necessita, sem dúvida, de investimentos que possibilitem a criação de estratégias para a formação de profissionais em alto nível, com capacidade para atuar no gerenciamento de dados primários, assimilando questões contextuais – de domínio –, reconhecendo a existência de um novo paradigma que deve ser observado com a lucidez e atitude de um gerenciador de informação em um mundo dinâmico. (MEDEIROS; CAREGNATO, 2012, p. 320).

No entanto, sabe-se que nem sempre o pesquisador possui tempo hábil para a realização da curadoria dos dados produzidos em suas pesquisas, logo no cenário da *e-Science*, dentro do “dilúvio de dados”, torna-se evidente a necessidade de um profissional capacitado para gerenciar esses dados, facilitando, agilizando, assim como permitindo melhor acesso por parte do pesquisador:

Howe et al. (2008) ao discutirem o crescimento exponencial na quantidade de dados científicos em Biologia, apontam que são necessárias medidas revolucionárias para a gestão, análise e acessibilidade destes dados. Os autores propõem três ações para a solução desta problemática: 1) autores, revistas e curadores devem começar imediatamente a trabalhar em conjunto para facilitar o intercâmbio de dados entre as publicações de periódicos e bases de dados; 2) nos próximos cinco anos, os curadores, pesquisadores e administrações universitárias devem desenvolver uma estrutura de reconhecimento para facilitar os esforços de curadoria de base comunitária; 3) curadores, pesquisadores, instituições acadêmicas e agências de financiamento devem nos próximos dez anos, aumentar a visibilidade e apoio de curadoria científica como uma carreira profissional (BUFREM et al. 2016, p. 46).

Nesse contexto de crescimento, em um mundo onde a competitividade e a inovação seguem com a mesma proporcionalidade, o bibliotecário, dado as suas habilidades, principalmente no que se refere à organização da informação, deve assumir essa função como bibliotecário(a) de dados. Com o avanço da *Big data*, a curadoria de dados de pesquisa se torna mais complexa devido à imensa geração desses dados de forma exponencial, associada ao seu compartilhamento, exigindo novas habilidades de um profissional com perfil para organização, acesso e uso de informações, sendo extensivo ao gerenciamento de dados. Sobre isso, temos que

De fato, a gestão de dados se complexificou com a emergência do, que aponta para novos horizontes científicos e pragmáticos conduzindo para a exigência de novas habilidades e conhecimento para os profissionais da área [Ciência da Informação] visando o enquadramento no perfil para a gerência dos dados. À medida que o tratamento dados se intensifica e indica desafios, competitividade e necessidade de inovação, na mesma proporção, exige-se um

profissional que esteja apto a gerenciar informação neste novo contexto (RODRIGUES; DUARTE; DIAS, 2017).

Rodrigues e Dias (2017) demonstram em um diagrama de Venn, criado por Conway (2010) (Figura 2), as competências que o profissional, para gerir os dados gerados e compartilhados na *Data Science*, deve reunir organizando o conhecimento em três grandes áreas: estatística e matemática, para fazer uma análise qualitativa em uma quantidade excessiva de dados, executando a mineração para extrair informação (no sentido de atribuir valor semântico aos dados); habilidades de *hacking*, com conhecimento em linguagem de programação e conhecimento em privacidade e segurança de dados, trabalhando com dados obtidos eletronicamente; e, *expertise* substantiva que lhe conferirá o *know-how* para ser um bom profissional na área.

Na Figura 2, pode-se verificar também áreas menores advindas do cruzamento entre essas grandes áreas, assim: entre a habilidades de *hacking* e conhecimento de matemática e estatística emerge a aprendizagem de máquina; entre o conhecimento de matemática e estatística e *expertise* substantiva surge a pesquisa tradicional; entre a *expertise* substantiva e a habilidade de *hacking* há a zona de perigo; e por fim, ao centro, com o cruzando de todas três grandes áreas (habilidades *hacking*, conhecimento de matemática e estatística e *expertise* substantiva) a **Ciência de Dados**. (CONWAY, 2010; RODRIGUES; DIAS, 2017, p. 222).

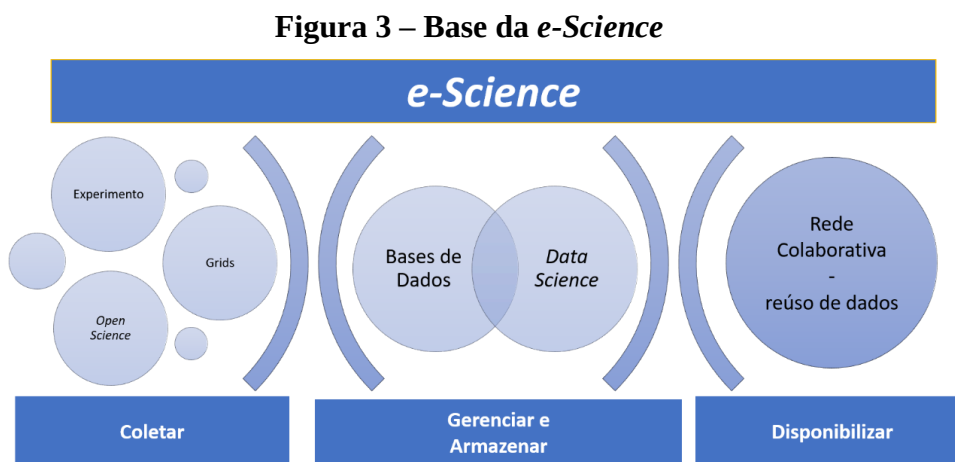
Figura 2 – Diagrama de Venn



Fonte: Rodrigues e Dias (2017, p.222).

Em análise da figura 02, observa-se que a *e-Science*, ou quarto paradigma, advém da evolução e aglutinação dos três paradigmas científicos anteriores, dado principalmente pelo surgimento e aperfeiçoamento das TDICs.

Em sequência, baseado nos principais conceitos de *e-Science* vistos até o momento, no intuito de ilustrar graficamente essa nova forma de se fazer pesquisa, criou-se a seguinte figura:



Fonte: Baseado em Medeiros; Caregnato, 2012; Albagli; Appel; Maciel, 2014; Bell; et al, 2011.

Na Figura 3, é possível notar as três fases da *e-Science*: num primeiro momento há a coleta dos dados gerados pelas pesquisas, em experimentos ou disponibilizados pela Ciência Aberta (*Open Science*), com acesso pelos *grids* (tecnologias de compartilhamento em rede); num segundo momento, há a necessidade primordial do gerenciamento e armazenamento, quando se evidencia principalmente a intersecção entre as Bases de Dados e a Ciência de Dados (*Data Science*), esse momento compreende o conceito de gestão de dados e curadoria de dados; por fim, com a junção dos processos, os dados são disponibilizados e reusados na Rede Colaborativa, estabelecida através de todo esse processo.

Por fim, ressalta-se que um dos principais desdobramentos do quarto paradigma é a evolução da Ciência de Dados, isto é, a pesquisa científica fundamentada no uso intensivo de dados criados e/ou disponibilizados de forma colaborativa através de uma ciência aberta e com enormes quantidades de dados, desdobramento teórico que será mais aprofundado na próxima seção.

3 DOS DADOS AOS METADADOS

Em 1963, Derek J. de Solla Price funda os estudos métricos da produção científica e através deles categoriza a ciência em *Grande Ciência* e *Pequena Ciência*, preocupado como a posição da ciência perante a sociedade:

De acordo com essa metáfora [*Big Science, Little Science*], minha primeira palestra se refere ao volume da ciência, a segunda à distribuição de velocidade de suas moléculas, a terceira à maneira pela qual as moléculas interagem com outras e a quarta à derivação das propriedades políticas e sociais deste gás⁶ [gás como termo geral para caos] (PRICE, 1986, p. xvii, tradução nossa).

Mais adiante, Solla Price (1986) evidencia as mudanças ocorridas na ciência no momento após Segunda Guerra, quando se deixa de ser praticada por pequeno(a) cientista, visto pela sociedade como “[...] gênio solitário e de cabelos compridos, moldando em uma oficina no sótão ou no porão, desprezado pela sociedade como não-conformista, existindo em um estado de quase pobreza, motivado pela chama ardente dentro dele [...]”⁷ (PRICE, 1986, p. 2, tradução nossa) e passa a ser praticada com projetos de grande porte (tanto em tamanho do grupo de cientistas quanto pelo tamanho do financiamento e da produção científica), geralmente financiados por governos ou grupos governamentais e em quantidade e velocidade cada vez maior.

Apoiada nos preceitos teóricos de Price (1986), Borgman demonstra a consolidação do *Big data* e o que se é esperado dele:

O “*Big data*” adquiriu a hipérbole que a “*big science*” fez cinquenta anos atrás. O *big data* está nas capas das revistas *Science*, *Nature*, *The Economist* e *Wired* e nas primeiras páginas do *Wall Street Journal*, no *New York Times* e em muitas outras publicações, tanto de grande quanto de menor porte. Assim como a grande ciência deveria revelar os segredos do universo, espera-se que o *big data* revele os tesouros enterrados no pequeno fluxo da vida⁸ (BORGMAN, 2015, p. 4, tradução nossa).

⁶ Texto original: According to this metaphor, my first lecture is concerned with the volume of science, the second with the velocity distribution of its molecules, the third with the way in which the molecules interact with one another, and the fourth in deriving the political and social properties of this gas.

⁷ Texto original: [...] long-haired genius, moldering in an attic or basement workshop, despised by society as a nonconformist, existing in a state of near poverty, motivated by the flame burning within him [...].

⁸ Texto original: “Big data” has acquired the hyperbole that “big science” did fifty years ago. Big data is on the covers of *Science*, *Nature*, the *Economist*, and *Wired* magazine and the front pages of the *Wall Street Journal*, *New York Times*, and many other publications, both mainstream and minor. Just as big

O termo *Big data* surge com Cox e Ellsworth (1997), pesquisadores da NASA que tentavam solucionar o problema de como armazenar o volume imenso de dados, estruturados e não-estruturados, gerados a cada segundo pelo experimento que realizavam.

Arelado ao *Big data*, a publicação de acesso aberto é geralmente datada a partir da Declaração de Budapeste em 2002, que tem raízes em experimentos de publicação eletrônica iniciados na década de 1970, porém possui raízes ainda mais longínquas. O sistema do *World Data Center* foi criado na década de 1950 para arquivar e distribuir dados coletados dos programas observacionais do Ano Geofísico Internacional de 1957-1958. Em 2007, os princípios para o acesso aos dados da pesquisa, a partir de financiamento público, foram codificados pela Organização para Cooperação e Desenvolvimento Econômico (OCDE) (BORGMAN, 2015, p. 7).

Desse contexto surgem a Ciência de Dados (*Data Science*) e a Ciência Orientada por Dados (*data-driven science*). Dependem, além da estrutura física, de sistemas de informação que integrem arquivamento e compartilhamento de coleções de dados com a atribuição de definir a sua utilidade, tanto para curto prazo quanto para longo, reconhecidamente de valor permanente ou de referenciais universais (SAYÃO; SALES, 2016).

Para ilustrar as possibilidades alcançadas através do quarto paradigma, da Ciência de Dados no *big data*, pode-se observar a recente “fotografia” de um buraco negro. Para fazê-la, devido à distância da Terra com o fenômeno, seria preciso de uma antena do tamanho da própria Terra. Nessa impossibilidade, foram dispostas várias antenas espalhadas pelo planeta capitando pontos diferentes do fenômeno, durante dez dias, gerando uma quantidade astronômica de dados, impossíveis de serem transmitido pela internet. “Os cientistas tiveram que esperar o verão do hemisfério sul para, depois do degelo, trazer quase 500 kg de discos rígidos de um telescópio da Antártica” (SCHULTZ, 2019, *online*).

Uma outra perspectiva é apresentada por Ziviani, Porto e Ogasawara, (2015, *online*) que compreendem a Ciência de Dados como um grupo de ações utilizadas em uma coleção de dados (*dataset*) para guiar o pesquisador à descoberta de novos conhecimentos.

Rodrigues e Dias (2017, p. 221) elucidam que a Ciência de Dados é estruturada com teorias e métodos em processo de amadurecimento e se constitui em ferramentas

science was to reveal the secrets of the universe, big data is expected to reveal the buried treasures in the bit stream of life.

teóricas e metodológicas, ainda em andamento, fundamentais para promoverem inovações disruptivas das estruturas visuais no ambiente digital, através da construção das visualizações de dados científicos futuros.

Dessa forma, percebe-se a Ciência de Dados num aspecto mais ferramental e tecnológico, aproximando-se mais das engenharias no que tange às linguagens de programação, forte uso de matemática e estatísticas, além de habilidades de *hacking*. No entanto, este trabalho a distingue da ciência orientada a dados, pois esta possui uma amplitude maior e de caráter mais epistemológico.

Bufrem, *et al* (2016, p. 42) trazem em um trabalho de pesquisa na base *Scopus* e na *Web of Science*, a então produção internacional sobre ciência orientada a dados e *e-Science*. Eles encontram uma edição da *Revista Informação & Informação* especial sobre este tema e pode-se verificar a amplitude desse objeto, quando a edição contempla artigos sobre

Curadoria de Dados de Pesquisa (*Research data curation*); Mineração de Dados no Âmbito Científico (*scientific research*); Compartilhamento e Reúso de Dados de Pesquisa (*Research data sharing e Research data reuse*); Ciberinfraestrutura de Apoio à *Data Science* (*cyberinfrastructure*); Dados Interligados (*Linked Data*); Publicações Ampliadas (*Enhanced Publications*) e Ferramentas para transformação e visualização de dados (BUFREM; *et al*, 2016, p. 42).

Nessa perspectiva, autores como Semeler (2017) e Murriete (2017) apontam como equivalentes a *e-Science*, ciência orientada a dados, computação fortemente orientada a dados, dados científicos, dados de pesquisa ou quarto paradigma. Entretanto, essas terminologias ainda estão sendo discutidas e, sem entrar no mérito da mais adequada, é possível vislumbrar um único cenário: o uso intensivo de dados.

3.1 Dados

Tendo em vista as discussões aqui propostas, é preciso esclarecer que dados, no plural, na visão de Borgman (2015) e Bryson (2008), é mais usual na comunicação acadêmica quando usado para se referir aos dados como entidades, e no singular quando se referir ao conceito, sendo esse segundo conceito de dados adotado por este trabalho.

Sobre o termo, Christine Borgman (2015) traz que *dados* já era utilizado em teologia desde 1646, e o seu conceito foi discutido na área até o século XVIII. Para a autora, os dados foram invocados como o conjunto de princípios aceitos como base de um argumento ou então como fatos, particularmente aqueles retirados das escrituras, e somente no século XVIII, eles passam a ser discutidos também nas ciências implicando

em fatos na forma de evidência científica reunida por experimentos, observações e outras investigações.

Atualmente, o termo *dados* é muito utilizado em pesquisas científicas, em diversos domínios, porém, dificilmente se encontra uma definição consensual, tendo em vista que a literatura mostra com frequência que os termos “dados”, “informação” e “conhecimento” se confundem, pois, suas concepções semânticas são muito próximas entre si (SETZER, 2015; BORGMAN, 2015).

Muito se discutiu no meio científico, e ainda se discute, sobre a melhor definição para *dado*, direcionando que cada domínio científico acaba utilizando um sentido que melhor se adequa a sua realidade. Em Ciência da Informação, é necessário chegar a um consenso, principalmente por ser um domínio interdisciplinar (SARACEVIC, 1995; PINHEIRO; LOUREIRO, 1998; PINHEIRO, 2005). Entretanto, este trabalho não tem o intuito de findar a discussão ou incrementá-la, mas buscar-se-á uma definição que aponte uma melhor adequação tanto no contexto amplo da Ciência da Informação, quanto ao mais específico da Ciência Orientada a Dados que, assim como a Ciência da Informação, também perpassa, entre outras, pela computação e comunicação.

Em vista disso, é importante trazer a contribuição posta por Silva e Gomes (2015) que discutem os principais conceitos em Ciência da Informação, evidenciando que *dado* aparece como coadjuvante na área, reconhecido como um meio para se alcançar informação e/ou conhecimento. Porém, no contexto atual da *e-Science* e Ciência de Dados, é preciso uma definição mais aprofundada, pois **dado** se torna uma temática em evidência na Ciência da Informação com sua interdisciplinaridade com a *e-Science*.

Num conceito para administradore(a)s de base de dados, FRY e SIBLEY, (1976, p. 9, tradução nossa) apresentam *dados* como aquilo que “consiste em uma declaração dos nomes dos elementos, suas propriedades [...] e seu relacionamento com outros elementos [...] que compõem a base de dados”.⁹

Numa outra perspectiva, Davis e Rush (1979 apud SANTOS; SANT’ANA, 2013, p. 201) preconizam que “a forma mais simples de definir dados é dizendo que eles são o resultado de mensuração ou observação”.

Em complemento, dados podem ser considerados como elementos representativos utilizados pelo pesquisador para apontar como chegou aos seus

⁹ Texto original: consists of a statement of the names of elements, their properties (such as character or numerical type), and their relationship to other elements (including complex groupings) which make up the data base.

resultados, “[...] os dados são representações de observações, objetos ou outras entidades [algo que tem uma existência real] usadas como evidência de fenômenos para fins de pesquisa [...]”¹⁰ (BORGMAN, 2015, p. 28, tradução nossa).

Pensando em uma definição possível em Ciência da Computação, *dado* seria a representação mais bruta, plausível, que seja legível pela máquina. “Em informática chama-se dado a representação convencional, codificada, de uma informação sob uma forma que permite seu processamento eletrônico” (LE COADIC, 1996, p. 6). Para Santos e Sant’Ana (2015), de forma mais abrangente,

[...] dado é uma unidade de conteúdo necessariamente relacionada a determinado contexto e composta pela tríade entidade, atributo e valor, de tal forma que, mesmo que não esteja explícito o detalhamento sobre contexto do conteúdo, ele deverá estar disponível de modo implícito no utilizador, permitindo, portanto, sua plena interpretação. (SANT’ANA, 2015, p. 205)

Buckland (1991) corroborado por Capurro e Hjørland (2007) deu um grande contributo para a Ciência da Informação ao diferenciar os usos do termo “informação”, dividindo-o em três vertentes: como processo, quando se refere ao ato de informar; como conhecimento, quando não é possível de mensurá-la por ser intangível; e como informação como coisa, que é considerada informativa e, por isso, é referenciada como informação, mas são na verdade objetos, indicando, em geral, *dado* como um exemplo de informação como coisa, que seria o registro físico da informação.

Já Furner (2016) diferencia documentação de dados através de uma revisão bibliográfica histórica desses conceitos e relaciona os dois conceitos vinculando um ao outro. Assim, ele conclui que

De fato, não é o caso que os documentos sejam compostos de dados, nem que o documento seja uma espécie de conjunto de dados: é o contrário, em ambos os aspectos. Um conjunto de dados é composto de documentos; e o conjunto de dados é uma espécie de documento.¹¹. (FURNER, 2016, p. 303, tradução nossa).

Semeler e Pinto (2019), com o propósito de uma definição mais atual, caracteriza *dado* como um objeto nascido em um contexto computacional ou

¹⁰ Texto original: [...] data are representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research [...]

¹¹ Texto original: It is not in fact the case that documents are made up of data, nor that the document is a species of dataset: rather it is the other way round, in both respects. A dataset is made up of documents; and the dataset is a species of document.

transformado para uso nesse meio. Esses teóricos indicam *dado* como “todo objeto criado em formato digital ou convertido para o formato digital que possa ser usado para geração de *insights* de informação e conhecimento” (SEMELER; PINTO, 2019 p.115).

Outra definição em Ciência da Informação para *dado* assinala-o como registro do atributo de um ente, objeto ou fenômeno, compreendendo o termo *registro* como gravação ou impressão de caracteres ou símbolos, com significado, em qualquer tipo de documento ou suporte físico; ao termo atributo como referente a uma característica (uma propriedade) de um ente, objeto ou fenômeno; e o termo fenômeno para incluir o registro de um fato, de algo que ocorreu em um determinado intervalo de tempo, e que efetivamente não está ligado a um objeto ou ente específico (GOMES; PIMENTA; SCHNEIDER, 2019). Percebe-se aqui a o zelo em determinar *dado* como a ação de gravar alguma característica de algo (que os autores classificam como um ente, objeto ou fenômeno) em algum suporte, sendo esse registro em qualquer tipo de linguagem.

No Quadro 1, é possível verificar todas as definições de dados existentes neste referencial teórico, de forma sintética e objetiva.

Quadro 1 – Quadro de Definições sobre Dados

AUTOR	DEFINIÇÃO
FRY e SIBLEY, (1976, p. 9, tradução nossa)	“consiste em uma declaração dos nomes dos elementos, suas propriedades [...] e seu relacionamento com outros elementos [...] que compõem a base de dados”.
Davis e Rush (1979 apud SANTOS; SANT’ANA, 2013, p. 201)	“a forma mais simples de definir dados é dizendo que eles são o resultado de mensuração ou observação”.
Borgman (2015, p. 28, tradução nossa)	“[...] os dados são representações de observações, objetos ou outras entidades [algo que tem uma existência real] usadas como evidência de fenômenos para fins de pesquisa [...]”.
Le Coadic (1996, p. 6)	“Em informática chama-se dado a representação convencional, codificada, de uma informação sob uma forma que permite seu processamento eletrônico”.
Santos e Sant’Ana (2015, p. 205)	“[...] dado é uma unidade de conteúdo necessariamente relacionada a determinado contexto e composta pela tríade entidade, atributo e valor, de tal forma que, mesmo que não esteja explícito o detalhamento sobre contexto do conteúdo, ele deverá estar disponível de modo implícito no utilizador, permitindo, portanto, sua plena interpretação”.
Furner (2016, p. 303, tradução nossa)	“De fato, não é o caso que os documentos sejam compostos de dados, nem que o documento seja uma espécie de conjunto de dados: é o contrário, em ambos os aspectos. Um conjunto de dados é composto de documentos; e o conjunto de dados é uma espécie de documento.”
Semeler e Pinto (2019 p.115)	“todo objeto criado em formato digital ou convertido para o formato digital que possa ser usado para geração de insights de informação e conhecimento”

Gomes, Pimenta e Schneider (2019)	registro do atributo de um ente, objeto ou fenômeno, tendo o termo registro designado como ato de registrar, ou seja, é a gravação ou a impressão de caracteres ou símbolos que tenham um significado em algum documento ou suporte físico; ao termo atributo como referente a uma característica (uma propriedade) de um ente, objeto ou fenômeno; e o termo fenômeno para incluir o registro de um fato, de algo que ocorreu em um determinado intervalo de tempo, e que efetivamente não está ligado a um objeto ou ente específico.
Setzer (2015) e Borgman (2015)	fatos, números, letras e símbolos que podem ser quantificados ou quantificáveis.

Fonte: Dados da pesquisa, 2021.

Em suma, no contexto desse estudo, dados podem ser considerado como fatos, números, letras e símbolos que podem ser quantificados ou quantificáveis (SETZER, 2015; BORGMAN, 2015), principalmente por demonstrar ser o elemento mais “puro” ou “bruto” da Ciência da Informação. Porém, tais definições não apresentam consenso na área, principalmente por possuir diferentes tipos de valores que nem sempre são visíveis num primeiro momento, mas evidenciam seu impacto, tendo em vista que no cenário mundial contemporâneo o seu uso está promovendo mudanças radicais, especialmente no contexto acadêmico (BORGMAN, 2015).

3.2 Dados de pesquisa: os cuidados necessários

A definição de dados de pesquisa, por si só, indica uma atividade acadêmica, ou melhor, uma atividade de pesquisa acadêmica, como afirma Borgman (2015). Nessa perspectiva, é possível afirmar que dados de pesquisa são definidos como os dados obtidos através de uma pesquisa acadêmica?

Para obter essa resposta, reporta-se à OCDE (2007, p. 13, tradução nossa), que define dados de pesquisa como “registros de fatos usados como fontes primárias de pesquisa científica e que são comumente aceitos na comunidade científica como necessários para validar resultados da pesquisa”¹².

Em seu livro *Big data, Little Data, No Data: scholarship in the networked world*¹³, Borgman além de dissertar sobre dados e dados de pesquisa, faz uma associação às bolsas de estudos na academia, afirmando que “[...] dados são representações de observações, objetos ou outras entidades usadas como evidência de

¹² Texto original: factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings.

¹³ Texto original: Megadados, microdados, sem dados: bolsa de estudos no mundo em rede [tradução nossa baseada na proposta de Silva Segundo e Araújo (2019, p. 113)]

fenômenos para fins de pesquisa ou bolsa de estudos”¹⁴ (BORGMAN, 2015, p. 28, tradução nossa).

Outra definição de dados de pesquisa que acrescenta seu valor ao empirismo e à *e-Science*, reconhecendo-a como essa forma de pesquisa colaborativa em rede se encontra em:

Os dados são o resultado de avanços incrementais na tecnologia que serve ao empirismo. Eles informam a teoria, conduzem e validam simulações e são mais bem atendidos por uma comunicação padrão, bidirecional, com essas camadas da rede de conhecimento (WILBANKS, 2011, p. 219).

Arrazoando sobre o objetivo, por parte do pesquisador, de se obter e/ou gerar dados, Borgman (2015, p. xviii) afirma que eles “[...] são um meio para atingir um fim, que geralmente é o artigo de revista, livro, jornal de conferência ou outro produto digno de reconhecimento acadêmico.”¹⁵, evidenciando para a comunicação científica que o(a) cientista não trabalha somente para si, mas para toda uma rede que constitui a ciência.

Sobre isso, Borgman (2015) complementa ao afirmar que dados são “representações de observações, objetos e outras entidades utilizadas como evidências de um fenômeno para propósitos de pesquisa” (p.42, tradução nossa).

O pesquisador deverá se responsabilizar quanto à qualidade dos dados de pesquisa, qualidade esta que transpõe a capacidade desses dados gerar informação e/ou conhecimento, e recai sobre a qualidade de curadoria desses dados, permitindo que estes possam ser acessados e compreendidos futuramente com objetivo de uso e reúso não só na replicação do experimento para o qual foi gerado, mas também por outros experimentos a serem feitos. Nesse sentido:

Devemos pensar em nós mesmos como módulos conectados que precisam passar dados, testar teorias, acessar as simulações de cada um. E, considerando que todos os gráficos sobre a capacidade de coleta de dados estão disparando exponencialmente, também precisamos escalonar nossa capacidade de usar esses dados – e precisamos muito disso. Precisamos colocar nosso conhecimento e a nós mesmos em rede. Nada que tenhamos projetado até hoje, como humanos, conseguiu crescer tão rápido como uma rede aberta (WILBANKS, 2011, p. 221).

¹⁴ Texto original: data are representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research or scholarship.

¹⁵ Texto original: Data are a means to an end, which is usually the journal article, book, conference paper, or other product worthy of scholarly recognition.

Este trabalho concorda com o posto por Semeler e Pinto (2019, p. 115) na definição de que “dados de pesquisa científica são o resultado de qualquer investigação sistemática que envolva processos de observação, experimentação ou simulação de procedimentos de pesquisa científica”.

Ao gerar os dados de pesquisa, o(a) cientista deve ponderar que ele não é mais aquele que trabalha sozinho em sua garagem sem apoio, mas que agora a ciência é global, e as pesquisas científicas são realizadas em rede. “Grosso modo, o paradigma que precisa ser destruído é a ideia de que nós, cientistas, existimos como indivíduos fora da rede” (WILBANKS, 2011, p. 219).

Não obstante o surgimento do *big data*, os recursos disponíveis de tecnologias digitais na comunicação científica e os pressupostos da ciência aberta revelam que “Estimativas recentes indicam que mais de 50% dos achados científicos não aparecem na literatura publicada e residem em gavetas de arquivos e discos rígidos pessoais” (FERGUSON et al, 2014, p. 1443). Esse fenômeno prejudica todo o sistema de pesquisa científica, principalmente no contexto da *e-Science*, gerando um registro incompleto de dados e, muitas vezes, uma duplicação desnecessária de esforços científicos que seriam evitados ao propiciar acesso, contribuindo para falhas na replicação dos experimentos realizados. Essa ciência invisível se encontra na cauda longa dos dados e “[...] constitui um território onde os dados são altamente heterogêneos e dispersos entre instituições, projetos, laboratórios, pequenos grupos de pesquisa e pesquisadores individuais” (SALES, SAYÃO, 2018, p. 4185).

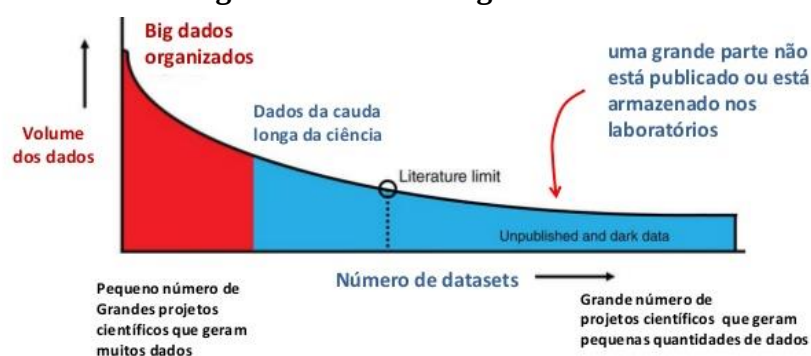
Dessa forma, a cauda longa nos permite visualizar a disparidade na distribuição dos megadados utilizados entre os campos científicos ou sociais. Esse formato de visualização de dados é um gráfico matemático que se popularizou em 2004 por Chris Anderson quando descreveu o mercado de bens em lojas físicas *versus* lojas on-line. Em sua descrição, o autor aponta que cerca de 15% (quinze por cento) da distribuição está na cabeça da curva e os demais 85% (oitenta e cinco por cento) estão distribuídos ao longo da cauda (BORGMAN, 2015). Quando aplicado à pesquisa científica, verifica-se que um grupo pequeno de pesquisadores trabalham com megadados e um grande número de pesquisas científicas são geradas com pequenos dados. Uma descrição mais detalhada, neste caso para a neurociência, é fornecida por Ferguson *et al.* (2014):

Os dados de cauda longa em neurociência podem ser definidos como conjuntos de dados pequenos e granulares, coletados por laboratórios individuais no decorrer da pesquisa do dia-a-dia. Esses dados consistem em pequenas unidades publicáveis (por exemplo, desfechos clínicos direcionados), bem como desfechos clínicos alternativos,

dados paramétricos, resultados de estudos-piloto e metadados sobre os dados publicados. A cauda longa dos dados também é composta de "dados escuros", dados não publicados que incluem resultados de experiências com falha e registros que são vistos como auxiliares de estudos publicados (por exemplo, registros de cuidados veterinários). Embora esses dados possam não ser considerados úteis no sentido tradicional, os esforços de compartilhamento de dados podem iluminar informações e descobertas importantes ocultas nesta longa cauda (FERGUSON *et al*, 2014, p. 1442).

De forma ilustrativa, na “cabeça” do gráfico ilustrado na Figura 4, encontra-se um pequeno grupo de grandes projetos científicos muito bem organizados e com enormes recursos financeiros, gerando consideráveis quantidades de dados, megadados, estruturados e homogêneos completamente inseridos na *e-Science*, trabalhando com Ciência Orientada a Dados e na Ciência Aberta; do outro lado, na cauda longa, há um grande número de projetos científicos, porém, com menos recursos e gerando pequenas quantidades de dados, esses dados são poucos estruturados e heterogêneos. Na cauda longa se localiza a ciência invisível, os dados científicos ficam “perdidos” nos computadores pessoais, HDs ou até *pendrives* do(a)s cientistas que ali atuam, quando mais esses grupos de pesquisas se afastam da cabeça, menores os recursos e mais invisíveis os dados ficam (BORGMAN, 2015; FERGUSON *et al*, 2014; SAYÃO, SALES, 2019).

Figura 4 – Cauda longa da ciência



Fonte: adaptado de Sayão e Sales (2017, p.8)

A localização da ciência na cauda longa não significa que os dados dispostos são de qualidade inferior, de menor relevância ou que podem ser descartados no meio científico, pelo contrário, dada a sua heterogeneidade resultam num “território de altos coeficientes de autonomia, que favorecem a inovação e a geração de conhecimentos multi e interdisciplinares” (SAYÃO, SALES, 2019, p. 42).

Sendo assim, essas características da cauda longa demonstram a necessidade de uma melhor gestão desses dados e a importância de uma curadoria de dados científicos, tendo em vista que a “curadoria digital envolve manter, preservar e agregar valor aos dados de pesquisa digital durante todo o seu ciclo de vida¹⁶” (Digital Curation Centre, *online*, tradução nossa).

Observando o funcionamento da cauda longa, nota-se que o investimento público em pesquisa soma grandes volumes, porém poucos dados dessas pesquisas são disponibilizados à sociedade, o que ressalta a importância de uma mudança dessa prática, considerando que “[...] a ciência aberta expressa um novo modelo do processo da produção e comunicação do conhecimento refletida nas relações entre ciência, tecnologia, informação e inovação” (HENNING *et al*, 2019, p. 390).

A Ciência Aberta é, pois, uma realidade cada vez mais reconhecida e praticada, assim, com ela e para ela se destaca a necessidade de princípios que facilitem e permitam a gestão de dados nesse universo. Partindo dessa perspectiva, o *Digital Curation Centre* (DCC) (Centro de Curadoria Digital) é um centro especializado e reconhecido internacionalmente em curadoria digital de dados de pesquisa, que foi criado em 1º de março de 2004, através de um consórcio formado pelas Universidades de Edimburgo e Glasgow UKOLN na *University of Bath* e *Science and Technology Facilities Council* (STFC), com o objetivo de estabelecer um centro nacional para resolver desafios de curadoria digital que não seria simples para nenhuma instituição ou disciplina.

Essa organização pontua a necessidade da curadoria digital para a pesquisa científica e, dessa forma, mostra sua importância, principalmente, para “desinvisibilizar” a ciência da cauda longa:

O gerenciamento ativo dos dados de pesquisa reduz as ameaças ao seu valor de pesquisa a longo prazo e reduz o risco de obsolescência digital. Enquanto isso, dados selecionados em repositórios digitais confiáveis podem ser compartilhados entre a comunidade de pesquisa mais ampla do Reino Unido.

Além de reduzir a duplicação de esforços na criação de dados de pesquisa, a curadoria aprimora o valor de longo prazo dos dados existentes, disponibilizando-os para pesquisas adicionais de alta qualidade¹⁷ (Digital Curation Centre, *online*, tradução nossa).

¹⁶ Texto original: Digital curation involves maintaining, preserving and adding value to digital research data throughout its lifecycle.

¹⁷ Texto original: The active management of research data reduces threats to their long-term research value and mitigates the risk of digital obsolescence. Meanwhile, curated data in trusted digital repositories may be shared among the wider UK research community.

Assim, a partir dessa necessidade, surgem na *e-Science* novos modelos científicos denominados de ciclo de vida dos dados (CVD), conhecidos na literatura internacional como *Data Lifecycle*. Sobre isso, Anjos (2019, p. 39) informa que são “[...] modelos aplicados ao processo de curadoria de dados científicos, sendo compostos por etapas de gestão e compartilhamento de dados que, em sua maioria, objetivam a preservação desses dados para gerações atuais e futuras de usuários”

Sayão e Sales (2015, p. 11) ponderam que “Os dados e as coleções de dados de pesquisa possuem um tempo de vida maior que os projetos de pesquisa que os criaram.” Portanto, eles podem ser utilizados mesmo depois de findados os projetos e financiamentos do qual são provenientes, sendo reutilizados em novos projetos de pesquisa, reiniciando um novo ciclo de vida para os futuros estudos. Sobre eles, entende-se que esse modelo é “[...] constituído em estágios que conduzem à gestão dos dados científicos e estimulam a colaboração entre pesquisadores e instituições através do compartilhamento de parcerias institucionais” (OLIVEIRA, 2016, p.19).

Sobre os modelos de Ciclo de Vida de Dados é importante saber que existem muitos modelos, a saber: *Digital Curation Centre*, *Curation Lifecycle Model*¹⁸, *DataONE Data Lifecycle*¹⁹, *DDI Combined Lifecycle Model*²⁰ e *Uk Data Archive Data Lifecycle*²¹, por exemplo. Anjos (2019), em sua dissertação, utiliza de três modelos: o *Dataone* o *Data Lifecycle*, o *Digital Curation Centre* e um modelo proposto por Sant’ana para uso pela Ciência da Informação. Para o propósito dessa investigação, harmoniza-se com os modelos referenciados por Anjos (2019) e, para efeitos dessa pesquisa, adotar-se-á os modelos da *DataONE*, *DCC* e de Sant’Ana.

Para o *Data Observation Network for Earth* (*DataONE*), o gerenciamento de dados tem o objetivo de produzir conjuntos de dados autoexplicativos, que permitam ao pesquisador, que for usar ou reusar, a capacidade de fazê-lo de forma eficaz e adequada, mesmo sem ter participado da pesquisa que os geraram. Nesse sentido, são apontadas as práticas fundamentais de gerenciamento de dados que permitirão desenvolver um PGD, além de como criar, organizar, gerenciar, descrever, preservar e compartilhar dados com

As well as reducing duplication of effort in research data creation, curation enhances the long-term value of existing data by making it available for further high quality research.

¹⁸ <http://www.dcc.ac.uk/resources/curation-lifecycle-model>

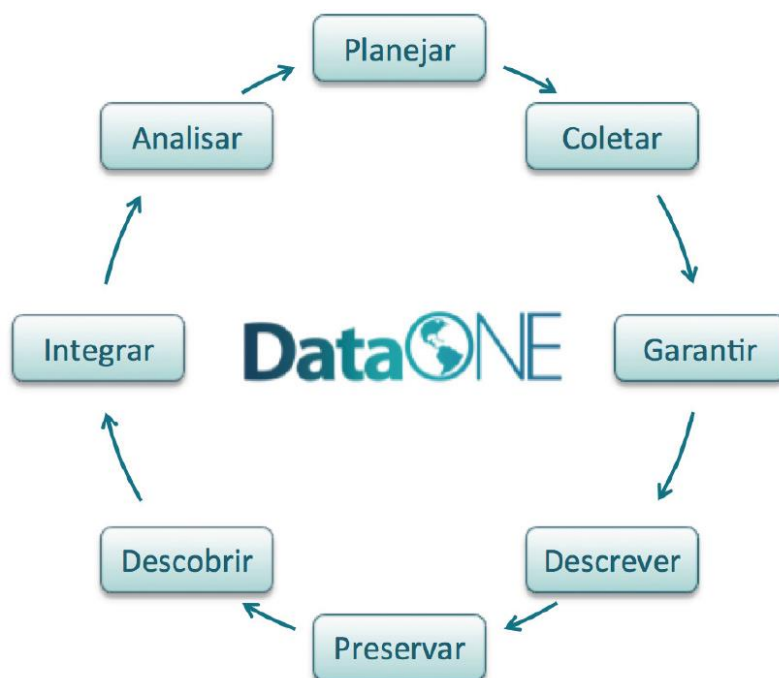
¹⁹ <https://escholarship.org/uc/item/7tf5q7n3>

²⁰ <https://ddialliance.org/Specification/DDI-Lifecycle/>

²¹ <https://www.data-archive.ac.uk/>

eficácia (STRASSER; COOK; MICHENER; BUDDEN, 2012), conforme se observa na imagem do Ciclo de vida dos dados:

Figura 5 – Ciclo de vida dos dados



Fonte: Córdula e Araújo (2019, p.198) traduzido de DataOne (2017, online).

O ciclo de vida dos dados do DataONE possui oito componentes:

Planejar: descrição dos dados que serão compilados e como os dados serão gerenciados e disponibilizados ao longo de sua vida útil;

Coletar: as observações são feitas manualmente ou com sensores ou outros instrumentos e os dados são colocados em um formato digital;

Garantir: a qualidade dos dados é garantida através de verificações e inspeções;

Descrever: os dados são descritos de maneira precisa e completa usando os padrões de metadados apropriados;

Preservar: os dados são enviados para um arquivo de longo prazo apropriado (ou seja, data center);

Descobrir: dados potencialmente úteis são localizados e obtidos, juntamente com as informações relevantes sobre os dados (metadados);

Integrar: dados de fontes diferentes são combinados para formar um conjunto homogêneo de dados que pode ser facilmente analisado;

Analisar: os dados são analisados²². (DATAONE, 2020, *online*, grifo nosso, tradução nossa)

Outro modelo de CVD é o proposto pelo DCC, no Reino Unido. Criado em 2011, a DCC trabalha com instituições de pesquisa com o objetivo de fornecer suporte personalizado para o desenvolvimento de serviços, desde o gerenciamento de dados de pesquisa, avaliação de necessidades até a implementação de políticas, suporte, infraestrutura e serviços. Concerne, sobre isso, afirma que "[...] um centro de especialização reconhecido internacionalmente em curadoria digital, com foco na criação de capacidades e habilidades para o gerenciamento de dados de pesquisa"²³ (DCC, 2020b, *online*, tradução nossa).

Para a DCC, a curadoria digital envolve as ações de manter, preservar e agregar valor aos dados da pesquisa digital durante todo o seu ciclo de vida, através de um gerenciamento desses dados, além de reduzir a duplicação de esforços na criação de dados de pesquisa. Para isso, a curadoria aprimora o valor de longo prazo dos dados existentes, disponibilizando-os para pesquisas adicionais de alta qualidade:

A curadoria digital e a preservação de dados são processos contínuos, exigindo considerações consideráveis e o investimento de tempo e recursos adequados. Você deve estar ciente e realizar ações para promover a curadoria e preservação ao longo do ciclo de vida dos dados. (DCC, 2020a, *online*, tradução nossa)

No XIV Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB 2013), Sant'Ana (2013) apresentou uma proposta de modelo de Ciclo de Vida dos Dados para a Ciência da Informação por acreditar que essa área científica possui um

²²Texto original: The DataONE data life cycle has eight components:

Plan: description of the data that will be compiled, and how the data will be managed and made accessible throughout its lifetime

Collect: observations are made either by hand or with sensors or other instruments and the data are placed a into digital form

Assure: the quality of the data are assured through checks and inspections

Describe: data are accurately and thoroughly described using the appropriate metadata standards

Preserve: data are submitted to an appropriate long-term archive (i.e. data center)

Discover: potentially useful data are located and obtained, along with the relevant information about the data (metadata)

Integrate: data from disparate sources are combined to form one homogeneous set of data that can be readily analyzed

Analyze: data are analyzed

²³ Texto original: an internationally-recognised centre of expertise in digital curation with a focus on building capability and skills for research data management.

papel preponderante na construção de novos arcabouços teóricos e na definição de caminhos para se obter “[...] novos conhecimentos e técnicas, cobrindo todas as fases do acesso a dados, desde o planejamento sobre a coleta e geração até a visualização e neste sentido ganha importância a participação de todas as áreas do conhecimento” (SANT’ANA, 2013, p. 2).

Sant’Ana (2013) pondera que se deve observar as três fases para o acesso aos dados em seu ciclo de vida para a Ciência da Informação: coleta, armazenamento e recuperação, sendo fatores que ocorrem durante todo o processo e não em uma fase específica do CVD como se observa na aformativa de que “ao analisar os exemplos de modelos de ciclo de vida de dados, percebe-se que existem aqueles que apresentam somente fases interligadas de forma linear e outros que apresentam fases permeadas por objetivos” (SANT’ANA, 2013, p. 15).

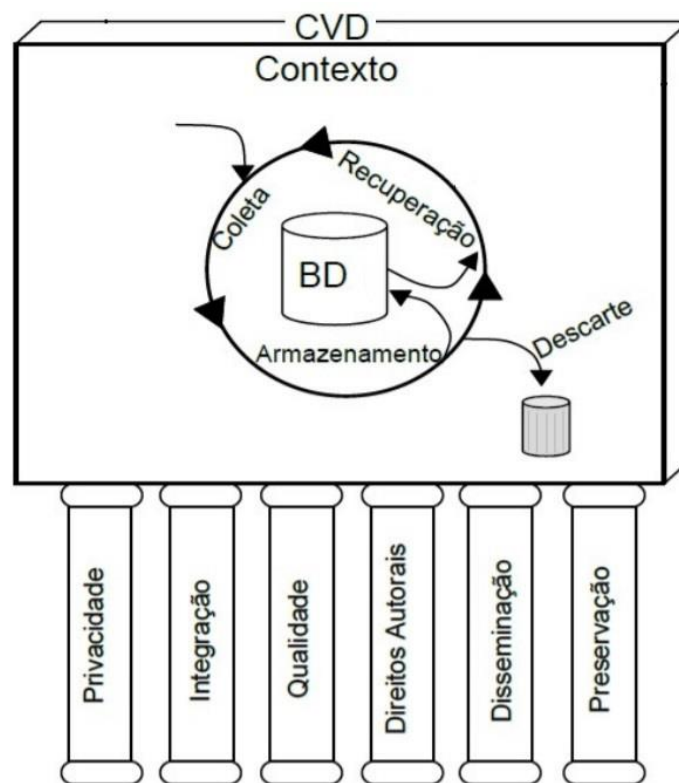
Em sua proposta, Sant’Ana (2013; 2019), considerando o estudo do comportamento da informação, seu fluxo e os meios para o seu acesso como objetivos da Ciência da Informação, propõe que as fases podem ser permeadas pelos objetivos de privacidade, qualidade, direitos autorais, integração, disseminação e preservação. Detalhadamente, as quatro fases são:

1. **Coleta** – definir os requisitos que os dados devem atender, delimitar o conteúdo a ser obtido e deliberar os procedimentos para a localização;
2. **Armazenamento** – com foco na manutenção dos dados coletados para que possam ser acessados em um momento futuro;
3. **Recuperação** – baseada nas questões relacionadas a como tornar estes dados acessíveis, incluindo o que poderá ser acessado, quem poderá fazê-lo e como este acesso será disponibilizado;
4. **Descarte** – esse processo pode possuir diversas motivações, desde aquelas relacionadas ao espaço de armazenamento, até as relacionadas ao direito dos envolvidos/referenciados.

Essas fases são profundamente entrelaçadas e interdependentes uma da outra e podem ocorrer concomitantemente. Na Figura 6, verifica-se que o forte vínculo entre as fases permite que o dado coletado possa ser armazenado e disponibilizado para recuperação, sendo que qualquer uma dessas fases pode desencadear na fase de descarte. Porém, existem alguns fatores que permeiam todas as fases, de forma

transversal, tais como: a privacidade, a integração, a qualidade, os direitos autorais, a disseminação e a preservação.

Figura 6 – Ciclo de vida dos dados para Ciência da Informação



Fonte: Sant'Ana (2019, p.20).

Diante da necessidade de gestão e curadoria dos dados, dentro de um leque de ferramentas, no que compete ao pesquisador, destaca-se o plano de gestão de dados, um documento pré-pesquisa que se junta a estes processos no intuito de consolidar a preservação dos dados para posterior compartilhamento, acesso, uso e reúso.

Sendo assim, o PGD “[...] envolve a gestão de dados de pesquisa desde a sua gênese, assegurando a preservação por longo prazo, descoberta, interpretação e reúso” (SAYÃO; SALES, 2012, p. 179). Embora possam diferir em formato e conteúdo, vários elementos básicos são essenciais para o gerenciamento eficaz dos dados. A princípio, os dados devem ser gerenciados para que qualquer pesquisador, coletor ou criador dos dados, seja capaz de descobrir, usar e interpretar os dados após um período de tempo. Dessa forma, “Um componente chave do gerenciamento de dados é a descrição abrangente dos dados e informações contextuais que futuros pesquisadores precisam entender para usar os dados” (STRASSER; COOK; MICHENER; BUDDEN, 2012, p. 2).

No Brasil, a Comissão Nacional de Energia Nuclear ligada ao Ministério da Ciência, Tecnologia, Inovações e Comunicações (BRASIL, 2021, *online*) sugere 10 razões para a criação de um Plano de Gestão de Dados de pesquisa (como uma ferramenta da curadoria):

1. Ajustar o seu projeto de pesquisa às **políticas mandatórias da sua instituição** e/ou dos **órgãos de fomento à pesquisa**;
2. Assegurar a **integridade** da pesquisa e o seu potencial de replicação;
3. Assegurar que os dados e demais registros de pesquisa sejam **acurados, completos, autênticos e confiáveis**;
4. Aumentar a sua **eficiência** como pesquisador – um plano que organize os dados e seu armazenamento permite que você foque na sua pesquisa. Você estará mais capacitado a localizar e usar os seus dados e compartilhá-los com os seus colaboradores;
5. Permitir que os seus dados sejam compreensíveis **agora e no futuro** – se os dados são bem documentados antes e durante a formação da coleção de dados, eles serão mais facilmente entendidos e reutilizados;
6. **Economizar tempo e recursos** a longo prazo;
7. Aumentar a **segurança dos dados** e minimizar os riscos de perda;
8. **Evitar a duplicação** de esforços na coleta ou regeneração dos dados, possibilitando que outros pesquisadores se beneficiem dos seus dados e os interprete em outros contextos e com novas visões;
9. **Aumentar a visibilidade da pesquisa** – se os seus dados foram planejados para estarem organizados e corretamente arquivados, eles poderão ser identificados, recuperados e citados, aumentando a visibilidade da sua pesquisa e o seu prestígio como pesquisador;
10. Tornar mais fácil a **preservação e o arquivamento** – ajustando antecipadamente a geração de dados tomando como referência as práticas, formatos e demais padrões mais adequados ao arquivamento e à preservação de longo prazo, torna a gestão de dados mais fácil e menos custosa; além do mais, tornam os dados mais aderentes aos requisitos dos repositórios e centros de dados. (BRASIL, 2021, *online*)

A FAPESP²⁴ expõe em sua página *web* a importância do PGD informando “que deve responder a duas perguntas básicas: 1. Quais dados serão gerados pelo projeto?; 2. Como serão preservados e disponibilizados, considerando questões éticas, legais, de confidencialidade e outras?”

Diante esse zelo para com os dados de pesquisa e a comunicação científica, visando apoiar uma gestão desses elementos fortemente evidenciados no quarto paradigma e na Ciência Orientada a Dados, é preciso ter como base um arcabouço tecnológico e gerencial que atendam às fases do CVD. No cerne desse arcabouço estão:

²⁴ SÃO PAULO. Secretaria de Desenvolvimento Econômico. Plano de Gestão de Dados - FAPESP. Elaborada por Fundação de Amparo à Pesquisa do Estado de São Paulo. Disponível em: <http://www.fapesp.br/gestaodedados/>. Acesso em: 26 ago. 2019.

[Os] **repositórios digitais de dados de pesquisa** que, por muitas razões e demandas, rapidamente se tornam parte essencial da infraestrutura mundial de pesquisa. Duas dessas demandas são determinantes para a ampliação das ações em torno da gestão de dados e para o seu ordenamento: as políticas mandatórias das agências financiadoras de pesquisa e a incorporação pelas instituições de pesquisa e pela sociedade em geral dos valores e princípios preconizados pela Ciência Aberta (SAYÃO; SALES, 2016, p. 92).

Quando se trabalha com um conjunto de dados, principalmente com uso intensivo de dados, segundo Wilkinson (2016), algumas questões que criam barreiras à descoberta e reutilização de dados, não apenas para humanos, mas ainda mais para máquinas devem ser levantadas: onde publicar o conjunto de dados? Como pesquisar dados em conjunto de dados? Como os dados podem ser baixados? Em quais formato(s) podem ficar disponíveis? Esses dados podem ser integrados a outros conjuntos de dados de outros repositórios? Essa integração pode ser de fácil execução, sem a necessidade do "copiar e colar", visando evitar erros? O pesquisador tem licença para utilizar esses dados?

Nesse sentido, para possibilitar a Ciência Aberta no contexto da *e-Science*, é preciso pensar na “criação de princípios e práticas que pudessem orientar a descoberta, o acesso, a integração e a reutilização da vasta quantidade de dados e informação gerada pela ciência” (HENNING *et al*, 2019, p. 390).

Dentre estes princípios, destacam-se o que se denomina de FAIR, acrônimo de: Encontráveis (*Findable*), Acessíveis (*Accessible*), Interoperáveis (*Interoperable*) e Reutilizáveis (*Reusable*) (DIAS; ANJOS; RODRIGUES, 2019, p; 179). Segundo esses princípios, cada dado terá em seu metadado um identificador único que o torna encontrável e acessível, a partir de uma linguagem comum formalizada com um bom mecanismo de busca, com amplas referências cruzadas que o torna interoperável e capaz de distinguir proveniências e licença, dando-lhe capacidade de ser reutilizável (RODRIGUEZ-IBLESIAS *et al*, 2016, p.1 *apud* DIAS; ANJOS; RODRIGUES, 2019, p; 179). Essas características são melhores descritas no Quadro 2.

Quadro 2 – Princípios FAIR

LOCALIZÁVEL (FINDABLE)

O primeiro passo para reusar os dados é encontrá-los. Metadados e dados devem ser fáceis de encontrar para humanos e computadores. Os metadados legíveis por máquina são essenciais para a descoberta automática de conjuntos de dados e serviços.

F1	(Meta)dados recebem um identificador persistente e globalmente exclusivo: o principal atributo, pois os demais atributos dependem deste, uma espécie de "CPF" do dado. O identificador único deve atender a duas condições: (1) deve ser globalmente exclusivo (ou seja, alguém não pode reutilizar / reatribuir o mesmo identificador sem se referir aos seus dados). Você pode obter identificadores globalmente exclusivos de um serviço de registro que usa algoritmos que garante a exclusividade dos identificadores recém-criados; (2) deve ser persistente. É preciso tempo e dinheiro para manter os <i>links</i> da <i>web</i> ativos; portanto, os <i>links</i> tendem a se tornar inválidos ao longo do tempo. Os serviços de registro garantem a resolução desse <i>link</i> no futuro, pelo menos até certo ponto.
F2	Dados são descritos com metadados avançados (definidos por R1 abaixo): os metadados podem (e devem) ser generosos e extensivos, incluindo informações descritivas sobre o contexto, qualidade e condição ou características dos dados. Os metadados ricos permitem que um computador realize automaticamente tarefas rotineiras e tediosas e priorize tarefas que atualmente exigem muita atenção dos pesquisadores. A lógica por trás desse princípio é que alguém deve ser capaz de encontrar dados com base nas informações fornecidas por seus metadados, mesmo sem o identificador dos dados.
F3	Metadados incluem clara e explicitamente o identificador dos dados que descreve: os metadados e o conjunto de dados que eles descrevem geralmente são arquivos separados. A associação entre um arquivo de metadados e o conjunto de dados deve ser explicitada, mencionando os identificadores globalmente exclusivos e persistentes de um conjunto de dados nos metadados.
F4	(Meta)dados são registrados ou indexados em um recurso pesquisável: identificadores e descrições completas de metadados não garantirão a 'localização' na Internet. Recursos de dados perfeitamente bons podem não ser utilizados simplesmente porque ninguém sabe que eles existem. Há muitas maneiras pelas quais os recursos digitais podem ser descobertos, incluindo a indexação. Por exemplo, o <i>Google</i> envia <i>spiders</i> ²⁵ que 'leem' as páginas da <i>web</i> e as indexam automaticamente, para que possam ser encontradas na caixa de pesquisa do <i>Google</i> . Para dados de pesquisas acadêmicas, a indexação é fundamental. Os princípios (F1) e (F3) fornecerão os elementos principais para a indexação refinada de alguns repositórios atuais e serviços futuros.
ACESSÍVEL (ACCESSIBLE) Quando o usuário encontra os dados necessários, ele precisa saber como eles podem ser acessados, possivelmente incluindo autenticação e autorização.	
A1	(Meta)dados são recuperáveis pelo seu identificador usando um protocolo de comunicação padronizado: a recuperação de dados FAIR deve ser mediada sem a necessidade de especialistas, ferramentas proprietárias ou métodos de comunicação. Concentra-se em como os dados e metadados podem ser recuperados de seus identificadores, a autorização de uso desses dados (quem pode utilizá-los) devem estar claramente estabelecidas permitindo, inclusive, se for o caso, contato direto com o pesquisador que gerou dos dados.

²⁵ *Spider* são programas usados por mecanismos de busca para explorar a internet automaticamente e fazer *download* de conteúdo *web* de sítios *web*. Metodicamente, não expõe o conteúdo que julga irrelevante no código fonte dos sítios, e armazena o resto em seu banco de dados.

A1.1	O protocolo é aberto, gratuito e universalmente implementável: para maximizar a reutilização de dados, o protocolo deve ser gratuito, aberto e, portanto, implementável globalmente para facilitar a recuperação de dados. Qualquer pessoa com um computador e uma conexão à internet pode acessar pelo menos os metadados. Portanto, esse critério afetará a escolha do repositório em que você compartilhará seus dados.
A1.2	O protocolo permite um procedimento de autenticação e autorização, quando necessário: <i>Accessible</i> não significa a obrigatoriedade de os dados estarem abertos ao público em geral. Implica que se deve fornecer as condições exatas sob as quais os dados estão acessíveis. Portanto, mesmo dados altamente protegidos e privados podem ser FAIR. Idealmente, a acessibilidade é especificada de tal maneira que uma máquina possa entender automaticamente os requisitos e, em seguida, executá-los automaticamente ou alertar o usuário sobre os requisitos. Geralmente, faz sentido solicitar aos usuários que criem uma conta de usuário para um repositório. Isso permite autenticar o proprietário (ou colaborador) de cada conjunto de dados e potencialmente definir direitos específicos do usuário. Portanto, esse critério também afetará a escolha do repositório em que você compartilhará seus dados.
A2	Metadados são acessíveis, mesmo quando os dados não estão mais disponíveis: os conjuntos de dados tendem a se degradar ou desaparecer com o tempo, porque há um custo em manter uma presença <i>online</i> para recursos de dados. Quando isso acontece, os <i>links</i> se tornam inválidos e os usuários perdem tempo procurando dados que podem não estar mais nesse espaço. Armazenar os metadados geralmente é muito mais fácil e mais barato. Portanto, os metadados devem persistir, mesmo quando os dados não são mais sustentados. Está relacionado aos problemas de registro e indexação descritos em (F4).
INTEROPERÁVEL (INTEROPERABLE) Os dados geralmente precisam ser integrados a outros dados. Além disso, os dados precisam interoperar com aplicativos ou fluxos de trabalho para análise, armazenamento e processamento.	
I1	(Meta)dados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para representação do conhecimento: tantos humanos como máquinas devem ser capazes de trocar e interpretar os dados uns dos outros, portanto não se deve utilizar idiomas mortos e os dados devem ser legíveis por máquinas sem a necessidade de algoritmos, tradutores ou mapeamentos especializados ou <i>ad hoc</i> . Para que isso ocorra e para garantir a localização e interoperabilidade automáticas dos conjuntos de dados, é essencial usar (1) vocabulários, ontologias controlados comumente usados, tesauros (com identificadores persistentes e únicos globalmente resolvíveis, consulte (F1)) e (2) um bom modelo de dados (uma estrutura bem definida para descrever e estruturar (meta)dados).
I2	(Meta)dados usam vocabulários que seguem os princípios FAIR: o vocabulário controlado usado para descrever conjuntos de dados precisa ser documentado e resolvido usando identificadores persistentes e únicos em todo o mundo. Esta documentação precisa ser facilmente encontrada e acessível por qualquer pessoa que use o conjunto de dados.
I3	(Meta)dados incluem referências qualificadas a outros (meta)dados: uma referência qualificada é uma referência cruzada que explica sua intenção. O objetivo, portanto, é criar o maior número possível de <i>links</i> significativos entre os recursos (meta) de dados para enriquecer o conhecimento contextual sobre os dados, equilibrado com o tempo/energia envolvido na criação de um bom modelo de dados. Deve-se especificar se um conjunto de dados se baseia em outro conjunto de dados, se são necessários conjuntos de dados adicionais para concluir os dados ou se informações complementares são armazenadas em um conjunto de dados diferente. Em particular, os <i>links</i> científicos entre os conjuntos de dados precisam ser descritos. Além disso, todos os conjuntos de dados precisam ser citados corretamente (ou seja, incluindo seus identificadores persistentes e únicos globalmente).

REUSÁVEL (REUSABLE) O objetivo final do FAIR é otimizar a reutilização de dados. Para conseguir isso, os metadados e os dados devem ser bem descritos para que possam ser replicados e / ou combinados em diferentes configurações.	
R1	Meta(dados) são ricamente descritos com uma pluralidade de atributos precisos e relevantes: este princípio está relacionado a (F2), mas concentra-se na capacidade de um usuário (máquina ou humano) de decidir se os dados são realmente ÚTEIS em um contexto específico. Para tomar essa decisão, o publicador de dados deve fornecer não apenas metadados que permitam a descoberta, mas também metadados que descrevam detalhadamente o contexto em que os dados foram gerados. Além disso, (R1) afirma que o editor de dados não deve tentar prever a identidade e as necessidades do consumidor de dados. O autor dos metadados deve ser o mais generoso possível ao fornecer metadados, incluindo informações que possam parecer irrelevantes.
R1.1	(Meta)dados são liberados com uma licença de uso de dados clara e acessível: sai da área técnica e entra no campo jurídico. Os direitos de uso que você atribui aos seus dados deve ser descrito claramente. A ambiguidade pode limitar severamente a reutilização de seus dados por organizações que lutam para cumprir as restrições de licenciamento. A clareza do status de licenciamento se tornará mais importante com pesquisas automatizadas que envolvem mais considerações sobre licenciamento. As condições sob as quais os dados podem ser usados devem ser claras para máquinas e para seres humanos.
R1.2	(Meta)dados estão associados à proveniência detalhada: para que outras pessoas reutilizem seus dados, eles devem saber de onde vieram, quem citar e/ou como você deseja ser reconhecido. Inclua uma descrição do fluxo de trabalho que levou aos seus dados: quem os gerou ou coletou, como foi processado, se já foi publicado antes, se contém dados de outra pessoa que você pode ter transformado ou concluído. Idealmente, esse fluxo de trabalho é descrito em um formato legível por máquina.
R1.3	(Meta)dados atendem aos padrões da comunidade relevantes para o domínio: é mais fácil reutilizar conjuntos de dados se eles forem semelhantes: mesmo tipo de dados, dados organizados de maneira padronizada, formatos de arquivo bem estabelecidos e sustentáveis, documentação (metadados) seguindo um modelo comum e usando vocabulário comum.

Fonte: GO FAIR (2020, tradução nossa).

Com as práticas da *e-Science* de tratamento e uso intensivo de dados cada vez mais consolidada, em janeiro de 2014, se reuniram representantes de várias partes interessadas, a pedido do Centro de *e-Science* da Holanda e do *Dutch Techcentre for the Life Sciences*²⁶ (DTL), para discutirem como fornecer dados legíveis por máquina. Desse encontro surge a noção de que, através da definição e do amplo suporte de um conjunto mínimo de princípios e práticas orientadores, acordados pela comunidade, os provedores de dados e os consumidores de dados – máquinas e humanos – poderiam descobrir, acessar, interoperar e facilmente reutilizar, com citação adequada, as vastas quantidades de informações geradas pela ciência contemporânea. Desse acordo, com

²⁶ Centro de Tecnologia Holandês para Ciências da Vida

base nos princípios necessários direcionados ao uso intensivo de dados, surge os “FAIR Guiding Principles” (FORCE11, 2020, *online*).

Nessa nova forma de pesquisa, o(a) cientista deve coletar os dados de seu experimento, gerenciar, guardar e disponibilizar esses dados para que ele mesmo e os demais pesquisadores, no presente e/ou no futuro, possam usufruir desses dados através das tecnologias digitais de comunicação. Assim, a *e-Science* envolve pesquisadores de diversas localidades dando um lastro maior às suas investigações. Sobre esse processo, sabe-se que “Os resultados de um bom gerenciamento e administração de dados, portanto, são publicações digitais de alta qualidade que facilitam e simplificam esse processo contínuo de descoberta, avaliação e reutilização em estudos posteriores”²⁷ (WILKINSON et al., 2016, p. 1)

Considerando essa ampliação no alcance das pesquisas, é importante ressaltar que a estrutura central que possibilita a gestão de dados e a curadoria desses dados é o repositório digital de dados de pesquisa que, segundo Sayão e Sales (2016, p. 95) “têm um papel importante nas interações que envolvem a validação do trabalho de pesquisa e na própria dinâmica social da comunicação científica”, assim, seguindo os princípios FAIR, possibilitará avigorar a revisão por pares, alcançando uma comunidade de pesquisa mais ampla e conectada em rede. Consequentemente, “Um repositório permite exame, prova, revisão, transparência de resultados de pesquisa por outros especialistas que vão além da revisão por pares do artigo acadêmico publicado” (UZWYSHYN, 2016, p. 1 apud SAYÃO; SALES 2016, p. 95). Na perspectiva do uso e reúso intensivo de dados de pesquisa, no contexto de *e-Science* e uso de modelos como o FAIR e de repositório de dados, uma ferramenta essencial é a representação desses dados de pesquisa, o uso de metadados.

3.3 Metadados para dados de pesquisa

Para Borgman (2015), as pessoas possuem o hábito de classificar as coisas, quando o primeiro passo é geralmente nomeá-las para depois categorizá-las organizadamente em hierarquias, redes, gráficos ou outros relacionamentos com diferentes graus de formalidade. Nessa perspectiva, a autora traz uma definição simples

²⁷ Texto original: The outcomes from good data management and stewardship, therefore, are high quality digital publications that facilitate and simplify this ongoing process of discovery, evaluation, and reuse in downstream studies.

para metadados, expondo que “‘dados sobre dados’, são um meio de nomear coisas, representar dados e representar relacionamentos”²⁸ (BORGMAN, 2005, p. 66).

Caplan (2003) versa que a Biblioteconomia e a Ciência da Informação possuem grande interesse em metadados, eles advêm da Computação, são usados para descrever outros dados, no qual o prefixo "meta" é usado para significar "sobre". Caplan (2003) evidencia que esse termo foi usado pela primeira vez no sentido de dados sobre dados, na primeira edição do Manual de Formato de Intercâmbio de Diretórios da NASA, publicado em 1988.

No entanto, para Pomerantz (2015), definir metadados como “dados sobre dados” é insuficiente, o autor pondera que uma definição que mais se aproxima da realidade contemporânea é que os metadados, basicamente, são dados que descrevem dados:

Os dados devem ser entendidos não como um conceito abstrato, mas como objetos potencialmente informativos. Em seguida, os metadados podem ser definidos como ‘um objeto potencialmente informativo que descreve outro objeto potencialmente informativo’²⁹ (POMERANTZ, 2015, p. 26, tradução nossa).

Pomerantz (2015) afirma que, ao contrário da gramática em que o objeto de uma frase é a entidade sobre a qual o sujeito age, para os metadados o assunto é a entidade que está sendo descrita e o objeto é outra entidade que está sendo usada para descrever o assunto. Ou seja, na gramática há um objeto e uma entidade, nos metadados há duas entidades, uma que está sendo usada para descrever o assunto, e outra que está descrevendo a primeira entidade. Nesse sentido, é possível afirmar que, inclusive, os metadados podem ser representação dos dados de pesquisa ou ser usado como dados de pesquisa.

A aplicabilidade de metadados depende de um determinado conjunto de regras que normatizam e padronizam uma estrutura de descritivos sobre os dados na forma de um esquema e “Um esquema de metadados é, portanto, uma linguagem muito simples, com um pequeno número de regras”³⁰ (POMERANTZ, 2015, p. 30). Ponderando a

²⁸ Texto original: “data about data,” are a means to name things, to represent data, and to represent relationships

²⁹ Texto original: Data must be understood not as an abstract concept but as objects that are potentially informative. Then metadata can be defined as “a potentially informative object that describes another potentially informative object.”

³⁰ Texto original: is the set of rules according to which a language operates. A metadata schema is therefore a very simple language, with a small number of rules.

partir dessa perspectiva, os metadados podem ser aplicados tanto para humanos quanto para máquinas, observando-se a seguinte definição de metadados:

[pode-se] aplicar a qualquer tipo de dados, eletrônicos ou não, sem se importar se os metadados são destinados para compreensão humana ou de máquinas, porém pondera duas restrições, primeiro as informações devem ser estruturadas ou seja, registradas de acordo com algum esquema de metadados documentados, e segundo os metadados devem descrever um recurso de informação onde "Metadados aqui são usados para significar informações estruturadas sobre um recurso de informações de qualquer tipo ou formato de mídia"³¹ (CAPLAN, 2003, p. 3).

Percebe-se a importância dos metadados posto que “em metadados, a primeira idéia que nos ocorre é inspirada no seu uso no ambiente da biblioteca; no seu papel de um esquema formal para descrição de todo tipo de objetos informacionais, digitais e não digitais” (SAYÃO, 2010, p. 2). Nesse estudo, Sayão (2010) evidencia também uma noção expandida de metadados no intuito de apoiar a gestão de objetos digitais, cujo escopo inclui os processos de preservação digital de longo prazo.

No contexto da *e-Science*, nota-se a presença dos metadados em todos os momentos da pesquisa científica, tanto no ciclo de vida dos dados, quanto nos princípios FAIR, além de verificar que, no que compete à organização e à representação da informação, tanto no aspecto profissional, como na organização realizada por bibliotecários,—observa-se que os metadados sempre existiram, embora sem essa denominação.

Na contemporaneidade, principalmente com a ascensão da importância dos dados de pesquisa advindo da Ciência Aberta, a necessidade cada vez maior de um consenso de definições e melhor qualidade na geração desses metadados fazem surgir organizações como a Metadata 2020, formada por diversos profissionais dos mais diferentes domínios científicos. “Metadados 2020 é uma colaboração que defende metadados abertos mais ricos, conectados e reutilizáveis para todas as saídas de pesquisa, o que promoverá atividades acadêmicas em benefício da sociedade.” (METADATA, 2020, 2020a, *online*).

Essa organização ressalta a importância do(a) bibliotecário(a) nesse contexto, trazendo uma página exclusiva para informar as dificuldades, desafios e oportunidades desse profissional: “O(a)s bibliotecário(a)s enfrentam desafios ao se comunicar com os

³¹ Texto original: Metadata is here used to mean structured information about an information resource of any media type or format.

pesquisadores sobre metadados; ao usar metadados criados por outras pessoas; e ao aplicar metadados a objetos digitais³²” (METADATA, 2020, 2020b, *online*).

Por fim, a partir do pressuposto de dados como números, letras e símbolos que podem ser quantificados ou quantificáveis, e de dados de pesquisa científica que têm como finalidade validar os resultados da pesquisa, compreende-se que, no contexto da *e-Science*, os dados científicos devem ser encontráveis, acessíveis, interoperáveis e reutilizáveis nesse cenário colaborativo de prática científica. No que tange à Ciência da Informação, principalmente na Biblioteconomia, os dados de pesquisa científica devem ser organizados e representados se utilizando, principalmente, de metadados, classificando-os e indexando-os.

³² Texto original: Librarians face challenges when communicating with researchers about metadata;

4 A ATUAÇÃO DO(A) BIBLIOTECÁRIO(A) E OS DADOS NA CIÊNCIA DA INFORMAÇÃO

Num período em que o acesso ao conhecimento laico era controlado pela Igreja, a primeira biblioteca que se tem conhecimento no Brasil foi aberta em 1568, no Colégio da Bahia, dentro de uma instituição de ensino dos Jesuítas no Brasil Colonial, porém, as atividades bibliotecárias iniciam-se somente em 1604 por Antônio Gonçalves (FONSECA, 1979). No entanto, o primeiro curso de Biblioteconomia foi criado apenas em 1911, na Biblioteca Nacional.

O primeiro curso de Biblioteconomia no Brasil foi instituído pela Biblioteca Nacional no Rio de Janeiro, em 1911, basicamente com quatro disciplinas: Bibliografia, Paleografia e Diplomática, Iconografia e Numismática. Foi extinto em 1922 e restaurado em 1933, mantendo o caráter humanista, no entanto, incluía ensino de técnicas bibliotecárias como catalogação e classificação (antes ensinadas como parte da disciplina de Bibliografia) (MUELLER, 1985, CASTRO, 2000). Ainda sobre o curso, “O ensino da Biblioteca Nacional era influenciado pela escola francesa *École de Chartes* com forte característica humanística e voltada para os funcionários daquela biblioteca” (ALMEIDA; BAPTISTA, 2013, p. 2).

O colégio Mackenzie foi criado em 1870, em São Paulo, com ideias de ensino revolucionárias para aquele tempo, abolindo castigos como a palmatória, aceitando alunos à época marginalizados, como filhos de abolicionistas, republicanos, protestantes, judeus e mulheres. É nesse ambiente que, em 1929, é implantado o segundo curso de Biblioteconomia do Brasil, seguindo um modelo pragmático de ensino de Biblioteconomia e de organização de bibliotecas sob influência estadunidense, sendo marcado pelas disciplinas de caráter técnico, orientado pela bibliotecária americana Dorothy Muriel Gedds Gropp (CASTRO, 2000, p. 64-65).

Também em São Paulo, Adelpha Figueiredo e Rubens Borba de Moraes criam o Curso de Biblioteconomia do Departamento de Cultura da Prefeitura Municipal de São Paulo, em 1936, mesma época em que se encerra o curso do Mackenzie. Preocupados com o fato das bibliotecas brasileiras serem antiquadas e os cursos de Biblioteconomia existentes até então se dividirem em duas vertentes, a tecnicista do Mackenzie e a humanista da BN, este novo curso tinha um enfoque maior no livro e em outras técnicas bibliotecárias como: serviços de preservação, instalação e funcionamento das bibliotecas modernas, sistemas de classificação e catalogação, voltados mais ao

pragmatismo. (ALMEIDA; BAPTISTA, 2013; CASTRO, 2000). Sobre o curso é apresentado que

Portanto, podemos afirmar que [Rubens Borba] Moraes e [Adelfa] Figueiredo, ao criarem o primeiro Curso de Biblioteconomia, de natureza pragmática, retomam os conceitos e debates sobre as bibliotecas e a Biblioteconomia americana que, desde os anos 20, Eurico de Góes e Braz de Souza apresentavam à imprensa paulista (CASTRO, 2000, p. 73).

A partir dessa discussão, em 05 de janeiro 1937, é sancionada no Estado de São Paulo a Lei 2.839, conhecida como a *Lei das Bibliotecas*, fundindo a biblioteca estadual com a municipal, e em seu artigo 13º indicando a exigência de concursos para Biblioteconomia, condicionava “[...] que apresentassem diploma de Curso Superior e de Biblioteconomia, exceção para cargos em pequenas bibliotecas; neste caso seria exigido diploma de curso secundário” (CASTRO, 2000, p. 73).

Na década de 1940, a BN também adota o modelo pragmático estadunidense, além da concessão de bolsas para quem fosse de fora do Rio de Janeiro. Houve uma ampliação do acesso ao ensino com a fundação da Curso da Escola Livre de Sociologia e Política de São Paulo (ELSP) e a oferta de bolsas a quem fosse de outros estados com o intuito de que ao regressarem as suas bibliotecas as reorganizassem e/ou criassem novas. Essa ampliação de acesso ao curso de Biblioteconomia permitiu uma descentralização da profissão e consequente criação de novos cursos em outros lugares, tais como Bahia, Minas Gerais, Campinas e São Carlos (SP), Rio Grande do Sul, Pernambuco e o curso na Universidade de São Paulo (USP) (CASTRO, 2000).

Com o Decreto-Lei nº 6.440, de 27 de abril de 1944, que dava nova organização aos cursos de Biblioteconomia e o Decreto-Lei 6. 732, de 24 de julho, de 1944, que tratou das finalidades da BN, fez surgir a necessidade de formulação de diretrizes que regulamentassem a profissão e o ensino de Biblioteconomia no Brasil (CASTRO, 2000, p. 170). A Lei nº 4.084, de 30 de junho de 1962, dispõe sobre a regulamentação da profissão, juntamente ao Decreto nº 51.624, de 17 de dezembro de 1962, que concede aos funcionários federais uma gratificação especial de nível universitário, incluindo o profissional da Biblioteconomia, contribuem com a consolidação da profissão, mediante uma legislação implementada a partir da efetiva atuação da Federação Brasileira de Associações de Bibliotecários (FEBAB)³³. A FEBAB, fundada em 26 de julho de 1959,

³³ Atualmente, o significado de FEBAB é Federação Brasileira de Associações de Bibliotecários, Cientistas da Informação e Instituições.

tinha como objetivo fortalecer a classe bibliotecária e era, à época, apoiada por sete associações individuais regionalizadas, formando uma associação federativa (CASTRO, 2000).

Após a criação da FEBAB, com base na Lei 4.084/62, e com discussões no IV Congresso Brasileiro de Biblioteconomia e Documentação (CBBD), em 1963, criou-se o Código de Ética Profissional Bibliotecário Brasileiro (CEPBB), indicando que a profissão “[...] Requer conhecimento e estudo especial, adquiridos para benefício da coletividade” (RUSSO 1961, p.2, *apud* CASTRO, 2000, p. 189-90), e pontuando, assim, a profissão também na esfera social, a fim de edificar o cunho liberal e humanista da profissão e evitar a penetração de outras profissões nos afazeres bibliotecários. Ainda seguindo a Lei 4.084, em 1965, é criado o Conselho Federal de Biblioteconomia (CFB).

Castro (2000), baseado em Prado (1961, p. 2), pondera que, no que tange ao ensino de Biblioteconomia, entre as décadas de 1950 e 1960, alguns autores discutiam a estagnação da profissão graças ao ensino puramente técnico, necessitando assim de profissionais de outras áreas, como a administrativa, por exemplo, requerendo um ensino de um caráter mais humanista e abrangente. O ensino de Biblioteconomia carecia de “Saberes que tinham a finalidade de capacitar o bibliotecário a dirigir, desenvolver e manter bibliotecas universitárias, nacionais e especializadas” (CASTRO, 2000, p. 202).

Sobre essa questão, Castro (2000) aponta que na década de 1960, com um *déficit* educacional na qualidade e instrumentação da Biblioteconomia, fortalecia-se a discussão de um currículo mínimo e máximo, além de maiores e menores aportes financeiros, conforme a região geográfica de localização dos cursos. Foram debatidas temáticas como a pesquisa em Biblioteconomia, o currículo e a duração dos cursos e Pós-Graduação em Biblioteconomia, evidenciando uma preocupação crescente com as pesquisas científicas na área.

Com base nessas discussões o CFB, em 16 de novembro de 1962, determina o currículo mínimo para os cursos de Biblioteconomia com uma grade curricular estruturada nas seguintes disciplinas: História do Livro; História da Literatura; História da Arte; Introdução aos Estudos Históricos; Evolução do Pensamento Filosófico e Científico; Organização e Administração de Bibliotecas; Catalogação e Classificação; Bibliografia e Referência; Documentação; e Paleografia, dando enfoque maior às disciplinas técnicas e menor às disciplinas humanísticas (CASTRO, 2000, p. 208).

Almeida e Baptista (2013) apontam que o currículo mínimo, em 1962, foi elaborado por um grupo de professores, porém a proposta inicial foi alterada pelo Conselho Federal de Educação (CFE), e com o passar dos anos, as disciplinas culturais

eram muito genéricas "[...] ocasionando em geral um conhecimento superficial e pouco significativo, frustrando a intenção da formação humanística" (ALMEIDA; BAPTISTA, 2013, p. 6).

No Brasil, fundou-se em 1954 o Instituto Brasileiro de Bibliografia Documentação (IBBD) para que junto com a Fundação Getúlio Vargas (FGV) e o Departamento Administrativo do Serviço Público (DASP) promovessem "Ações que pudessem reverter o caos documentário existente no Brasil, propiciado pelo aumento da produção científica em nível nacional e internacional [...]." (CASTRO, 2000, p. 242). Logo, o IBBBD teve forte influência na qualificação do(a) bibliotecário(a), principalmente no que se refere à Organização da Informação através de cursos de Pós-Graduação promovidos pelo Instituto:

Em linhas gerais, ao IBBBD cabia levantar toda a produção científica nacional, a fim de facilitar o seu acesso e, para tanto, deveria publicar bibliografias nas diversas áreas do conhecimento. Entretanto, este trabalho requeria pessoal qualificado (CASTRO, 2000, p. 247)

Após o golpe militar, o IBBBD passa a ser controlado por militares que o associam com C&T e, assim, na década de 1970, é transformado no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), concentrando seus esforços nos serviços de informação do país, com a finalidade de apoiar o regime ditatorial.

Em 1982, com as mudanças no cenário informacional, sobreveio a necessidade de uma atualização do currículo mínimo. O CFE em parceria com a Associação Brasileira de Escolas de Biblioteconomia e Documentação (ABEBD) e a participação de professores de diversos cursos de Biblioteconomia elaboraram o 2º Currículo Mínimo de Biblioteconomia, com disciplinas divididas em três grupos: de fundamentação geral, instrumentais e de formação profissional; mas não agradou por completo a classe acadêmica, pois embora tenha havido algumas mudanças, os dois currículos (de 1962 e de 1982) eram equivalentes (ALMEIDA; BAPTISTA, 2013).

Outro marco histórico significativo é quando a Secretaria de Educação Superior do Ministério da Educação e do Desporto (SESu/MEC), no ano de 1998, elabora duas propostas de Diretrizes Curriculares Nacionais (DCN) para os cursos de Arquivologia, Biblioteconomia e Museologia a serem ministrados no Brasil. Para tanto, foram consideradas as sugestões enviadas pelas IES em atendimento ao Edital SESu/MEC nº 4, de 10 de dezembro de 1997, bem como as disposições da Lei nº 9.131, de 25 de novembro de 1995, da Lei nº 9.394, de 20 de dezembro de 1996, do Parecer nº 776, de 3 de dezembro de 1997 (Câmara de Ensino Superior do Conselho Nacional de

Educação/MEC), e da legislação privativa das três profissões. Foi elaborada uma versão preliminar das diretrizes da área, divulgada em setembro de 1998, que recebeu críticas e contribuições das entidades representativas da Biblioteconomia, especialmente da ABEBD, da ANCIB e do CFB, cujas sugestões foram incorporadas à segunda proposta. O mesmo ocorreu com os subsídios oferecidos, posteriormente, por consultores *ad hoc* das áreas de Arquivologia e Museologia, culminando na proposta final em 1999 (ABECIN, 2020, *online*).

A seguir, o Conselho Nacional de Educação/Câmara de Educação Superior (CNE/CES) emitiu em julho de 2001 o Parecer CNE/CES 492/2001³⁴ definindo as DCN para alguns cursos superiores, entre eles o de Biblioteconomia, estabelecendo: perfil dos formandos, competências e habilidades, conteúdos curriculares, estágios e atividades complementares, estrutura do curso e avaliação institucional.

Assim, o conteúdo dos cursos passou a se dividir em: formação geral, que visa oferecer referências cardeais externas aos campos de conhecimento próprios da Biblioteconomia; e de formação específica, que são nucleares em relação a cada uma das identidades profissionais em pauta. O desenvolvimento de determinados conteúdos, entre eles as Tecnologias em Informação, foi indicado como objeto de itens curriculares formalmente constituídos para este fim ou de atividades praticadas no âmbito de um ou mais conteúdo.

Conforme as Diretrizes Curriculares estabelecidos para os Cursos de Biblioteconomia, os formandos em Biblioteconomia deveriam estar

[...] preparados para enfrentar com proficiência e criatividade os problemas de sua prática profissional, produzir e difundir conhecimentos, refletir criticamente sobre a realidade que os envolve, buscar aprimoramento contínuo e observar padrões éticos de conduta [...]. (BRASIL, 2001, p. 32).

Ainda nesse contexto, o CNE/CES enumerou em gerais e específicas as competências e habilidades dos graduados em Biblioteconomia. Sobre as competências e as habilidades definidas, consideram-se como gerais:

Gerar produtos a partir dos conhecimentos adquiridos e divulgá-los;
Formular e executar políticas institucionais;
Elaborar, coordenar, executar e avaliar planos, programas e projetos;
Utilizar racionalmente os recursos disponíveis;
Desenvolver e utilizar novas tecnologias;

³⁴ Disponível em <http://portal.mec.gov.br/cne/arquivos/pdf/CES0492.pdf> Acesso em 17 de mar de 2020.

Traduzir as necessidades de indivíduos, grupos e comunidades nas respectivas áreas de atuação;
 Desenvolver atividades profissionais autônomas, de modo a orientar, dirigir, assessorar, prestar consultoria, realizar perícias e emitir laudos técnicos e pareceres;
 Responder a demandas sociais de informação produzidas pelas transformações tecnológicas que caracterizam o mundo contemporâneo (BRASIL, 2001, p. 32).

Já como competências e habilidades específicas tem-se:

Interagir e agregar valor nos processos de geração, transferência e uso da informação, em todo e qualquer ambiente;
 Criticar, investigar, propor, planejar, executar e avaliar recursos e produtos de informação;
 Trabalhar com fontes de informação de qualquer natureza;
 Processar a informação registrada em diferentes tipos de suporte, mediante a aplicação de conhecimentos teóricos e práticos de coleta, processamento, armazenamento e difusão da informação;
 realizar pesquisas relativas a produtos, processamento, transferência e uso da informação (BRASIL, 2001, p. 32-33).

Dessa forma, as diretrizes apontam para um ensino humanístico, com um currículo estruturado pelo colegiado do curso, de acordo com as demandas locais, além de apoiar atividades extras à sala de aula,

[...] atividades complementares de monitoria, pesquisa, participação em seminários e congressos, visitas programadas e outras atividades acadêmicas e culturais, igualmente orientadas por docentes (de preferência em regime de tutoria) a serem computadas como carga horária (BRASIL, 2001, p. 33).

Ao longo do processo do ensino da Biblioteconomia no Brasil, bem como os processos de formação do profissional da área, promoveram-se discussões sobre a ênfase do ensino, em um caráter mais humanístico ou mais tecnicista, porém em toda formação curricular as disciplinas como Catalogação, Classificação e Indexação sempre foram mantidas por comporem um espaço nuclear para a área (GUIMARÃES, 2017), que trabalha com a informação e, atualmente, em uma organização mais contemporânea, com dados.

Enfaticamente, os estudos de organização e representação da informação se desenvolvem no país por meio da institucionalização dos cursos de Biblioteconomia, e se aprimoraram a partir da adoção das tecnologias de informação e comunicação e da implantação dos cursos de pós-graduação, expandindo as questões de atividade profissional para a pesquisa científica (MARTINS, 2014).

Portanto, os sistemas de organização do conhecimento se desenvolveram, com a adoção de diversas tecnologias, especialmente as computacionais, e com as novas formas de representação como os metadados e as linguagens de máquinas, dando um enfoque semântico e estabelecendo a necessidade de organizar informações e dados de forma compreensiva por humanos e por máquinas.

Com a chegada da internet e a popularização da criação e compartilhamento de dados e informações de forma exponencial, a capacidade operacional de organizar este gigantesco volume se torna cada vez mais desafiador. “Neste contexto, a Ciência da Informação estabelece-se como uma ciência de organização e recuperação do relevante em meio ao excessivo, evidenciada por Jesse Hauk Shera”, ainda entre as décadas de 1950 e 1960 (MARTINS, 2014, p. 92).

Fundamentadamente, este percurso histórico evidencia o desenvolvimento do ensino da Biblioteconomia, uma das disciplinas fundamentais na interdisciplinaridade da Ciência da Informação. “[...] A ciência da informação nasceu da Biblioteconomia, tomando, portanto, como objeto de estudo a informação [...]” (LE COADIC, 1996, p. 2), e, para além do objeto, valeu-se das habilidades de classificar, representar e disponibilizar informações. Destarte, a Biblioteconomia é a disciplina que desenvolveu estudos especializados “[...] reconhecida como uma das forças impulsionadoras da organização do conhecimento na sociedade moderna, em transição para a pós-modernidade” (MARTINS, 2014, p. 97).

Nessa perspectiva, Ribeiro (2017) aponta que é possível propor novas abordagens para representação no cenário atual, a partir das mudanças impostas pelas tecnologias, associadas ao dilúvio de dados e informação que, presente na sociedade atual, fazem surgir novas demandas de organização e representação. A partir dessa concepção, a discussão sobre informação e dados aproxima-se da importância e da necessidade do perfil especializado do(a) bibliotecário(a) de dados na sociedade atual, baseado no uso intensivo de dados pela *e-Science*.

Por fim, este estudo compreende esse perfil ser um nicho de atuação bibliotecária dentro da Ciência da Informação, dentre tantas outras características intrínsecas ao profissional. No entanto, embora na Biblioteconomia esse seja um trabalho essencial para a pesquisa científica contemporânea, não é o único. A partir dessa premissa, este estudo não compreende um ensino que contemple Biblioteconomia de Dados como uma característica atestadora de qualidade do curso, contudo, é preciso que existam cursos que forneçam capacitação de atuação nesse contexto, com alcance em todo território nacional, permitindo a mobilidade profissional no país.

4.1 Biblioteconomia de Dados

Tradicionalmente a biblioteca facilitou o acesso aos documentos e, quando se pensa nos processos de organização da informação, como a catalogação e as classificações, é possível afirmar que desde muito tempo os bibliotecários se dedicam à representação da informação por meio do uso de dados, mesmo que estes não se enquadrem necessariamente nas atuais definições, destarte, as bibliotecas têm se dedicado, comumente, a uma nova gama de informações e dados, em especial, as bibliotecas acadêmicas, responsáveis pelos repositórios das instituições de ensino e pesquisa.

Nesse ínterim, há uma convergência entre ciência intensiva em dados, avanços tecnológicos e experiência dos bibliotecários, o que pode torná-los mais visíveis nesse processo de organização e curadoria de dados. “[...] Os bibliotecários estão familiarizados com as necessidades de dados de pesquisa dos pesquisadores e estão entre os apoiadores de modelos inovadores de publicação, incluindo a publicação de acesso aberto.”³⁵ (KOLTAY, 2016, p. 6, tradução nossa).

Sobre a nomenclatura, embora a atuação seja de importância extremamente superior, o termo é importante quando se pensa em pesquisa e em estudos sobre o tema, nesse sentido, este trabalho apoiou-se na fundamentação de que a definição que melhor se adequou à necessidade de compreensão desse “novo” perfil é a de “Biblioteconomia de Dados” e “bibliotecário(a) de dados”.

As origens da Biblioteconomia de Dados e o título bibliotecário de dados remontam ao gerenciamento de dados das Ciências Sociais (particularmente conjuntos de dados publicamente disponíveis e dados geoespaciais) e, depois, evoluíram no campo da bioinformática. O conceito e o título agora cobrem o trabalho relacionado à biblioteca com dados digitais adquiridos e produzidos localmente em qualquer domínio³⁶ (SOEHNER *et al.*, 2010 *apud* KOLTAY, 2016, p.7, tradução nossa).

³⁵ Texto original: that librarians are familiar with the research data needs of researchers and have been among the supporters of innovative publishing models, including open access publishing.

³⁶ Texto original: The origins of data librarianship and the title data librarian go back to social science data management (particularly publicly available data sets and geospatial data), and they then evolved in the bioinformatics field. The concept and the title now cover library-related work with both purchased and locally produced digital data in any subject domain (Soehner et al., 2010)

O espaço da biblioteca historicamente surge como uma espécie de “repositório de conhecimento”, com acesso extremamente restrito às determinadas pessoas. Com a sua evolução, ampliou-se sua atuação para diferentes segmentos e assumiu novos formatos, expandindo o acesso para uma maior diversidade de usuários. No que se refere às bibliotecas universitárias ou de pesquisas acadêmicas, ela está envolvida principalmente com a pesquisa e/ou pesquisador(a), e “[...] começa a ser entendida não apenas como algo que apoia a descoberta e o acesso aos títulos ou recursos de informação publicados, mas também como algo que se envolve com o *processo* de pesquisa e investigação acadêmica” ³⁷(RICE; SOUTHALL, 2016, p. 2, grifo do autor, tradução nossa).

As bibliotecas universitárias, no contexto da *e-Science*, assumem um novo direcionamento no atendimento dos usuários, que originado principalmente pelo desenvolvimento das TDICs, evidenciam novas necessidades informacionais. Porém, demandas que antes requereriam o apoio bibliotecário e atualmente são supridas pelo próprio usuário, irão redirecionar a forma de pesquisa, tendo em vista que “Trabalhar com os leitores para acessar, manipular ou compartilhar dados de pesquisa é uma maneira de demonstrar a capacidade de resposta das bibliotecas às necessidades acadêmicas.”³⁸ (RICE; SOUTHALL, 2016, p. 2, tradução nossa).

Essa competência ao perfil bibliotecário é atestada pela Associação Europeia de Bibliotecas de Pesquisa LIBER – *Ligue des Bibliothèques Européennes de Recherche*³⁹ – que existe há quase cinquenta anos com a missão de “Habilitar a Pesquisa de Classe Mundial” representando os interesses de instituições a ela vinculadas, suas universidades e pesquisadores em várias áreas-chave. Em 2012, um grupo de trabalho da LIBER declarou que

[...] as bibliotecas deveriam auxiliar os professores na integração do gerenciamento de dados no currículo. Para isso, foram publicadas 10 recomendações sobre gerenciamento de dados de pesquisa, que destacam, entre outras, a importância de refazer a qualificação dos bibliotecários⁴⁰ (KOLTAY, 2016, p. 6-7, tradução nossa).

³⁷ Texto original: Librarianship then begins to be understood not simply as something that supports discovery of and access to published titles or information resources but also as something that engages with the conduct of research and academic enquiry.

³⁸ Texto original: Working with readers to access, manipulate or share research data is a way to demonstrate libraries’ responsiveness to academic needs.

³⁹ Em inglês: Association of European Research Libraries

⁴⁰ Texto original: libraries should assist faculty with the integration of data management into the curriculum. To enable this, 10 recommendations on research data management have been published, which underline, among others, the importance of re-skilling librarians.

Considerando a atuação contemporânea das bibliotecas, Koltay (2016) faz uma síntese da pesquisa de Lyon (2012) e desenvolve dez possíveis serviços para uma biblioteca acadêmica no contexto de pesquisa atual:

1. Coleta de requisitos para gerenciamento de dados de pesquisa (com departamentos acadêmicos);
2. Planejamento de PGD, incluindo defesa e orientação para pesquisadores (com centros de formação de doutorado);
3. Assessoria técnica sobre formatos de dados e metadados;
4. Citação de dados;
5. Treinamentos em PGD (com centros de formação de doutoramento);
6. Licenciamento de dados de pesquisa;
7. Avaliar dados de pesquisa;
8. Armazenamento de dados de pesquisa (com serviços de TI);
9. Acessar dados de pesquisa;
10. Avaliar o impacto dos dados de pesquisa (com pesquisa escritórios de apoio). (LYON, 2012 apud KOLTAY, 2016, p.7, tradução nossa)

Deve-se compreender que a Biblioteconomia de Dados “[...] como o suporte de dados baseado em TI – foi moldada por normas culturais relacionadas ao financiamento de pesquisas, expectativas institucionais e prática acadêmica”⁴¹ (RICE; SOUTHALL, 2016, p. 17, tradução nossa), sendo que esse campo de atuação existe nos EUA desde 1958, com serviço de dados na biblioteca da Universidade de Berkeley (RICE; SOUTHALL, 2016, p. 11).

Para compreender melhor essa afirmação e ampliação do uso de dados pelo mundo, é preciso entender o seu percurso histórico de surgimento e de uso. Primeiramente, o desenvolvimento de máquinas computacionais surge ainda na década de 1930, para alguns pesquisadores, com Alan Turing, dando origem à existência de máquinas de cartões perfurados, muito utilizadas nas décadas seguintes, principalmente nas Ciências Sociais. Nas décadas 1940 e 1950, ocorreram muitas transformações no mundo, foram fundados o Banco Mundial, o Fundo Monetário Internacional e as Nações Unidas, a IBM produziu seu primeiro computador “real”, o Instituto de Pesquisa Social de Michigan e o Bureau de Pesquisa Social Aplicada de Columbia produziam muitos dados de pesquisas, e o Roper Center arquivou mais de 3000 pesquisas de 70 países (RICE; SOUTHALL, 2016, tradução nossa)

⁴¹ Texto original: like IT-based data support – has been shaped by cultural norms to do with research funding, institutional expectations and academic practice.

Já Semeler (2017) indica que a Biblioteconomia de Dados teve início em torno da década de 1960, nos EUA, quando “As Ciências Sociais foram um dos primeiros campos disciplinares a adotar e a compartilhar dados de pesquisa para o reúso, isso por razões de economia e eficiência científica” (SEMELER 2017, p. 66). Assim, as bibliotecas e os arquivos de dados originados entre 1960 e 1970, principalmente na área de ciências sociais – que, embora geralmente considerados mais leves que as ciências duras, em parte por causa de suas metodologias aparentemente abstratas, em parte por ter assuntos relacionados aos comportamentos individuais ou coletivos do ser humano, o que faz ser muito difícil estabelecer previsões –, resultaram na utilização do método quantitativo, com as estatísticas aplicadas às amostras de populações para descrever, explicar e prever comportamentos.

Nesse sentido, as ciências sociais, como primo pobre das ciências físicas mais bem-dotadas, foram o primeiro grupo disciplinar a adotar o compartilhamento de dados e a reutilização de dados por razões de economia ou eficiência. No entanto, dados sobre seres humanos, como outros dados observacionais, como condições climáticas, nunca podem ser replicados nas mesmas circunstâncias – outra forte justificativa para o compartilhamento.⁴² (RICE; SOUTHALL, 2016, p. 4, tradução nossa).

Em sequência, no ano de 1974 é fundada a *International Association for Social Science Information Service and Technology* (IASSIST), em uma reunião patrocinada pelo Congresso Mundial de Sociologia em Toronto, alcançando, além da cidade sede, Londres, Edimburgo, Cocoa Beach, Itasca, Uppsala e Ottawa, quando foi indicado que uma profissão de arquivamento de dados estava em vigor, demonstrando também que, devido a gigantesca amostra de dados do censo dos EUA, as bibliotecas também tiveram forte influência. Ainda na década de 1970, a *American Library Association* (ALA) constituiu um subcomitê para recomendar regras para catalogar arquivos de dados legíveis por máquina (ROWE, 1999 *apud* RICE; SOUTHALL, 2016 p. 5, tradução nossa).

Esse processo de edificação da área Biblioteconomia de Dados é lento, mas com o desenvolvimento largamente na rede mundial, popularmente conhecida como internet. Durante as décadas de 1980 e 1990, há uma certa aceleração, impulsionada pelo estabelecimento do protocolo HTML e navegadores baseados em clientes, que “tornou

⁴² Texto original: In this sense the social sciences, as the poor cousin of the more well endowed physical sciences, were the first disciplinary group to embrace data sharing and the re-use of data for reasons of economy or efficiency. However data about human subjects, like other observational data such as weather conditions, can never be replicated in the same circumstances – another Strong rationale for sharing.

muito mais eficaz não apenas fornecer dados aos usuários, mas também lidar com descrições de coleta de dados, catalogação, descoberta e suporte à pesquisa”⁴³ (RICE; SOUTHALL, 2016, p. 9, tradução nossa). Esses avanços se deram, essencialmente, por essas tecnologias permitirem capacidade cada vez maior de armazenamento *online*, pelas melhorias nas velocidades de processamento da computação e com o setor de telecomunicação fornecendo serviços que permitissem mais acesso e mais velocidade dos sistemas.

Contudo, somente após os anos 2000 que se inicia uma forte discussão sobre Biblioteconomia de Dados, quando as “[...] políticas de compartilhamento de dados foram adotadas por várias agências de financiamento, em que os pesquisadores começaram a examinar as novas funções dos bibliotecários no gerenciamento de dados e nos serviços de dados”⁴⁴ (XIA; WANG, 2014, p. 362, tradução nossa).

Também nesse contexto, evidencia-se que nos anos 2000 são estabelecidas as DCN para os cursos de Biblioteconomia no Brasil, com uma revisão curricular que, a partir de então, inseria na formação o uso das tecnologias de informação e de comunicação, mas não há, nessas diretrizes legais, previsão para o contexto de atuação. Acredita-se que o(a)s bibliotecário(a)s de dados que assumiram esse perfil, o fizeram associando a formação aos processos continuados de aprendizagem, obtendo expertise para lidar com esse universo. Entretanto, passados 20 anos, investiga-se se este perfil se encontra em disciplinas de formação nos currículos, considerando possíveis atualizações. Para isso, na próxima seção serão delineadas as exigências que configuram o perfil do(a) bibliotecário(a) de dados.

⁴³ Texto original: has made it much more effective not only to deliver data to users, but also to deal with data collection descriptions, cataloguing, discovery and researcher support.

⁴⁴ Texto original: data-sharing policies were adopted by various grant-funding agencies that researchers started examining librarians’ new roles in data management and data services.

4.2 Bibliotecário(a) de dados

O suporte à recuperação e ao acesso de coleções de material publicado através das técnicas de catalogação, organização, serviços de referência e treinamento ao usuário são funções tradicionalmente ligadas ao bibliotecário(a). Todavia, este profissional da informação deve garantir o acesso à informação ao maior número possível de usuários, além de se dedicar com a preservação da informação, inclusive no contexto atual da *e-Science*.

O momento mostra-se oportuno tanto para profissionais da informação, como para profissionais da tecnologia de informação, dentre outros. Os limites dessa nova área ainda não estão definidos, por isso faz-se necessário que o bibliotecário posicione-se como um profissional que apresenta capacidades para lidar com o tratamento de dados oriundos da *e-Science* (COSTA, 2017, p. 89).

Essas funções se mantêm atuais, entretanto os suportes informacionais se atualizaram, e se mantêm em constante atualização, assim, novos métodos e técnicas surgem para que os profissionais da Biblioteconomia continuem atendendo aos princípios de sua profissão, que segundo a Classificação Brasileira de Ocupações:

Disponibilizam informação em qualquer suporte; gerenciam unidades como bibliotecas, centros de documentação, centros de informação e correlatos, além de redes e sistemas de informação. Tratam tecnicamente e desenvolvem recursos informacionais; disseminam informação com o objetivo de facilitar o acesso e geração do conhecimento; desenvolvem estudos e pesquisas; realizam difusão cultural; desenvolvem ações educativas. Podem prestar serviços de assessoria e consultoria (BRASIL, 2020, *online*)

A revista *The Economist*, conceituado periódico inglês, em sua edição de maio de 2017, traz em sua capa a manchete: "O recurso mais valioso do mundo não é mais petróleo, mas os dados"⁴⁵ (tradução nossa), ou seja, faz uma analogia ponderando que os dados atualmente equivalem ao que o petróleo valia há um século, como uma fonte de poder e riqueza (THE ECONOMIST, 2017, *online*). Sem adentrar nas questões políticas e/ou econômicas que este fato carrega consigo, mas observando o valor competente ao meio científico, isto é, de pesquisa e de desenvolvimento, a matéria ressalta a importância dos dados de pesquisa na contemporaneidade que permanecem

⁴⁵ Texto original: The world's most valuable resource is no longer oil, but data.

vinculadas à responsabilidade bibliotecária indicada pela Classificação Brasileira de Ocupações (CBO) como uma profissão em que os profissionais da informação “Disponibilizam informação em qualquer suporte; gerenciam unidades como bibliotecas, centros de documentação, centros de informação e correlatos” (BRASIL, 2020, *online*). Nesse atual contexto, desponta um “novo” perfil de atuação na Biblioteconomia por meio da “Biblioteconomia de Dados” e “bibliotecário(a) de dados”.

No contexto dos repositórios de dados de pesquisa como sistemas contemporâneos e heterogêneos de recuperação da informação emergem dois conceitos: Biblioteconomia de Dados e bibliotecário de dados. Tais concepções estão em consonância com as discussões no campo da Organização do Conhecimento referentes à defesa de instituições eficazes e sustentáveis, tendo em vista a necessidade do estabelecimento de métodos e práticas específicos para a análise de dados de pesquisa, especialmente no ambiente *web* (TARTAROTTI; DAL'EVEDOVE; FUJITA, 2019, p. 210).

O termo “*Data Librarian*” ou “bibliotecário(a) de dados” surge pela primeira vez no artigo “*The Data Librarian: introducing the Data Librarian*”, de Joe Liscouski (1997), que pode ser compreendido como a atribuição de características ao bibliotecário(a) de dados, no entanto, refere-se a um *software* cujo nome faz uma analogia à profissão, sendo chamado *Data Librarian*, com funções desse *software* análogas ao do(a) bibliotecário(a):

O *Data Librarian* é um pacote de *software* projetado para gerenciar arquivos de dados criados por instrumentação de laboratório. Dados padrões de formato estão sendo desenvolvidos para permitir pessoal para exportar dados da análise instrumental sistemas. Isso vai resultar na criação de um grande número de arquivos, com potencial para perda de dados, incompatibilidade nomes de arquivos e o que se espera neles. (Liscouski, 1997, p. 199, tradução nossa).

Assim, o campo Biblioteconomia de Dados e a atividade de bibliotecário(a) de dados surgem com a evolução da bioinformática, principalmente após as publicações de conjunto de dados, *datasets*, das ciências sociais, como dito anteriormente. Atualmente, as bibliotecas acadêmicas assumem esse papel ao fornecerem suporte aos pesquisadores, aos dados digitais adquiridos e produzidos em qualquer domínio da ciência e à gerência de repositórios institucionais de ensino e pesquisa.

Rice e Southall (2016) aludem que, em 1966, na Universidade de Wisconsin (em Madison nos EUA), existia o “Programa de Serviços de Dados da Biblioteca” com

funções de adquirir, armazenar, manter e preservar arquivos de dados legíveis por máquina e pelos programas de computador. Com a sucessão de bibliotecários nesse setor, formou-se a Escola de Biblioteconomia e Estudos da Informação. Outro exemplo de biblioteca de dados é a *Northwestern University Library*, que oferece "Serviços de dados e informações geográficas e governamentais"⁴⁶ (RICE; SOUTHALL, 2016, p. 12, tradução nossa).

Para Rice e Southall (2016), o(a) bibliotecário(a) de dados é uma combinação da evolução das habilidades bibliotecárias tradicionais e das tecnologias atuais, com competências “como assessorar na descoberta de recursos, por um lado, e fornecer informações básicas sobre questões específicas de formatos de dados ou obsolescência digital, por outro” (RICE; SOUTHALL, 2016, p. 16, tradução nossa). A evidência desse novo perfil aponta que o ambiente de pesquisa está mudando impulsionado pela tecnologia, e que as bibliotecas têm a chance de recriar suas funcionalidades para dar suporte e serviços a esse volume informacional e seus pesquisadores.

Considerando as definições apresentadas e as demandas apontadas, o(a) bibliotecário(a) é capacitado para gerir dados, e dessa forma, nesse mundo competitivo, esse profissional reúne competências diante dos demais profissionais para assumir este posto frente ao panorama do chamado “tsunami” de dados.

[...] Sendo assim, no panorama dos grandes volumes de dados, o cientista de dados desponta como uma das profissões promissoras e desafiantes da atualidade diante do fenômeno do *Big Data* e da gestão desses mega dados no contexto da Ciência da Informação (RODRIGUES; DUARTE; DIAS, 2017, *online*).

Deve-se atentar que a atual gestão de dados vai além das ferramentas tecnológicas, que são essenciais, mas que, no contexto atual, exige domínio das estratégias voltas ao livre acesso, compartilhamento, uso e reúso.

[...] considerando que as tecnologias da informação e da comunicação possibilitam ampla produção, acesso e distribuição de dados digitais, gerir dados, objetivando acesso livre, é mais do que simplesmente adotar qualquer plano de gestão e instalar um *software* de processamento e de análise (SILVA SEGUNDO; ARAÚJO, 2019, p. 131).

⁴⁶ Texto original: Government and Geographic Information and Data Services

O profissional da informação, emergido neste contexto, deve estar preparado para trabalhar com dados gerados, tanto com finalidade comercial, quanto com a pesquisa científica (nosso enfoque) e atentar-se que “Essa explosão de dados alimenta outra característica dos negócios e da pesquisa científica, que possibilita vislumbrar novas tendências e identificar padrões” (SILVA SEGUNDO; ARAÚJO, 2019, p. 118). No entanto, para atender às tendências e identificar padrões, precisará de uma formação que garanta habilidades e competências para o atendimento das necessidades mercadológicas e científicas que despontam.

4.2.1 Habilidades e competências

Os avanços tecnológicos criam a necessidade de todo profissional se atualizar e se aperfeiçoar constantemente. Com o(a) bibliotecário(a) a perspectiva é a mesma, principalmente com o avanço da TDICs, e a necessidade de especificidades cada vez mais profundas por parte dos usuários, aumentando a complexidade da gestão e mediação da informação, exigindo do profissional habilidades e competências em consonância com aspectos como desenho dos sistemas de informação, operação, divulgação e avaliação, tanto para o âmbito empresarial como para o educacional e de pesquisa.

A profissão bibliotecária enfrenta constantes desafios, as Tecnologias de Informação e de Comunicação requerem um conhecimento de sua aplicação e de seu funcionamento, os usuários da informação demandam serviços e produtos especializados, a gestão de um ambiente de mediação da informação funciona de maneira complexa e atende a uma grande variedade de solicitações. Os profissionais requerem uma formação na qual os conhecimentos (processos inteligentes da informação, que servem como base para o agir), habilidades (o conhecimento em ação) e atitudes (estímulo e empatia com a qual afrontamos uma situação, assim como a pré-disposição para fazer algo de uma determinada maneira) estejam em função de diversos aspectos, tais como o desenho dos sistemas de informação, operação, divulgação e avaliação; o que adverte um aumento no valor da informação, tanto para as empresas como para as instituições educacionais (PRADO, 2016, p.13).

Nessa perspectiva, Rodrigues, Duarte e Dias, em uma pesquisa publicada em 2017, pontuam cinco áreas fundamentais para um(a) cientista de dados:

Deste modo, a pesquisa exploratória permitiu apontar quatro áreas que fundamentam os aspectos do cientista de dados:

- a) Tecnologia: Dentre o arsenal que compõem o rol de habilidades que o profissional deve ter, maior manejo e familiaridade no aparato tecnológico, mais vantagem terá na hora de gerenciar dados. Domínios em computação em nuvem, linguagem de programação, segurança e privacidade dos dados, tecnologia e infraestrutura conceitos de Map Reduce, manejo de programas e *software*;
- b) Análise de dados: Domínios em: método científico, conhecimento em Ciência de Dados (Data Science), análise de gráficos, mineração de dados e textos, conhecimento em estatística e aprendizado por máquina;
- c) Gestão de dados: Domínios em: Big Data (Hadoop, SQL, Spark), gestão de dados estruturados (RDB, XML), gestão de dados não-estruturados (bancos de dados NoSQL), recuperação dos dados, inteligência competitiva;
- d) Design e Comunicação: Domínios em: comunicação interpessoal, conhecimento em visualização de dados, habilidades de contar histórias visuais, proatividade, poder de negociação, flexibilidade;
- e) Matemática e estatística: Domínios em: Conhecimento em estatística avançada, gerenciamento de grandes volumes de dados, relacionar dados e descobrir padrões, cruzar dados estatísticos, data mining, otimização. (RODRIGUES; DUARTE; DIAS, 2017, p. 13).

É preciso enfatizar que as habilidades de um(a) cientista de dados não são obrigatórias ao (à) bibliotecário(a) de dados, embora esse conhecimento possa ser de grande utilidade ao serviço prestado. A Biblioteconomia de Dados atua em diversos momentos, coincidentes ou não, à Ciência de dados. Conforme indica o próprio termo, o bibliotecário(a) de dados atua na intersecção entre o contexto da ciência orientada aos dados com a Biblioteconomia, como aponta Semeler e Pinto (2017, p. 19):

[...] *e-Science* como o novo paradigma das ciências orientadas aos dados, o que inclui habilidades, métodos, técnicas e tecnologias de *Data Science* como uma fonte para novas práticas em Biblioteconomia de Dados. O bibliotecário(a) de dados é aquele que deve aplicar a CI, a *e-Science* e a *Data Science* em bibliotecas; esse profissional precisa resolver os problemas relacionados a coleta, manipulação, análise e visualização de dados de pesquisa em repositórios digitais.

Em um artigo publicado pela *The Medical Library Association Data Service*, visando suprir as necessidades de usuários das bibliotecas de saúde e biomédicas ao trabalhar com seus próprios dados e com os de outras pessoas na área da saúde, Federer *et al* (2020) propõem cinco habilidades essenciais para o bibliotecário(a) de dados:

1. **Aplicar princípios de competência de dados:** os bibliotecários devem estar familiarizados com as habilidades para gerenciar adequadamente os conjuntos de dados, incluindo nomeação e versão de arquivos,

preservação e armazenamento de dados, escrita de planos de gerenciamento e compartilhamento de dados que atendam aos requisitos do financiador;

2. **Estabelecer e aprimorar serviços de dados:** os bibliotecários devem possuir a capacidade de desenvolver serviços de dados com base nas necessidades estabelecidas, bem como de aprofundar e aprimorar esses serviços ao longo do tempo, são duas habilidades necessárias;
3. **Oferecer suporte às práticas recomendadas de dados de pesquisa em todo o ciclo de vida dos dados:** os bibliotecários devem possuir um entendimento das melhores práticas que garantem a qualidade dos dados em cada ponto do processo e para a variedade de resultados de pesquisas, incluindo dados, *software* e código;
4. **Adquirir conhecimento dos métodos de pesquisa, ética e rigor da pesquisa e práticas científicas abertas:** os bibliotecários devem ter o entendimento dos métodos e práticas da ciência em que o usuário pesquisa para que possa fornecer orientações mais significativas para seus usuários, porém não é necessário um conhecimento aprofundado na área do conhecimento em que esteja atuando;
5. **Fornecer treinamento e consulta para tópicos relacionados a dados:** os bibliotecários devem ter as habilidades de ensino e consulta que lhes permitam efetivamente transferir esse conhecimento para seus usuários, individualmente ou em sessões de instrução formais.

No que se refere à atuação do(a) bibliotecário(a) de dados no auxílio à pesquisa científica, é preciso que esse profissional se faça presente em todas as fases da pesquisa, desde o planejamento até a sua publicação dos dados digitais, bem como em sua preservação.

A participação dos bibliotecários deve ocorrer já nas fases que precedem a pesquisa, no planejamento, na elaboração de plano de gestão de dados, no apoio na geração/coleta, na identificação de dados de pesquisa de outros pesquisadores, que podem ser reutilizados pelos pesquisadores locais, e na seleção de repositórios adequados à publicação dos dados gerados, incluindo o controle de versionamento, armazenamento e *backups*, e, na pesquisa finalizada, auxiliando na publicação dos dados e na criação de ambientes confiáveis para armazenamento por longo prazo, caso seja necessário (SALES; SAYÃO; MARANHÃO; DRUMOND; SILVA, 2019, p. 307)

Sales, Sayão, Maranhão, Drumond e Silva (2019), baseados em pesquisa bibliográfica em três bases de dados (*Scopus*, *Lista* e *Lisa*) criam um quadro de competências dos bibliotecários na gestão de dados de pesquisa (Quadro 3/Quadro 3).

Quadro 3 – Competências dos bibliotecários na gestão de dados de pesquisa

ANTES DA PESQUISA	DURANTE A PESQUISA	DEPOIS DA PESQUISA
Auxiliar pesquisadores na elaboração do plano de gestão de dados	Tipificar dados de pesquisa	Auxiliar na publicação de dados (identificação de repositórios ou outras formas de publicação ex: <i>data journal</i> , periódico de resultado negativo etc)
Planejar a curadoria	Conhecer a estrutura informacional do dado de pesquisa e o seu ciclo de vida	Auxiliar na contextualização, isto é, na documentação de conjuntos de dados (definições, metodologia de coleta etc)
Identificar fluxos de trabalhos (ou mais especificamente o fluxo da pesquisa)	Administrar o ciclo de vida dos dados de pesquisa, desde sua geração/coleta, bem como seleção e desenvolvimento de coleção.	
Identificar recursos e infraestruturas para manutenção e promoção de dados de pesquisa	Organizar dados de pesquisa / atribuir metadados gerais e disciplinares	
Orientar a organização de arquivos de dados e o uso de ferramentas de gestão de dados de pesquisa		
Apoiar na identificação e escolha de ferramentas adequadas para análise, processamento e visualização	Apoiar a análise de dados e o processamento, indicando ferramentas e promovendo treinamentos	Apoiar a visualização de dados, indicando ferramentas e provendo treinamentos
Apoiar a adoção de práticas de gestão de dados de pesquisa em parceria com departamentos, grupos de pesquisa, comissões etc.;		
Entender e promover preservação digital, isto é, elaboração de política de preservação	Entender e promover preservação digital, isto é, gerenciamento de versões, armazenamento e <i>backup</i>	Entender e promover preservação digital de longo prazo, isto é, desenvolvimento de ambientes confiáveis para preservação
Contribuir para a elaboração de políticas institucionais de dados de pesquisa	Gerenciar sistemas de armazenamento de dados	Promover o reúso de dados, através de divulgação e seleção de dados adequados
Conhecer aspectos legais dos dados de pesquisa, bem como as leis de direitos autorais	Promover a capacitação para o desenvolvimento da competência em gestão de dados de pesquisa (<i>research data literacy</i>)	Auxiliar na elaboração de citação e referência de dados
Definir políticas de acesso	Criar e oferecer tutoriais sobre a elaboração de planos de gestão de dados	

Fonte: Sales, Sayão, Maranhão, Drumond e Silva (2019, p. 311).

Com base nas pesquisas evidenciadas na Ciência da Informação, cabe ressaltar que este quadro (Quadro 3) orientará a identificação nos currículos de Biblioteconomia no Brasil, de conteúdos de formação que evidenciem o ensino e o desenvolvimento de competência para atuar como bibliotecário(a) de dados.

Antes de encerrar essa seção é importante ressaltar que, conforme referenciado até este ponto, a Biblioteconomia tem se desenvolvido conforme as necessidades sociais e científicas, assim este trabalho percebe a Biblioteconomia de Dados como um campo em desenvolvimento e de atuação possível ao bibliotecário(a), a partir dos processos, produtos e instrumentos oriundos da Biblioteconomia. Compreende-se, neste contexto, que a Biblioteconomia de Dados é uma área de atuação que exige especialização, como os diversos outros campos de atuação igualmente significativos, que se entrelaçam com demais áreas de especificidade, como os estudos de Competência Informacional, Regime de Informação e as aplicações do novo contexto de atuação nas Bibliotecas Escolares, dentre outros.

5 PROCEDIMENTOS METODOLÓGICOS

Todo projeto de pesquisa deve pautar-se na definição dos métodos e técnicas que conduzirão a coleta dos dados, sua análise e o alcance dos resultados, disciplinando a execução do trabalho de acordo com os prazos estabelecidos.

Para fundamentação da pesquisa aqui proposta, os estudos tiveram início na contextualização da formação em Biblioteconomia, a fim de aferir a relação do profissional com a pesquisa científica, mais especificamente no contexto da *e-Science*, e assim diligenciar evidências nos currículos através das ementas das disciplinas neles apontadas (tanto obrigatórias, quanto optativas). A busca pelos ementários seu deu nos PPC dos cursos, quando o ementário das disciplinas não se encontrava nos PPC, buscou-os nos sítios institucionais dos cursos.

Nessa fase para a fundamentação teórica foi realizada uma pesquisa em livros, artigos, teses e dissertações. Com exceção dos livros, a maioria dos documentos foram localizados no portal de periódicos CAPES⁴⁷. Para delinear as contextualizações históricas, buscou-se referências mais antigas, e nas buscas por definições e conceitos mais novos, evidenciou-se as buscas em produções compreendidas entre os anos de 2015 a 2020. Quanto ao idioma, quando se refere à Biblioteconomia de Dados e/ou Bibliotecário(a) de dados, a maioria do referencial é em língua inglesa, predominantemente advindos dos EUA e Europa.

No que tange ao tratamento de dados de pesquisa científica como prática tradicional e pioneira na área, seguido para uma discussão sobre o surgimento e consolidação da *e-Science*, buscou-se discutir sobre o desenvolvimento da ciência a partir dessa nova prática científica, pontuando os principais conceitos sobre a Ciência de Dados – no âmbito da pesquisa científica, *e-Science* e suas aproximações com as práticas bibliotecárias —, práticas essas que apontam para trabalhos relacionados aos dados de pesquisa científica.

O recorte sobre os cursos analisados e seus respectivos PPCs se deu através de uma consulta realizada no mês de março/2020, ao sítio do Ministério da Educação, e-MEC⁴⁸, estabelecendo como estratégia de busca: 1º) acessar a guia “busca avançada”; 2º) utilizar o termo “Biblioteconomia” no campo “Curso de Graduação”; 3º) a partir dos

⁴⁷ <https://www.periodicos.capes.gov.br/>

⁴⁸ Cadastro e-MEC de Instituições e Cursos de Educação Superior. Disponível em <https://emec.mec.gov.br>. Acesso em: 09 jul. 2020.

filtros, refinou-se a pesquisa pelo campo "Situação", selecionando “em atividade”, tendo em vista que os cursos “extintos” ou “em extinção” são irrelevantes para esse levantamento.

Como resultado dessa pesquisa, foram recuperados pelo sistema 55 (cinquenta e cinco) Instituições de Ensino Superior (IES) públicas e privadas que oferecem cursos de bacharelado presencial e a distância e uma licenciatura presencial em Biblioteconomia. Destes, verificou-se que 7 (sete) cursos ainda não tinham iniciado, embora em situação de atividade. Como recorte, utilizou-se somente os que constavam como iniciados, totalizando 48 (quarenta e oito) cursos, sendo 42 (quarenta e três) presenciais e 6 (seis) na modalidade à distância.

Apoiado neste recorte, fez-se um levantamento nos PPCs dos cursos de Biblioteconomia, buscando-os inicialmente por meio de seus sítios institucionais. Na ausência desses documentos, para a complementação de informações, foram contactadas dezoito coordenações de cursos, sendo que apenas seis retornaram e, dentre essas, somente três enviaram seus PPCs; da amostra de 48 (quarenta e oito) cursos pesquisados, teve-se acesso a 33 (trinta e três) PPCs, ou seja, 69% das instituições.

Com base nos objetivos elencados esta pesquisa, estabelece-se o estudo como documental, tendo em vista que:

[...] trabalha com base em documentos que não receberam tratamento de análise e síntese. Embora se identifique com a pesquisa bibliográfica, está só se realiza sobre documentos analisados e pertencentes a aurores que deram o estudo pronto e acabado (SANTOS, 2005, p. 174).

Nesse momento, foi utilizado como corpus de análise documentos como projetos pedagógicos de curso em busca do perfil dos egressos, competências e habilidades e a descrição das matrizes curriculares e ementas disciplinares, para evidenciar as bases da formação bibliotecária. O PPC é, segundo a Coordenação Geral de Avaliação de Cursos de Graduação e Instituições de Educação Superior (CGACGIES), um:

Documento que representa o planejamento e organização do curso, sendo insumo formal e estruturante da oferta de serviço de ensino. Possui parâmetros que orientam o cotejamento entre o realizado e o almejado para um curso de graduação, em diferentes aspectos. Deve refletir as condições concretas de oferta de um curso de graduação, observados seus elementos constituintes e previsões estabelecidas no âmbito do curso, obedecidas as Diretrizes Curriculares nacionais (CGACGIES, 2019, p. 87).

Quanto à análise de dados, utilizou-se a técnica de Análise de Conteúdo, tendo em vista que "o ponto de partida da Análise de Conteúdo é a mensagem, seja ela verbal (oral ou escrita), gestual, silenciosa, figurativa, documental ou diretamente provocada" (FRANCO, 2008, p.12). Nessa perspectiva, aplicou-se a Análise de Conteúdo proposta por Bardin (2011) para realização de uma pesquisa exploratória sobre os PPCs selecionados, baseando-se nas três fases de sua organização que envolve: 1) a pré-análise, 2) a exploração do material, e 3) o tratamento dos resultados obtidos e a interpretação, utilizando a categorização para classificar os documentos da amostra durante o processo de análise. A categorização dos resultados tem como propósito o "[...] armazenamento sob uma forma variável e a facilitação do acesso ao observador, de tal forma que este obtenha o máximo de informação (aspecto quantitativo), com o máximo de pertinência (aspecto qualitativo). (BARDIN, 2011, p.51).

Essa primeira categorização foi estruturada definindo: 1) PPCs com conteúdo que descrevem objetivamente itens de formação voltados para a *e-Science*, ciência de dados e/ou *Big data*, seja na definição de perfil e/ou formação, seja nas ementas e seus conteúdos disciplinares; 2) PPCs que apresentam conteúdos que indicam estudos nessas áreas, mas de maneira não explícita, e; 3) PPCs que não apresentam nenhuma menção aos conteúdos voltados para essa formação.

Nesse sentido, foram criadas três categorias: *disciplinas ligadas à e-Science*, em que foi inserida a quantidade de disciplinas cujo ementário descrevia objetivamente habilidades e competências ligadas à *e-Science*; *disciplinas possivelmente ligadas à e-Science*, em que foi inserida a quantidade de disciplinas cujo ementário descrevia a possibilidade do docente abordar, no conteúdo do ementário, sobre habilidades e competências ligadas à *e-Science*; e *disciplinas sem ligação e-Science*, em que foi inserida a quantidade de disciplinas sem ligação com a *e-Science*.

Após identificados os PPCs nas categorias supracitadas, analisou-se os itens da estrutura curricular dos cursos, verificando os conteúdos de formação bibliotecária que apresentassem elementos e definições enfáticas relacionadas ao contexto da *e-Science*. Em busca dessas características, a Análise de Conteúdo teve como intenção proporcionar o "tratamento semântico-estrutural dos textos (produtos de uma comunicação) que pretende, por um processo de normatização da diversidade de um grande conjunto de documentos, torná-los comparáveis" (ARAÚJO, 2009, p 62). Essa normatização permitiu que o conteúdo analisado tivesse como parâmetro categorias semânticas com base no "Quadro de competências dos bibliotecários na gestão de dados

de pesquisa”, de Sales, Sayão, Maranhão, Drumond e Silva (2019, p. 311) disposto nesse trabalho no Quadro 3, seção 5.

Baseado nesse quadro foram elencados os termos: dados de pesquisa, pesquisa científica, Ciência de Dados, *Data Science*, ciclo de vida, dados abertos, *open data*⁴⁹, preservação digital, gestão de dados, reúso de dados, colaborativo⁵⁰, curadoria e metadados. Esses termos foram procurados nos ementários contidos nos PPCs para então serem analisados apenas as ementas que apresentaram algum desses termos. Adotou-se este procedimento devido às dificuldades encontradas no contexto atual em que, somado ao curto tempo de pesquisa, o mundo atravessa uma pandemia, a COVID-19, obrigando que a pesquisa fosse feita de forma remota, dividindo o tempo e espaço entre pesquisa e atenção à família, esta última demandando uma atenção mais específica e sensível pelo próprio cenário pandêmico.

Optou-se por usar o *MS-Excel* para salvar, categorizar e analisar os dados encontrados, mesmo sabendo da existência de ferramentas com maior potencialidade para esse tipo de análise, em razão de enorme heterogeneidade entre os PPCs, e o tempo escasso para aquisição do domínio sobre essas ferramentas, já que havia um conhecimento existente deste pesquisador sobre o *MS-Excel*.

É importante destacar que a localização dos PPCs centrou-se nos sítios institucionais, recorrendo-se por e-mail às Coordenações/Departamentos de Curso quando na impossibilidade dessa localização, e na busca de informações complementares, como a oferta de disciplinas que não estejam explícitas no PPC.

Após a coleta dos PPCs, na pré-análise, verificou-se a falta de homogeneidade indicada por Bardin (2011, p. 128) na qual “[...] os documentos retidos devem ser homogêneos, quer dizer, devem obedecer a critérios precisos de escolha e não apresentar demasiada singularidade fora destes critérios de escolha”. A princípio, acreditou-se que os PPCs, por advirem da DCN para os Cursos de Biblioteconomia, deveriam apresentar em suas estruturas: perfil dos formandos, competências e habilidades, conteúdos curriculares, estágios e atividades complementares, estrutura do curso e avaliação institucional; estes apresentam homogeneidade, uma vez que obrigatoriamente precisam seguir as diretrizes estabelecidas pelo Conselho Nacional de Educação/Câmara de Educação Superior (BRASIL, 2001). Os documentos seguem os princípios impostos pela DCN para os Cursos de Biblioteconomia, porém a estrutura em

⁴⁹ Este termo foi inserido na lista ao ser encontrado em uma das disciplinas pesquisadas.

⁵⁰ Termo localizado durante a pesquisa em ambientes colaborativos ou serviços colaborativos, assim investigou-se sua relação com a pesquisa colaborativa praticada pela *e-Science*.

que são postos nos documentos e a profundidade de informações contidas diferem muito.

Nessa etapa da pesquisa, após a pré-análise, os PPCs foram distribuídos de acordo com as regiões brasileiras em que a instituição mantenedora se encontra. Findada essa fase, que, segundo Bardin (2011, p. 125), é longa e “consiste essencialmente em operações de codificação, decomposição ou enumeração, em função de regras previamente formuladas”, iniciou-se a exploração do material.

Com a coleta das informações nos conteúdos e a categorização definida, iniciou-se a fase de tratamento dos resultados e a interpretação (BARDIN, 2011), dispondo de operações estatísticas, sínteses e seleções dos resultados, inferências e interpretação, permitindo a visualização do panorama de ensino nos cursos de Biblioteconomia. A partir dessas inferências, a pesquisa visou apontar variáveis como PPCs estruturados com observâncias para essa formação, seja ela explícita ou implícita, número de disciplinas que contemplam conteúdos de formação na temática, a distribuição dos cursos de acordo com sua região e, dentre estes, quais evidenciavam tendências de formação para um perfil que atenda às competências necessárias para atuação do(a) bibliotecário(a) de dados.

Esta pesquisa entende que o pesquisador deve se responsabilizar quanto à forma de obtenção, armazenamento e disponibilização dos dados de sua pesquisa para que ela possa ter um alcance tanto por quem os gerou como por outros pesquisadores. Dessa forma, buscando uma aproximação da proposta aqui apresentada e as tendências advindas da *e-Science*, em relação ao tratamento com os dados coletados e gerados durante a pesquisa, desenvolveu-se um Plano de Gerenciamento de Dados, a partir da ferramenta *DMPTool*⁵¹, para subsidiar a coleta e organização dos dados dessa investigação, disposto no APÊNDICE B.

Vale ressaltar a importância do PGD para um trabalho científico, principalmente no atual contexto de *e-Science*, de tal modo que já existem no Brasil agências de fomento que o exige como FAPESP por exemplo.

Os dados coletados e tratados estão reunidos em duas planilhas do tipo *.ODS*: *PPCs.ods*, contendo todas as disciplinas e as ementas que se teve acesso; e *Cursos Biblioteconomia - PPCs.ods*, contendo a lista de cursos pesquisados e um breve diário de anotações, com observações de conquistas e dificuldades encontradas na procura por PPCs. Embora a ferramenta usada tenha sido o Excel, optou-se em exportar para o

⁵¹ Ver <https://dmptool.org/>

formato *.ODS* por ser um formato de acesso aberto. Há também arquivo compactado no tipo *.rar* contendo todos os documentos localizados na pesquisa referentes aos ementários dos cursos. Inicialmente os dados foram compartilhados na plataforma *Google Drive*, após aprovação e conclusão desta pesquisa está disponibilizada na plataforma *Zenodo*⁵² no endereço <https://zenodo.org/record/4778986>, repositório de dados abertos multidisciplinar e gratuito associado ao *DataCite*⁵³ sob o DOI 10.5281/zenodo.4778986.

⁵² Ver <https://zenodo.org/>

⁵³ Ver <http://www.datacite.org/>

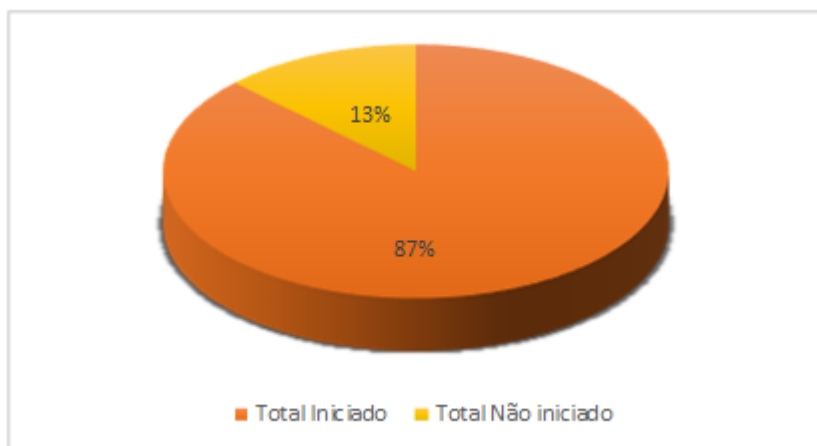
6 ANÁLISE DOS CURSOS DE BIBLIOTECONOMIA

Historicamente os cursos de Biblioteconomia brasileiros trazem duas vertentes de ensino, uma mais humanista abordando conceitos mais teóricos e/ou filosóficos, e outra mais tecnicista, trazendo uma abordagem mais aplicada, se utilizando de técnicas bibliotecárias, mas em ambas as abordagens a questão da organização, acesso e uso da informação estão presentes, questão também existente ao se referir à gestão de dados de pesquisa.

Nesta seção serão feitas as análises dos dados obtidos nessa pesquisa, ressaltando que, do universo pesquisado, se recuperou 33 PPCs. Em 19 instituições não foram encontrados disponíveis seus PPCs, sendo enviado uma mensagem de solicitação e que obteve o seguinte resultado: um curso não estava disponível até a completude da análise dos dados; em duas outras IES, os cursos não eram mais ofertados; um dos cursos respondeu à mensagem, mas não enviou o PPC; sete cursos não responderam à solicitação; por fim, não foi possível fazer contato com quatro delas, conforme disponível no APÊNDICE B.

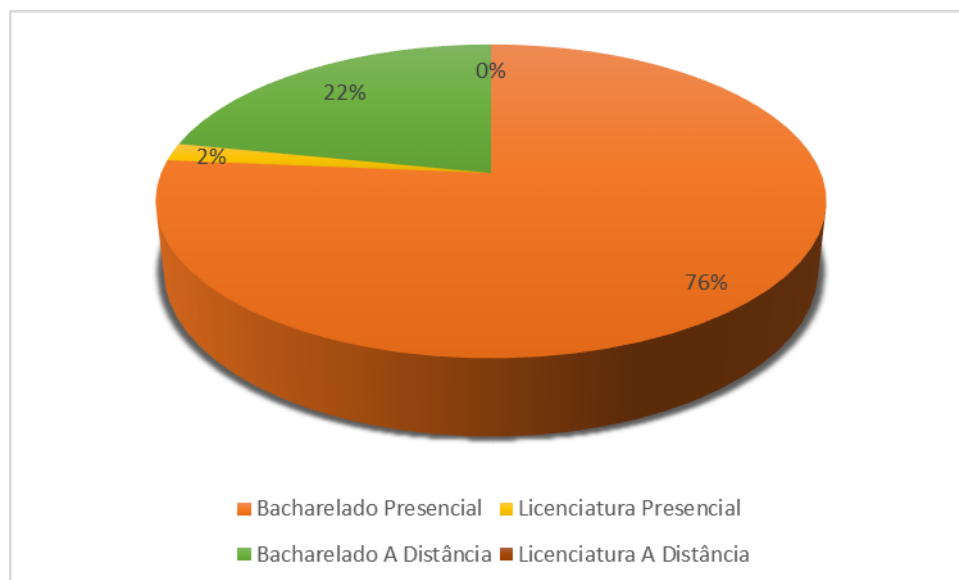
6.1 Panoramas Gerais

A partir do recorte recuperado no sistema do e-MEC, é possível notar que 07 cursos de Biblioteconomia se encontram cadastrados na plataforma, no entanto terão início de atividades até julho/2020, dentre os quarenta e oito cursos ativos. Partindo da perspectiva de que os cursos estão sendo criados em meio a esse contexto do *e-Science*, a possibilidade de abordagem do tema em seus conteúdos de formação é supostamente maior. No Gráfico 1, observa-se o percentual dos cursos de Biblioteconomia ativos e não iniciados:

Gráfico 1 – Panorama dos cursos ativos e não iniciados.

Fonte: Dados da pesquisa, 2021.

No gráfico 02, pode-se visualizar a habilitação e a modalidade de ofertas, verificando-se cinquenta e quatro cursos com a habilitação de Bacharelado, dentre estes quarenta e dois presenciais e doze à distância e apenas um curso com a habilitação em licenciatura, na modalidade presencial.

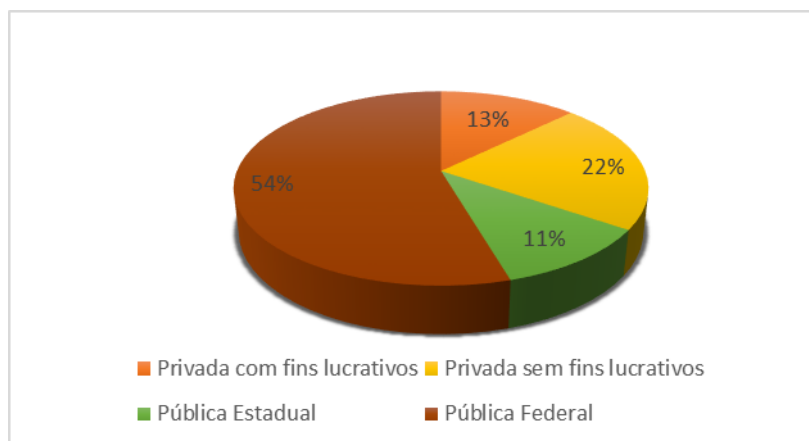
Gráfico 2 – Habilitação e modalidade de ofertas dos Cursos de Biblioteconomia.

Fonte: Dados da pesquisa, 2021.

No que se refere à categoria administrativa, os cursos de Biblioteconomia se classificam da seguinte forma: privada com fins lucrativos, com 7 cursos; privada sem fins lucrativos, com 12 (doze) cursos; pública estadual, com 6 cursos; e pública federal, com 30 (trinta) cursos. Assim, conforme se verifica no Gráfico 3, evidencia-se a oferta preponderante do ensino de Biblioteconomia em IES públicas federais em relação as

demais, e mesmo entre as privadas, as instituições sem fins lucrativos são majoritárias. Nessa perspectiva, aventa-se que o acesso principal à área no Brasil se dá pela rede pública ou sem fins lucrativos.

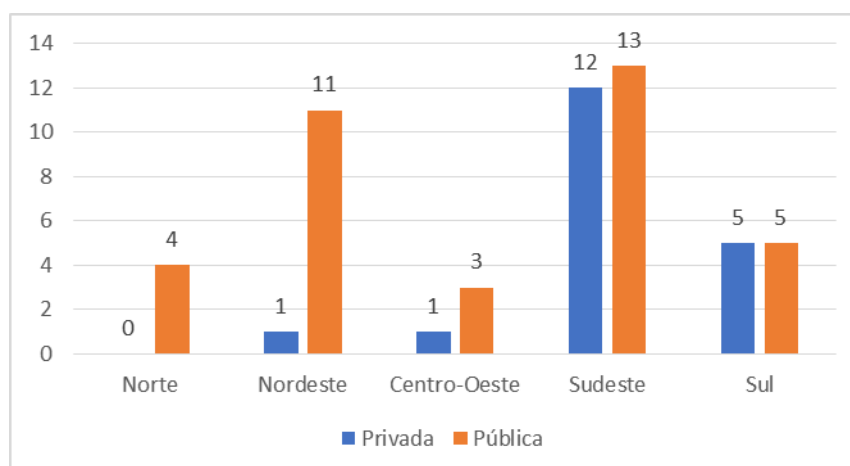
Gráfico 3 – Tipo de categoria administrativa a qual pertence os Cursos de Biblioteconomia



Fonte: Dados da pesquisa, 2021.

Para uma melhor compreensão da forma em que os cursos de Biblioteconomia se dividem pelas regiões no país, o gráfico 04 permite a comparação da quantidade de cursos de Biblioteconomia ofertada pelas instituições privadas e públicas nas regiões geográficas, verificando-se que as instituições privadas se concentram nas regiões Sul, com cinco cursos e na região Sudeste, com doze cursos. Já as instituições públicas alcançam todas as regiões, mas concentram-se no Sudeste, com treze cursos e no Nordeste com onze cursos. A maioria dos cursos são oferecidos por instituições públicas, trinta e seis, em contrapartida com as 19 ofertas das instituições privadas.

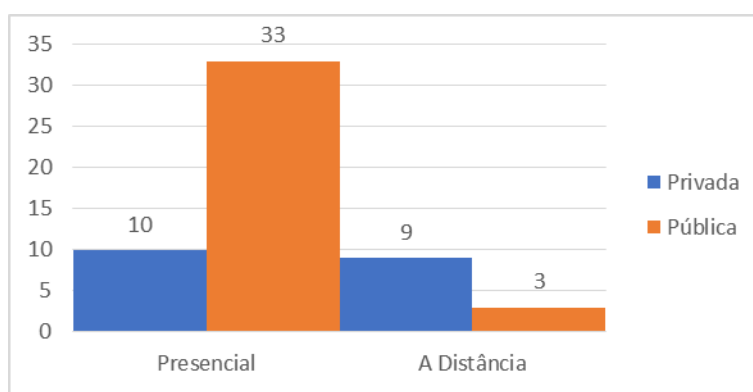
Gráfico 4 – Quantidade de cursos de Biblioteconomia por Região, divididos entre públicos e privados



Fonte: Dados da pesquisa, 2020.

A seguir, no Gráfico 5, pode-se perceber que a maioria dos cursos de Biblioteconomia no Brasil são oferecidos na modalidade presencial, totalizando quarenta e três cursos, já a modalidade à distância, oferta doze cursos. No entanto, essa disparidade melhor se evidencia nas instituições públicas com trinta e três cursos presenciais e três cursos na modalidade à distância, já nas instituições privadas a divisão aponta dez cursos presenciais e nove cursos à distância, demonstrando um equilíbrio entre as modalidades.

Gráfico 5 – Cursos de Biblioteconomia, segundo a modalidade



Fonte: Dados da pesquisa, 2020.

Num panorama geral, verifica-se que os cursos de Biblioteconomia no Brasil, em sua maioria, estão em instituições públicas na modalidade presencial e, em quase sua totalidade, oferecendo o grau de Bacharelado (há oferta atualmente de apenas um curso de licenciatura). Nota-se, também, que enquanto as instituições públicas alcançam todas as regiões do Brasil, as privadas se concentram nas regiões Sul e Sudeste, economicamente mais desenvolvidas.

Vale ressaltar que os cursos na modalidade à distância oferecidos pelas instituições públicas, embora estejam na situação “Em Atividade”, ainda não iniciaram sua atuação, sendo que dois destes cursos (FURG e UFF) entraram nessa situação ainda no ano de 2020.

6.2 Análise curricular

Para melhor compreensão dos resultados obtidos e suas posteriores análises, estabelecemos que a análise curricular relativa às disciplinas associativas a alguma

competência do(a) bibliotecário(a) de dados, inerentes a cada matriz curricular analisada, será realizada pela divisão regional para, posteriormente, apresentar uma análise genérica.

Partindo dos pressupostos discutidos no referencial teórico, há habilidades e competências tradicionais em Biblioteconomia que certificam o profissional a atuar com dados de pesquisa científica, principalmente em relação à organização, ao acesso e ao uso. Tais disciplinas são relacionadas à administração, classificação, indexação, assim, mesmo que não haja em suas ementas uma relação direta com os dados de pesquisa e/ou *e-Science*, sabe-se que, mesmo indiretamente, há essa relação, contudo, essas disciplinas não serão analisadas, pois este trabalho pretende analisar o esforço dos cursos em se adaptarem ao novo cenário da pesquisa científica colaborativa.

Nesse sentido, baseado no Quadro 3, seção 4, elencou-se os seguintes termos: **dados de pesquisa, pesquisa científica, Ciência de Dados, Data Science, ciclo de vida, dados abertos, open data, preservação digital, gestão de dados, reúso de dados, colaborativo, curadoria e metadados**; sendo desenvolvida uma pesquisa a partir destas expressões nos títulos e nas ementas para verificar se o curso oferta componentes curriculares que tenham como objetivo a capacitação do(a) bibliotecário(a) de dados.

6.2.1 Nordeste

A Região Nordeste compreende nove estados e todos ofertam pelo menos um curso de Biblioteconomia, dentre destes oito cursos foram analisados (no estado do Piauí o curso é ofertado pela UESPI, porém este trabalho não teve acesso ao seu PPC). Dentre os cursos pesquisados foram analisados os PPC dos cursos dos seguintes estados: Paraíba (UFPB), Ceará (UFCA) e (UFC), Pernambuco (UFPE), Alagoas (UFAL), Sergipe (UFS), Bahia (UFBA) e Rio Grande do Norte (UFRN). No Maranhão há o curso na UFMA, porém não se obteve acesso ao seu PPC e nem à matriz curricular.

O curso de Biblioteconomia da Universidade Federal da Paraíba (UFPB) foi criado em 06 de janeiro de 1969, obteve seu reconhecimento em 01 de setembro de 1975, possui os turnos: vespertino⁵⁴ e noturno, com carga horária 2760 h/a (184 créditos), e seu último PPC foi formulado em 2007 (UFPB, 2007).

⁵⁴ Segundo informação na página 12, este turno está suspenso desde 2004, e conforme verificado permanece assim atualmente.

É ressaltado neste documento um caráter transigente, capaz de se moldar às novas realidades que se apresentarem no ensino de Biblioteconomia, dessa forma, segundo o PPC da UFPB, “[...] não o consideramos como um documento definitivo, ao contrário, é dotado de um caráter dinâmico que possibilita a revisão constante de seus planos de ensino, de programas e de conteúdo, sempre que se fizer necessária uma renovação [...]” (UFPB, 2007, p. 6).

Neste PPC, há uma seção específica para “competências, atitudes e habilidades do bibliotecário” (UFPB, 2007, p. 24), que pontua algumas competências objetivadas em seu ensino, entre elas, tendo em vista o objetivo deste estudo, destacam-se a: aptidão para criar, desenvolver e utilizar técnicas de coleta, tratamento, recuperação e disseminação da informação; domínio no desenvolvimento e utilização de tecnologias, a fim de responder às demandas sociais de informação produzidas pelas transformações tecnológicas que caracterizam o mundo contemporâneo; capacidade para trabalhar em equipes multidisciplinares; porém, mas não há referência a um estudo específico aos dados de pesquisa científica, conforme evidenciado no Quadro 4:

Quadro 4 – Termos relacionados a Bibliotecário(a) de dados na UFPB

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	1
Ciência de Dados	0	0
<i>Data Science</i>	0	0
Ciclo de vida	0	0
Dados abertos	0	0
<i>Open data</i>	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	1

Fonte: dados da pesquisa, 2021.

Dos termos pesquisados, foram encontrados *pesquisa científica* na disciplina *Pesquisa Aplicada à Ciência da Informação*, referenciando os paradigmas da pesquisa científica, que potencialmente pode contemplar a introdução do conhecimento à *e-Science*; já o termo *metadados* na disciplina *Tecnologia da Informação II* pode indicar a presença de estudos sobre a disseminação da informação por processos eletrônicos,

arquivos abertos, padrões para intercâmbio de informação eletrônica e metadados, tendências em tecnologias da informação. Nessa disciplina, fica aberta a possibilidade de abordagem sobre Biblioteconomia de Dados, porém de forma subjetiva, ficando exclusivamente a critério do docente que lecioná-la a sua abordagem.

Quadro 5 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFPB

DISCIPLINAS	EMENTAS
PESQUISA APLICADA A CIÊNCIA DA INFORMAÇÃO	Paradigmas da pesquisa científica. Métodos e técnicas nas Ciências Sociais Aplicadas. Investigação científica e produção do conhecimento na área da Ciência da Informação. Informação e pesquisa como prática. Enfoques analíticos alternativos. A construção do corpus na pesquisa científica.
TECNOLOGIA DA INFORMAÇÃO II	Disseminação da informação por processos eletrônicos. Bases de Dados. Bibliotecas eletrônicas, digitais e virtuais. Sistemas de gerenciamento de periódicos eletrônicos. Arquivos abertos, padrões para intercâmbio de informação eletrônica e metadados. Tendências em tecnologias da informação.

Fonte: elaborado pelo autor, 2021.

Dessa forma, constata-se que o curso de Biblioteconomia da UFPB não está formalmente preparado para formar um profissional habilitado em atuar como bibliotecário(a) de dados, porém há abertura em ao menos duas disciplinas para docentes, por iniciativa própria, introduzir o tema aos discentes.

O próximo curso de Biblioteconomia a ser analisado pertence à Universidade Federal do Cariri (UFCA), cujo PPC foi aprovado em 03 de fevereiro de 2006, visando a criação do curso de Biblioteconomia na instituição. Vale ressaltar que a UFC-Cariri, criada em 2006, fez parte do projeto Reuni das IES, e em 2013 recebeu autonomia e desmembrando-se da UFC, o que originou o PPC utilizado pela UFCA até hoje.

A integralização curricular da UFCA é composta por um total de 200 créditos correspondentes a 3.200 horas, devendo, este total, ser cursado no prazo de oito semestres. Além da carga horária em disciplinas de formação, o discente deverá cumprir no mínimo 200 horas em Atividades Complementares, em conformidade ao que estabelece as Diretrizes Curriculares do Curso de 2005 (UFC/Campus Cariri 2006).

A missão do curso contempla a atenção à formação de um profissional capaz de gerenciar unidades de informação, utilizando recursos tecnológicos para tal, bem como

habilitar esse profissional a exercer sua função, conforme o desenvolvimento da sociedade contemporânea no exercício de sua cidadania.

[...] espera-se formar um profissional com habilidades gerenciais e tecnológicas, capazes de desenvolver a função técnica, política e social com capacidade para interpretar a realidade dessa sociedade contemporânea no que compete a gestão da informação para subsidiar o exercício da cidadania, assim como para a tomada de decisão em diversos estádios [estágios] organizacionais da sociedade, como forma de promover os seus crescimentos e desenvolvimentos (UFC/Campus Cariri, 2006, p. 7).

Quadro 6 – Termos relacionados a Bibliotecário(a) de dados na UFCA

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	1
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

No PCC da UFCA não se localizou nenhuma disciplina que se relacionasse diretamente com a formação de um profissional habilitado a trabalhar com dados de pesquisa. No Quadro 6, verifica-se que em uma das ementas encontra-se o termo *colaborativo*, porém refere-se à educação no ensino à distância.

Quadro 7 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFCA

DISCIPLINAS	EMENTAS
-------------	---------

TECNOLOGIAS DA INFORMAÇÃO II	Reflexões teórica sobre as tecnologias da informação e sua prática nas modernas organizações. A compreensão do uso eficaz das TÍ's, a fim de obter de garantir vantagens competitivas. Reflexão sobre a importância das tecnologias da informação e as exigências do perfil moderno dos profissionais da área de Ciência da Informação. Compreensão da aplicação das tecnologias da informação como suporte estratégico nos ambientes colaborativo e cooperativo de aprendizagem via Educação à Distância.
------------------------------	--

Fonte: elaborado pelo autor, 2021.

Verifica-se o distanciamento do curso de Biblioteconomia da UFCA na formação de um(a) Bibliotecário(a) de dados, atentando que este trabalho não faz uma análise nas já referidas disciplinas cujo ensino traz uma capacitação tradicional na formação desse novo profissional.

O curso de Biblioteconomia da Universidade Federal do Ceará (UFC) foi criado em 1965, com ensino mais próximo ao tecnicismo. Em 1985, seguindo o CFE, inicia a implantação de um currículo mais direcionado a “[...] um compromisso social que respondesse aos anseios informacionais da sociedade.” (UFC, 2006, p. 2). A integralização curricular é composta por um total de 3.072 horas, devendo ser cursado integralmente em um prazo de oito semestres. O PPC analisado desse curso foi atualizado em outubro de 2006.

Vale ressaltar que diversas disciplinas não possuíam ementas, dessa forma, procurou-se completar os dados com uma pesquisa no sítio institucional do curso, na página de Planos de Ensino,⁵⁵ porém, mesmo assim, algumas ainda permaneceram sem ementas; ressalta-se que duas disciplinas que estão na matriz curricular não estão no PPC.

Quadro 8 – Termos relacionados a Bibliotecário(a) de dados na UFC

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0

⁵⁵<https://biblioteconomia.ufc.br/pt/sobre-o-curso/ementas-das-disciplinas-do-curso-de-biblioteconomia/>
Acesso em 01/04/2021.

Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	1
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

No PCC do curso de Biblioteconomia da UFC não se localizou nenhuma disciplina que direcionasse em seu escopo a formação de um profissional capacitado a trabalhar com dados de pesquisa. No Quadro 8, verifica-se que na ementa da disciplina *Tecnologias da Informação II*, encontra-se o termo *colaborativo* como tecnologias informacionais para apoio de aprendizagem colaborativa, dessa forma, o curso de Biblioteconomia da UFC, tal qual a UFCA, se encontra em igual distanciamento na formação de um(a) Bibliotecário(a) de dados.

Quadro 9 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFC

DISCIPLINAS	EMENTAS
Tecnologias da Informação II	Compreensão do uso eficaz das TI's, a fim de garantir vantagens competitivas. Estudo sobre a importância das tecnologias da informação e as exigências do perfil moderno dos profissionais da área de Biblioteconomia e Ciência da Informação. Compreensão da aplicação das tecnologias da informação como suportes estratégicos em ambientes colaborativos e cooperativos de aprendizagem.

Fonte: elaborado pelo autor, 2021.

O curso de Biblioteconomia da Universidade Federal de Pernambuco (UFPE) foi criado em 1948 e teve seu reconhecimento em 1966. A integralização curricular é composta por um total de 2450 horas, devendo ser cursado totalmente no prazo de oito semestres. O PPC analisado desse curso teve uma atualização em dezembro de 2018 (UFPE, 2018).

O curso de Biblioteconomia da UFPE neste PPC salienta a relação entre o profissional bibliotecário(a) e o trabalho com a tecnologia, a rede e os dados:

Ainda, no que concerne à tecnologia, os profissionais bibliotecários estão, cada vez mais, lidando com os seguintes aspectos: curadoria digital, preservação e distribuição; criação, gestão e análise de conteúdo nas redes sociais; análise de dados e de plataformas robustas de armazenamento e disseminação. Esses fatores exigem uma formação interdisciplinar e transversal.

Sob a perspectiva econômica, há um crescente número de negócios mediados pela tecnologia e que exigem conhecimento amplo e diversificado. Os profissionais bibliotecários diariamente são desafiados a trabalhar com uma quantidade enorme de dados e de informações dispersas em diversos pontos da rede mundial de computadores. (UFPE, 2018, p. 11).

A princípio, verifica-se uma ação do curso de Biblioteconomia da UFPE no sentido de atualizar seu PPC para a realidade de ensino contemporâneo, embora o documento atual possua apenas dois anos de efetivação e ainda esteja aberto para um maior desenvolvimento de seu conteúdo. Nota-se a atenção do colegiado do curso em relação às demandas do profissional bibliotecário(a) com *e-Science* e ao tratamento de Dados de Pesquisa, embora não cite essas áreas diretamente, percebe-se uma atenção e um cuidado com o tratamento de enorme quantidade de dados, curadoria digital e análise de dados.

Quadro 10 – Termos relacionados a Bibliotecário(a) de dados na UFPE

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	1
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	1
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

No curso de Biblioteconomia da UFPE também não se localizou nenhuma disciplina que contribuísse, de modo direto, na formação de um profissional capacitado a trabalhar com dados de pesquisa. No Quadro 10, verifica-se que, na ementa da disciplina *Métodos e Técnicas Aplicados à Ciência da Informação*, se encontra o termo

pesquisa científica, como sentido atribuído de se utilizar e métodos e técnicas para pesquisa; e na ementa da disciplina *Gestão Documental*, encontra-se o termo *preservação digital* mais próximo pragmaticamente em relação à gestão documental do que a dados de pesquisa.

Quadro 11 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFPE

DISCIPLINAS	EMENTAS
Métodos e Técnicas Aplicados à Ciência da Informação	Principais métodos e técnicas de pesquisa científica aplicados à Ciência da Informação. Pesquisa quantitativa e pesquisa qualitativa. Etapas da pesquisa e operacionalização de modelos técnicos e metodológicos na Ciência da Informação.
Gestão Documental	Princípios fundamentais do processo de gerenciamento de documentos em meio digital, incluindo critérios e metodologias para avaliação de documentos e elaboração de projetos de digitalização e preservação digital.

Fonte: elaborado pelo autor, 2021.

Dessa forma, é possível atentar que embora este PPC do curso de Biblioteconomia da UFPE, mesmo que indiretamente, pondere a necessidade do profissional bibliotecário(a) atuar no contexto da *e-Science* em atenção aos dados de pesquisa, em suas disciplinas, efetivamente, não há essa aplicação. De toda forma, é preciso ponderar o processo de desenvolvimento em que este documento se encontra, bem como a existência da uma observação demonstrando que o quadro de disciplinas ainda não foi atualizado.

O curso de Biblioteconomia da Universidade Federal de Alagoas (UFAL) foi criado em 1998 e teve seu reconhecimento em 2005. A proposta inicial apresentava “uma visível preocupação em romper com a formação tradicional adotada em alguns cursos de Biblioteconomia do país, embora fosse necessário atender às exigências do currículo mínimo, em vigência” (UFAL, 2019, p. 16). Em 2000, dois anos após sua criação, o curso já passou por atualizações visando ampliar o horizonte: “A finalidade era formar um(uma) profissional capacitado(a) para atuar na organização e disseminação da informação e na gestão de acervos, utilizando técnicas modernas e tecnologias de informação e comunicação contemporâneas” (UFAL, 2019, p. 16).

A integralização curricular é composta por um total de 2400 horas, devendo ser cursado no prazo de oito semestres. Cabe ressaltar que o PPC analisado foi atualizado em 03 setembro de 2019.

Destarte, é possível notar que o curso de Biblioteconomia da UFAL, já em sua criação, possuía o intuito de ser um curso voltado a suprir necessidades contemporâneas, conforme elas surgissem, avançando em seu ensino para além do tradicional e conservador, alcançando assim as atualizações da profissão quando elas surgissem.

Quadro 12 – Termos relacionados a Bibliotecário(a) de dados na UFAL

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	1
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

No curso de Biblioteconomia da UFAL não se localizou nenhuma disciplina que mantivesse uma relação direta com a formação de um profissional capacitado a trabalhar com dados de pesquisa. No Quadro 13, verifica-se que, na ementa da disciplina *Tecnologias de Informação, Comunicação e Inovação*, se encontra o termo *preservação digital*, referindo-se mais à gestão documental do que a dados de pesquisa.

Quadro 13 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFAL

DISCIPLINAS	EMENTAS
Tecnologias de Informação, Comunicação e Inovação	Sociedade da informação. Revolução digital. Divulgação científica, tecnológica e de inovação. Acessibilidade tecnológica. Gerenciamento eletrônico de documentos e preservação digital. Inovação e processos tecnológicos em unidades de informação.

Fonte: elaborado pelo autor, 2021.

Logo, infere-se que, embora o intuito do curso seja estar sempre atualizado em relação às atualizações da profissão da Biblioteconomia, e seu PPC tenha sido atualizado recentemente, ainda não contempla disciplinas que atendam à capacitação para atividades voltadas ao perfil do(a) bibliotecário(a) de dados.

Não foi encontrado o PPC do curso de Biblioteconomia da Universidade Federal de Sergipe (UFS), somente a resolução 95/2011/CONEPE que o aprovou em outubro de 2011, porém nessa resolução e no sítio institucional do curso foi possível encontrar os dados necessários.

O curso foi criado em 2008, a partir do Programa de Expansão da Universidade, em consonância com o REUNI. A sua integralização curricular é composta por um total de 2640 horas, devendo ser cursado no prazo mínimo de oito semestres.

Quadro 14 – Termos relacionados a Bibliotecário(a) de dados na UFS

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	8
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	1

Fonte: dados da pesquisa, 2021.

No curso de Biblioteconomia da UFS não se identificou nenhuma disciplina que colaborasse diretamente na formação de um profissional capacitado a trabalhar com dados de pesquisa. No Quadro 14, verifica-se duas disciplinas que se utilizam do termo *pesquisa científica*: *Pesquisa Aplicada à Biblioteconomia*, que discorre sobre métodos, técnicas e tipos de pesquisa científica; e a disciplina *Metodologia Científica Aplicada à Biblioteconomia*, voltada para as Ciências Sociais Aplicadas, com ênfase na Ciência da Informação, além da normalização de trabalhos de grau.

Dentre as optativas, nota-se que possui mais seis disciplinas que apresentam o termo *pesquisa científica*, mas todas direcionadas à metodologia, com ênfase em outras linhas das Ciências Sociais Aplicadas.

Quadro 15 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFS

DISCIPLINAS	EMENTAS
Pesquisa Aplicada à Biblioteconomia	Métodos, técnicas e tipos de pesquisa científica.
Metodologia Científica Aplicada à Biblioteconomia	Métodos e Técnicas de Pesquisa Científica voltadas para as Ciências Sociais Aplicadas, com ênfase na Ciência da Informação. Normalização de Trabalhos de Grau.
Metodologia Científica Aplicada às Ciências Humanas	Métodos e Técnicas de Pesquisa Científica voltadas para as Ciências Sociais Aplicadas, com ênfase em Ciências Humanas. Normalização de Trabalhos de Grau.
Metodologia Científica Aplicada às Ciências Biológicas e Saúde	Métodos e Técnicas de Pesquisa Científica voltadas para as Ciências Sociais Aplicadas, com ênfase em Ciências Biológicas e Saúde. Normalização de Trabalhos de Grau.
Metodologia Científica Aplicada às Ciências Exatas	Métodos e Técnicas de Pesquisa Científica voltadas para as Ciências Sociais Aplicadas, com ênfase em Ciências Exatas. Normalização de Trabalhos de Grau.
Metodologia Científica Aplicada às Ciências Humanas	Métodos e Técnicas de Pesquisa Científica voltadas para as Ciências Sociais Aplicadas, com ênfase em Ciências Humanas. Normalização de Trabalhos de Grau.
Metodologia Científica Aplicada às Ciências Biológicas e Saúde	Métodos e Técnicas de Pesquisa Científica voltadas para as Ciências Sociais Aplicadas, com ênfase em Ciências Biológicas e Saúde. Normalização de Trabalhos de Grau.
Metodologia Científica Aplicada às Ciências Exatas	Métodos e Técnicas de Pesquisa Científica voltadas para as Ciências Sociais Aplicadas, com ênfase em Ciências Exatas. Normalização de Trabalhos de Grau.
Formato de Intercâmbio MARC 21	Estruturas automatizadas de descrição bibliográfica, metadados, padrões de metadados. Automação de unidades de informação.

Fonte: elaborado pelo autor, 2021.

Verifica-se no curso de bacharelado em Biblioteconomia da UFS a ausência de disciplinas que contemplem habilidades e competências necessárias ao bibliotecário(a) que atue com dados de pesquisa científica no contexto de *e-Science*.

O curso de Biblioteconomia da Universidade Federal da Bahia (UFBA), segundo seu sítio institucional⁵⁶, foi criado em 1942 e, em 1954, recebeu a nomenclatura *Biblioteconomia e Documentação*, passando a fazer parte dos cursos ofertados pela Universidade da Bahia, atual Universidade Federal da Bahia. Atualmente, o curso tem o propósito de preparar o profissional bibliotecário(a) a “(...) atuar em bibliotecas públicas, escolares, universitárias e especializadas, com acervos físicos e/ou digitais, bem como na gestão de informação, na pesquisa, na inteligência tecnológica, na consultoria e editoração” (UFBA, 2021, *online*).

Este estudo não teve acesso ao PPC deste último curso, porém foi obtida a matriz curricular, contendo as disciplinas obrigatórias e optativas. Deste universo, doze disciplinas não continham suas ementas, eram elas: *Introdução à Filosofia*, *Lógica I*, *Introdução à Sociologia II*, *Inglês Instrumental II*, *Organização Temática da Informação III*, *Psicologia das Relações Humanas*, *História da Literatura I*, *História da Arte III*, *Literatura Brasileira VI*, *História da Civilização Brasileira*, *Paleografia e Ecdótica I*, *Literatura Universal I*. Dessas disciplinas, somente a *Lógica I* poderia ter relação com Biblioteconomia de Dados, caso abarcasse lógica algorítmica visando tratamento com dados de pesquisa científica, porém sem a ementa não foi possível fazer uma análise fidedigna à natureza de sua aplicação.

Quadro 16 – Termos relacionados a Bibliotecário(a) de dados na UFBA

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	1
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

⁵⁶ SEAD - SUPERINTENDÊNCIA DE EDUCAÇÃO A DISTÂNCIA (Bahia). Universidade Federal da Bahia. **Bacharelado em Biblioteconomia**. 2021. Disponível em: <https://sead.ufba.br/cursos/bacharelado-em-biblioteconomia>. Acesso em: 01 fev. 2021.

Na falta de acesso ao PPC, não foi possível analisar o direcionamento adotado pelo curso de Biblioteconomia da UFBA. Observando a matriz curricular com o ementário das disciplinas, localizou-se apenas uma disciplina com o termo *pesquisa científica*, porém sem relação com *e-Science*, pois trata de noções gerais sobre conhecimento, método e pesquisa científica, baseadas em métodos e técnicas tradicionais de pesquisa, conforme exposto no Quadro 17.

Quadro 17 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFBA

DISCIPLINAS	EMENTAS
Metodologia e Técnica de Pesquisa em Biblioteconomia e Ciência da Informação.	Noções gerais sobre conhecimento, método e pesquisa científica. O processo de pesquisa, etapas do projeto e do desenvolvimento de uma pesquisa, desde a definição do problema até a apresentação dos resultados. A pesquisa em Biblioteconomia e Ciência da Informação. Métodos e técnicas qualitativos e quantitativos aplicados à Biblioteconomia e à Ciência da Informação. Elaboração e desenvolvimento de um projeto de pesquisa.

Fonte: elaborado pelo autor, 2021.

No que foi possível investigar no curso de Biblioteconomia da UFBA, não se encontrou nenhuma disciplina ligada à *e-Science*.

O Curso de Bacharelado em Biblioteconomia da Universidade Federal do Rio Grande do Norte (UFRN) tem seu PPC atual datado de 2017. Funciona em turno vespertino, com uma carga horária de 2730 horas e oferta 35 vagas por semestre. O Departamento de Ciência da Informação (DECIN) ainda proporcionou três cursos *lato sensu*: o Curso de Especialização em Gestão Estratégica de Sistemas de Informação; o curso de aperfeiçoamento em Planejamento, Criação e Manutenção de Repositórios Digitais; e o curso de especialização em Gestão Documental (UFRN, 2017).

A estrutura curricular deste curso, segundo o seu PPC, capacita o profissional bibliotecário(a) tanto no âmbito social quanto no mercado de trabalho, em diversos campos relacionados à informação:

A estrutura curricular do curso capacita o profissional bibliotecário(a) a atuar de acordo com as necessidades reais da sociedade e do mercado, uma vez que no decorrer da sua formação o aluno tem a oportunidade de perpassar pelas diversas áreas que compõem o curso: Fundamentos Teóricos da Biblioteconomia e Ciência da Informação;

Quadro 18 – Termos relacionados a Bibliotecário(a) de dados na UFRN

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	1
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	1
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	1	1
Curadoria	0	1
Metadados	0	3

Fonte: dados da pesquisa, 2021.

No PPC do curso de Biblioteconomia da UFRN, os termos relacionados ao Bibliotecário de dados, selecionados para uso nessa pesquisa, foram encontrados em sete disciplinas diferentes, sendo que em uma delas o termo se encontra tanto no título quanto na ementa da disciplina. Nem todos os casos havia realmente relação com o contexto de dados, mas, no contexto geográfico dos cursos situados no Nordeste, este é o perfil de curso de Biblioteconomia que demonstra maior preocupação na preparação do(a) bibliotecário(a) para atuar com Biblioteconomia de Dados.

Na ementa da disciplina *Metodologia da Pesquisa em Biblioteconomia e Ciência da Informação* o termo *pesquisa científica* está presente, contudo, sem relação com *e-Science*, visto que trata de noções gerais sobre conhecimento, método e pesquisa científica baseadas em métodos e técnicas tradicionais de pesquisa.

Já na disciplina *Preservação e Conservação de Documentos Impressos e Digitais* o termo *preservação digital* é usado em sua ementa, cujo conteúdo aponta a abordagem da conservação preventiva e curativa, ou reparativa de documentos, tanto analógicos como digitais. Sendo assim, aqui pode haver uma breve introdução sobre CVD, hipótese que não é obrigatória, já que não está explícito na ementa, portanto, não sendo possível afirmar que a disciplina versa sobre a temática, mas compreendendo que há abertura para esse tema ser trazido à tona nas discussões e estudos.

O termo *colaborativo* aparece tanto no título quanto na ementa da disciplina *Recursos e Serviços Colaborativos*. Nessa disciplina aparece também o termo *curadoria*, no caso, indicando forte aproximação com *e-Science* e Biblioteconomia de Dados, porém, tal qual na disciplina *Preservação e Conservação de Documentos Impressos e Digitais*, não há como alegar se é/será abordada a pesquisa colaborativa exercida pela *e-Science*, nem a curadoria existente no CVD.

Em sequência, observa-se que o termo *metadados* aparece em três disciplinas: *Catálogo Automatizada*, em que se verifica uma discussão sobre os metadados gerados na realização de Catalogação, principalmente no formato MARC21; na disciplina *Redes e Sistemas de Informação*, com ementa fazendo relação direta à Biblioteconomia de Dados, uma vez que aborda tecnologias emergentes no âmbito da Biblioteconomia, avaliação de sistemas de informação e ambientes informacionais digitais, além de Repositórios Digitais, assuntos diretamente envolvidos com dados de pesquisa.

Por fim, na disciplina *Serviço de Informação e Referência*, cuja ementa aborda o Serviço de Referência e Competência Informacional, pode existir relação com a Biblioteconomia de Dados, porém apenas através dos dados levantados não foi possível afirmar que o faça.

Quadro 19 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFRN

DISCIPLINAS	EMENTAS
Metodologia da Pesquisa em Biblioteconomia e Ciência da Informação (Equivalente à Metodologia da Pesquisa em Biblioteconomia II)	Fundamentos teóricos e práticos da pesquisa científica. Métodos e técnicas de pesquisa em Biblioteconomia e Ciência da Informação. Elaboração e desenvolvimento do projeto de pesquisa.
Preservação e Conservação de Documentos Impressos e Digitais	Teoria e métodos de preservação e conservação de documentos em suporte analógico e digital. Memória e sua Materialização na perspectiva da preservação documental. Conservação preventiva e curativa. Preservação Digital.
Recursos e Serviços Colaborativos	<i>Web</i> e recursos colaborativos: conceitos, características e aplicações. Biblioteca 2.0: conceitos, fundamentos e gestão. Curadoria de Conteúdos: fundamentos, processos e ferramentas.

Catálogo Automatizada (Equivalentes: Representação Descritiva III)	Catálogo de recursos de informação em ambientes digitais. O intercâmbio de informações: formatos, regras e padrões. Metadados. MARC (<i>Machine Readable Catalogin</i>): origem e histórico. Formato MARC: conceitos, finalidades, composições, campos e indicadores. Formato MARC: título uniforme e entidade coletiva. Catálogo descritiva e o RDA (<i>Resource Description Access</i>).
Redes e Sistemas de Informação	Redes de informação: protocolos de comunicação, metadados e interoperabilidade. Sistemas de informação para gestão de bibliotecas e de repositórios digitais. Tecnologias emergentes no âmbito da Biblioteconomia. Avaliação de sistemas de informação e ambientes informacionais digitais.
Repositórios Digitais	Repositórios digitais: tipologia, características e iniciativas. Gestão de repositórios digitais: políticas, metadados e <i>software</i> para implementação. Atuação do profissional da informação na construção e gestão de repositórios digitais.

Fonte: elaborado pelo autor, 2021.

Logo, é possível afirmar que o curso de Biblioteconomia da UFRN, embora se observe que em suas disciplinas não há uma abordagem direta ao tema Bibliotecário(a) de dados, oferta uma disciplina que versa sobre algumas habilidades e competências desse profissional, e outras duas disciplinas com grande potencial para abordar tais habilidades e competências.

Dessa forma, no que tange aos cursos de Biblioteconomia da Região Nordeste, é possível afirmar que apenas o pertencente à UFRN, se aproxima mais da formação de competências e habilidades do profissional em Biblioteconomia, que o habilita para a gestão de dados no contexto da *e-Science*. Já o curso integrante da UFPB tem uma disciplina que, possivelmente, estabelece alguma relação com a *e-Science*. Os demais cursos, seis ao todo, não possuem atenção a este tema.

6.2.2 Norte

A Região Norte compreende sete estados, possuindo, ao todo, dois cursos de Biblioteconomia analisados, ofertados nos estados de Rondônia (UNIR) e Pará (UFPA).

O curso de Biblioteconomia da Universidade Federal de Rondônia (UNIR) foi criado em 18 de novembro de 2008, sob a Resolução nº 198/CONSEA e teve seu início em 27 de julho de 2009, com reconhecimento a partir da Portaria nº 544, de 23 de

setembro de 2016. O curso oferece 50 vagas presenciais, com carga horária total de 2.520 horas e seu PPC datado em 2018 (UNIR, 2018).

[...] faz a opção em privilegiar a formação do bibliotecário para o trabalho com ferramental tecnológico, em sistemas de informações tecnológicas e empresariais, sem descuidar do campo social, uma vez que Rondônia se encontra em fase de grandes mudanças urbanas e sociais (UNIR, 2018, p. 5).

O projeto pedagógico deste curso visa formar bibliotecários aptos a atuarem com ferramentas de tecnologias informacionais, tanto no campo social quanto no empresarial:

Quadro 20 – Termos relacionados a Bibliotecário(a) de dados na UNIR

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	1	2
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	1
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

No PPC do curso de Biblioteconomia da UNIR, os termos relacionados às habilidades do(a) bibliotecário(a) de dados, selecionados para uso nesta pesquisa, foram encontrados em quatro disciplinas.

O termo *pesquisa científica* foi encontrado em três disciplinas, uma no título da disciplina *Introdução ao Trabalho Científico*, que trata de forma mais epistemológica a pesquisa científica, e não indicando relação direta com Biblioteconomia de Dados; outra na ementa da disciplina *Metodologia da Pesquisa Científica em Biblioteconomia* e

*Ciência da Informação*⁵⁷, que evidencia metodologias de pesquisa em Biblioteconomia e Ciência da informação visando à elaboração do anteprojeto de pesquisa, não estando diretamente ligada à Biblioteconomia de Dados, mas, potencialmente, o docente pode incluir métodos e técnicas de pesquisa relacionados com a *e-Science*; e, por fim, na disciplina optativa *Comunicação e Produção Científica*, que aborda a pesquisa científica e a produção do conhecimento, também não relacionada diretamente à Biblioteconomia de Dados, mas apontando possibilidades ao docente para abordar *e-Science* e/ou Ciência de Dados.

Outro termo encontrado foi *preservação digital*, na disciplina optativa *Fundamentos em Arquivologia*, a qual aborda preservação digital e pode versar sobre o CVD, mas também ficando a cargo do docente a decisão da abordagem, sem encontrar expressa a abordagem em sua ementa.

Quadro 21 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UNIR

DISCIPLINAS	EMENTAS
Introdução ao Trabalho Científico	Filosofia da Ciência. História do pensamento científico. Estudo dos diferentes enfoques e vertentes teóricas adotadas na pesquisa científica. Estrutura e apresentação de pesquisa científica e trabalho acadêmico. Comunicação científica.
Metodologia da Pesquisa Científica em Biblioteconomia e Ciência da Informação*	Estudo das principais etapas do trabalho científico, com ênfase nas metodologias de pesquisa em Biblioteconomia e Ciência da informação. Elaboração de anteprojeto de pesquisa.
Comunicação e produção científica	Introdução aos fundamentos da Ciência, Tecnologia e Sociedade. A pesquisa científica e a produção do conhecimento. Modelos e instrumentos de aferição e avaliação do conhecimento científico. Produção científica em Biblioteconomia e Ciência da Informação. Periódicos científicos em Biblioteconomia e Ciência da Informação: política e critérios para admissão e permanência em bases de dados.

⁵⁷ A disciplina Metodologia da Pesquisa Científica em Biblioteconomia e Ciência da Informação não consta no PPC, por este faltar uma página, mas a ementa foi recuperada no sítio institucional do curso: <http://www.daci.unir.br/pagina/exibir/11799>

Fundamentos em Arquivologia	Origens da Arquivologia. Fundamentos epistemológicos da arquivologia (Princípio da proveniência). Paradigmas da arquivologia. Gestão de Documentos. Teoria das Três Idades. Análise da produção científica e o desenvolvimento arquivístico no Brasil e no mundo. Arquivo na sociedade contemporânea. Princípios e técnicas de arquivamento. A prática arquivística em tempos de gestão do conhecimento. Preservação digital.
-----------------------------	---

Fonte: elaborado pelo autor, 2021.

Logo, é possível verificar uma ausência de disciplinas direcionadas à formação do profissional bibliotecário(a) para atuar na *e-Science*, mesmo este PPC sendo relativamente recente, por ter sido efetivado em 2018. Mesmo assim, há duas disciplinas em que os docentes podem realizar a abordagem do tema.

O PPC do curso de Biblioteconomia da Universidade Federal do Pará (UFPA) é datado de 2009, o curso surgiu a partir da Resolução nº 1 A, de 28/01/1963. Segundo este PPC, o curso oferece sessenta vagas, divididas em dois turnos, com carga horária mínima de 2.840h. (UFPA, 2018).

Em seu escopo, traz trinta e oito itens descrevendo as competências e as habilidades que deverão possuir os(as) bacharéis formados(as) por esta instituição, mas em sua maioria são descrições genéricas que, embora não abarquem o tema *e-Science*, e/ou Biblioteconomia de Dados- -mesmo porque esses temas se fortalecem no debate da área mais recentemente-, por serem gerais, há abertura para que esses temas sejam incluídos no ensino (UFPA, 2018).

Quadro 22 – Termos relacionados a Bibliotecário(a) de dados na UFPA

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	1
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	3

Fonte: dados da pesquisa, 2021.

No PPC do curso de Biblioteconomia da UFPA, quanto aos termos relacionados ao bibliotecário(a) de dados, selecionados para uso nessa pesquisa, foram encontrados dois termos em um total de quatro disciplinas. No entanto, vale ressaltar a existência de vinte e seis disciplinas com ementas indisponíveis.

O termo *pesquisa científica* foi encontrado na disciplina *Elaboração do Trabalho Acadêmico*, cuja ementa trata mais de uma introdução à elaboração de trabalho acadêmico, compreendendo as técnicas, os métodos e a normalização, não abrangendo a área de *e-Science* e/ou Biblioteconomia de Dados.

O termo *metadados* foi encontrado em três disciplinas: *Fontes de Informação II*, com a ementa direcionada para redes de informação e base de dados, abarcando estrutura e pesquisa, aqui há abertura para discorrer sobre repositório de dados e/ou dados de pesquisa relacionando-os com metadados; *Planejamento de Bases de Dados*, com sua ementa abordando de forma mais aprofundada a base de dados e as tecnologias envolvidas, bem como um maior aprofundamento em metadados, discutindo diretamente competências e habilidades de um(a) bibliotecário(a) de dados; e, por fim, em *Representação Descritiva da Informação II*, cuja ementa aborda com maior profundidade a catalogação, incluindo formatos utilizados para o controle bibliográfico universal, como MARC, catálogos abertos *online* e interoperabilidade entre eles, através de estruturas como os metadados e também incutindo habilidades e competências de um(a) bibliotecário(a) de dados.

Quadro 23 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFPA

DISCIPLINAS	EMENTAS
Elaboração do Trabalho Acadêmico	O uso de métodos para eficiência nos estudos. História, divisão e classificação do conhecimento científico. Introdução ao estudo da pesquisa científica: planejamento, conceitos, finalidades, fontes, tipos, fases, entre outros aspectos. Elaboração do pré- projeto de pesquisa. Aplicação das normas da ABNT: trabalhos acadêmicos, referências, citações, resumo, sumário, numeração progressiva das seções de um documento.

Fontes de Informação II	Sistemas informatizados de informação. Redes de informação e bases de dados. Tipos de sistemas de informação e bases de dados. Estrutura de bases de dados. Metadados. Pesquisa online e em CD-ROM. Principais bases de dados bibliográficos nacionais e internacionais. Acesso a bases de dados.
Planejamento de Bases de Dados	Caracterizar Bases de Dados – Bases de Dados Relacionais – <i>Data Warehouse</i> Bases de Referências e Fontes – <i>World Wide Web</i> . Metadados – Conceito de Metadados – Padrões de Metadados para Catálogos <i>On-line</i> e BD Digitais. Interoperabilidade de Informação – Modelo <i>Open Archives</i> – Arquivos Abertos. Construção de Catálogo <i>on-line</i> – Sistema Win ISIS. Modelagem de Bases de Dados Relacional – Modelo Entidade Relacionamento.
Representação Descritiva da Informação II	Cabeçalhos para pessoas e entidades; Legislação; Controle bibliográfico universal; Formatos de intercâmbio: IBICT, CALCO, MARC, UNISIST etc.; Redes de catalogação cooperativa; Sistemas automatizados de catalogação. <i>On-line Public Access Catalogs</i> – OPACs (Catálogo on-line): histórico, principais recursos, estrutura; Metadados; Dublin Core; Arquivos abertos; Conversão retrospectiva.

Fonte: elaborado pelo autor, 2021.

Com esses dados, é possível verificar que embora este PPC tenha mais de dez anos de sua efetivação, quando ainda estava nascendo nas discussões em Biblioteconomia e Ciência da Informação as questões inerentes à formação de bibliotecário de dados, já se percebe que duas disciplinas abordam diretamente algumas das habilidades e competências necessárias ao bibliotecário(a) de dados, e uma disciplina que o docente pode abordar potencialmente a formação do perfil para esse profissional.

Nos cursos de Biblioteconomia da Região Norte, o que possui PPC mais atual (2018) é a UNIR, porém, não foi possível encontrar disciplinas diretamente ligadas a alguma habilidade ou competência do(a) bibliotecário(a) de dados. No curso da UFPA, embora o PPC seja de 2009, possui duas disciplinas que abordam diretamente algumas dessas habilidades, sendo, no Norte, a instituição com melhor indicativo de fluxograma para o discente com desejo de trabalhar com dados de pesquisa científica.

6.2.3 Centro-Oeste

A Região Centro-Oeste compreende três estados e o Distrito Federal possuindo, ao todo, três cursos de Biblioteconomia analisados, os quais encontram assim ofertados: Mato Grosso (UFMT), Goiás (UFG) e Distrito Federal (UNB), o estado de Mato Grosso do Sul não oferta mais o curso de Biblioteconomia.

O curso de Biblioteconomia da Universidade Federal do Mato Grosso (UFMT) tem seu PPC datado de 2011, no entanto, sua estrutura curricular atual foi implementada pela Resolução nº 4 de 22/01/2007. O curso se iniciou em 2000, com o campus de Rondonópolis, sendo autorizado com a Resolução nº 63 de 2/8/1999 e reconhecido pela Portaria do MEC nº 4398, de 29/12/2004, oferecendo 40 vagas com carga horária mínima de 2.625h.

Quadro 24 – Termos relacionados a bibliotecário(a) de dados na UFMT

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	1
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

No PPC do curso de Biblioteconomia da UFMT, após a pesquisa com os termos relacionados a Bibliotecário(a) de dados, foi encontrado em apenas uma disciplina, a saber, o termo *pesquisa científica*, na disciplina *Pesquisa em Biblioteconomia e Ciência da Informação*, cuja ementa trata da introdução à elaboração de métodos e técnicas de pesquisa científica sem alcançar para a temática da *e-Science* e/ou Biblioteconomia de Dados.

Quadro 25 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFMT

DISCIPLINAS	EMENTAS
Pesquisa em Biblioteconomia e Ciência da Informação	Métodos científicos de investigação. A aplicação da pesquisa científica à Biblioteconomia e documentação. Tipos de Pesquisa. Projeto de Pesquisa. Coleta e análise de dados. Hipóteses. Variáveis.

Fonte: elaborado pelo autor, 2021.

O PPC do curso de Biblioteconomia da Universidade Federal de Goiás (UFG), criado em 1980, até 2004 se baseou numa formação generalista de bibliotecários possuindo uma grade curricular pautada em uma perspectiva tecnicista. O curso oferece cinquenta vagas anuais com carga horária mínima de 3.080h e tem seu PPC atual datado de 2016.

Em seus objetivos específicos, há um eixo tecnológico o qual propõe que o bacharel em Biblioteconomia da UFG seja capaz de “Dominar as tecnologias e os processos envolvidos na sua produção, organização, uso e recuperação de informação;” (UFG, 2016, p. 11), o que poderia contemplar algumas habilidades e competências do(a) bibliotecário(a) de dados.

Quadro 26 – Termos relacionados a Bibliotecário(a) de dados na UFG

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	1

Fonte: dados da pesquisa, 2021.

Embora nos objetivos do PPC do curso de Biblioteconomia da UFG aponte para um domínio tecnológico voltado para os agentes aos quais a informação é o cerne, em

que o bibliotecário(a) de dados se enquadra, não há disciplina diretamente ligada a esse tema. Ressalta-se que, embora a disciplina *Representação Descritiva III* inclua o termo *metadados*, esta versa sobre seu uso na catalogação e no domínio bibliográfico, mas não na organização de dados de pesquisa.

Quadro 27 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFG

DISCIPLINAS	EMENTAS
Representação Descritiva III	Requisitos funcionais para registros bibliográficos. Padrões de metadados no domínio bibliográfico. Estudo e aplicação do código de catalogação RDA. Tendências em representação descritiva da informação.

Fonte: elaborado pelo autor, 2021.

O PPC do curso de Bacharelado em Biblioteconomia da Universidade de Brasília (UnB) está em processo de revisão, mesmo assim, foi fornecido seu acesso, via e-mail, pela então coordenadora do curso, Professora Dra. Fernanda de Souza Monteiro. O curso se iniciou em março de 1966, foi autorizado e reconhecido legalmente em novembro de 1972, sendo ofertado no período diurno, com carga horária de 2700h. (UnB, 2019, p. 6).

Esse curso é pautado pela interdisciplinaridade, alcançando diversas áreas do conhecimento, entre elas a Administração e a Ciência da Computação, ambas necessárias para contemplar algumas das habilidades do(a) bibliotecário(a) de dados.

[...] o curso é engrandecido com disciplinas de áreas como a Ciência da Computação, Letras tradução, Administração e as áreas conceitualmente mais próximas como Arquivologia e Museologia. Todas essas trazem conhecimentos interdisciplinares, capazes de complementar a formação do bibliotecário (UnB, 2019, p. 38).

Quadro 28 – Termos relacionados a Bibliotecário(a) de dados na UnB

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	1	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0

Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	1	1

Fonte: dados da pesquisa, 2021.

O termo *pesquisa científica* aparece no título da disciplina *Introdução à Pesquisa Científica*, e seu conteúdo trata de uma introdução ao conceito e à natureza científica, discutindo métodos e técnicas de pesquisa científica, sem abordar a *e-Science* e/ou a Biblioteconomia de Dados.

Observa-se na disciplina intitulada por *Tópicos Especiais: Introdução ao Estudo de Metadados*, o termo *metadados*. Essa disciplina traz uma introdução geral ao estudo dos padrões de metadados aplicáveis aos objetos de informação textuais, videográficos, sonoros, iconográficos, museológicos e arquivísticos, mostrando também correlação direta com habilidades e competências do(a) bibliotecário(a) de dados.

A disciplina *Redes de Informação e Transferência de Dados* não possui nenhum dos termos buscados em sua ementa, mas o título evoca sua relação com Biblioteconomia de Dados e *e-Science*, deste fato, buscou-se no programa da disciplina e foram localizados conceitos e técnicas para montar infraestruturas e arquitetura de redes e de comunicações de dados, trabalhando com quatro sistemas operacionais, diversos padrões de metadados, e com de modelos de redes bibliográficas: catálogo, biblioteca digital, repositório institucional e revista eletrônica, além de redes de biblioteca como *Pergamum*, BDTD-IBICT, BVS-FIOCRUZ, RVBI, CCN etc. Outra disciplina que se enquadra de maneira efetiva no quadro de habilidades e competências de um(a) bibliotecário(a) de dados.

Quadro 29 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UnB

DISCIPLINAS	EMENTAS
Introdução à Pesquisa Científica	EMENTA: A leitura como método, a compreensão do conceito de ciência, a natureza do conhecimento científico, o método científico, as normas para a apresentação de trabalhos científicos.

Tópicos Especiais – Introdução ao estudo de metadados	Estudo de temas de caráter geral ou específico complementares às demais disciplinas do curso e relacionados com a Biblioteconomia ou com a Ciência da Informação. A disciplina compreende uma introdução geral ao estudo dos padrões de metadados aplicáveis a objetos de informação textuais, videográficos, sonoros, iconográficos, museológicos e arquivísticos. Serão, abordados, também: os padrões de metadados de interoperabilidade METS e OAI-PMH; aspectos introdutórios das linguagens HTML e XML; aspectos introdutórios do padrão CSS; aplicações de padrões de metadados em bibliotecas, arquivos, museus e repositórios institucionais.
Redes de Informação e Transferência de Dados	EMENTA: Sistemas de informações cooperativos. Redes de bibliotecas. Infraestruturas e arquitetura de redes de comunicações de dados. Protocolos de comunicação e transferência de dados. Estratégias de acesso ao documento primário. Interfaces e formatos de intercâmbio de informação.

Fonte: elaborado pelo autor, 2021.

Assim, verifica-se que o curso de Biblioteconomia oferecido pela UnB prepara seus discentes para atuarem no contexto da Biblioteconomia de Dados com enfoque temático em duas disciplinas que trabalham diretamente com habilidades e competências desse perfil profissional.

No Centro-Oeste, foram encontrados e analisados três cursos de Biblioteconomia: na UFMT, onde não foi localizado nenhuma disciplina relacionada à Biblioteconomia de Dados; na UFG, cujo PPC do curso indica uma atenção à área tecnológica relacionada à gestão de dados, mas não se encontrou nenhuma disciplina que habilitasse os discentes às novas demandas relacionadas ou que mantivesse relação com o objeto de estudo deste trabalho, ou seja, disciplinas que permitam ir conteudisticamente além das tradicionais; e, no curso da UnB, onde observou-se uma tendência à inovação, sendo localizadas ao menos duas disciplinas que abordam algumas das habilidades e competências exigidas pela Biblioteconomia de Dados.

6.2.4 Sudeste

A Região Sudeste compreende quatro estados possuindo, ao todo, quatorze cursos de Biblioteconomia analisados, os quais se encontram assim distribuídos: Rio de Janeiro (UNIRIO, UFRJ e UNIVERSO), São Paulo (FABCI, PUC-CAMPINAS, USP,

USP - Ribeirão Preto, UNESP, UFSCAR e CEUCLAR), Minas Gerais (UFMG e UNIFORMG) e Espírito Santos (UFES)

O curso de Biblioteconomia da Faculdade de Biblioteconomia e Ciência da Informação da Fundação Escola de Sociologia e Política de São Paulo (FaBCI / FESPSP) foi criado em 1940 e é um dos cursos com antigos em destaque na história da Biblioteconomia brasileira. Possui duração mínima de três anos, com 2880 horas distribuídas em seis semestres, com PPC datado de 2020.

Este PPC aponta para os profundos efeitos das inovações científico-tecnológicas na sociedade, exigindo atualizações em seus cursos, como evidenciado neste trecho: “Isto requer discernimento de novas habilidades profissionais, suscitando constantes reflexões sobre a atual formação profissional em todas as áreas do conhecimento” (FESPSP, 2020, p.9), e destaca novos campos de atuações do(a) bibliotecário(a), dentre eles a Ciência de Dados: “Atualmente, os bibliotecários têm ampliado seus espaços de atuação, caracterizam-se também, como gestores da informação, cientistas de dados, curadores de informação etc.” (FESPSP, 2020, p.10).

Quadro 30 – Termos relacionados a Bibliotecário(a) de dados na FESPSP

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	1
Pesquisa científica	1	2
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	2

Fonte: dados da pesquisa, 2021.

O termo *dados de pesquisa* apareceu na disciplina *Normalização Documentária Aplicada à Produção Acadêmica e à Pesquisa Científica*, e aqui já aborda a atuação bibliotecária no auxílio ao pesquisador no que se refere aos seus dados de pesquisa, mostrando, assim, uma relação direta com serviços prestados pelo bibliotecário(a) de dados.

Sobre o termo *pesquisa científica*, percebeu-se que ele aparece também na ementa da disciplina *Prática em Pesquisa I*, discorrendo sobre a prática da pesquisa científica a ser realizada pelos discentes sem conexão com o objeto de estudo deste trabalho.

Por fim, o termo *metadados* foi encontrado em duas ementas: a disciplina *Representação Descritiva II*, que versa sobre tipologias de metadados com ênfase nos processos descritivos, visando a representação descritiva em ambiente digital que é uma das habilidades necessárias ao bibliotecário(a) contemporâneo, incluindo os que trabalham com dados científicos; e a outra disciplina é *Modelagem de Bancos de Dados*, que trabalha com estrutura de bancos de dados e seus sistemas gerenciadores, enfatizando, assim, o conhecimento tecnológico de metadados, necessário também no leque de conhecimentos do(a) bibliotecário(a) de dados.

Quadro 31 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na FESPSP

DISCIPLINAS	EMENTAS
Normalização Documentária aplicada à produção acadêmica e à pesquisa científica	Aborda a importância da normalização no contexto do desenvolvimento acadêmico e da pesquisa científica. Destaca o contexto da atuação profissional do bibliotecário, no âmbito da realização da pesquisa científica, no que diz respeito à padronização de processos e dados de pesquisa. Apresenta os principais sistemas de normalização: Associação Brasileira de Normas Técnicas (ABNT), <i>American Psychological Association</i> (APA) e Vancouver. Instrui para o uso e para a aplicação das principais normas da ABNT para a área de Informação e Documentação, com enfoque na reflexão crítica para a utilização das convenções estabelecidas.
Prática em Pesquisa I	Retomada de pontos significativos em Metodologia da Pesquisa Científica voltados para aplicação. Prática de definição e descrição de um problema e objeto de pesquisa, abordagens e respectivos instrumentos, a partir da temática da Biblioteconomia e Ciência da Informação enquanto Ciência Social Aplicada, e voltados para o enfrentamento de problemas, previamente delimitados, abrangendo situações encontradas na região metropolitana de São Paulo. Estabelecimento de programa de orientação discente e gerenciamento de calendário de atividades de pesquisa aplicada.

Representação Descritiva II	Estudo teórico e prático da Representação Descritiva no ambiente digital. Tendências da catalogação e dos catálogos e suas relações com a recuperação da informação na <i>web</i> . Tipologias de metadados com ênfase nos metadados descritivos. Criação de registros em sistemas automatizados de informação, importação de registros e trabalho em rede. Catálogos bibliográficos, catálogos de autoridade digitais e recuperação da informação. Articulação com as disciplinas: Representação Temática I e Representação Temática II.
Modelagem de Bancos de Dados	Estudo dos conceitos de metadados, estrutura de bancos de dados e seus sistemas gerenciadores. Princípios de banco de dados relacional e não-relacional. Estudo de modelos e modelagem de dados (Entidade-relacionamento), normalização e estrutura de dados. Introdução a linguagem SQL (<i>Structured Query Language</i> / Linguagem de Consulta Estruturada). <u>Articulação com as disciplinas:</u> Arquitetura e Design de Informação e Informatização de Serviços de Informação.

Fonte: elaborado pelo autor, 2021.

Logo, a partir da análise de seu PPC, pode-se constatar uma atualização recente no curso de Biblioteconomia da FESPSP, e em ao menos quatro disciplinas, percebe-se uma abordagem objetiva com as habilidades e competências necessárias ao bibliotecário(a) de dados, apontando que o curso prepara o futuro profissional para atuar nesse novo perfil da profissão ligado à *e-Science*.

O curso de bacharelado em Biblioteconomia da Universidade Federal do Estado do Rio de Janeiro (UNIRIO), assim como o curso da FESPSP, faz parte da história da Biblioteconomia do Brasil e do mundo, já que este foi o primeiro curso criado no Brasil, em 1911, oferecido então pela BN, e que em 1969, passa a integrar a Federação das Escolas Isoladas do Estado da Guanabara (FEFIEG), federalizando-se em 1979, e tornando-se UNIRIO (UNIRIO, 2009a).

O PPC deste curso é de 2010, produzido por meio de pesquisa junto à literatura oferecida pela ABEBD, visando harmonizar com demais cursos do MERCOSUL, sempre tendo com norte um caráter humanista (UNIRIO, 2010). O curso é oferecido em dois turnos, diurno e noturno, com carga horária total de 3.095 horas e 200 vagas anuais (100 diurnas e 100 noturnas). A diferença entre o curso diurno e o noturno é apenas da distribuição das disciplinas, uma vez que o diurno tem um período mínimo de duração de oito períodos e o noturno de nove, por este motivo serão analisados como sendo iguais em formação (UNIRIO, 2009a).

Quadro 32 – Termos relacionados a Bibliotecário(a) de dados na UNIRIO (bacharelado)

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	1	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	1
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

No curso de bacharelado da UNIRIO, diferente do curso da FESPSP não foram encontradas disciplinas que tratem diretamente de características necessárias ao bibliotecário(a) de dados. Dentre os termos elencados para pesquisa, o único encontrado foi *preservação digital*, na disciplina *Políticas de Preservação de Acervos Bibliográficos*, utilizado no sentido de conservação de acervos bibliográficos e não de dados de pesquisa científica.

Quadro 33 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UNIRIO

DISCIPLINAS	EMENTAS
Políticas de Preservação de Acervos Bibliográficos	Apresentação dos conceitos, princípios envolvidos nas políticas de preservação de acervo bibliográfico. Conceitos e princípios do patrimônio bibliográfico. Políticas de preservação do patrimônio bibliográfico. Noção do instrumental básico para a identificação dos problemas de preservação do acervo bibliográfico e para o planejamento de ações de conservação. Preservação digital. A política de preservação e o desenvolvimento de coleções.

Fonte: elaborado pelo autor, 2021.

Em 1986, disciplinas de licenciatura foram inseridas no curso de Bacharelado em Biblioteconomia da UNIRIO com carga horária de 840 horas, e durou regularmente

até 1991. O curso de licenciatura em Biblioteconomia se efetiva em 2009, até hoje sem atualizar seu PPC. Atualmente possui carga horária total de 3.405 horas no período noturno com 80 vagas anuais (UNIRIO, 2009b)

No caso desse curso, em nenhuma das disciplinas foram encontradas as palavras-chave elencadas. Este curso tem como objetivo a formação docente, no entanto, diante do cenário em que a *e-Science* e Ciência de Dados estão cada vez mais consolidados, esperava-se que houvesse alguma indicação de estudo do tema. Mas seus conteúdos estão direcionados para as práticas de licenciatura e as disciplinas de formação tradicional.

O curso de bacharelado em Biblioteconomia da Universidade Federal do Rio de Janeiro (UFRJ) oferece o mínimo de 3.300 horas. O seu PPC é de 2008, informando que o curso inicia o ano oferecendo 30 vagas e tem entre seus objetivos tratar tecnicamente a informação registrada, em qualquer tipo de suporte, por meio do desenvolvimento e da utilização de novas tecnologias (UFRJ, 2008).

Quadro 34 – Termos relacionados a bibliotecário(a) de dados na UFRJ

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	1
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	1
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	1

Fonte: dados da pesquisa, 2021.

Foram encontrados três dos termos elencados com a temática do(a) bibliotecário(a) de dados, na UFRJ. O termo *pesquisa científica* está presente na disciplina *Metodologia da Pesquisa*, porém, sem relação com *e-Science*, trata de noções sobre conhecimento, método e pesquisa.

Outro termo encontrado foi *ciclo de vida*, na disciplina *Gerenciamento Eletrônico de Documentos*, porém relacionado a documentos digitais, dando abertura ao docente para abordar sobre o CVD.

O último termo encontrado foi *metadados*, na disciplina *Indexação e Resumo*, visando a indexação de documentos eletrônicos via metadados que, mesmo indiretamente, capacita, por meio da representação da informação, o bibliotecário(a) a atuar com dados de pesquisa científica.

Quadro 35 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFRJ

DISCIPLINAS	EMENTAS
Metodologia da Pesquisa	Pesquisa científica: voltada para a compreensão e explicação da realidade. Pesquisa pragmática: voltada para a ação imediata. Etapas da pesquisa. Técnicas de seleção: tipos de amostragem. Técnicas de observação: observação ordinária; observação participativa; questionário escrito; entrevista; sondagens de opinião. Técnicas de descrição: técnicas qualitativas e quantitativas; escalas de atitude e de opinião; análise de conteúdo. Técnicas de análise: elaboração de hipóteses; mensuração de variáveis; operacionalização de conceitos; análise e explicação de fenômenos e sua generalização empírica. Redação e apresentação de relatório de pesquisa.
Gerenciamento Eletrônico de Documentos	O conhecimento e a informação nas organizações. Ciclo de vida dos documentos. Conceituação de gerenciamento eletrônico de documentos - GED. Justificativas para a implementação do GED. Tecnologias correlatas ao GED: ferramentas de <i>hardware e de software</i> .
Indexação e Resumo	Teorias de indexação. A indexação como processo de representação temática da informação e sua interação com o processo de recuperação. Princípios e mecanismos básicos, características, linguagens e medidas. A indexação manual e automática. O processo de indexação por metadados e a recuperação eletrônica da informação. O resumo como processo de representação temática do documento: tipos e funções. A redação do resumo. O sistema de avaliação da indexação e do resumo.

Fonte: elaborado pelo autor, 2021.

Isto posto, é possível verificar que no curso de Biblioteconomia da UFRJ, considerando que seu PPC não está atualizado há mais de dez anos, não aborda disciplinas direcionadas à capacitação de um(a) bibliotecário(a) de dados, porém possui duas disciplinas que o docente pode realizar a abordagem, se assim optar.

O curso de bacharelado em Biblioteconomia da Universidade Salgado de Oliveira (UNIVERSO), iniciado em 2014, realizado na modalidade à distância, possui polos nas Regiões Sudeste, Centro-Oeste e Nordeste do Brasil. Com duração mínima de quatro anos e carga horária de 2.655 horas, seu PPC é de 2018, organizado em parceria entre o corpo docente e a Gestão Acadêmica do EaD (UNIVERSO, 2018). Vale ressaltar que a UNIVERSO se trata de “[...] uma associação civil de direito privado, com finalidade educacional, cultural, assistencial, social, filantrópica, sem fins lucrativos [...]” (UNIVERSO, 2018, p. 8).

O egresso do Curso de Biblioteconomia tratará a informação independente do suporte informacional e a tornará acessível ao usuário final. Por isso, precisa ter a capacidade de disponibilizar informação contida em vários tipos de instrumentos, gerenciar unidades, redes e sistemas de informação, tratar tecnicamente e desenvolver recursos informacionais, prestar serviços de assessoria e consultoria, realizar difusão cultural e desenvolver ações educativas. (UNIVERSO, 2018, p. 68)

Este PPC descreve seu egresso de forma ampla, abordando habilidades e competências necessárias ao bibliotecário(a) de dados:

Quadro 36 – Termos relacionados a Bibliotecário(a) de dados na UNIVERSO

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	1
Open data	0	0
Preservação digital	0	1
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	2

Fonte: dados da pesquisa, 2021.

Foram encontrados quatro termos em ementas desse curso: *dados abertos* na disciplina *Fontes de Informação II*, que busca compreender os dados abertos como fonte de informação, compreensão esta necessária ao bibliotecário(a) de dados; *preservação digital* na disciplina *Preservação de Documentos*, que traz o conceito de preservação digital também necessário para um(a) bibliotecário(a) de dados; *metadados*, nas ementas de duas disciplinas diferentes: *Representação Descritiva III*, utilizando metadados na catalogação de acervo bibliográfico, não tratando diretamente com dados científicos, mas podendo ser usado em analogia; e na disciplina *Publicações Digitais*, a qual trata de diversos tipos de publicações digitais, não especificamente sobre dados de pesquisa científica, porém deixa margem para tratar desse assunto que está intimamente ligado às habilidades e às competências do(a) bibliotecário(a) de dados.

Quadro 37 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UNIVERSO

DISCIPLINAS	EMENTAS
Fontes de Informação II	Apresentar e compreender as principais fontes de informação da área educacional, empresarial, saúde e jurídica. Assim como possibilitar ao aluno entender a utilização das patentes e dos dados abertos como fonte de informação.
Preservação de Documentos	Preservação de documentos: conceitos e referenciais teóricos. O acervo: degradação, fatores internos e externos. Políticas e estratégias de preservação. Rotinas de atividades no ambiente de trabalho a serem evitadas. Deterioração física e química do acervo. Medidas preventivas. Preservação digital: conceito. Estratégias de acesso a informação. Perspectivas da preservação de acervos bibliográficos.
Representação Descritiva III	Pontos de acesso: terminologia, regras, entradas principais e secundárias. Remissivas: conceito, nomes pessoais, geográficos, entidades coletivas, títulos uniformes. Formato Marc: estrutura, campos, subcampos e indicadores. Compartilhamento de informações bibliográficas automação de acervos. RDA (<i>Resource Description and Access</i>): e principais características e diferenças com o AACR. Metadados no domínio bibliográfico. Organização do conhecimento em ambiente eletrônico.

Publicações Digitais	Informação Digital. A Web Semântica. Ontologia. As bibliotecas digitais. O profissional bibliotecário. A arquitetura da informação. Os Metadados. O Perfil de aplicação. O serviço de referência digital. Os livros digitais. Os aspectos jurídicos da informação digital. Os aspectos éticos da informação digital. Os Desafios dos profissionais da informação.
----------------------	---

Fonte: elaborado pelo autor, 2021.

Dessa forma, o curso de bacharelado em Biblioteconomia da UNIVERSO, embora não se aprofunde no objeto de estudo dessa dissertação, traz referência em suas ementas que remetem a um arcabouço teórico razoável para que seus egressos possam adentrar à Biblioteconomia de Dados e estejam aptos a buscar a teoria necessária para trabalhar na área e, assim, adquirir as habilidades e as competências para atuar no contexto da *e-Science*.

No caso da Pontifícia Universidade Católica de Campinas (PUC-Campinas), o curso de bacharelado em Biblioteconomia foi criado em 1964, advindo da Escola de Biblioteconomia, fundada em 1945. A atualização curricular deste curso é uma prática recorrente e o PPC atual é de 2019, com 2.856 horas/aula ou 2.524 horas/relogio, oferecendo 60 vagas anualmente (PUC-CAMPINAS, 2019).

Ainda sobre esse PPC, há a constatação da inserção na matriz curricular de novos conceitos, teorias e práticas contemporâneas no intuito de “abordar a informação enquanto fenômeno capaz de gerar novas percepções na mente do indivíduo” (PUC-CAMPINAS, 2019, p. 38), bem como, ao se referir ao cenário das organizações, “num mundo globalizado e competitivo as quais precisam fazer uso de conhecimentos, ferramentas e tecnologias diversificados tendo em vista a eficácia de seus processos organizacionais, operacionais e produtivos” (*idem, Ibidem*). Assim, revelando disponibilidade à visão de nichos novos ou emergentes, porém não indica, formalmente, nenhum destes ligado Biblioteconomia de Dados.

Entre os objetivos do curso, encontra-se a gestão voltada para a “identificação das necessidades e demandas informacionais, tratamento, organização, análise, mensuração e monitoramento da informação científica, tecnológica” (PUC-CAMPINAS, 2019, p. 52), além do incentivo ao “[...] uso das tecnologias de informação e comunicação no desenvolvimento de produtos e serviços de informação” (*Ibidem*, p. 53), reforçando a abertura para o objeto de estudo deste trabalho.

Quadro 38 – Termos relacionados a bibliotecário(a) de dados na PUC-Campinas

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	1	1
Reúso de dados	0	0
Colaborativo	0	0

Fonte: dados da pesquisa, 2021.

Dos termos selecionados, foi encontrado apenas o *colaborativo*, na disciplina *Software Colaborativos*, mas relacionado ao uso de ferramentas de códigos abertos para uso na construção de espaços colaborativos. Aqui há a possibilidade de o docente abordar a *e-Science*.

Quadro 39 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na PUC-Campinas

DISCIPLINAS	EMENTAS
SOFTWARE COLABORATIVOS	Estudo teórico-prático de ferramentas abertas para construção de espaços colaborativos, visando o armazenamento e o mapeamento de informações em contextos variados.

Fonte: elaborado pelo autor, 2021.

Conquanto o PPC do curso de bacharelado em Biblioteconomia da PUC-Campinas esteja completando dois anos de reformulação, e possua na descrição do curso atenção ao uso de tecnologias da comunicação e informação no contexto contemporâneo, não foram encontradas disciplinas que abordem diretamente a atuação bibliotecária no contexto da *e-Science*.

Ao analisar o PPC do curso de Biblioteconomia oferecido pela Escola de Comunicações e Artes (ECA) da USP, este se diferenciou dos demais por manter uma

estrutura desde a fundação, com um currículo pautado na corrente europeia, com disciplinas mais voltadas à ação cultural. O Projeto Pedagógico⁵⁸ atual visa

[...] fornecer as competências e habilidades necessárias ao profissional para atuar tanto no âmbito da memória escrita, da cultura e dos processos de organização da informação tradicionais, como na construção de conhecimentos multimídia em formatos digitais distribuídos em rede, garantindo sua inserção na vanguarda da sociedade contemporânea conectada, com suas novas lógicas de produção, circulação, distribuição e apropriação da informação. (USP, [2013], p. 2).

Em seu texto, ressalta-se a importância da instituição à sociedade com a ponderação: “[...] cabe à USP manter-se na fronteira do conhecimento e colocar o resultado desse trabalho a serviço da população” (USP, [2013], p.3) e possui o intuito de capacitar o discente com habilidades de planejar, implementar e gerir redes, sistemas, unidade e recursos de informação, além de identificar, analisar e traduzir necessidades informacionais em contextos sociais específicos (USP, [2013], p.6).

Neste PPC, constam as disciplinas oferecidas, porém não o ementário, dessa forma, buscou-se os dados no sítio institucional do curso⁵⁹.

Quadro 40 – Termos relacionados a bibliotecário(a) de dados na USP

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	1
Reúso de dados	0	0
Colaborativo	0	1
Curadoria	0	1
Metadados	1	1

Fonte: dados da pesquisa, 2021.

⁵⁸ O documento usado para análise, nesse caso, é o Projeto Político Pedagógico (PPP), acredita-se que feito em 2013 e, por isso, a nomenclatura está desatualizada.

⁵⁹ Disponível em:

<https://uspdigital.usp.br/jupiterweb/listarGradeCurricular?codcg=27&codcur=27501&codhab=4&tipo=N>
Acesso em 31 mar. 2021.

Embora o documento possua mais de sete anos, foram encontradas quatro disciplinas que utilizam os termos elencados por este estudo. O termo *gestão de dados* usado na ementa da disciplina *Inteligência Artificial no Tratamento da Informação Registrada*, que se refere ao iminente uso da Inteligência Artificial nos produtos e serviços oferecidos pelo Bibliotecário, incluindo gestão de dados, que estão ligadas às habilidades pautadas. Outro termo localizado foi *colaborativo*, na disciplina *Documentação Audiovisual*, que discute sobre documentos em suporte audiovisual, entre eles, os localizados de forma colaborativa na *web*, não havendo relação com habilidade e competências do(a) bibliotecário(a) de dados.

Foi encontrado também o termo *curadoria* na disciplina *Serviços ao Usuário*, fazendo relação com serviços de referência, dentre eles, curadoria digital, que pode abordar uma ou mais fases do CVD e está diretamente ligada aos serviços oferecidos pelo bibliotecário(a) de dados.

O último termo identificado foi *metadados*, tanto no nome da disciplina quanto em sua ementa: *Catálogo Descritiva: introdução aos metadados Descritivos*, que visa proporcionar aos discentes de Biblioteconomia o conhecimento e aplicação dos formatos de metadados descritivos utilizados para representação e/ou codificação de conteúdos digitais, outra habilidade estritamente necessária ao bibliotecário(a) de dados.

Quadro 41 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na USP

DISCIPLINAS	EMENTAS
-------------	---------

Inteligência Artificial no Tratamento da Informação Registrada	Inteligência Artificial (IA) é um campo que tem longa história, mas que prossegue crescendo e mudando. Principalmente, segue influenciando diversas áreas do conhecimento, incluindo a Biblioteconomia e Ciência da Informação. Neste aspecto, a competência de pensar os serviços e produtos, bem como os processos de intercambiar e comunicar ideias, informações e conhecimento continuará a ser importante. Entretanto, o bibliotecário(a) do futuro terá de saber mais do que isso. Precisarão interagir com sistemas inteligentes de automação, análise de conteúdo e gestão de dados e informações. A mudança, na forma como os serviços de informação bibliográficos são prestados, já ocorre e é irreversível. Usuários querem serviços rápidos e consistentes. Assim, considerando o cenário, pretende inserir o estudante nos princípios teóricos e conceituais da inteligência artificial e nos seus principais processos de aplicação dentro do ambiente dos serviços bibliotecários, em especial nos processos catalográficos. Permitir também o conhecimento de algumas das principais pesquisas em desenvolvimento. Tem também por objetivo motivar a Iniciação Científica, Trabalhos de Conclusão de Curso e fortalecer a ligação da graduação com a pós-graduação do departamento, incentivando alunos de graduação a participarem.
Documentação Audiovisual	Discutir as especificidades da documentação audiovisual com relação aos seus formatos e suportes, necessidades de conservação e preservação e as novas formas para representação e acesso possibilitadas pelas novas tecnologias de comunicação e informação. Ampliar a visão sobre as possibilidades de organização, tratamento e uso dos documentos audiovisuais em sistemas de informação colaborativos na web e sua inserção no campo das Humanidades Digitais. Enfatizar a especificidade da informação neles contida e a necessidade da conservação e preservação física e digital dos referidos documentos.
Serviços ao Usuário	1) Apresentar e discutir fundamentos relacionados ao tema “Serviços ao Usuário” no âmbito dos Serviços de Informação visando condições objetivas para os alunos refletirem sobre os desafios atuais e perspectivas futuras para os Serviços de Referência em dispositivos de informação. 2) Discutir o "estado da arte" do assunto, em contextos nacionais e internacionais. 3) Completar o estudo básico, incluindo discussão de conhecimentos interdisciplinares na área de Psicologia; Comunicação Visual e Marketing;

	Curadoria Digital.
Catálogo Descritiva: Introdução aos Metadados descritivos	Introduzir os alunos de Biblioteconomia no conhecimento e aplicação dos formatos de metadados descritivos utilizados para representação e/ou codificação de conteúdos digitais.

Fonte: elaborado pelo autor, 2021.

Dessa forma, a despeito do documento ser, provavelmente, de 2013, quando os assuntos sobre dados científicos no contexto da *e-Science* ainda eram insipientes, o curso oferece ao menos três disciplinas que abordam diretamente habilidades e competências do(a) bibliotecário(a) de dados.

O curso de Biblioteconomia e Ciências da Informação e da Documentação, da Faculdade de Filosofia Ciências e Letras de Ribeirão Preto (FFCLRP), um campus da USP, teve seu primeiro Projeto Pedagógico aprovado em 18 de fevereiro de 2002, passando a ser oferecido em 2003, e o atual aprovado em 2016. Exige o comprimento mínimo de 2.970 horas, em no mínimo oito semestres, entre disciplinas obrigatórias e optativas.

Quadro 42 – Termos relacionados a Bibliotecário(a) de dados na USP - Ribeirão Preto

Termos	Nome das Disciplinas	Ementas
dados de pesquisa	0	0
pesquisa científica	0	1
Ciência de Dados	0	0
<i>Data Science</i>	1	0
ciclo de vida	0	0
dados abertos	0	0
<i>open data</i>	0	0
preservação digital	0	0
gestão de dados	0	0
reúso de dados	0	0
colaborativo	0	0
Curadoria	0	0
metadados	0	0

Fonte: dados da pesquisa, 2021.

Nessa instituição, foram encontradas duas disciplinas com os termos sugeridos por esta pesquisa, a *Metodologia Científica: processos e técnicas de pesquisa*, na qual apresenta-se o termo *pesquisa científica* em sua ementa, ponderando de forma genérica

as etapas da pesquisa científica e abordagem de processos, métodos e técnicas metodológicas, ficando subjetivo ao docente abordar ou não sobre a forma de pesquisa colaborativa através da *e-Science*.

A outra disciplina é a *Data Science para Humanas, Introdução à análise de dados, modelos e algoritmos de Machine Learning* que aborda a Ciência de Dados, principalmente, no contexto da indústria 4.0⁶⁰, a qual não necessariamente está ligada aos dados científicos acadêmicos, mas com forte ligação à *e-Science* e à Biblioteconomia de Dados, logo essa disciplina deve trazer uma introdução do tema ao bibliotecário(a) egresso desse curso.

Quadro 43 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na USP - Ribeirão Preto

DISCIPLINAS	EMENTAS
Metodologia Científica: Processos e Técnicas de Pesquisa	Estudo dos processos, métodos e técnicas, bem como das etapas de realização de pesquisa científica.
<i>Data Science para Humanas, Introdução a análise de dados, modelos e algoritmos de Machine Learning</i>	Discutir os conceitos e técnicas introdutórias para o entendimento do contexto da Indústria 4.0 e da Ciência dos Dados, e como elas podem impactar diretamente na vida das pessoas, em ambientes culturais, nos aspectos jurídicos e nas tomadas de decisão em negócios. Demonstrar as características e procedimentos do processo de Análise de Dados, <i>Big data</i> e os conceitos e recursos de algoritmos de <i>Machine Learning</i> .

Fonte: elaborado pelo autor, 2021.

Percebe-se neste PPC relativamente novo, aprovado em 2016, uma disciplina diretamente ligada à *e-Science* e ao bibliotecário(a) de dados, com forte ênfase nas habilidades e competências necessárias, todavia mais direcionadas ao mercado de trabalho fora do contexto acadêmico, configurando-se numa disciplina em que pode abordar o tema.

O curso de Biblioteconomia da UNESP iniciou suas atividades no campus de Marília, em 06 de abril de 1977, e foi reconhecido pelo MEC em 11 de fevereiro de 1981. Desde sua criação até o final da década de 1990, o curso atualiza periodicamente seus projetos pedagógicos seguindo as evoluções da profissão. Na primeira década de 2000, houve reformulações nas diretrizes curriculares, com evidente aumento das TIC.

⁶⁰ Indústria 4.0, também conhecida como Quarta Revolução Industrial, visa melhorar a eficiência e produtividade de seus processos se valendo principalmente da troca de dados através da ciberinfraestrutura, internet das coisas e computação em nuvem.

Atualmente, o curso oferece trinta e cinco vagas por ano, com carga horária total de 2.670 horas, em turno diurno.

Quadro 44 – Termos relacionados a Bibliotecário(a) de dados na UNESP

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	1	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	1	2
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	1
Metadados	1	4

Fonte: dados da pesquisa, 2021.

Este PPC possui mais de nove anos, ainda assim foram encontradas sete disciplinas contendo quatro dos termos pesquisados: *pesquisa científica* encontrado no título da disciplina *Metodologia da Pesquisa Científica*, que trata da elaboração de projetos de pesquisa, tendo sua ementa mais genérica e dando possibilidade ao docente a introdução à *e-Science*. A segunda disciplina, *Automação de Unidades de Informação*, contempla um dos termos pesquisados em sua ementa, *preservação digital*, abordando a automação de biblioteca e a gestão eletrônica de documentos, alcançando a preservação digital, mas não está diretamente ligada ao tema deste trabalho. Também foi identificada uma disciplina com um dos termos em seu título e em sua ementa: *Preservação Digital*, possuindo uma abordagem mais ampla, que traz intrínseca uma melhor conceituação e o uso da preservação digital, abarcando inclusive outro termo, *metadados*, ao indicar análise e aplicação de metadados de preservação digital.

A quarta disciplina, *Cultura Digital*, possui em sua ementa o termo *curadoria*, ao destacar curadorias de acervos em ambientes digitais. Nessa disciplina, o docente potencialmente pode trabalhar diretamente com uma das habilidades necessárias ao bibliotecário(a) de dados, podendo abordar sobre CVD, por tratar -se de uma de suas fases.

A disciplina *Catálogo*, a quinta a possuir um dos termos, *metadados*, aborda a descrição de metadados para catalogação de acervos, não tratando diretamente com dados científicos, mas podendo, por analogia, alcançar esse assunto.

Na sexta disciplina encontrada, há o termo *metadados* tanto no título quanto na ementa: *Metadados de Objetos Digitais*, indicando um aprofundamento no tema, o que se aplica perfeitamente aos dados científicos, portanto, estando diretamente ligada à *e-Science* e à formação do(a) bibliotecário(a) de dados.

Por fim, a disciplina *Web Semântica* também utiliza o termo *metadados*, discutindo a interoperabilidade através do uso de elementos dos metadados na *web* semântica, por analogia, é compreendido que seu uso contempla os dados científicos, contribuindo com as habilidades e competências necessárias ao bibliotecário(a) de dados.

Quadro 45 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UNESP

DISCIPLINAS	EMENTAS
Metodologia da Pesquisa Científica	Elaboração de projeto de pesquisa: natureza e objetivos. Definição do objeto de estudo. Estudo de linhas de pesquisa.
Automação de Unidades de Informação	Revisão de conceitos de Informática. Histórico da automação de bibliotecas. O processo de Automação e Gerenciamento Eletrônico de Documento (GED). Web Semântica. Certificação e preservação Digital. Repositórios institucionais e <i>Open Archives</i> .
Preservação Digital	Conceituação de preservação digital. Apresentação dos elementos para preservação do acesso a longo prazo de informações digitais. Estudo sobre as estratégias de preservação digital. Análise e aplicação de metadados de preservação digital.
Cultura Digital	Apresentação da cultura contemporânea e modelos de redes sociais digitais e de compartilhamento. Estudos e avaliações sobre coletivos de geração e produção de informações. Acervos, coleções, fundos e curadorias em ambientes digitais. Design visual, estrutura e organização de ambientes digitais.
Catálogo	Estudo das teorias, princípios, esquemas, estruturas e formatos de descrição de recursos informacionais e acesso. Análise de sistemas e práticas que atendam às necessidades do usuário, de forma prática e ética. Experiência, em nível inicial, com instrumentos e códigos de catalogação

	para a descrição de metadados e esquemas de codificação, escolha de pontos de acesso e controle de autoridade.
Metadados de Objetos Digitais	Estudo dos princípios e aplicação de metadados para a organização de recursos informacionais em rede a partir da concepção de esquemas de metadados específicos para atender as demandas informacionais da comunidade. Instruções sobre a aplicação e implementação de esquemas de metadados. Criação de registros de metadados, analisando o uso de elementos de metadados.
Web Semântica	Conceituação e evolução da <i>World Wide Web</i> (Web). Definição de Web Semântica. Fundamentação básica da Web Semântica. Estudos sobre os níveis de interoperabilidade, os conjuntos de recursos, os elementos de metadados, esquemas de codificação e linguagens de marcação.

Fonte: elaborado pelo autor, 2021.

Dessa forma, observa-se a inserção do objeto de estudo deste trabalho em seis disciplinas do curso de graduação de Biblioteconomia da UNESP, mesmo seu PPC sendo do ano de 2012, quando os debates sobre *e-Science* e dados científicos eram emergentes. Infere-se que, pelas análises feitas, este curso preparava, e prepara, o egresso em Biblioteconomia para atuar como bibliotecário(a) de dados.

Criado em 1994, com docentes oriundos das áreas de Letras, Ciências Sociais e Metodologia de Ensino, o curso de Biblioteconomia e Ciência da Informação, da Universidade Federal de São Carlos (UFSCar), tem seu PPC estabelecido em 2012 e atualizado em 2014, ofertando 48 vagas anuais, com turno de funcionamento noturno, e carga horária total 2.880 horas.

Quadro 46 – Termos relacionados a Bibliotecário(a) de dados na UFSCar

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	1	1
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	1
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0

Colaborativo	0	2
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

O termo *pesquisa científica* aparece na ementa da disciplina *Introdução ao Trabalho Científico*, que traz enfoques introdutórios à realização de trabalhos científicos direcionados à estrutura e à apresentação de pesquisas científicas e trabalhos acadêmicos, normalização e apresentação de trabalhos científicos. Aparece também no título da disciplina *Metodologia da Pesquisa Científica* para BCI⁶¹, que aborda o trabalho científico mais voltado para etapas e metodologias de pesquisa do trabalho científico. Ambos os casos tratam a pesquisa científica de forma introdutória e, provavelmente, não abordam a *e-Science*.

Outro termo encontrado é o *ciclo de vida*, porém sem relação direta com dados científicos. O termo foi localizado na disciplina *Introdução à Análise de Sistemas*, cuja ementa discorre sobre ciclo de vida de sistemas de informação, apoiado em análise e projeto estruturado.

Já o termo *colaborativo* foi encontrado em duas ementas, a saber: na disciplina *Linguagens Documentárias III*, que versa, entre outros assuntos, sobre redes sociais e construção de sistemas colaborativos para *web*; e na disciplina *Gestão da Informação e Gestão de Redes de Pessoas e Organizações*, em que é utilizado para indicar repositórios digitais abertos e interoperáveis e recursos interativos e colaborativos da *web*, diretamente ligado à *e-Science* e algumas habilidades e competências do(a) bibliotecário(a) de dados.

Quadro 47 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFSCar

DISCIPLINAS	EMENTAS
Introdução ao trabalho científico	Estudos introdutórios dos diferentes enfoques e vertentes teóricos adotados na pesquisa científica. Estrutura e apresentação de pesquisa científica e trabalho acadêmico, normalização e apresentação de trabalhos científicos.
Metodologia da pesquisa científica para BCI	Estudo das principais etapas do trabalho científico com ênfase nas metodologias de pesquisa em Biblioteconomia e Ciência da informação. Elaboração de projeto de pesquisa.

⁶¹ Biblioteconomia e Ciência da Informação

Introdução à análise de sistemas	Abordagem sistêmica e organização, sistemas de informação: conceitos, ciclo de vida de sistemas de informação apoiado em análise e projeto estruturado, sistemas de gerenciamento, operação e tomada de decisão, noções de planejamento.
Linguagens documentárias III	Websemântica e os aspectos interdisciplinares na representação de recursos informacionais. Apresentação e caracterização das linguagens de representação na <i>web</i> . As ontologias e as folksonomias: conceitos, estruturas e aplicações na representação e recuperação da informação na <i>web</i> . As redes sociais como ambientes colaborativos na construção e atualização de linguagens de representação na <i>web</i> . Inovações em linguagens de representação na <i>web</i> e em outros ambientes informacionais.
Gestão da informação e gestão de redes de pessoas e organizações	Conceitos e fundamentos do ciclo informacional. Conceitos, propósitos e importância da rede. Uso e acesso da informação científica e tecnológica. Redes de informação e sua influência no planejamento e uso de fontes de informação. Repositórios digitais abertos e interoperáveis. Recursos interativos e colaborativos da <i>web</i> . Cooperação e intercâmbio entre unidades de informação. Clusters de organizações e competitividade. Mapeamento de redes para o compartilhamento de informações e conhecimento.

Fonte: elaborado pelo autor, 2021.

Dessa forma, se verifica cinco disciplinas em que constam os termos pesquisados, porém, apenas uma oferece amplo espaço para discussão sobre algumas habilidades e competências do(a) bibliotecário(a) de dados, bem como o contexto da *e-Science*, e em duas disciplinas, caso seja de iniciativa do docente, é possível também introduzir o tema.

O Curso de bacharelado de Biblioteconomia e Gestão da Informação da UFMG oferece o curso nos turnos diurno e noturno, num total de 122 vagas anuais, oferecidos no período da manhã, tarde ou noite, com seu PPC datado de 2008. Para finalizar o curso, precisa-se integralizar 2.400 horas/aula (UFMG, 2008).

Quadro 48 – Termos relacionados a Bibliotecário(a) de dados na UFMG

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	0
Ciência de Dados	0	0

Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	1

Fonte: dados da pesquisa, 2021.

Dos termos procurados, foi encontrado somente um, *metadados* na disciplina *Representação da Informação Digital*, mas que foi utilizado como padronização de formatos de registros bibliográficos para internet, não tratando diretamente sobre representação de dados científicos.

Quadro 49 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFMG

DISCIPLINAS	EMENTAS
Representação da informação digital	Bibliotecas digitais. Representação do conhecimento no contexto digital. Hipertexto. Categorização, taxonomias e ontologias. A Web semântica. Uso e a adequação de formatos de registro bibliográfico e de normas para tratamento de recursos informacionais da Internet (padrões de metadados).

Fonte: elaborado pelo autor, 2021.

Dessa forma, o curso de bacharelado de Biblioteconomia e Gestão da Informação da UFMG não possui nenhuma disciplina direcionada diretamente à organização, ao acesso e ao uso de dados científicos, apenas às práticas tradicionais que podem ser, por analogia, utilizadas para este fim.

O curso de bacharelado em Biblioteconomia da Centro Universitário de Formiga (UNIFORMG) oferece 45 vagas para ensino presencial com duração mínima de 3 anos com no mínimo 2.400 horas, com PPC datado de 2018 (UNIFORMG, 2018).

Quadro 50 – Termos relacionados a Bibliotecário(a) de dados na UNIFORMG

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	1

Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	1
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

Nessa busca, foram encontradas apenas duas disciplinas contendo algum dos termos elencados. A primeira é a *Metodologia Científica*, que contém na ementa a expressão *pesquisa científica*, mas sem nenhuma indicação do ensino sobre *e-Science*; e a segunda é a Conservação e Preservação do Acervo, com o termo *preservação digital* também sem indicar relação com os dados científicos, somente com outros tipos de documentos digitais.

Quadro 51 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UNIFORMG

DISCIPLINAS	EMENTAS
Metodologia Científica	A natureza da ciência e os níveis do conhecimento. O registro do conhecimento científico. A pesquisa científica.
Conservação e Preservação do Acervo	Noções básicas sobre conservação e preservação de documentos. O papel como suporte da escrita, sua composição, fatores de degradação intrínsecos e extrínsecos. Estrutura, problemas, soluções e processos de reparos em livros. Técnicas de higienização e reestruturação do suporte papel. Métodos de armazenamento e práticas de manuseio de livros, folhetos, fotografias. Preservação digital. Educação ambiental.

Fonte: elaborado pelo autor, 2021.

Logo, verifica-se que o curso de bacharelado de Biblioteconomia da UNIFORMG não indica que seus discentes terão a abordagem de habilidades e competências para a atuação enquanto bibliotecário(a) de dados.

O curso de bacharelado em Biblioteconomia da Universidade Federal do Espírito Santo (UFES) é integralizado com o total de 2.490 horas/aula, possui 80 vagas

anuais e funcionamento noturno. Apresenta, entre seus objetivos específicos o de desenvolver e utilizar novas tecnologias de informação e comunicação, demandadas na contemporaneidade, ao promover programas de competência em informação (UFES, 2018, p. 6).

Tendo em vista o alcance de tais missões o curso de Biblioteconomia/Ufes, atualmente opera com dois grupos de pesquisa registrados no CNPQ: 1) Educação e Trabalho em Arquivologia e Biblioteconomia no qual integra as seguintes linhas de pesquisa: Formação e Atuação de Arquivistas e Bibliotecários; Mineração de Dados Educacionais; Organização de Documentos Multimídia; Sociedade, Informação e Cultura(s); Mediação, compartilhamento e apropriação da informação; Competência em informação em variados ambientes profissionais da informação; Biblioteca escolar e 2) Memória e Tecnologia tendo como linha de pesquisa: Memória e Tecnologia. (UFES, 2018, p. 25).

É importante ressaltar a existência de linhas de pesquisa e extensão, vinculadas ao curso, voltadas para tecnologias e que abrangem algumas habilidades e competências do(a) bibliotecário(a) de dados:

Quadro 52 – Termos relacionados a bibliotecário(a) de dados na UFES

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	2
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

Nessa instituição, foram encontradas duas disciplinas que possuem o termo *pesquisa científica* em suas ementas: *Estatística Aplicada*, cuja ementa integra os procedimentos estatísticos à pesquisa científica, e, conforme visto, estatística é uma das competências do(a) cientista de dados, porém o docente responsável por essa disciplina não necessariamente pode associar o seu uso tendo em vista a atuação do discente como

bibliotecário(a) de dados; outra disciplina que usa o mesmo termo é *Tópicos Especiais em Biblioteconomia II*, cuja ementa contempla pesquisa científica e tecnológica nos setores público e privado, não possuindo ligação direta com a Biblioteconomia de Dados, mas o docente poderá fazer alguma introdução ao tema.

Quadro 53 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFES

DISCIPLINAS	EMENTAS
Estatística Aplicada	Estatística descritiva com utilização de pacotes estatísticos. Noções de probabilidade e curva normal. Conceitos em inferência estatística. Integração dos procedimentos estatísticos à pesquisa científica e ao processo de tomada de decisão.
Tópicos Especiais em Biblioteconomia II	Informação em Ciência e Tecnologia. Direito e acesso à informação científica e tecnológica e as políticas públicas de informação e comunicação no Brasil. Transferência de tecnologia. Informação para o setor produtivo. Informação e produtividade. Pesquisa científica e tecnológica nos setores público e privado.

Fonte: elaborado pelo autor, 2021.

Por meio da leitura, percebe-se a ausência, no curso de Biblioteconomia da UFES, às temáticas que contemplem as habilidades e competências necessárias ao bibliotecário(a) de dados. No entanto, há duas disciplinas as quais o tema, se aspirado pelo docente responsável pela disciplina, pode ser introduzido.

O curso de bacharelado em Biblioteconomia Claretiano Centro Universitário (CEUCLAR) é relativamente novo, iniciado em 2017, por funcionar na modalidade EaD e pela estrutura oferecida e pelo número de polos em todas as regiões do Brasil, tem alcance nacional. Nesse interim, oferece 1.200 vagas anuais, em regime semestral, com carga horária de 2.400 h e duração mínima de 3 anos (CEUCLAR, 2017).

Entre os objetivos deste curso, no que tange aos conhecimentos adquiridos por seus egressos, está “Ser competente para lidar com tecnologias da informação, acompanhando os avanços nesse setor e os incorporando aos serviços biblioteconômicos” (CEUCLAR, 2017, p. 63), e quando se refere às competências, busca que dominem as tecnologias de comunicação e de informação.

Quadro 54 – Termos relacionados a bibliotecário(a) de dados na CEUCLAR

Termos	Nome das Disciplinas	Ementas
--------	----------------------	---------

Dados de pesquisa	0	0
Pesquisa científica	1	2
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	1
Curadoria	0	0
Metadados	0	0

Fonte: dados da pesquisa, 2021.

Neste curso, o termo *pesquisa científica* foi encontrado em duas disciplinas: no título e na ementa da disciplina *Metodologia da Pesquisa Científica*, a qual não aborda diretamente pesquisa na estrutura da *e-Science* e/ou geração, uso e reúso de dados científicos de forma colaborativa, e na ementa da disciplina *Pesquisa Bibliográfica e Normalização*, quando versa sobre a normalização de trabalhos acadêmicos/científicos e introdução à pesquisa científica, mas sem indicar concepção e uso de dados científicos de forma colaborativa.

Verificou-se que o termo *colaborativo* é encontrado na ementa da disciplina *Linguagens Documentárias*, quando o termo alude sobre a construção colaborativa através de mídias sociais na *web* e na atualização de linguagens de representação na *web*, assunto no qual pode abordar o objeto desta dissertação.

Quadro 55 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na CEUCLAR

DISCIPLINAS	EMENTAS
-------------	---------

Pesquisa Bibliográfica e Normalização	A disciplina Pesquisa Bibliográfica e Normalização tem como propósito subsidiar os alunos no que diz respeito à compreensão e aplicação do processo da Pesquisa Bibliográfica e da utilização de Normas Brasileiras, com ênfase na formatação do trabalho acadêmico e domínio de parte das normas que versam sobre documentação. Para tanto, serão objetos de estudo da disciplina: Estudo introdutório da ciência e da pesquisa científica; Introdução ao trabalho acadêmico e as fontes de informação; Aspectos teóricos das normas brasileiras e internacionais de documentação e instituições normativas; Estudo da normalização para apresentação de trabalhos científico; Aplicação da normalização para apresentação de trabalhos científico.
Metodologia da Pesquisa Científica	A disciplina Metodologia da Pesquisa Científica tem como propósito contribuir para a formação do estudante para o exercício da pesquisa, nos campos teórico e prático, preparando-o para a autonomia com relação à solução de situações-problema em sua área de atuação, amparado na reflexão que estabeleça relação com o conhecimento já produzido. Em consonância com Projeto Educativo e os Princípios do Claretiano – Rede de Educação, os eixos temáticos da disciplina perpassam a ética na pesquisa - no que tange ao tratamento de propriedade intelectual veiculada pela comunidade científica – bem como a ética da alteridade no tratamento para com a Pessoa Humana, possível sujeito participante de uma pesquisa. A disciplina ainda aborda noções da linguagem empregada na esfera científica, a atuação dos Comitês de Ética em Pesquisa com Seres Humanos, a construção dos elementos clássicos de um Projeto de Pesquisa, e a elaboração do Artigo Científico como gênero discursivo de maior abrangência na seara acadêmica.

Linguagens Documentárias	A disciplina “Linguagens Documentárias” visa no contexto do curso compreender os fundamentos, as funções e a importância das linguagens documentárias pré e pós-coordenadas no processo de representação para recuperação da informação. Saber utilizar as normas, os procedimentos e as metodologias de construção, atualização, avaliação e gestão de linguagens documentárias; diferenciar o uso e a aplicação das linguagens documentárias nos processos documentários de indexação e recuperação da informação; conhecer novas tendências teóricas e metodológicas que norteiam a temática linguagens documentárias, vistas como sistemas de organização do conhecimento. Compreender os fundamentos da <i>web</i> semântica e os aspectos interdisciplinares na representação de recursos informacionais. Apresentação e caracterização das linguagens de representação na <i>web</i> . As ontologias e as folksonomias: conceitos, estruturas e aplicações na representação e recuperação da informação na <i>web</i> . As redes sociais como ambientes colaborativos na construção e atualização de linguagens de representação na <i>web</i> . Inovações em linguagens de representação na <i>web</i> e em outros ambientes informacionais.
--------------------------	--

Fonte: elaborado pelo autor, 2021.

Dessa forma, constatou-se a possibilidade do concluinte do curso da CEUCLAR adquirir algum conhecimento sobre Biblioteconomia de Dados no contexto da *e-Science*, porém nenhuma competência e habilidade para atuar nesse nicho.

Por fim, verificou-se que a maior parte dos cursos de Biblioteconomia estão localizados no sudeste brasileiro. Dentre os quatorze cursos analisados, três desses cursos não abordam, direta ou indiretamente, *e-Science* e/ou Biblioteconomia de Dados, e naqueles que possuem alguma ligação, suas disciplinas estão divididas igualmente entre as que estão diretamente e indiretamente ligadas a este tema.

6.2.5 Sul

A Região Sul compreende três estados possuindo, ao todo, seis cursos de Biblioteconomia: Santa Catarina (UFSC, UDESC e UNOCHAPECÓ), Paraná (UEL) e Rio Grande do Sul (UFRGS e FURG).

O curso de Biblioteconomia da Universidade Federal do Rio Grande do Sul (UFRGS) teve seu último PPC efetivado em 2012. O curso oferece 70 vagas anuais no

turno matutino e carga horária mínima de 2.835h, para serem concluídas em quatro anos. (UFRGS, 2012).

Este PPC afirma que o egresso de seu curso deverá ser capaz de interagir, entre outras coisas, com o “[...] desenvolvimento da produção científica e tecnológica e aumento dos produtos de inovação” (UFRGS, 2012, p.23). E entre suas competências e habilidades encontra-se o de “Desenvolver e utilizar tecnologias de informação e comunicação [...] processar a informação em diferentes suportes, mediante a aplicação de conhecimentos teóricos e práticos de coleta, processamento, armazenamento e difusão da informação” (UFRGS, 2012, p.24-25), mostrando, assim, a princípio, consonância com a Biblioteconomia de Dados em contexto de *e-Science*.

É preciso destacar a diferença nas quantidades de disciplinas descritas no PPC e as encontradas na Matriz Curricular disponível em seu sítio institucional, a Matriz Curricular possui 59 disciplinas a mais, por isso este trabalho analisou ambos os conteúdos.

Quadro 56 – Termos relacionados a bibliotecário(a) de dados na UFRGS

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	1	1
Metadados	0	0

Fonte: dados da pesquisa, 2021.

O único termo encontrado foi *curadoria*, tanto no título, quanto na ementa da disciplina *Curadoria Digital, Fundamentos e Aplicações*, referindo-se ao tratamento específico de dados digitais diretamente ligado à Biblioteconomia de Dados e à *e-Science*.

Quadro 57 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UFRGS

DISCIPLINAS	EMENTAS
Curadoria Digital, Fundamentos e Aplicações	Apresentação do contexto histórico de Curadoria digital, partindo de princípios fundamentais que dizem respeito aos processos de seleção, preservação, descrição, manutenção, coleção e arquivamento de ativos ou dados digitais, considerando a continuidade e compatibilidade semântica e ontológica de conteúdo, na perspectiva de unidades de informação da internet.

Fonte: elaborado pelo autor, 2021.

Observa-se, portanto, que o curso da UFRSG oferece apenas uma disciplina ligada diretamente a alguma habilidade e competência necessária ao bibliotecário(a) de dados no contexto da *e-Science*.

Não houve acesso ao PPC do curso de Biblioteconomia da Universidade Estadual de Londrina (UEL), mas em seu sítio institucional foram encontradas as disciplinas com as ementas e a pesquisa foi realizada sob esses dados.

No caso dessa instituição, no momento de tratamento dos dados, foi encontrado o termo *compartilhamento dos dados* que não estava sendo pesquisado. Nesse sentido, foi feita uma procura em toda pasta de trabalho do *Excel*, variando o termo para *compartilhamento de dados* também, e em nenhum dos casos foi encontrado nas demais instituições, logo, decidiu-se analisá-lo, nesse momento, mas não se viu a necessidade de inseri-lo nos demais quadros de termos dos demais cursos relacionados a bibliotecário(a) de dados investigados.

Quadro 58 – Termos relacionados a bibliotecário(a) de dados na UEL

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	1
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0

Metadados	0	0
Termos	0	1

Fonte: dados da pesquisa, 2021.

Destarte, foram encontrados dois temas: *preservação digital*, na ementa da disciplina *Gestão da Informação na web*, inserido no contexto da gestão e segurança na rede e preservação digital. Por meio da análise da ementa, não fica clara a ligação com Biblioteconomia de Dados, mas há forte indício dessa ligação; e, *compartilhamento dos dados*, na ementa da disciplina *Tecnologias Aplicadas à Representação Descritiva*, abordando além da representação descritiva também o compartilhamento dos dados nas redes, sendo identificado aqui um conhecimento extremamente necessário ao bibliotecário(a) de dados atuante na *e-Science*.

Quadro 59 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UEL

DISCIPLINAS	EMENTAS
Gestão da Informação na <i>web</i>	Tecnologias, arquiteturas e aplicações na gestão da informação na <i>web</i> . Segurança na rede e preservação digital
Tecnologias Aplicadas à Representação Descritiva	A representação descritiva de documentos em sistemas automatizados e a descrição dos documentos eletrônicos. O compartilhamento dos dados nas redes.

Fonte: elaborado pelo autor, 2021.

Verifica-se aqui duas disciplinas, uma que está provavelmente ligada ao objeto desse estudo e a outra que está diretamente ligada a ele, mostrando que o curso de Biblioteconomia da UEL fornece uma introdução à Biblioteconomia de Dados ligada à *e-Science*.

O curso de bacharelado em Biblioteconomia da Universidade Federal de Santa Catarina (UFSC) empreendeu a reestruturação de seu PPC em 2015 e oferece 60 vagas anuais, funcionando no período noturno, com duração mínima de 4 anos e 2.490 horas/aula. (UFSC, 2015)

Entres as capacidades do egresso deste curso estão elencadas a de “Reunir e avaliar dados e proceder ao seu armazenamento. Buscar, registrar, avaliar e difundir a informação com fins acadêmicos e profissionais” (UFSC, 2015, p. 17), mostrando assim disposição com habilidades e competências do(a) bibliotecário(a) de dados. Além disso, indica ainda a capacitação para “Planejar, constituir, administrar e participar de redes

regionais e globais de informação” (UFSC, 2015, p. 17), relacionando-se, assim, com a *e-Science*.

Quadro 60 – Termos relacionados a bibliotecário(a) de dados na UFSC

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	1
Ciência de Dados	0	0
Data Science	1	0
Ciclo de vida	0	1
Dados abertos	0	0
Open data	0	1
Preservação digital	1	1
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	1

Fonte: dados da pesquisa, 2021.

Em sua análise foram encontrados seis termos em cinco disciplinas: *pesquisa científica*, na ementa da disciplina *Pesquisa Bibliográfica*, ponderando sobre o fazer científico como um todo; *Data Science*, no título da disciplina homônima, que versa sobre acesso e uso de dados científicos, estando, pois, intrinsecamente ligada à Biblioteconomia de Dados; o termo *ciclo de vida* encontrado na ementa da disciplina *Gerenciamento de Projetos*, que trata da gestão de projetos, competência que tangencia Biblioteconomia de Dados, no entanto, não faz parte do rol procurado neste estudo.

O termo *preservação digital* foi encontrado tanto no título como na ementa da disciplina *Preservação Digital*, ementa esta que também possui o termo *metadados* e aborda um ensinamento amplo sobre a preservação digital, evidenciando uma ou mais habilidades e competências necessárias ao bibliotecário(a) de dados.

Outro termo encontrado foi o *open data*, na ementa da disciplina *Linked Data* (dados vinculados), abordando criação, compartilhamento e utilização de dados ligados, sobre os *datasets* (conjunto de dados) e o uso de *triplestores*, banco de dados próprio para o armazenamento e recuperação de dados.

Quadro 61 – Ementas dos conteúdos curriculares relacionados com bibliotecário de dados na UFSC

DISCIPLINAS	EMENTAS
Pesquisa Bibliográfica	Comunicação científica. Introdução às fontes de informação. Caracterização da pesquisa. Pesquisa científica. Pesquisa Bibliográfica: métodos, técnicas e normalização. Elaboração de projeto e relatório de pesquisa bibliográfica.
<i>Data Science</i>	Extração de informação a partir de dados. Definição de domínios de interesse. Segurança de dados. Mineração de dados científicos. <i>Data Warehouse</i> .
Gerenciamento de Projetos	Metodologia de gerenciamento de projetos. Ciclo de vida da gestão de projetos. PMBOK. <i>Project management body of knowledge</i> . PMI. Ferramentas de planejamento e acompanhamento de projetos. <i>MS Project</i> . Administração do tempo e reuniões. Gerência por processos.
Preservação Digital	Preservação digital. Estratégias estruturais: adoção de padrões; elaboração de normas e manuais; metadados de preservação digital; montagem de infraestrutura; formação de consórcios e parcerias. Estratégias Operacionais: migração; emulação; preservação de tecnologia; encapsulamento. Modelos de referência e projetos.
<i>Linked Data</i>	Práticas de <i>Linked Data</i> . Criação, compartilhamento e utilização de dados e informações semanticamente conectadas. <i>Datasets</i> e <i>triplestores</i> . <i>Linked Open Data</i> .

Fonte: elaborado pelo autor, 2021.

A partir da leitura do quadro, é possível averiguar a atenção dada pelo curso de Biblioteconomia da UFSC à Biblioteconomia de Dados, na *e-Science*, ao aprovisionar disciplinas que abordam uma ou mais dessas habilidades/competências, além de trazer mais uma com possibilidade de amplitude para esse tema e uma que tangencia esse objeto de estudo.

O curso de Biblioteconomia da Universidade Federal do Rio Grande (FURG) estruturou seu PPC atual em 2015, oferecendo 35 vagas anuais, e integralização mínima de quatro anos (FURG, 2015).

Quadro 62 – Termos relacionados a bibliotecário(a) de dados na FURG

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	2
Ciência de Dados	0	0

Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0
Open data	0	0
Preservação digital	0	1
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	1

Fonte: dados da pesquisa, 2021.

Foram encontrados, neste curso, três termos nas ementas de quatro disciplinas: *pesquisa científica*, na ementa da disciplina *Metodologia da Pesquisa em Ciência da Informação I*, que aborda uma introdução ampla da pesquisa e metodologia científica, porém não oferece indícios de abordagem sobre uso e reúso de dados científicos e/ou pesquisa no contexto da *e-Science*; esse termo aparece também na disciplina *Metodologia da Pesquisa em Ciências da Informação II*, versando sobre “Diretrizes e execução de pesquisa científica em Biblioteconomia e Ciência da Informação” (FURG, 2015, p. 39).

Outro termo encontrado foi *preservação digital* na disciplina *Gestão da Informação nas Redes de Computadores*, dando enfoque à gestão da informação e tecnologia de redes de computadores, abordando segurança de rede e preservação digital. Compreende-se que esta disciplina pode introduzir a gestão de dados científicos.

O quarto termo encontrado foi o *metadados*, na disciplina *Tópicos Especiais em Organização do Conhecimento* abordando, entre outros, metodologias de aplicação no tratamento e organização da informação, preservação digital, metadados e ontologias que são conhecimentos necessários para o bibliotecário(a) de dados atuar no contexto de *e-Science*.

Quadro 63 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na FURG

DISCIPLINAS	EMENTAS
Metodologia da Pesquisa em e Ciência da Informação I	Ciência e conhecimento científico: definições e diferenças. Classificação das ciências. Pesquisa científica: tipos e fases. Método científico e técnicas de pesquisa. Metodologia do trabalho científico. Pesquisa científica em Ciência da Informação. Planejamento da pesquisa em Ciência da Informação: projeto de pesquisa.

Metodologia da Pesquisa em Ciências da Informação II	Diretrizes e execução de pesquisa científica em Biblioteconomia e Ciência da Informação.
Gestão da informação nas redes de computadores	Tecnologias, arquiteturas e aplicações na gestão da informação nas redes de computadores. Segurança na rede e preservação digital.
Tópicos especiais em organização do conhecimento	Estudo teórico acerca das metodologias de aplicação no tratamento e organização da informação. Web semântica. Preservação digital. Metadados. Ontologias.

Fonte: elaborado pelo autor, 2021.

Dessa forma, verifica-se no curso de bacharelado de Biblioteconomia da FURG uma disciplina com conteúdo diretamente ligado às habilidades e competências do(a) bibliotecário(a) de dados, e uma segunda disciplina oferecendo ao docente a escolha de abordar, ou introduzir, noções sobre a *e-Science*.

O curso de bacharelado em Biblioteconomia – Habilitação em Gestão da Informação da Universidade Estadual de Santa Catarina (UDESC) tem seu PPC atualizado em 2017, oferece 40 vagas, com 2880 horas/aula a serem integralizadas no mínimo em três anos (UDESC, 2017).

Este Projeto possui entre seus objetivos “Atuar como gestor de unidades, serviços e sistemas de informação” (UDESC, 2017, p.11) incluindo, então, uma das habilidades necessárias ao bibliotecário(a) de dados, além de trazer outro objetivo que propõe ao futuro egresso “Utilizar tecnologias de informação e comunicação, bem como aplicar técnicas especializadas para gestão dos processos de organização, tratamento, recuperação e disseminação da informação” (UDESC, 2017, p.11), outra capacitação necessária para que esse profissional atue com dados científicos.

Quadro 64 – Termos relacionados a bibliotecário(a) de dados na UDESC

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	0
Pesquisa científica	0	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	1
Dados abertos	0	0
Open data	0	0
Preservação digital	0	1
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0

Curadoria	0	0
Metadados	0	3

Fonte: dados da pesquisa, 2021.

Em sua leitura foram encontrados três termos nas ementas de cinco disciplinas diferentes: *ciclo de vida* na disciplina Gestão de Projetos, abordando sobre ciclo de vida de projetos de gestão, não envolvendo dados de pesquisa; e, *preservação digital*, na ementa da disciplina Bibliotecas Digitais II, versando sobre preservação digital, arquivos abertos, repositórios digitais e tecnologia para interoperabilidade, estando diretamente ligada às habilidades e competências do(a) bibliotecário(a) de dados.

O último termo é *metadados*, destacado em ementas de três disciplinas: *Modelagem de Informação*, que diferencia dado de informação, além de trazer conceitos de banco de dados, dados relacionais e os padrões de metadados, disciplina diretamente ligada aos conhecimentos necessários para atuação do(a) bibliotecário(a) de dados; já a disciplina *Representação Descritiva I* versa sobre metadados utilizados em catalogação, seguindo padrões como o MARC; por fim, tem-se a disciplina *Representação Descritiva III* sendo um complemento da disciplina anterior.

Quadro 65 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UDESC

DISCIPLINAS	EMENTAS
Gestão de Projetos	Conceitos, características e tipos de projetos. Etapas, monitoramento e ciclo de vida de projetos. Planejamento e execução de projetos. Papel, competências e perfil do gestor de projetos. Metodologias e ferramentas para planejamento e acompanhamento de projetos. Gestão de projetos e captação de recursos.
Bibliotecas Digitais II	Padrões, formatos e protocolos. <i>Software</i> para Bibliotecas e Repositórios Digitais. Tecnologias para interoperabilidade. Arquivos abertos. Preservação Digital.
Modelagem de Informação	Conceito de dado e informação. A informação e os processos de negócio. Tipos de informação: estruturada e não-estruturada. Conceitos de banco de dados relacionais. Modelo entidade-relacionamento. Modelo conceitual e lógico. Abstração. Normalização. Metadados e seus padrões.

Representação Descritiva I	Fundamentos da catalogação. Catálogos: tipologias e estruturas. CBU. Princípios Internacionais. Catalogação/IFLA. ISBD. Catalogação Cooperativa. Formatos de intercâmbio. Formato MARC: história, evolução, variação, estrutura. “Família MARC”. Metadados. Dublin Core. FRBR; FRAD; RDA. Fontes de Informação para Catalogadores. Catalogação Social.
Representação Descritiva III	Padrões Bibliográficos. Conversão retrospectiva. Padrões de Metadados. MARC21 bibliográfico e MARCXML. MODS. FRBR – MARC. Panorama de representação da informação em formato e ambiente digital. RDA. Determinação dos Pontos de Acesso. RDA - RDA toolkit.

Fonte: elaborado pelo autor, 2021.

Sendo assim, foram encontradas duas disciplinas abordando assuntos relacionados às habilidades e competências necessárias para atuação do(a) bibliotecário(a) nesse cenário de *e-Science* e exploração de dados científicos, fazendo desse curso uma preparação de seu egresso para atuar com um perfil profissional direcionado à Biblioteconomia de dados.

O curso de bacharelado em Biblioteconomia da Universidade Comunitária da Região de Chapecó (UNOCHAPECÓ) oferece 200 vagas anuais na modalidade EaD, com duração mínima de três anos, possuindo 2600 horas/aula.

Sobre este curso, tem-se que foi implementado em 2017 e é distribuído em seis áreas, uma delas aborda conteúdos concomitantes com os conhecimentos necessários para o egresso em Biblioteconomia possa atuar como bibliotecário(a) de dados:

Tecnologia da Informação: Aplicações da tecnologia da informação e comunicação nas unidades de informação. Análise, avaliação e desenvolvimento (hardware e *software*). Gestão de bases de dados e bibliotecas virtuais. análise e avaliação de sistemas[sic] e redes de informação. Informatização das unidades de informação. (UNOCHAPECÓ, 2017, p. 9).

Quadro 66 – Termos relacionados a Bibliotecário(a) de dados na UNOCHAPECÓ

Termos	Nome das Disciplinas	Ementas
Dados de pesquisa	0	1
Pesquisa científica	0	0
Ciência de Dados	0	0
Data Science	0	0
Ciclo de vida	0	0
Dados abertos	0	0

Open data	0	0
Preservação digital	0	0
Gestão de dados	0	0
Reúso de dados	0	0
Colaborativo	0	0
Curadoria	0	0
Metadados	0	1

Fonte: dados da pesquisa, 2021.

Foram encontrados dois termos nas ementas de duas disciplinas: *dados de pesquisa* na disciplina *Projeto de Pesquisa II* (TCC), que remete ao uso de dados de pesquisa, mas não especificando quanto ao uso e reúso no atual contexto de *e-Science*, logo o assunto se dará conforme disponibilidade do docente; e o termo *metadados* na disciplina *Linguagem Documentária II: Indexação*, em que os metadados se relacionam com a indexação documentária, novamente se observa que não há uma ligação direta com o objeto de estudo deste trabalho.

Quadro 67 – Ementas dos conteúdos curriculares relacionados com bibliotecário(a) de dados na UNOCHAPECÓ

DISCIPLINAS	EMENTAS
Projeto de Pesquisa II (TCC)	Desenvolvimento do projeto de pesquisa: coleta/construção de dados. Organização, interpretação e análise dos dados de pesquisa. Relatório. Contemplar a finalização do relatório de acordo com a especificidade dos cursos que terão como trabalho final um produto.
Linguagem Documentária II: Indexação	Políticas de indexação em unidades e sistemas de informação. Vocabulário controlado como instrumento de indexação: Tesauro. Cabeçalho de Assunto. Índices e indexação. Resumo: tipos, funções e prática. Indexação automática. Metadados.

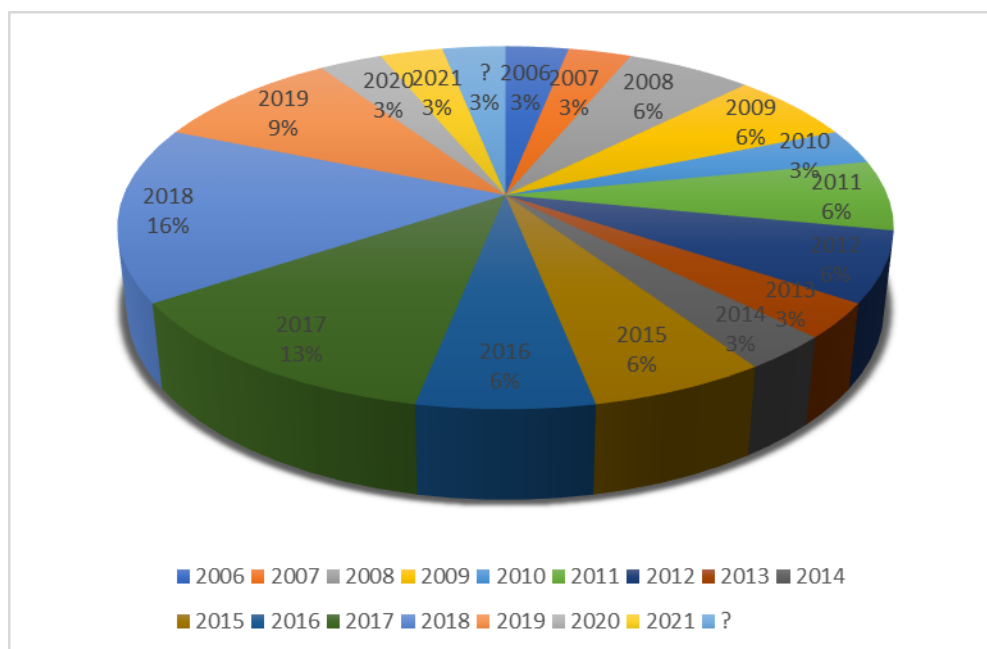
Fonte: elaborado pelo autor, 2021.

Logo, foi possível verificar que, embora o PPC aponte conteúdos potencialmente relacionados ao contexto da *e-Science* e Biblioteconomia de Dados, não há nenhuma disciplina que esteja diretamente ligada a esse cenário, embora tenham sido encontrados dois dos termos pesquisados.

6.3 Panorama geral dos currículos

Nessa subseção estão expostas e descritas graficamente análises quantitativas, a partir das análises qualitativas, dos dados obtidos, através das ementas das disciplinas dos cursos de Biblioteconomia no Brasil, possibilitando uma análise mais ampla do panorama de formação desses cursos, frente às demandas informacionais no contexto da *e-Science*.

Gráfico 6 – Ano da última atualização dos PPCs das IES

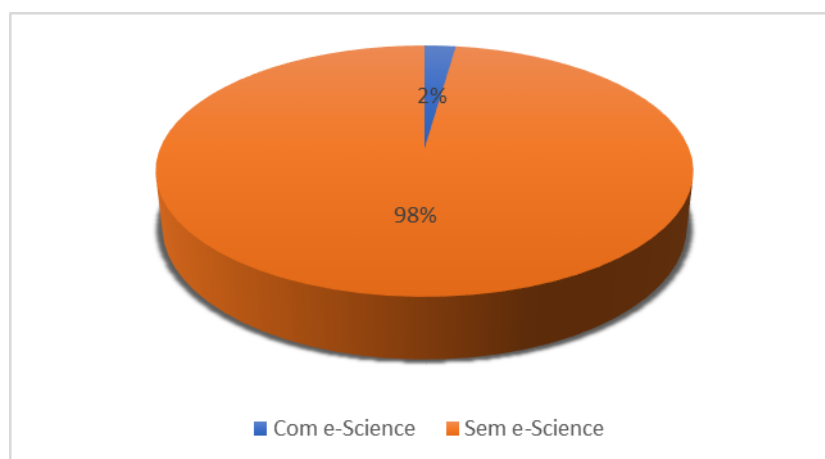


Fonte: dados da pesquisa, 2021.

No gráfico 6, é possível identificar que o PPC mais atual é o da FABCI, datado de 2020, porém a matriz curricular mais atual é da UEL que foi atualizada em 2021, uma vez que não houve acesso ao seu PPC.

O mais antigo é o da UFC, datado de 2004, três anos após a elaboração das DCN para os cursos de Biblioteconomia. É possível constatar uma distribuição homogênea nas atualizações dos PPCs entre os anos de 2004 e 2014, porém não é possível, com esses dados, identificar a causa desse fenômeno. Ainda assim, é notório que grande parte dos cursos possuem seus PPCs com menos de cinco anos de atualização, sendo dezessete, dos trinta e quatro, configurando 52% do total analisado. No ano de 2018, apenas cinco cursos tiveram seus PPCs atualizados. Destaca-se, dentre os PPCs analisados, a impossibilidade de identificar a data do PPC da UFBA.

Gráfico 7 – Disciplinas que abordam as temáticas relativas à Biblioteconomia de Dados/e-Science.

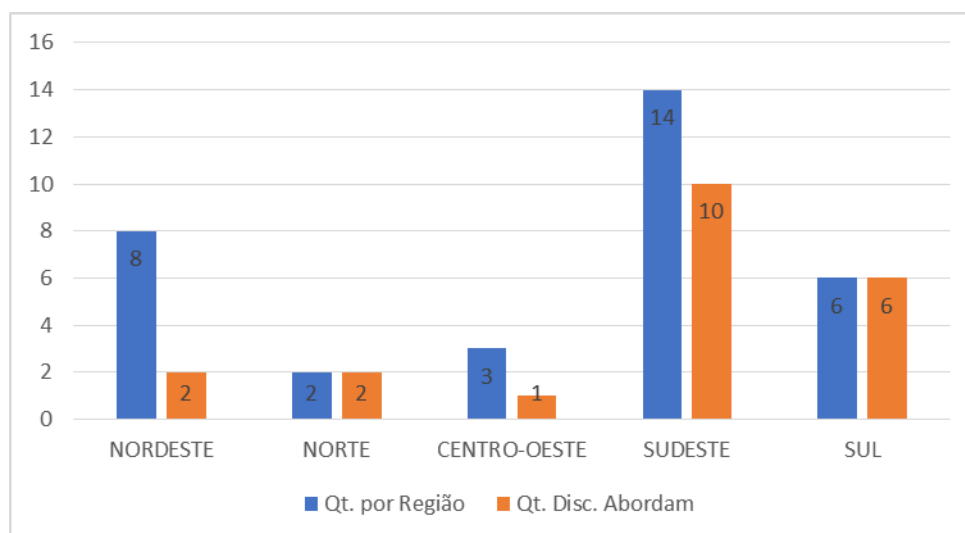


Fonte: dados da pesquisa, 2021.

Do ponto de vista da relação entre as disciplinas oferecidas pelos cursos do país, baseado no Quadro 3 – Competências dos bibliotecários na gestão de dados de pesquisa, em comparação com aquelas que abordam temáticas voltadas à *e-Science* e Biblioteconomia de Dados, é possível constatar que apenas 2% das disciplinas evidenciaram a temática, ou seja, num universo de 2017 disciplinas, 45 abordam este tema, conforme exposto no gráfico 7.

Com essa constatação, pode-se afirmar que é um número baixíssimo, dos trinta e três cursos pesquisados apenas doze possuem disciplinas diretamente ligadas ao tema e metade delas com apenas uma disciplina. A falta de atualização nos PPCs pode explicar essa ausência de abordagens para a formação no contexto da *e-Science*. Destaca-se ainda que os cursos que foram identificados com conteúdo que abordam o tema tiveram seus PPCs atualizados nos últimos anos.

Gráfico 8 – Cursos que abordam Biblioteconomia de Dados/*e-Science* por região brasileira

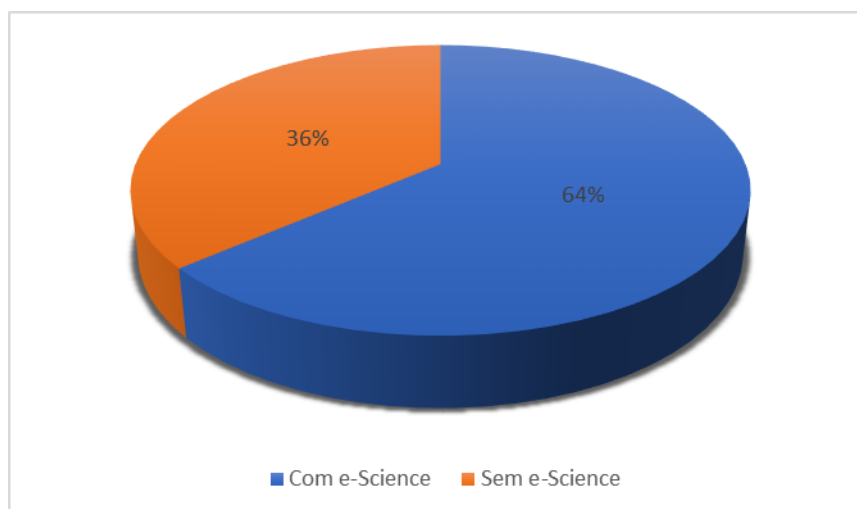


Fonte: dados da pesquisa, 2021.

A partir de uma análise pelas regiões brasileiras, no gráfico 8, fica latente a concentração de cursos que contemplam as temáticas em seus conteúdos, nas regiões Sudeste e Sul, considerando quantitativamente os cursos oferecidos por região.

Ao todo, foram encontradas disciplinas que contemplam o tema em 21 instituições, sendo a maioria no Sudeste e no Sul, contudo, embora seja notório que os cursos com maior inclinação para se cursar uma graduação com o objetivo de atuação enquanto bibliotecário de dados estejam nessas regiões, em todas as demais regiões há pelo menos um curso que possua no mínimo uma disciplina indicando esse campo de atuação.

Ponderando que a formação do(a) Bibliotecário(a) de Dados não finda e não é o escopo principal da graduação de Biblioteconomia (de formação mais abrangente) é possível verificar que, em alguma medida, estes conteúdos estão presentes nos cursos em todas as regiões brasileiras, de forma que esse futuro profissional possa, potencialmente, atuar no contexto de *e-Science* no Brasil, advindo de qualquer região nacional, mesmo que essa possibilidade se intensifique no Sul e Sudeste.

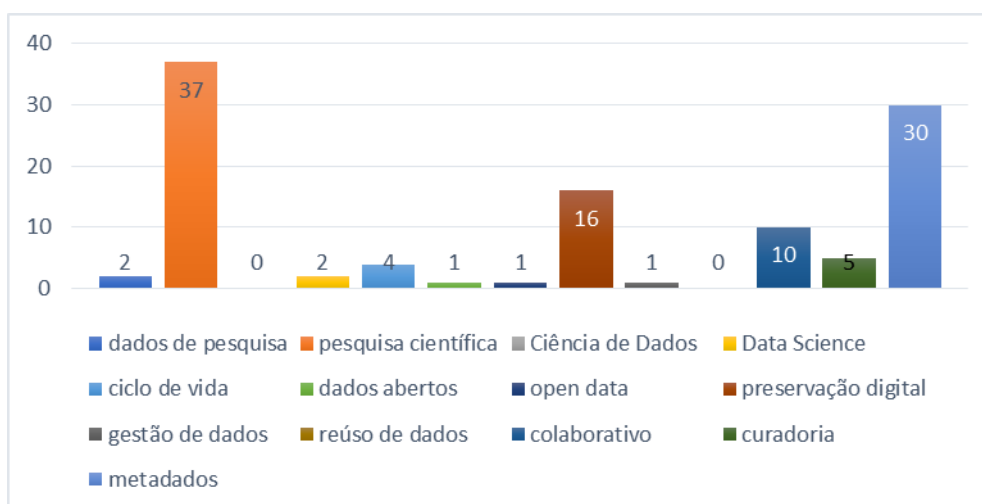
Gráfico 9 – Segundo IES a quais abordam Biblioteconomia de Dados/e-Science.

Fonte: dados da pesquisa, 2021.

Estabelecendo como base as IES investigadas, é perceptível que, em sua maioria, ao menos uma disciplina aborde aproximadamente alguma(s) habilidade(s) e competência(s) necessárias ao bibliotecário(a) de dados, a exemplo do observado no Gráfico 9, quando se verifica que 64% das instituições abordam o tema, ou seja, das 33 (trinta e três) IES, 21 (vinte e uma) tratam de alguma forma o tema e 12 (doze) não apresentam menção em seus conteúdos. Complementarmente, destaca-se a existência de disciplinas técnicas no núcleo de formação profissional que, por analogia, podem habilitar o profissional bibliotecário(a) a atuar com dados de pesquisa científica, no contexto da *e-Science*, porém o intuito deste trabalho é examinar as disciplinas expressas e atualizadas nesse sentido.

Com a consolidação da *e-Science* e a íntima relação entre os dados de pesquisa científica e a Biblioteconomia, os cursos de graduação devem indicar esse campo de atuação para seus egressos, principalmente pelo fato da Biblioteconomia se relacionar diretamente com os pesquisadores científicos. Embora o gráfico 9 demonstre que a maioria dos cursos aborde de alguma forma o tema, 36% dos cursos não possuem nenhuma relação, o que se considera como um alto índice dado ao atual contexto da pesquisa científica e sua relação direta com as tecnologias de informação e a comunicação.

Como forma de verificar o emprego dos termos elencados nos conteúdos, por meio do uso da análise do conteúdo, no Gráfico 10, indica-se a frequência em que os termos ocorreram entre os conteúdos das disciplinas.

Gráfico 10 – Frequência de ocorrências dos termos

Fonte: dados da pesquisa, 2021.

Dos termos identificados nas matrizes curriculares, o mais frequente foi *pesquisa científica*, presente 37 vezes, aproximadamente 34%, próximo dele o termo *metadados* com 30 frequências, configurando aproximadamente 27%. Os termos *Ciência de Dados* e *reúso de dados* não apareceram em nenhuma ementa. Ambos os termos estão diretamente ligados ao tema, o que demonstra a existência de uma grande lacuna na orientação dos egressos em Biblioteconomia a atuarem no contexto de *e-Science*.

Os termos pesquisados possuem relação direta com o tema desta pesquisa, no entanto, muitos não possuem relação exclusiva, mas todos foram investigados justamente para averiguar se havia ou não uma relação entre a disciplina a qual o continha em sua ementa e o tema, contudo é evidente que alguns termos são diretamente ligados, tais como *dados de pesquisa*, *Data Science*, *dados abertos*, *open data* e *gestão de dados*, que aparecem, e outros apresentam uma proximidade do tema, tais como *preservação digital*, *colaborativo* (no sentido de pesquisa colaborativa) e *metadados*, que aparecem com maior frequência que os diretamente ligados. Constatase que, diante das práticas profissionais tradicionais da biblioteconomia e que acompanharam o desenvolvimento das TICs, há grandes possibilidades de os cursos, a partir de seus conteúdos, abordarem vários aspectos e pontos da *e-Science*, todavia essa possibilidade não se confirma diretamente na análise e é preciso que se avance nas atualizações dos PPCs e se implemente mais disciplinas sobre o tema.

Vale ressaltar que o termo *pesquisa científica* foi encontrado em grande parte nas disciplinas relacionadas à metodologia e técnicas de pesquisa, que pode perfeitamente abordar *e-Science*, por ser uma forma de se fazer pesquisa científica. No

entanto, este estudo debruçou-se sobre as ementas das disciplinas, que, em sua maioria, são genéricas, e que apenas apontam a possibilidade de abordagem dessas temáticas, apenas sendo possível confirmar em um estudo mais aprofundado por meio da análise das bibliografias adotadas nas disciplinas.

Dessa forma, foi possível notar que embora o tema seja relativamente novo, em torno de dez anos do início de sua consolidação, e pelo referencial teórico encontrado fazendo a relação no contexto da *e-Science*, que em sua maioria advém de teóricos estrangeiros, grande parte dos cursos de Biblioteconomia brasileiros possuem alguma disciplina que faz menção ao tema, ou aos aspectos relativos a estes temas. Salienta-se que, a Biblioteconomia de Dados, no sentido de campo de atuação, é uma entre várias possibilidades de formação e atuação e que em nenhum momento este estudo possuiu o ímpeto de mensurar a qualidade do curso, mas sim, dadas as perspectivas no fazer pesquisa na contemporaneidade, compreender que quanto mais instituições oferecerem disciplinas habilitando seus egressos a atuarem nesse campo, melhor para os indivíduos que nele tencionam de formar e atuar.

A partir das análises apresentadas, aponta-se a estreita relação entre a Ciência da Informação e o cenário da *e-Science*, na perspectiva do estudo dos processos informacionais inerentes aos dados de pesquisa. Dessa forma, os(as) Bibliotecários(as) de Dados podem encontrar sua efetiva capacitação nas pós-graduações, no entanto é nas graduações que este profissional da informação deve descobrir este campo de atuação e assim ter um melhor direcionamento de suas pesquisas ou melhor profissionalização de suas atuações. Diante do exposto, cabe indicar que, uma investigação nos programas de pós-graduação em Ciência da Informação, poderão dar maiores indícios de como estes conteúdos estão sendo suprimidos do nível de graduação e estão contemplados nos currículos de pós-graduação, tendo em vista a atualização destes, por sua natureza, acompanhar mais diretamente os avanços científicos.

7 DISPOSIÇÕES FINAIS

Dado todo o trabalho de investigação, foi possível verificar a consolidação do uso intensivo de dados através do atual contexto científico e por meio de pesquisa colaborativa, oriunda da Ciência Aberta e da evolução da Ciência da Dados, o chamado por Gray, em 2007, de quarto paradigma científico ou *e-Science*.

No que tange à competência natural do(a) bibliotecário(a) em lidar com os processos de organização, acesso e uso da informação, bem como na sua atuação como profissional auxiliar ao pesquisador, constatou-se suas habilidades para o trato dos dados de pesquisa científica, além de, a partir destas habilidades, promover assistência ao pesquisador na curadoria e/ou gestão dos dados oriundos de pesquisas, atuando assim como bibliotecário(a) de dados.

Dentre os apontamentos desta pesquisa, vale ressaltar que, pela busca de cursos com oferta de disciplinas que capacitassem o egresso para atuar em ambientes de dados, apenas no curso de Biblioteconomia da UFRGS, entre as diversas disciplinas não constantes em seu PPC, mas que são ofertadas no curso, foi identificada a disciplina *Curadoria Digital, Fundamentos e Aplicações*, indicando ligação direta com o perfil de formação para a Biblioteconomia de Dados.

Dentre os resultados obtidos, tem-se que a maioria dos cursos, a partir de 2015, atualizaram seus PPCs (dezessete cursos entre os trinta e três analisados), no entanto, ainda há muitos que precisam passar por atualizações. Destes, há cursos que não passam por atualização há mais de 10 anos.

No processo de construção da pesquisa fora levantada a hipótese de que os cursos recém-criados, em especial os na modalidade EaD, poderiam contemplar um perfil mais atualizado de formação. No entanto, nos três cursos novos EaD, apenas em um havia uma disciplina diretamente ligada à *e-Science*. Dentre os cursos analisados, também é importante salientar a atualização recente de PPCs, como o da FABC, criado em 1948 e atualizado em 2020, o da UNB, criado em 1966 e em processo de atualização desde 2019, e o da PUC-CAMPINAS, criado em 1945 e atualizado em 2019. Ao mesmo tempo, também há cursos igualmente antigos e sem atualização, como o da UFMG criado em 1950 e atualizado em 2008, o da UFPB criado em 1963 e atualizado em 2007, e o da UFC criado em 1948 e atualizado em 2004. Contudo, é possível notar que a maioria dos cursos, mesmo que vagarosamente, estão em processo de atualização.

Dentre os objetivos elencados neste trabalho, todos foram alcançados: a partir do referencial teórico foi possível discutir o surgimento e confirmar a consolidação da *e-Science*; com base no referencial internacional, delineou-se o perfil do(a) bibliotecário(a) de dados nesse contexto da *e-Science*; na sequência, foram identificadas as ementas que se relacionam com a formação voltada para o Bibliotecário de Dados, mediante a investigação dos Projetos Pedagógicos de Curso; e, com os dados coletados, por fim, foi possível revelar o panorama de formação dos cursos de Biblioteconomia do Brasil, frente às demandas informacionais no contexto da *e-Science*. Panorama este que ainda está aquém do necessário para a profissão diante da proposta de formação que contempla o perfil do(a) Bibliotecário(a) de Dados.

Nos últimos cinco anos foi identificado um pequeno avanço em direção da ampliação do campo de atuação bibliotecária com dados de pesquisa científica, no contexto da *e-Science*, na graduação. Nas análises, fazendo uma aproximação dos conteúdos ministrados com as competências para lidar com dados de pesquisa, constata-se indícios de que o tema tem sido abordado, ainda que de maneira tímida, por pelo menos um curso nas regiões brasileiras investigadas.

A inquietação relativa à formação bibliotecária e o universo de tratamento de dados se dá diante da consolidação da *e-Science* e dos rumos na operacionalização, curadoria e gestão de dados abertos, tanto na ciência como na sociedade mundial, haja vista as políticas públicas e as legislações criadas para a proteção de dados, como a *General Data Protection Regulation* (GDPR) na Europa e a Lei Geral de Proteção de Dados (LGPD) no Brasil.

No campo científico nacional, pode-se exemplificar também essa consolidação com o recém lançado Edital de Chamada Aberta do IBICT 01/07/2021⁶², para incubação de repositórios de dados, em conjunto com a Rede Nacional de Ensino e Pesquisa (RNP) e o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), que visa ao incentivo e à implantação de Repositório de Dados Institucionais e à criação de um Consórcio Nacional para Ciência Aberta (CoNCienciA). O referido Edital tem como exigência que a administração do repositório de dados e a gestão da oferta como um serviço para os pesquisadores das instituições de ciência e tecnologia, incluindo as universidades, sejam tipicamente realizadas por um profissional com formação em Ciência da Informação ou Biblioteconomia. Assim, reforça-se que, uma

62

http://www.propesq.ufpb.br/propesq/contents/downloads/edital_incubacao_repositorios_2021_rerratificacao1.pdf

vez que a Biblioteconomia é tradicionalmente ligada em diversas práticas existentes na gestão de dados e começa a ser evidenciada para atuação desse campo, logo é latente a necessidade de maior avanço na formação para atender a tais demandas.

Com este estudo foi possível averiguar a consolidação da Biblioteconomia de Dados em países Europeus e nos EUA, no entanto, no Brasil há pesquisas nesse sentido, não obstante, ainda são poucas e apontam a necessidade de aprofundamento destas discussões formativas que alcancem os currículos de graduação.

Os cursos de graduação têm acompanhado, em alguma medida, os avanços tecnológicos e científicos, uma vez que 64% das instituições abordam o tema em suas disciplinas, o que pode sinalizar, para um futuro próximo, que os estudos sobre a gestão de dados sejam adotados para expansão do que já tem sido ofertado.

Conforme visto, a última reforma curricular estabelecida pelo CNE/CES é datada de 2001. Esta acompanha a sequência e a periodicidade de atualizações anteriores, como a do primeiro currículo mínimo, em 1962, e segundo currículo mínimo, em 1982, apontando um tempo médio de 20 anos entre as revisões curriculares. Caso seja mantida esta linearidade, diante das inovações no mundo do trabalho, pode-se ter uma perspectiva de que uma atualização esteja próxima e o vislumbre de uma revisão dos currículos, que potencialmente poderá contemplar as competências do perfil do(a) bibliotecário(a) de dados nas graduações. Não se afirma com esta pesquisa que tal formação não exista. Acredita-se que ela não é percebida nas graduações, no entanto, ela possivelmente deva se efetivar nas pós-graduações de áreas correlatas com a temática, principalmente por seu caráter inter, multi e pluridisciplinar.

A partir das análises levantadas com base na lacuna formativa constatada, aponta-se diversas possibilidades de estudos, que deem continuidade a estas discussões, com investigações tais como: o ensino da *e-Science* nas pós-graduações em Ciência da Informação; o ensino de *e-Science* nas graduações de Biblioteconomia da América Latina, as áreas do conhecimento que mais atuam no contexto de *e-Science*; entre outras.

Dessarte, após analisar disciplinas possivelmente ligadas às habilidades e competências do(a) Bibliotecário(a) de Dados, entre quase 70% dos cursos de Biblioteconomia do Brasil, foi possível verificar a existência de atualização nos respectivos PPCs, mas não de forma homogênea em todos os cursos e não na velocidade necessária para a orientação dos graduandos à possibilidade de atuação na Biblioteconomia de Dados no contexto de *e-Science*. Com este resultado e com o extenso referencial teórico abordado, espera-se que este estudo possa servir não só de

incentivo aos cursos nacionais para abordarem o tema e iniciar assim a formação do(a) Bibliotecário(a) de Dados no Brasil, mas também que possa servir como apoio na implantação destes conteúdos em suas diretrizes curriculares.

REFERÊNCIAS

- ABECIN. **Documentos Oficiais**. Associação Brasileira de Educação em Ciência da Informação. São Paulo. Disponível em: <http://www.abecin.org.br/documentos/documentos-oficiais/>. Acesso em: 21 jul. 2020.
- ALBAGLI, Sarita; APPEL, Andre Luiz; MACIEL, Maria Lucia. **E-Science e ciência aberta**: questões em debate. XIV ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB). 2013. GT 5 – Política e Economia da Informação. Disponível em: <http://ridi.ibict.br/handle/123456789/465>. Acesso em: 23 mai. 2021.
- ALMEIDA, Neilia Barros Ferreira de; BAPTISTA, Sofia Galvão. Breve histórico da Biblioteconomia brasileiro: formação do profissional. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTO E CIÊNCIA DA INFORMAÇÃO, 25., 2013, Florianópolis. **Anais [...]**. São Paulo: Febab, 2013. p. 3450 - 3462. Disponível em: <https://portal.febab.org.br/anais/article/view/1508>. Acesso em: 19 mar. 2020
- ANJOS, Renata Lemos dos. **Análise dos Componentes do Ciclo de Vida dos Dados na Matriz Curricular dos Cursos de Biblioteconomia**: um estudo comparado. 2019. 282 f. Dissertação (Mestrado) - Curso de Ciência da Informação, CCSA, Universidade Federal da Paraíba, João Pessoa, 2019.
- ARAÚJO, Ronaldo Ferreira. **Apropriações de Bruno Latour pela ciência da informação no Brasil**: descrição, explicação e interpretação. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação. Belo Horizonte, 139f., 2009.
- BARDIN, Laurence. **Análise de conteúdo** Lisboa: Edições 70, 2009?.
- BELL, Gordon. Prefácio. In: HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin. **O quarto paradigma**: descobertas científicas na era da e-Science. São Paulo: Oficina de Textos, 2011. p. 7-8.
- BELL, Gordon; HEY, Tony; SZALAY, Alex. Computer Science: beyond the data deluge. **Science**, [S.L.], v. 323, n. 5919, p. 1297-1298, 6 mar. 2009. American Association for the Advancement of Science (AAAS). <http://dx.doi.org/10.1126/science.1170411>.
- BERTIN, Patrícia Rocha Bello; VISOLI, Marcos Cezar; DRUCKER, Debora Pignatari. A gestão de dados de pesquisa no contexto da e-science: benefícios, desafios e oportunidades para organizações de p&d. **Ponto de Acesso**, v. 11, n. 2. DOI:10.9771/rpa.v11i2.21449. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/21449>. Acesso em: 23 mai. 2021.
- BORGMAN, Christine L. **Big data, little data, no data**: scholarship in the networked world. Cambridge, Massachusetts; London, England: MIT Press Books, 2015. 383 p.
- BOYD, Danah; CRAWFORD, Kate. Critical questions for big data. **Information, Communication & Society**, v. 15, n. 5, p. 662-679, 2012. Disponível em: <https://www->

tandfonline.ez15.periodicos.capes.gov.br/doi/full/10.1080/1369118X.2012.678878.
Acesso em: 16 abr 2020.

BRASIL. Ministério da Educação. CNE/CES 492/2001. Diretrizes Curriculares Nacionais dos cursos de Filosofia, História, Geografia, Serviço Social, Comunicação Social, Ciências Sociais, Letras, Biblioteconomia, Arquivologia e Museologia. **Diário Oficial da República Federativa do Brasil**, Brasília, 9 jul. 2001, Seção 1e, p. 50. Disponível em: <http://portal.mec.gov.br/cne/arquivos/pdf/CES0492.pdf> Acesso em: 26 mar. 2020.

BRASIL. Ministério do Trabalho. **Classificação Brasileira de Ocupações:** Profissionais da informação. Disponível em:
<http://www.mtecbo.gov.br/cbosite/pages/pesquisas/BuscaPorTituloResultado.jsf>.
Acesso em: 22 maio 2020.

BRASIL. Ministério da Ciência, Tecnologia, Inovações e Comunicações. **10 razões para você criar um plano de gestão de dados de pesquisa**. Elaborada por Comissão Nacional de Energia Nuclear. Disponível em: <http://antigo.cnem.gov.br/centro-de-informacoes-nucleares/sobre-o-cin/78-cin/dicas-academicas/232-10-razoes-para-voce-criar-um-plano-de-gestao-de-dados-de-pesquisa>. Acesso em: 02 jul. 2021.

BRYSON, Bill. **Bryson's Dictionary for Writers and Editors**. New York: Broadway Books, 2008.

BUCKLAND, Michael. Information as thing. **Journal of the American Society for Information Science**, v. 42, n. 5, p. 351-36, 1991. Disponível em:
<http://people.ischool.berkeley.edu/~buckland/thing.html>. Acesso em: 14 ago. 2014.

BUFREM, Leilah Santiago et al. Produção Internacional Sobre Ciência Orientada a Dados: análise dos termos Data Science e E-Science na Scopus e na Web of Science. **Informação & Informação**, Londrina, v. 21, n. 2, p.40-67, ago. 2016. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/26543>. Acesso em: 27 dez. 2019.

CAPLAN, Priscilla. **Metadata Fundamentals for All Librarians**. Chicago: American Library Association, 2003.

CAPURRO, Rafael. Epistemologia e ciência da informação. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB, 5., 2003. Belo Horizonte. **Anais[...]** Belo Horizonte: UFMG, 2003.

CASTRO, César. **História da Biblioteconomia brasileira**. Brasília: Thesaurus, 2000. 287 p.

CGACGIES. INEP. **Avaliação IN Loco**: glossário dos instrumentos de avaliação externa. 4. ed. Brasília: Inep, 2019. 108 p. Disponível em:
http://download.inep.gov.br/educacao_superior/avaliacao_institucional/apresentacao/glossario_4_edicao.pdf. Acesso em: 22 jul. 2020.

CONEGLIAN, Caio Saraiva; SANTAREM SEGUNDO, José Eduardo; SANT'ANA, Ricardo Cesar Gonçalves. A. *Big Data*: fatores potencialmente discriminatórios em análise de dados. **Em Questão**, v. 23, n. 1., DOI:10.19132/1808-5245231.62-86. Disponível em: <http://seer.ufrgs.br/index.php/EmQuestao/article/view/62122/0>. Acesso em: 23 mai 2021.

CONWAY, Drew. (2003). **The data science venn diagram**. Disponível em: <http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram> Acesso em: 20 abr 2020

COSTA, Maíra Murrieta. **Diretrizes para uma Política de Gestão de Dados Científicos no Brasil**. 2017. 288 f. Tese (Doutorado) - Curso de Ciência da Informação, Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, 2017.

COX, Michael; ELLSWORTH, David. **Application-Controlled Demand Paging for Out-of-Core Visualization**. Moffett Field: Nasa Ames Research Center, 1997. 11 p.

DATAONE (ed.). **Data Life Cycle**. Disponível em: <https://www.dataone.org/data-life-cycle>. Acesso em: 07 maio 2020.

DELANEY, John R.; BARGA, Roger S.. Uma visão de 2020 para a Oceanografia. In: HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin. **O quarto paradigma: descobertas científicas na era da eScience**. São Paulo: Oficina de Textos, 2011. Cap. 4. p. 53-63.

DIGITAL CURATION CENTRE. Edimburgo: DCC. (ed.). **About the DCC** Disponível em: www.dcc.ac.uk/about-us. Acesso em: 05 maio 2020a.

DIGITAL CURATION CENTRE. Edimburgo: DCC. (ed.). **What is digital curation?** Disponível em: <http://www.dcc.ac.uk/digital-curation/what-digital-curation>. Acesso em: 05 maio 2020b.

FEDERER, Lisa et al. The Medical Library Association Data Services Competency: a framework for data science and open science skills development. : a framework for data science and open science skills development. **Journal Of The Medical Library Association**, Pittsburgh, v. 108, n. 2, p. 304-309, 1 abr. 2020. University Library System, University of Pittsburgh. <http://dx.doi.org/10.5195/jmla.2020.909>. Disponível em: <http://jmla.pitt.edu/ojs/jmla/article/view/909>. Acesso em: 23 maio 2020.

FERGUSON, Adam R; NIELSON, Jessica L; CRAGIN, Melissa H; BANDROWSKI, Anita e; MARTONE, Maryann e. Big data from small data: data-sharing in the 'long tail' of neuroscience. : data-sharing in the 'long tail' of neuroscience. **Nature Neuroscience**, [S.l.], v. 17, n. 11, p. 1442-1447, 28 out. 2014. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/nn.3838>. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4728080/citedby/>. Acesso em: 01 maio 2020.

FERREIRA, Valdinéia.Barreto. **E-science e políticas públicas para ciência, tecnologia e inovação no Brasil** [online]. Salvador: EDUFBA, 2018, 256 p. ISBN: 978-85-232-1865-2..

FONSECA, Edson Nery da. **A Biblioteconomia brasileira no contexto mundial**. Rio de Janeiro: Tempo brasileiro, 1979. 247 p.

FORCE11. Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing version b1.0. La Jolla: Force11. Disponível em: <https://www.force11.org/fairprinciples>. Acesso em: 11 maio 2020

FRANCO, Maria Laura Puglisi Barbosa. **Análise de Conteúdo**. 3. ed. Brasília: Liber Livros, 2008. 80 p. (Pesquisa V. 6).

FREIRE, Gustavo Henrique Araújo. Ciência da Informação: temática, histórias e fundamentos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.11, n. 1, p. 6-19, jan./jun. 2006.

FRY, James P.; SIBLEY, Edgar H.. Evolution of Data-Base Management Systems. **Acm Computing Surveys** (csur), [S.l.], v. 8, n. 1, p. 7-42, mar. 1976. Association for Computing Machinery (ACM). <http://dx.doi.org/10.1145/356662.356664>.

FURNER, Jonathan. “Data”: the data. **Information Cultures In The Digital Age**, [S.L.], p. 287-306, 2016. Springer Fachmedien Wiesbaden. http://dx.doi.org/10.1007/978-3-658-14681-8_17.

GO FAIR. **FAIR Principles**. Disponível em: <https://www.go-fair.org/fair-principles/>. Acesso em: 11 maio 2020.

GUIMARÃES, José Augusto C.; FERREIRA, Gustavo M.; FREITAS, Maria Fernanda M.. Correntes teóricas do tratamento temático da informação: uma análise de domínio da presença da catalogação de assunto e da indexação nos congressos de isko-españa. In: ANTONIO, Y Otros Garcia Gutierrez de. **Actas Del X Congreso De Isko-España**: ferrol 30 de junio-1 de julio 2011. Coruña: Universidade La Coruña, 2013. p. 181-194. Disponível em: https://ruc.udc.es/dspace/bitstream/handle/2183/11647/CC_132_art_12.pdf?sequence=1&isAllowed=y. Acesso em: 21 jul. 2021.

GOMES, Josir Cardos PIMENTA, Ricardo Medeiros; SCHNEIDER, Marco. **Mineração de Dados na Pesquisa em Ciência da Informação**: desafios e oportunidades. XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB). 2019. GT 1 – Estudos Históricos e Epistemológicos da Ciência da Informação. Disponível em: <https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/546>. Acesso em: 16 jul. 2020.

GRAY, Jim. **EScience**: um método científico transformado. Transcrição de palestra ministrada por Jim Gray no Conselho Nacional de Pesquisa (EUA), 11 Jan. 2007. In: HEY, Tony;

GUIMARÃES, José Augusto Chaves. Organização do conhecimento: passado, presente e futuro sob a perspectiva da ISKO. **Informação & Informação**, [s.l.], v. 22, n. 2, p. 84-98, 29 out. 2017. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/31443>. Acesso em: 21 jul. 2020.

HENNING, Patricia Corrêa; RIBEIRO, Claudio José Silva; SANTOS, Luiz Olavo Bonino da Silva; SANTOS, Paula Xavier dos. GO FAIR e os princípios FAIR: o que representam para a expansão dos dados de pesquisa no âmbito da Ciência Aberta. **Em Questão**, [S.l.], v. 25, n. 2, p.389-412, 26 abr. 2019. Faculdade de Biblioteconomia Comunicação.

HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin. **O quarto paradigma**: descobertas científicas na era da eScience. São Paulo: Oficina de Textos, 2011. 263 p. Tradução de Leda Maria Marques Dias Beck.

HEY, Tony; TREFETHEN, Anne E.. Cyberinfrastructure for e-Science. **Science**, Washington, v. 308, n. 5723, p. 817-821, 6 maio 2005. American Association for the

Advancement of Science (AAAS). <http://dx.doi.org/10.1126/science.1110410>. Disponível em: <https://science.sciencemag.org/content/308/5723/817>. Acesso em: 13 abr. 2020.

HEY, Tony; TREFETHEN, Anne. E-Science and Its Implications. **Philosophical Transactions: Mathematical, Physical And Engineering Sciences**, Nova York, v. 361, n. 1809, p. 1809-1825, 15 ago. 2003. Disponível em: https://www.jstor.org/stable/3559222?seq=1&cid=pdf-reference#references_tab_contents. Acesso em: 14 abr. 2020.

JANKOWSKI, Nicholas W.. Exploring e-Science: an introduction. **Journal Of Computer-Mediated Communication**, [S.L.], v. 12, n. 2, p. 549-562, jan. 2007. Disponível em: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1083-6101.2007.00337.x>. Acesso em: 22 jul. 2020.

KOLTAY, Tibor. Data literacy for researchers and data librarians. **Journal Of Librarianship And Information Science**, [S.L.], v. 49, n. 1, p. 3-14, 10 jul. 2017. SAGE Publications. <http://dx.doi.org/10.1177/0961000615616450>.

KUHN, Thomas. **A estrutura das revoluções científicas**. São Paulo: Perspectiva, 2000.

LE COADIC, Yves-françois. **A ciência da informação**. Brasília: Briquet de Lemos, 1996. 115 p.

LISCOUSKI, Joe. The Data Librarian: introducing the Data Librarian. **Journal Of Automatic Chemistry**, [S.L.], v. 19, n. 6, p.99-204, nov 1997.

MARTINS, Gracy Kelli. **Institucionalização cognitiva e social da Organização e Representação do Conhecimento na Ciência da Informação no Brasil**. 2014. 184 f. TESE (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

MASTERMAN, Margareth. **A natureza de um paradigma**. In: LAKATOS, I.; MUSGRAVE, A. A crítica e o desenvolvimento do conhecimento. São Paulo, Cultrix, 1979.

MEDEIROS, Jackson da Silva; CAREGNATO, Sônia Elisa. Compartilhamento de dados e e-Science: explorando um novo conceito para a comunicação científica? data sharing and e-Science: exploring a new concept for scientific communication. **Liinc em revista**, v. 8, n. 2, 2012. Disponível em: <http://revista.ibict.br/liinc/article/view/3368>. Acesso em: 18 abr 2020.

METADATA 2020 (ed.). **About**. Disponível em: <http://www.metadata2020.org/about/>. Acesso em: 17 maio 2020a.

METADATA 2020 (ed.). **Librarians**. Disponível em: <http://www.metadata2020.org/communities/librarians/>. Acesso em: 17 maio 2020b.

MUELLER, Suzana Pinheiro Machado. O ENSINO DE BIBLIOTECONOMIA NO BRASIL. **Ciência da Informação**, Brasília, v. 1, n. 14, p.3-15, jun. 1985. Semestral

OECD – ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **OECD Principles and Guidelines for Access to Research Data from Public Funding**. Paris: Oecd Publishing, 2007. 22 p.

OLIVEIRA, Adriana Silva de Oliveira. **Desvendando a autoridade colaborativa na e-science sob A ótica dos direitos de propriedade intelectual**. 2016. 300 f. tese (Doutorado em Ciência da Informação) - Curso de Ciência da Informação, CCSA, Universidade Federal da Paraíba, João Pessoa, 2016.

PINHEIRO, Lena Vania Ribeiro; LOUREIRO, José Mauro Matheus. Traçados e limites da ciência da informação. **Ciência da Informação**, Brasília, v. 24, n. 1, p. 42-53, 1998. Disponível em: <http://revista.ibict.br/ciinf/article/view/609/611>. Acesso em: 23 mai. 2021.

POMERANTZ, Jeffrey. **Metadata**. London: The Mit, 2015. 239 p.

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS. **Projeto Pedagógico Curso de Biblioteconomia**. Campinas, 2019. Disponível em: <https://www.puc-campinas.edu.br/wp-content/uploads/2019/04/Projeto-Pedagogico-2019-Biblioteconomia.pdf> Acesso em 04 ago. 2020.

PRADO, Jorge de. (Org.). **Ideias emergentes em Biblioteconomia**. São Paulo: Febab, 2016. 116 p. Disponível em: <https://ideiasemergentes.wordpress.com/2016/03/24/download-do-livro/>. Acesso em: 29 jul. 2018.

PRICE, Derek John de Solla. **Little science, big science ... And beyond**. New York: Columbia University, 1986. 301 p.

RESEARCH COUNCILS UK. **RCUK Review of e-Science 2009**. Report of the International Panel for the 2009 Review of the UK Research Councils e-Science Programme, 2009. Disponível em: <https://epsrc.ukri.org/newsevents/pubs/rcuk-review-of-e-science-2009-building-a-uk-foundation-for-the-transformative-enhancement-of-research-and-innovation/>. Acesso em: 14 maio 2020.

RIBEIRO, Cláudio José Silva. Desafios para a Representação Documental no Âmbito da Web Semântica. In: ZAFALON, Zaira Regina; DAL'EVEDOVE, Paula Regina. **Perspectivas da Representação Documental: discussão e experiências**. São Carlos: Cpoi/ufscar, 2017. p. 287-302.

RICE, Robin; SOUTHALL, John. **The Data Librarian's Handbook**. London: Facet Publishing, 2016. 177 p.

RODRIGUES, Adriana Alves; DIAS, Guilherme Ataíde. Estudos sobre visualização de dados científicos no contexto da *Data Science* e do *Big data*. **Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, João Pessoa, v. 12, n. 1, p. 219-228, 2017. Disponível em: <http://www.periodicos.ufpb.br/ojs/index.php/pbcib/article/view/34774/17950>. Acesso em: 20 abr. 2020.

RODRIGUES, Adriana Alves; DUARTE, Emeide Nóbrega; DIAS, Guilherme Ataíde. Desafios da gestão de dados na era do *Big data*: perspectivas profissionais. **Encontro Nacional de Pesquisa em Ciência da Informação**. 2017. Disponível em: <https://periodicos.ufpb.br/index.php/itec/article/view/40538>. Acesso em 23 mai. 2021.

SALES, Luana Farias; SAYÃO, Luís Fernando. **A Ciência Invisível: revelando os dados da cauda longa da pesquisa.** XIX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB). 2018. GT 7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação. Disponível em: <https://www.brapci.inf.br/index.php/res/v/103678>. Acesso em: 01 mai. 2020.

SALES, Luana Farias; SAYÃO, Luís Fernando. Há futuro para as bibliotecas de pesquisa no ambiente de e-science?. **Informação & Tecnologia**, v. 2, n. 1, 2015. Disponível em: https://www.researchgate.net/publication/291945846_Ha_Futuro_para_as_Bibliotecas_de_Pesquisa_no_Ambiente_de_eScience. Acesso em: 14 ago. 2018

SALES, Luana Farias; SAYÃO, Luís Fernando. Uma proposta de taxonomia para dados de pesquisa. **Revista Conhecimento em Ação**, Rio de Janeiro, v. 4, n. 1, p. 31-48, 2019. Disponível em: <https://revistas.ufrj.br/index.php/rca/article/view/26337>. Acesso em: 28 abr. 2020.

SALES, Luana Farias; SAYÃO, Luis Fernando; MARANHÃO, Ana Maria Neves; DRUMOND, Geisa Meirelles; SILVA, Maria Helena Ferreira Xavier da. Competências dos bibliotecários na gestão dos dados de pesquisa. **Ciência da Informação**, Brasília, v. 48, n. 3, p. 303-313, set. 2019.

SANT'ANA, Ricardo César Gonçalves. Campo informacional resultante da interação de ciclo de vida dos dados. In: DIAS, G. A.; OLIVEIRA, B. M. J. F. de. **Dados científicos: perspectivas e desafios.** João Pessoa: Editora UFPB, 2019. Cap. 1. p. 13-32.

SANT'ANA, Ricardo César Gonçalves. Ciclo de vida dos dados e o papel da Ciência da Informação, 2013. In: XIV Encontro Nacional de Pesquisa em Ciência da Informação, 2013, Florianópolis. **Anais [...].** Florianópolis: XIV ENANCIB, 2013. Disponível em: <http://enancib.ibict.br/index.php/enancib/xivenancib/paper/view/4383/3506>. Acesso em 08 mai. 2020.

SANT'ANA, R. C. G. Reflexões sobre a representação no ciclo de vida dos dados, 2017. In: ENCONTRO DE REPRESENTAÇÃO DOCUMENTAL, 1, 2017, São Carlos. **Anais [...].** São Carlos: UFSCar, 2017. Disponível em: <http://www.telescopium.ufscar.br/index.php/enredo/enredo/paper/viewFile/109/96>. Acesso em: 22 jul. 2020.

SANTOS, Izequias Estevam dos. **Manual de métodos e técnicas de pesquisa científica.** 5. ed. rev. e atual. Niterói: Impetus, 2005.

SANTOS, Plácida L. V. Amorim da Costa; SANT'ANA, Ricardo César Gonçalves. Dado e granularidade na perspectiva da informação e tecnologia: uma interpretação pela ciência da informação: uma interpretação pela ciência da informação. **Ciência da Informação**, Brasília, v. 42, n. 2, p. 199-209, ago. 2015. Disponível em: <http://revista.ibict.br/ciinf/article/view/1382/1560>. Acesso em: 24 abr. 2020.

SARACEVIC, Tefko. Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.1, n.1, p.41-62, jan.\jun., 1996.

SAYÃO, LUÍS FERNANDO; SALES, Luana Farias. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade:**

Estudos, v. 22, n. 3, p. 179-191, set./dez. 2012. Disponível em: <https://periodicos.ufpb.br/ojs/index.php/ies/article/view/12224/8586>. Acesso em: 16 abr. 2020.

SAYÃO, LUÍS FERNANDO; SALES, Luana Farias. A ciência invisível: os dados da cauda longa da pesquisa científica. In: DIAS, G. A.; OLIVEIRA, B. M. J. F. de. **Dados científicos: perspectivas e desafios**. João Pessoa: Editora UFPB, 2019. Cap. 2. p. 33-52.

SAYÃO, Luís Fernando; SALES, Luana Farias. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, v. 21, n. 2, p. 90-115, 2016. DOI:10.5433/1981-8920.2016v21n2p90. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/27939>. Acesso em: 27 jul 2018.

SAYÃO, Luis Fernando; SALES, Luana Farias. **Dados de pesquisa: infraestrutura para publicação, compartilhamento e preservação**. Recife: Liber Ufpe, 2017. 76 slides, color. Disponível em: <https://pt.slideshare.net/LiberUFPE/dados-de-pesquisa-infraestrutura-para-publicacao-compartilhamento-e-preservao/8>. Acesso em: 02 maio 2020.

SAYÃO, Luis Fernando; SALES, Luana Farias. **Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015. 90 p.

SCHULTZ, Daniel. **Relógios atômicos e 500 kg de dados: como o buraco negro foi fotografado?**. 2019. Disponível em: <https://paraondeomundovai.blogosfera.uol.com.br/2019/04/15/relogios-atomicos-e-500-kg-de-dados-como-o-buraco-negro-foi-fotografado/>. Acesso em: 23 mai. 2021.

SEMLER, Alexandre Ribas. **Ciência da Informação em Contextos de E-Science: bibliotecários de dados em tempos de data science**. 2017. 175 f. Tese (Doutorado) - Curso de Ciência da Informação, Programa de Pós-Graduação em Ciência da Informação do Centro de Ciências da Educação, Universidade Federal de Santa Catarina, Florianópolis, 2017.

SEMLER, Alexandre Ribas; PINTO, Adilson Luiz. Os diferentes conceitos de dados de pesquisa na abordagem da Biblioteconomia de dados. **Ciência da Informação**, Brasília, v. 48, n. 1, p. 113-129, mai. 2019. Disponível em: <http://revista.ibict.br/ciinf/article/view/4461>. Acesso em: 28 abr. 2020.

SETZER, Valdemar W.. **Dado, Informação, Conhecimento e Competência**. 2015. Depto. de Ciência da Computação, Universidade de São Paulo. Disponível em: <https://www.ime.usp.br/~vwsetzer/dado-info.html>. Acesso em: 30 ago. 2019.

SILVA SEGUNDO, Sanderli José da Silva; ARAÚJO, Wagner Junqueira de. Curadoria e ciclo de vida dos dados. In: DIAS, Guilherme. Ataíde.; OLIVEIRA, Bernadina Maria Juvenal Freire de. **Dados científicos: perspectivas e desafios**. João Pessoa: Editora UFPB, 2019. Cap. 6. p. 113-152.

SILVA, Jonathas Carvalho; GOMES, Henriette Ferreira. Conceitos de informação na Ciência da Informação: percepções analíticas, proposições e categorizações: percepções analíticas, proposições e categorizações. **Informação & Sociedade: estudos**, João Pessoa, v. 25, n. 1, p. 145-157, abr. 2015. Disponível em: <https://periodicos.ufpb.br/index.php/ies/article/view/145>. Acesso em: 28 abr. 2020.

SOMASUNDARAM, G.; SHRIVASTAVA, Alok. **Armazenamento e gerenciamento de informações**: como armazenar, gerenciar e proteger informações digitais. Porto Alegre: Bookman, 2011.

STRASSER, Carly; COOK, Robert; MICHENER, William; BUDDEN, Amber. **Primer on Data Management**: what you always wanted to know. what you always wanted to know. 2012. UC Office of the President: California Digital Library. Disponível em: <https://escholarship.org/uc/item/7tf5q7n3>. Acesso em: 22 jul. 2020 TARTAROTTI, Roberta Cristina Dal'evedove; DAL'EVEDOVE, Paula Regina; FUJITA, Mariângela Spotti Lopes. BIBLIOTECONOMIA DE DADOS EM REPOSITÓRIOS DE PESQUISA: perspectivas para a atuação bibliotecária. : PERSPECTIVAS PARA A ATUAÇÃO BIBLIOTECÁRIA. **Informação & Informação**, [S.l.], v. 24, n. 3, p. 207-226, set. 2019. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/38732>. Acesso em: 22 maio 2020

TANSLEY, Stewart; TOLLE, Kristin. **O quarto paradigma**: descobertas científicas na era da escience. descobertas científicas na era da eScience. São Paulo: Oficina de Textos, 2011. p. 17-29.

TENOPIR, Carol; SANDUSKY, Robert J.; ALLARD, Suzie; BIRCH, Ben. Research data management services in academic research libraries and perceptions of librarians. **Library & Information Science Research**, [S.l.], v. 36, n. 2, p.84-90, abr. 2014. Elsevier BV. <http://dx.doi.org/10.1016/j.lisr.2013.11.003>.

THE ECONOMIST (ed.). **The world's most valuable resource is no longer oil, but data**: The data economy demands a new approach to antitrust rules. The Economist. Londres, *online*. 6 maio 2017. Disponível em: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>. Acesso em: 22 maio 2020.

UNIVERSIDADE COMUNITÁRIA DA REGIÃO DE CHAPECÓ. **Proposta Pedagógica Curso de Biblioteconomia Modalidade Educação a Distância: EaD**. Chapecó: Unochapecó, [2017]. Disponível em: <https://www.unochapeco.edu.br/static/data/portal/sites/ppc/130.pdf>. Acesso em: 03 ago. 2021.

UNIVERSIDADE DE BRASÍLIA. Faculdade de Ciência da Informação. **Projeto Pedagógico do Curso de Biblioteconomia**. Brasília: UnB, 2019. No prelo enviado por correspondência eletrônica.

UNIVERSIDADE DE SÃO PAULO. **[O Projeto Político Pedagógico do Curso de Biblioteconomia]**. São Paulo: USP, [2013]. Disponível em: <http://www2.eca.usp.br/cbd/site/wp-content/uploads/2013/01/Projeto-Pedag%C3%B3gico.pdf>. Acesso em: 02 ago. 2021.

UNIVERSIDADE DE SÃO PAULO. **Projeto Pedagógico do Curso Bacharelado em Biblioteconomia e Ciência da Informação**. Ribeirão Preto: USP, 2016. Disponível <http://www.ffclrp.usp.br/down.php?id=133&d>. Acesso em: 31 mar. 2021.

UNIVERSIDADE ESTADUAL PAULISTA. **Projeto Político Pedagógico Do Curso De Biblioteconomia**. Marília: UNESP, 2012. Disponível

https://www.marilia.unesp.br/Home/Graduacao/Biblioteconomia/ppp_06jul2017compri_mido.pdf. Acesso em: 04 ago. 2020.

UNIVERSIDADE ESTADUAL DE SANTA CATARINA. **Reformulação Curricular e Projeto Pedagógico do Curso de Biblioteconomia:** habilitação gestão da informação. Florianópolis: Núcleo Docente Estruturante, 2012. Disponível https://www.udesc.br/arquivos/faed/id_cpmenu/139/PPC_Biblio___2018___Res_62_2018_15695156827205_139.pdf. Acesso em: 04 ago. 2020.

UNIVERSIDADE FEDERAL DA BAHIA. Instituto de Ciência da Informação. Biblioteconomia e Documentação. **[Matriz curricular]**. Salvador: UFBA, [201-?]. Disponível em: https://blog.ufba.br/ici/files/2011/08/Biblioteconomia_matriz.pdf. Acesso em: 03 fev. 2021.

UNIVERSIDADE FEDERAL DA PARAÍBA. **Projeto Político-Pedagógico**. João Pessoa, 2007. Disponível em: http://www.ccsa.ufpb.br/biblio/contents/documentos/ppp%20do%20curso%20de%20bi_blioteconomia. Acesso em 02 fev. 2020.

UNIVERSIDADE FEDERAL DE ALAGOAS. **Curso de Biblioteconomia:** bacharelado. Maceió, 2019. 245 p. Disponível em: <http://www.ichca.ufal.br/graduacao/biblioteconomia>. Acesso em: 13 nov. 2020.

UNIVERSIDADE FEDERAL DE MINAS GERAIS. **Projeto Pedagógico do Curso de Biblioteconomia e Gestão da Informação da Escola de Ciência da Informação da UFMG**. Belo Horizonte, 2008. Disponível em: http://colgradbiblio.eci.ufmg.br/o-curso/projeto-pedagogico-1/at_download/file. Acesso em: 05 ago. 2020

UNIVERSIDADE FEDERAL DE PERNAMBUCO. **Projeto Pedagógico do Curso de Graduação em Biblioteconomia**. Recife, 2018. Disponível em: https://www.ufpe.br/documents/39158/0/PPC_2018.pdf/9e500e3f-be15-42a9-b6da-4d9121001514. Acesso em 04 ago. 2020.

UNIVERSIDADE FEDERAL DE RONDÔNIA. **Projeto Pedagógico do Curso de Biblioteconomia**. Porto Velho, 2018. Disponível em: https://drive.Google.com/file/d/1fibzbbMdbw_YMWwSOu2EppbctBQ32mNw/view. Acesso em 02 ago. 2020.

UNIVERSIDADE FEDERAL DE SANTA CATARINA. **Projeto Pedagógico do Curso de Graduação em Biblioteconomia da Universidade Federal de Santa Catarina**. Florianópolis, 2015. Disponível em: http://biblioteconomia.paginas.ufsc.br/files/2014/10/BBD_PPC_2016.pdf. Acesso em 02 ago. 2020.

UNIVERSIDADE FEDERAL DE SERGIPE. Conselho do Ensino, da Pesquisa e da Extensão. **Resolução nº 95/2011, de 03 de outubro de 2011**. Aprova alteração no Projeto Pedagógico do Curso de Graduação em Biblioteconomia e Documentação - Bacharelado e dá outras providências. Aracaju: Conselho do Ensino, da Pesquisa e da Extensão, 2011. Disponível em: http://cienciainformacao.ufs.br/uploads/page_attach/path/6323/0952011.pdf. Acesso em: 05 ago. 2020.

UNIVERSIDADE FEDERAL DO CARIRI. **Projeto Pedagógico do Curso de Biblioteconomia**. Fortaleza, 2006. Disponível em:

<https://si3.ufc.br/sigaa/verProducao?idProducao=594710&key=b31d5e233b345fa1e102675a3666b001>. Acesso em 03 ago.2020.

UNIVERSIDADE FEDERAL DO CEARÁ. **Projeto Pedagógico do Curso de Biblioteconomia**. Cariri, 2006. Disponível em: <https://documentos.ufca.edu.br/wp-folder/wp-content/uploads/2019/08/BiblioteconomiaUFCA-Projeto-Pol%C3%ADtico-Pedag%C3%B3gico-2006.pdf>. Acesso em 04 ago.2020.

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO. **Projeto Pedagógico do Curso se Biblioteconomia**. Vitória, 2018. Disponível em: http://www.biblioteconomia.ufes.br/sites/biblioteconomia.ufes.br/files/field/anexo/ppc-biblio_2016_atualizacao_2018.pdf#overlay-context=grade-curricular. Acesso em 04 ago.2020.

UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO. **Projeto Pedagógico do Curso de Bacharelado em Biblioteconomia**. Rio de Janeiro, 2009a. Disponível em: <http://www.unirio.br/cchs/eb/arquivos/Projeto%20Politico%20Pedagogico%20Bacharelado%20-%2031.05.2010.pdf>. Acesso em 04 ago.2020.

UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO. **Projeto político pedagógico do curso de licenciatura em Biblioteconomia**. Rio de Janeiro, 2009b. Disponível em: <http://www.unirio.br/cchs/eb/arquivos/Projeto-Pedagogico-do-Curso-de-Licenciatura-em-Biblioteconomia.pdf>. Acesso em 03 ago.2020.

UNIVERSIDADE FEDERAL DO PARÁ. **Projeto Pedagógico do Curso de Biblioteconomia**. Belém, 2009. Disponível em: https://drive.Google.com/file/d/1Fgs2Z_R9eYGUhgrL-FdMZYpIJIM_H-xI/view. Acesso em 03 ago. 2020.

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO. **Proposta Político-Pedagógica de Implantação do Curso de Biblioteconomia e Gestão de Unidades de Informação (CBG/UFRJ)**. Rio de Janeiro, 2008. Disponível em: http://www.facc.ufrj.br/joomla/images/docs/Projeto_Pedag%C3%B3gico_CBG.pdf. Acesso em 05 ago. 2020.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE. **Projeto Pedagógico do Curso de Graduação em Biblioteconomia da Universidade Federal do Rio Grande do Norte**. Natal, 2017. Disponível em <https://sigaa.ufrn.br/sigaa/verProducao?idProducao=4909670&key=4b886b87c09f243f012822a16ccd598d>. Acesso em 04 de ago. de 2020.

UNIVERSIDADE FEDERAL DO MATO GROSSO. **Projeto Pedagógico do curso de Biblioteconomia 2007**. Rondonópolis, 2011. Disponível em: <http://sistemas.ufmt.br/ufmt.ppc/PlanoPedagogico/Download/601>. Acesso em 05 ago. 2020.

UNIVERSIDADE FEDERAL DO RIO GRANDE. **Projeto Pedagógico do Curso de Biblioteconomia da FURG**. Carreiros, 2015. Disponível em: <https://biblioteconomia.furg.br/images/importadoswordpress/ppp.pdf>. Acesso em 02 ago. 2020.

UNIVERSIDADE FEDERAL DE SÃO CARLOS. **Projeto Pedagógico Curso de Biblioteconomia**. Formiga, 2018. Disponível em:

https://www.uniformg.edu.br/images/Regulamentos/biblioteconomia/ppc_biblioteconomia.pdf. Acesso em 05 ago. 2020.

CENTRO UNIVERSITÁRIO DE FORMIGA. **Projeto Político Pedagógico do Curso de Bacharelado em Biblioteconomia e Ciência da Informação**. São Carlos, 2014. Disponível em: <https://www.dci.ufscar.br/arquivos/bci/projeto-pedagogico-bci.pdf>. Acesso em 05 ago. 2020.

CLARETIANO CENTRO UNIVERSITÁRIO. **Curso de Graduação em Biblioteconomia: bacharelado projeto político pedagógico**. Batatais, 2017. Disponível em: https://sgo.claretiano.edu.br/static/projetos/365/original/PPPC_Biblioteconomia_2017_2020.pdf. Acesso em 04 ago. 2020.

WILBANKS, John. Ver a mudança de paradigma: e somos nós. In: HEY, Tony; TANSLEY, Stewart; TOLLE, Kristin. **O quarto paradigma: descobertas científicas na era da e-Science**. São Paulo: Oficina de Textos, 2011. Cap. 25. p. 217-222. Tradução de Leda Maria Marques Dias Beck.

WILKINSON, Mark D. et al. The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**, [S.l.], v. 3, n. 1, p. 1-9, 15 mar. 2016. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/sdata.2016.18>. Disponível em: <https://www.nature.com/articles/sdata201618#article-info>. Acesso em: 09 maio 2020.

XIA, Jingfeng; WANG, Minglu. Competencies and Responsibilities of Social Science Data Librarians: an analysis of job descriptions. : An Analysis of Job Descriptions. **College & Research Libraries**, Chicago, v. 75, n. 3, p. 362-388, 1 maio 2014. American Library Association. <http://dx.doi.org/10.5860/crl13-435>. Disponível em: <https://crl.acrl.org/index.php/crl/article/view/16367>. Acesso em: 19 maio 2020.

ZIVIANI, Artur; PORTO, Fábio; OGASAWARA, Eduardo. Ciência de dados: desafio para a ciência, indústria e governo. **ComCiência: REVISTA ELETRÔNICA DE JORNALISMO CIENTÍFICO**, Campinas, v. 1, n. 170, 10 jul. 2015. Disponível em: <http://www.comciencia.br/comciencia/handler.php?section=8&edicao=115&id=1381>. Acesso em: 14 ago. 2018.

APÊNDICE A

LISTA DE IES QUE OFERTAM CURSOS DE BIBLIOTECONOMIA NO BRASIL

Ano de Início	Instituição (IES)	Nome do Curso	Grau	Modalidade	Região	Categoria	PPC
1910	UNIRIO	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Federal	NÃO
1942	UFBA	BIBLIOTECONOMIA E DOCUMENTAÇÃO	Bacharelado	Presencial	Nordeste	Federal	SIM
1945	PUC-CAMPINAS	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Privada	SIM
1947	UFRGS	BIBLIOTECONOMIA	Bacharelado	Presencial	Sul	Federal	SIM
1948	FABCI	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Privada	SIM
1950	UFPE	BIBLIOTECONOMIA	Bacharelado	Presencial	Nordeste	Federal	SIM
1950	UFMG	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Federal	SIM
1957	USU	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Privada	NÃO
1963	UFPA	BIBLIOTECONOMIA	Bacharelado	Presencial	Norte	Federal	SIM
1963	UFF	BIBLIOTECONOMIA E DOCUMENTAÇÃO	Bacharelado	Presencial	Sudeste	Federal	NÃO
1965	UFC	BIBLIOTECONOMIA	Bacharelado	Presencial	Nordeste	Federal	SIM
1966	UNB	BIBLIOTECONOMIA	Bacharelado	Presencial	Centro-Oeste	Federal	SIM
1966	USP	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Estadual	SIM
1966	UFAM	BIBLIOTECONOMIA	Bacharelado	Presencial	Norte	Federal	NÃO
1968	UNIFORMG	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Privada	SIM
1969	UFPB	BIBLIOTECONOMIA	Bacharelado	Presencial	Nordeste	Federal	SIM
1969	UFMA	BIBLIOTECONOMIA	Bacharelado	Presencial	Nordeste	Federal	NÃO

1973	UEL	BIBLIOTECONOMIA	Bacharelado	Presencial	Sul	Estadual	SIM
1974	UFSC	BIBLIOTECONOMIA	Bacharelado	Presencial	Sul	Federal	SIM
1975	FURG	BIBLIOTECONOMIA	Bacharelado	Presencial	Sul	Federal	SIM
1975	UFES	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Federal	SIM
1975	FATEA	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Privada	NÃO
1977	UNESP	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Estadual	SIM
1980	UFG	BIBLIOTECONOMIA	Bacharelado	Presencial	Centro-Oeste	Federal	SIM
1994	UFSCAR	BIBLIOTECONOMIA E CIÊNCIA DA INFORMAÇÃO	Bacharelado	Presencial	Sudeste	Federal	SIM
1997	UFRN	BIBLIOTECONOMIA	Bacharelado	Presencial	Nordeste	Federal	SIM
1999	UFAL	BIBLIOTECONOMIA	Bacharelado	Presencial	Nordeste	Federal	SIM
1999	UFMT	BIBLIOTECONOMIA	Bacharelado	Presencial	Centro-Oeste	Federal	SIM
2001	UDESC	BIBLIOTECONOMIA	Bacharelado	Presencial	Sul	Estadual	SIM
2003	UESPI	BIBLIOTECONOMIA	Bacharelado	Presencial	Nordeste	Estadual	NÃO
2003	USP	BIBLIOTECONOMIA E CIÊNCIA DA INFORMAÇÃO	Bacharelado	Presencial	Sudeste	Estadual	SIM
2005	UNIFAI	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Privada	NÃO
2006	UFRJ	BIBLIOTECONOMIA E GESTÃO DE UNIDADES DE INFORMAÇÃO	Bacharelado	Presencial	Sudeste	Federal	SIM
2006	UFCA	BIBLIOTECONOMIA	Bacharelado	Presencial	Nordeste	Federal	SIM
2009	UNIRIO	BIBLIOTECONOMIA	Licenciatura	Presencial	Sudeste	Federal	SIM
2009	UFS	BIBLIOTECONOMIA E DOCUMENTAÇÃO	Bacharelado	Presencial	Sudeste	Federal	SIM

2009	CENTRO UNIVERSITÁRIO UNIVEL	BIBLIOTECONOMIA	Bacharelado	Presencial	Sul	Privada	NÃO
2009	UNIR	BIBLIOTECONOMIA	Bacharelado	Presencial	Norte	Federal	SIM
2010	CENTRO UNIVERSITÁRIO UNIC	BIBLIOTECONOMIA	Bacharelado	Presencial	Centro-oeste	Privada	NÃO
2010	UFRJ	BIBLIOTECONOMIA E GESTÃO DE UNIDADES DE INFORMAÇÃO	Bacharelado	Presencial	Sudeste	Federal	NÃO
2010	UNIRIO	BIBLIOTECONOMIA	Bacharelado	Presencial	Sudeste	Federal	SIM
2013	UCS	BIBLIOTECONOMIA	Bacharelado	A Distância	Sul	Privada	NÃO
2014	UNIVERSO	BIBLIOTECONOMIA	Bacharelado	A Distância	Diversos	Privada	SIM
2016	UNOCHAP ECÓ	BIBLIOTECONOMIA	Bacharelado	A Distância	Sul	Privada	SIM
2016	CEUUN	BIBLIOTECONOMIA	Bacharelado	Presencial	Sul	Privada	NÃO
2017	CEUCLAR	BIBLIOTECONOMIA	Bacharelado	A Distância	Diversos	Privada	SIM
2017	UniFAJ	BIBLIOTECONOMIA	Bacharelado	A Distância	Sudeste	Privada	NÃO
2018	UNIASSELVI	BIBLIOTECONOMIA	Bacharelado	A Distância	Diversos	Privada	NÃO

Fonte: Dados da pesquisa, 2020.

APÊNDICE B

RELATÓRIO DE CONTATO COM CURSOS DE BIBLIOTECONOMIA NO BRASIL

Instituição(IES)	Sigla	Acesso ao PPC
UNIVERSIDADE SANTA ÚRSULA	USU	CURSO NÃO DISPONÍVEL
CENTRO UNIVERSITÁRIO UNIVEL	UNIVEL	NÃO É MAIS OFERTADA
CENTRO UNIVERSITÁRIO UNIC	UNIC	NÃO É MAIS OFERTADA
UNIVERSIDADE FEDERAL DO MARANHÃO	UFMA	NÃO ENVIOU
CENTRO UNIVERSITÁRIO LEONARDO DA VINCI	UNIASSELVI	NÃO RESPONDEU
UNIVERSIDADE DE CAXIAS DO SUL	UCS	NÃO RESPONDEU
CENTRO UNIVERSITÁRIO TERESA D'ÁVILA	FATEA	NÃO RESPONDEU
UNIVERSIDADE ESTADUAL DO PIAUÍ	UESPI	NÃO RESPONDEU
UNIVERSIDADE FEDERAL FLUMINENSE	UFF	NÃO RESPONDEU
CENTRO UNIVERSITÁRIO ASSUNÇÃO	UNIFAI	NÃO RESPONDEU
UNIVERSIDADE FEDERAL DO AMAZONAS	UFAM	NÃO RESPONDEU
CENTRO UNIVERSITÁRIO UNISEP	UNISEP	SEM CONTATO
INSTITUTO DE ENSINO SUPERIOR DA FUNLEC	IESF	SEM CONTATO
CENTRO UNIVERSITÁRIO UNISEP	CEUUN	SEM CONTATO
CENTRO UNIVERSITÁRIO DE JAGUARIÚNA	UniFAJ	SEM CONTATO
FACULDADE DE BIBLIOTECONOMIA E CIÊNCIA DA INFORMAÇÃO	FABCI	SIM
UNIVERSIDADE DE BRASÍLIA	UNB	SIM
UNIVERSIDADE SALGADO DE OLIVEIRA	UNIVERSO	SIM
UNIVERSIDADE FEDERAL DE SANTA CATARINA	UFSC	SIM
UNIVERSIDADE FEDERAL DE SERGIPE	UFS	SIM
UNIVERSIDADE FEDERAL DE MINAS GERAIS	UFMG	SIM
UNIVERSIDADE FEDERAL DE MATO GROSSO	UFMT	SIM

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO	UFES	SIM
UNIVERSIDADE FEDERAL DA PARAÍBA	UFPB	SIM
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO	UFRJ	SIM
UNIVERSIDADE DE SÃO PAULO (Ribeirão Preto)	USP	SIM
UNIVERSIDADE ESTADUAL DE LONDRINA	UEL	SIM
UNIVERSIDADE FEDERAL DE ALAGOAS	UFAL	SIM
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL	UFRGS	SIM
UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO	UNIRIO	SIM
UNIVERSIDADE DE SÃO PAULO	USP	SIM
UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO	UNIRIO	SIM
UNIVERSIDADE FEDERAL DO RIO GRANDE	FURG	SIM
UNIVERSIDADE FEDERAL DA BAHIA	UFBA	SIM
UNIVERSIDADE FEDERAL DO CARIRI	UFCA	SIM
UNIVERSIDADE FEDERAL DO PARÁ	UFPA	SIM
FUNDAÇÃO UNIVERSIDADE FEDERAL DE RONDÔNIA	UNIR	SIM
UNIVERSIDADE FEDERAL DE GOIÁS	UFG	SIM
CENTRO UNIVERSITÁRIO CLARETIANO	CEUCLAR	SIM
UNIVERSIDADE FEDERAL DO CEARÁ	UFC	SIM
UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE	UFRN	SIM
UNIVERSIDADE FEDERAL DE SÃO CARLOS	UFSCAR	SIM
UNIVERSIDADE ESTADUAL PAULISTA JÚLIO DE MESQUITA FILHO	UNESP	SIM
PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS	PUC-CAMPINAS	SIM
FUNDAÇÃO UNIVERSIDADE DO ESTADO DE SANTA CATARINA	UDESC	SIM
UNIVERSIDADE FEDERAL DE PERNAMBUCO	UFPE	SIM
CENTRO UNIVERSITÁRIO DE FORMIGA	UNIFORMG	SIM
UNIVERSIDADE COMUNITÁRIA DA REGIÃO DE CHAPECÓ	UNOCHAPECÓ	SIM

APÊNDICE C

BIBLIOTECONOMIA NO CONTEXTO DE E-SCIENCE: a formação do(a) bibliotecário(a) de dados em instituições de ensino superior brasileiras

Data Collection

What data will you collect or create?

Serão coletados os planos curriculares dos cursos de Biblioteconomias das Instituições de Ensino Superior do Brasil.

How will the data be collected or created?

Primeiramente acessando os sítios institucionais, em caso da não existência do plano curricular, diretamente com as coordenações do curso, preferencialmente por e-mail.

Documentation and Metadata

What documentation and metadata will accompany the data?

Será gerado um arquivo TXT como documentação orientando sobre meus metadados baseado no Dublincore, porém sem o uso de XML.

Ethics and Legal Compliance

How will you manage any ethical issues?

A princípio a pesquisa não se direcionará para a investigação com humanos, estando centrada na coleta de dados em documentos e publicações bibliográficas devidamente citadas no texto do trabalho.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

Attribution-NoDerivatives 4.0 International (CC BY-ND 4.0): Creative Commons

Storage and Backup

How will the data be stored and backed up during the research?

Será feito a partir de um *backup* automático pela ferramenta de computação nas nuvens, o *MS Onedrive*, e um *backup* extra pelo *Google Drive*.

How will you manage access and security?

Os dados antes de disponibilização permanecerão em HD próprio, com o *backup* automático para nuvens, no *Onedrive*, ambos acessos restritos aos pesquisadores envolvidos na pesquisa.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

Os dados referem-se a um momento atual, porém podem ser utilizados e reutilizados em pesquisas futuras de cunho histórico e/ou evolutivo.

What is the long-term preservation plan for the dataset?

A pesquisa e seus resultados estarão disponíveis no Repositório Institucional da Universidade Federal da Paraíba, e os dados serão disponibilizados no Serviço Experimental de Dados Abertos - RNP associado ao IBICT (caso esta ferramenta ainda esteja disponível, caso contrário procurar-se-á outro Repositório de Dados gratuito).

Data Sharing

How will you share the data?

A princípio pretende-se compartilhar os dados em Repositório de Dados gratuitos, devido a não existência de financiamento para este fim, porém se a Instituição pela qual a pesquisa está sendo realizada, Universidade Federal da Paraíba, dispor de um Repositório de Dados que possa ser utilizado, os Dados da Pesquisa serão disponibilizados nele.

Are any restrictions on data sharing required?

Os Dados da Pesquisa serão disponibilizados após a conclusão da pesquisa, ficando disponibilizados com a única exigência de citação da fonte.

Responsibilities and Resources

Who will be responsible for data management?

Durante toda a pesquisa o gerenciamento dos dados estará sob a responsabilidade do pesquisador/ autor. Após a conclusão da pesquisa, os dados serão compartilhados e ficarão disponíveis para acesso e gerenciamento pelo Repositório de Dados utilizado.

What resources will you require to deliver your plan?

Como os dados, a princípio, ficarão disponíveis em um Repositório de Dados gratuito, poderá ser utilizado recursos do próprio pesquisador.