

UNIVERSIDADE FEDERAL DE SÃO CARLOS (UFSCar)
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS (CECH)
CURSO DE BIBLIOTECONOMIA E CIÊNCIA DA INFORMAÇÃO

LARISSA ALVES DA SILVA

**A IMPORTÂNCIA DA CURADORIA DIGITAL PARA A GESTÃO DE DADOS DE
PESQUISA**

SÃO CARLOS/SP
2020

LARISSA ALVES DA SILVA

A IMPORTÂNCIA DA CURADORIA DIGITAL PARA A GESTÃO DE DADOS DE PESQUISA

Trabalho de Conclusão de Curso apresentado ao Curso de Biblioteconomia e Ciência da Informação como requisito para obtenção do título de Bacharela em Biblioteconomia e Ciência da Informação pela Universidade Federal de São Carlos (UFSCar).

Orientadora: Prof.^a Dr.^a Ana Carolina Simionato Arakaki.

Pesquisa financiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

Processo FAPESP: 2018/23441-6.



grupo de pesquisa
dados e metadados

SÃO CARLOS/SP
2020

Silva, Larissa Alves da

A importância da Curadora Digital para a Gestão de
Dados de Pesquisa / Larissa Alves da Silva -- 2020.
96f.

TCC (Graduação) - Universidade Federal de São Carlos,
campus São Carlos, São Carlos

Orientador (a): Ana Carolina Simionato Arakaki

Banca Examinadora: Ana Carolina Simionato Arakaki,
Camila Carneiro Dias Rigolin, Ariadne Chloë Mary

Furnival

Bibliografia

1. Curadoria Digital. 2. Gestão de Dados de Pesquisa. I.
Silva, Larissa Alves da. II. Título.

LARISSA ALVES DA SILVA

**A IMPORTÂNCIA DA CURADORIA DIGITAL PARA A GESTÃO DE DADOS DE
PESQUISA**

Trabalho de Conclusão de Curso apresentado ao Curso de Biblioteconomia e Ciência da Informação como requisito para obtenção do título de Bacharela em Biblioteconomia e Ciência da Informação pela Universidade Federal de São Carlos (UFSCar).

Data de defesa: 21/12/2020

Aprovado em: ____/____/____.

Banca Examinadora:

Orientadora:

Professora Doutora Ana Carolina Simionato Arakaki
Universidade Federal de São Carlos (UFSCar)

Avaliadoras:

Professora Doutora Camila Carneiro Dias Rigolin
Universidade Federal de São Carlos (UFSCar)

Professora Doutora Ariadne Chloë Mary Furnival
Universidade Federal de São Carlos (UFSCar)

São Carlos, 21 de dezembro de 2020.

A José Alves da Silva (*in memoriam*).
À Claudia Gabriela Rodrigues da Silva.
Amo vocês, para sempre.

AGRADECIMENTOS

A Deus,

pela vida, por honrar os meus sonhos, estar comigo nesta jornada e avante.

Aos meus avós,

Joana, pelas palavras de incentivo sobre meu futuro e zelo perene e, José (*in memoriam*), pelas palavras inteligentes, motivadoras e acolhedoras, pelo exemplo de felicidade em meio às limitações impostas, por ser o exemplo de superação, felicidade e amor; por não julgar a mulher como inferior ao homem em termos físicos, sociais e intelectuais, sendo exemplo do que é ser um homem não machista em sociedade e que por toda sua vida soube e expressou o valor e importância do conhecimento. E ainda me ensina: *memento vivere*.

Aos meus pais, Josenita e Natanael,

pessoas sábias que, em simplicidade e dignidade, foram pontos de luz para a condução da educação formal atribuída aos meus irmãos e a mim, sendo meus exemplos de superação e quebra de barreiras.

Aos meus irmãos, Jaqueline e Thiago,

pelo apoio em todas as circunstâncias e ciência do que apenas nós, como irmãos, vivemos e superamos a cada dia, pela união, laço de origem e fraternal.

À minha melhor amiga, Cláudia Gabriela,

pela lealdade, amizade, conexão, por seu lindo coração e, por ser – muito provavelmente, neste mundo – a pessoa que mais conhece a mim e minha trajetória; pela torcida, companhia e felicidade diária a cada pequena realização de sonhos.

À minha professora orientadora, Ana,

por me abrir as portas do mundo acadêmico e me acompanhar antes mesmo de conhecê-la em aula, neste caso, em atividade de extensão e caminho em pesquisa; sendo minha graduação proporcional ao tempo em que me acompanhou, do

primeiro ao último ano. Por sua trajetória admirável como pesquisadora a me servir de inspiração e pelos incentivos e conselhos sobre a vida acadêmica e pessoal.

Às professoras, Camila e Chloe,
pela gentileza de aceitarem participar da banca examinadora, pelos diálogos trocados em aula e pelo acolhimento.

Ao Artur, Mônica e Renan,
técnicos administrativos do departamento que muito fizeram minhas manhãs e tardes felizes de estágio no CCBCI/DCI, pelo acolhimento, apoio, companhia e conselhos. Guardo vocês três no meu coração para sempre.

À minha amiga de apartamento e vida, Thaís Baldez,
por todas as alegrias, conversas, desabafos, risadas, carinho e por tornar minha estadia em São Carlos muito mais leve e feliz pelo período em que estive ali e por se manter aqui, até hoje, em parceria e consideração eterna.

À Prof. Raquel,
pelo acolhimento, por todas as conversas, conselhos, apoio, carinho, hospitalidade, pela recepção de braços abertos e pelo tempo em que moramos juntas.

A Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo subsídio, apoio e por acreditarem no potencial de minha proposta de estudo.

A todos os funcionários da Biblioteca Comunitária (BCo) da UFSCar,
pelo acolhimento e edições de eventos em que mais me senti confortável e acolhida, a exemplo, o ConProduCi (Conversando Sobre a Produção Científica na UFSCar) e por fazerem parte de um dos lugares em que mais estive presente na graduação.

Aos professores e funcionários do Departamento de Ciência da Informação e ao Grupo de Pesquisa Dados e Metadados (GP-DM),
por todos os saberes adquiridos.

À Universidade Federal de São Carlos (UFSCar/São Carlos),
pelos aprendizados e superações, marcando uma trajetória bonita, bucólica,
reflexiva e desafiadora em minha vida.

Locus amoenus. Fugere urbem.

Muito obrigada.



MONET, C. Nenúfares (*Nymphéas*). Giverny, 1908.

CLUBE DA ESQUINA N.º 2

Porque se chamava moço
Também se chamava estrada
Viagem de ventania
Nem lembra se olhou pra trás
Ao primeiro passo-asso-asso-asso
Porque se chamavam homens
Também se chamavam sonhos
E sonhos não envelhecem
Em meio a tantos gases lacrimogêneos
Ficam calmos, calmos, calmos
E lá se vai
Mais um dia
E basta contar compasso
E basta contar consigo
Que a chama não tem pavio
De tudo se faz canção
E o coração
Na curva de um rio, rio, rio
E lá se vai
Mais um dia
E o rio de asfalto e gente
Entorna pelas ladeiras
Entope o meio-fio
Esquina mais de um milhão
Quero ver então a gente, gente, gente

**Canção: Clube da Esquina n.º 2,
Álbum: Clube da Esquina, de 1972**

**Composição:
Milton Nascimento, Lô Borges e Márcio Borges.**

RESUMO

Período de advento tecnológico proveniente da era da informação. Tempo-espço este que, concomitantemente conduz a modificações na ciência e na elaboração de pesquisas científicas. Dados gerados em massa de forma exponencial que abrangem todas as áreas do conhecimento. Por este parâmetro, a presente pesquisa se apresenta em escopo qualitativo de cunho teórico-aplicado, a fim de correlacionar conceitos de Curadoria Digital e Gestão de Dados de Pesquisa. De fato, tem-se que dados, quando gerenciados, compreendidos, tratados, preservados e reutilizados de maneira adequada, podem vir a servir de recurso informacional para a difusão da ciência e desenvolvimento de pesquisas e conhecimento científico. Busca-se assim, analisar e refletir sobre a questão: como as duas abordagens se correlacionam e em sua somatória, vinculam-se ao domínio “dados” para disseminação da informação e conhecimento científico? Por este domínio, em desenvolvimento, explana-se sobre a Curadoria Digital e sua importância para a Gestão de Dados de Pesquisa. Realizou-se mapeamento acerca dos conceitos de Curadoria Digital e Gestão de Dados de Pesquisa. Como resultado, inferiu-se que o conjunto de conceitos identificados ao longo do mapeamento apresenta que estes dialogam entre si e complementam-se, bem como, dentre as definições de Gestão de Dados de Pesquisa, destacou-se a forte vinculação com a atuação do bibliotecário como de alta relevância no que tange oferta de serviços, utilização de ferramentas, além de planos e métodos para a gestão de dados de pesquisa. Observou-se que as análises sugerem que há interesse crescente diante da comunidade científica e acadêmica em relação à Curadoria Digital e Gestão de Dados de Pesquisa, ainda que estejam como um conceito em evolução, manifestando-se aos poucos pela propagação científica. Concluiu-se que se faz constante o reconhecimento da Curadoria Digital e sua importância para a Gestão de Dados de Pesquisa que consiste ao fato de que os métodos utilizados pelo processo de Curadoria Digital, como etapas de arquivamento e preservação digital, incluem também ao processo, etapas que visam dar atenção a criação e efetiva gestão de dados de pesquisa, tornando dados passíveis de interpretação, uso, reuso e acesso e que, um efetivo gerenciamento e posterior compartilhamento de tais dados, impulsionam o avanço científico e tecnológico, auxiliando pesquisas e pesquisadores a progredirem cientificamente.

Palavras-chave: Curadoria Digital. Gestão de Dados de pesquisa. Ciência da Informação. Dados e Metadados. *e-Science*.

ABSTRACT

Period of technological advent from the information age. Time-space that, concomitantly leads to changes in science and in the elaboration of scientific research. Data generated in mass exponentially that encompass all areas of knowledge. By this parameter, the present research performs in qualitative scope of theoretical-applied nature, in order to correlate concepts of Digital Curation and Research Data Management. Indeed, data when managed, understood, treated, preserved and reused in a way proper, serve as informational resource for the dissemination of science and development of research and scientific knowledge. Intends to analyze and reflect on the question: how the two approaches correlate and together, link to the “data” domain for the dissemination of scientific information and knowledge? In that domain, under development, explains about Digital Curation and the importance for Research Data Management. Performs mapping about the concepts of Digital Curation and Research Data Management. As a result, it appears that the set of identified concepts along the mapping presents that these dialogue with each other and complement each other, as well as among the definitions of Research Data Management, highlights the great link with the performance of the librarian as highly relevant in relation to service offering, use of tools, in addition to plans and methods for research data management. It is observed that the analyzes suggest that there is growing interest in the scientific and academic community in relation to Digital Curation and Research Data Management, even though they are an evolving concept, manifesting in the course of scientific propagation. It is concluded that recognition of the Digital Curation and its importance for Research Data Management which consists of the fact that methods used by the Digital Curation process, as stages of archiving and digital preservation, they also include steps in the process with the purpose of giving attention to the creation and effective management of research data, making data subject to interpretation, use, reuse and access and effective management and subsequent data sharing, increase scientific and technological progress, assist research and researchers to progress scientifically.

Keywords: Digital Curation. Research Data Management. Information Science. Data and Metadata. *e-Science*.

LISTA DE FIGURAS E GRÁFICOS

Figura 1 - Estrutura dos dados científicos	36
Figura 2 - Ciclo de Vida dos Dados do DDI.....	46
Figura 3 - Ciclo de vida dos dados – DataONE	47
Figura 4 - Curation Lifecycle Model – DCC	54
Figura 5 - Ciclo de vida da Curadoria Digital do DCC	55
Figura 6 - Fases da gestão de dados de pesquisa.....	72
Gráfico 1 - Autorias coletadas referente à Curadoria Digital	60
Gráfico 2 - Número de publicações por ano.....	62
Gráfico 3 - Número de palavras-chave.....	65
Gráfico 4 - Autorias coletadas referente a Gestão de Dados de Pesquisa	74
Gráfico 5 - Número de publicações por ano.....	75
Gráfico 6 - Número de palavras-chave.....	78
Gráfico 7 - Área do Conhecimento coletada referente a Gestão de Dados de Pesquisa.....	80

LISTA DE QUADROS E TABELAS

Quadro 1 - Tipos de metadados quanto à sua função.....	42
Quadro 2 - Tópicos para Mapeamento dos Conceitos de Gestão de Dados de Pesquisa e Curadoria Digital	49
Quadro 3 - Benefícios da Curadoria Digital	58
Quadro 4 - Valor a longo prazo da Curadoria Digital	58
Quadro 5 - Número de palavras-chave	64
Quadro 6 - Número de palavras-chave	77
Tabela 1 - Número de publicações por ano	61
Tabela 2 - Número de publicações por ano	75

LISTA DE ABREVIATURAS E SIGLAS

BCI	Biblioteconomia e Ciência da Informação
Brapci	Base de Dados em Ciência da Informação
CI	Ciência da Informação
CD	Curadoria Digital
CDDC	Curadoria Digital de Dados Científicos
CDDP	Curadoria Digital de Dados de Pesquisa
CNEN	Comissão Nacional de Energia Nuclear
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
CV	Ciclo de Vida
CVCD	Ciclo de Vida da Curadoria Digital
CVD	Ciclo de Vida dos Dados
CVPC	Ciclo de Vida da Pesquisa Científica
DataONE	<i>Data Observation Network for Earth</i>
DCC	<i>Digital Curation Centre</i>
DCI	Departamento de Ciência da Informação
DDI	<i>Data Documentation Initiative</i>
DMP	<i>Data Management Plan</i>
EUA	Estados Unidos da América
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo
GD	Gestão de Dados
GDP	Gestão de Dados de Pesquisa
GP-DM	Grupo de Pesquisa Dados e Metadados

IEN	Instituto de Engenharia Nuclear
MIT	<i>Massachusetts Institute of Technology</i>
MDMC	<i>Material Data Management Consortium</i>
Npc	Número de palavras-chave
Npub	Número de publicações por ano
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
P&D	Pesquisa e Desenvolvimento
PGD	Plano de Gestão de Dados/Plano de Gerenciamento de Dados
ProEx	Pró-Reitoria de Extensão
RDM	<i>Research Data Management</i>
SciELO	<i>Scientific Electronic Library Online</i>
SIBiUSP	Sistema Integrado de Bibliotecas da Universidade de São Paulo
TCC	Trabalho de Conclusão de Curso
TI	Tecnologia da Informação/Tecnologias da Informação
TICs	Tecnologias da Informação e Comunicação
UCL	<i>University College London</i>
UFSCar	Universidade Federal de São Carlos
UNESCO	<i>United Nations Educational, Scientific and Cultural Organization</i>

SUMÁRIO

INTRODUÇÃO	16
1.1 Definição do problema	19
1.2 Objetivos	19
1.2.1 Objetivo geral	19
1.2.2 Objetivos específicos.....	19
1.3 Justificativa.....	20
1.4 Procedimentos Metodológicos	21
1.5 Estrutura do Trabalho.....	24
2 FUNDAMENTAÇÃO TEÓRICA	26
2.1 Conceito de dados	31
2.2 Dados de pesquisa ou dados científicos	35
2.3 Metadados.....	39
2.3.1 Tipos de metadados.....	41
2.3.2 Metadados para armazenar e preservar dados científicos ou de pesquisa	43
2.4 Ciclo de vida dos dados	44
3 MAPEAMENTO PARA ANÁLISE DOS CONCEITOS DE CURADORIA DIGITAL E GESTÃO DE DADOS DE PESQUISA	49
3.1 Estado da Arte e do Conhecimento.....	49
3.2 Mapeamento dos conceitos de Curadoria Digital	52
3.3 Mapeamento dos conceitos de Gestão de Dados de Pesquisa	68
4 CONSIDERAÇÕES FINAIS	86
REFERÊNCIAS	91

INTRODUÇÃO

O desenvolvimento da Ciência da Informação (CI) em tratamento e gestão da informação e do conhecimento emprega as perspectivas em Tecnologias da Informação e Comunicação, as TICs, em que se tratam de “[...] estruturas e modelos de sistemas computacionais atuantes nos processos de produção, de armazenamento, de preservação, de representação, de recuperação, de acesso, de (re)uso e de disseminação de conteúdos informacionais.” (SANTOS; VIDOTTI, 2009, p. 1).

Diante das tecnologias, ciberespaço e *Web*, a sociedade moderna passou a consumir e a produzir uma escala considerável de informações em formato digital, surgindo neste contexto, o domínio de dados. É desse modo que, em âmbito científico, no qual, recentemente expressaram-se maiores transformações, uma imensa quantidade de dados foi gerada por meio do “[...] desenvolvimento de novos aparatos científicos [...]”, culminando em inovações de metodologias para produção do conhecimento. (SAYÃO; SALES, 2016, p. 67).

Com o advento tecnológico, além da influência de todas as áreas do conhecimento, valendo mencionar para este contexto, o Quarto Paradigma Científico denominado *e-Science*, caracterizado por Sayão e Sales (2015) como:

A mudança mais significativa para a pesquisa científica contemporânea foi o surgimento e a evolução de um novo padrão intelectual de se produzir conhecimento científico – conhecido como o Quarto Paradigma Científico ou *eScience* – que é caracterizado pela produção e uso intensivo de dados e pelo compartilhamento e colaboração baseados em rede de computadores de alto desempenho. (SAYÃO; SALES, 2015, p. 30).

James “Jim” Gray (1944-2007), informático estadunidense, pesquisador da *Microsoft*, pioneiro do *software* de banco de dados e vencedor do Prêmio *Turing*, em 2007, descreveu a *e-Science* como uma “[...] ciência intensiva em dados [...]” no denominado “Quarto Paradigma da Ciência” e seus ramos empíricos, teóricos, computacionais e fundamentado em dados, trazendo ainda que, com o impacto das TICs e dilúvio de dados, a ciência passou a se modificar constantemente.

Markoff (2009) ao mencionar James Gray acerca do quarto paradigma científico, abarca que os três paradigmas anteriores eram: experimentais, teóricos e computacionais. Sendo essa a evolução dos paradigmas da pesquisa científica

proposta por James Gray no ano de 2007 trazendo, em sequência: Primeiro Paradigma Científico – Ciência Empírica; Segundo Paradigma Científico – Ciência Teórica; Terceiro Paradigma Científico – Ciência Computacional e o Quarto Paradigma Científico – *e-Science*.

O conceito de “[...] ciência intensiva em dados [...]” vinculados aos processos colaborativos, trabalhados em rede e orientados por dados, define a *e-Science* como: “[...] a síntese da tecnologia da informação e da ciência que permite enfrentar desafios em escalas antes inimagináveis.” (BELL; HEY; SZALAY, 2009, p. 1298).

Neste escopo, a *e-Science*, como uma ciência que utiliza métodos de computação intensiva – métodos de uso de sistemas computacionais complexos – que, cientificamente, opera volumes de dados que requerem computadores conectados em rede com funcionamento mútuo para análise de amplos conjuntos de dados, isto é, computação em grade; trata da aplicação das Tecnologias da Informação (TI) que utilizam recursos da computação a fim de produzir, armazenar, transmitir, acessar, assegurar, utilizar e reutilizar a informação.

Exploram-se os dados que são capturados por instrumentos ou gerados por um simulador, posteriormente processados por *software*, que armazena a informação em computadores e assim, cientistas analisam os bancos de dados fazendo uso do gerenciamento de dados e da estatística. Dessa maneira, a *e-Science* pode ser identificada como uma ciência centrada nos dados, sejam eles provenientes de observação ou simulação, objetivando unificar teoria, experimentação e simulação. (USP, 2007).

Para a prática da investigação científica em relação à preparação, experimentação, coleta de dados, análise de resultados, disposição e acessibilidade aos objetos de constatação do processo científico, o reuso de dados visa a disposição de uso e disseminação do conhecimento científico íntegro à produção do conhecimento e da pesquisa científica.

Na relação apresentada entre a Ciência, a *e-Science* e o mencionado volume de dados, passa-se a contextualizar o campo de Ciência da Informação, integrando-o aos conceitos citados, a fim de comunicar a importância de atividades do ramo enfatizadas pela proposta, sendo a Curadoria Digital (CD) e a prática de Gestão de Dados de Pesquisa (GDP), aplicações primordiais de planejamento, interpretação e uso de dados, desde seu nascimento até sua manipulação, formas de armazenamento, análise, preservação, acesso e disseminação para sua conversão

em informação e produção de conhecimento científico, a partir de um planejamento de gestão de dados, atividade integrante da Curadoria Digital de Dados Científicos (CDDC).

Define-se Curadoria Digital, segundo Abbott (2008), como a gestão e preservação de dados digitais a longo prazo, permitindo contínuo acesso aos dados, bem como a reutilização destes de modo que o mesmo permaneça significativo aos usuários presentes e futuros.

Nesse sentido, a curadoria digital de dados científicos, tem o intuito de preservar não apenas um conjunto de dados, como também de disseminar conhecimento para comunidades futuras, permitindo que futuros usuários possam reanalisar tais dados em diferentes contextos no ramo científico.

Faz-se essencial que o processo de preservação permita ao significado dos dados ser atemporal e que, comunidades futuras de usuários, sejam auxiliadas para uma efetiva utilização de dados, analisando tópicos vinculados a Curadoria Digital de Dados de Pesquisa (CDDP) e a importância dos dados científicos em padrões atuais de pesquisa científico-acadêmicas. (SAYÃO; SALES, 2012).

Para que os dados sejam compreendidos, exige-se que sua descrição seja precisa. Ball (2012), descreve a relação existente entre os modelos de Ciclo de Vida da Curadoria Digital (CVCD) como os que fornecem uma estrutura para considerar as diversas operações que devem ser executadas em um registro de dados ao longo de sua vida útil, garantindo uma cobertura suficiente de requisitos concernentes às principais operações da Gestão de Dados (GD).

As práticas de gerenciamento de dados merecem a atenção de um Plano de Gestão de Dados (PGD) e o estabelecimento de requisitos para a gestão destes, ao mesmo passo em que a configuração científica contemporânea ao reconhecer que tais dados, quando gerenciados, compreendidos, tratados, preservados e reutilizados de maneira adequada, podem servir de recurso informacional para a difusão da ciência e desenvolvimento de pesquisas e conhecimento científico.

Dessa maneira, atribui-se o devido valor aos dados e seus componentes de ciclo de vida (CV), além de como estes podem ser utilizados em auxílio à ciência de forma a não tonarem-se obsoletos em termos de significado, tecnologia, bem como, mídias digitais e que, a gestão adequada influencia em práticas de pesquisa em todas as áreas do conhecimento, assegurando benefícios para o avanço científico e tecnológico.

1.1 Definição do problema

Por este prisma, ao considerar as entrelinhas de complementação entre as duas diferentes temáticas apresentadas e interligadas entre si, a presente proposta busca refletir em pesquisa, acerca do conceito e aplicação de Curadoria Digital e Gestão de dados, entre eles, os dados de pesquisa.

Portanto, o principal questionamento desta pesquisa consiste na identificação dos preceitos da curadoria digital para a gestão de dados de pesquisa, abordando seu ciclo de vida, funcionalidades e assim, analisa-se sobre: como as duas abordagens, Curadoria Digital e Gestão de Dados de Pesquisa, se correlacionam e em sua somatória, vinculam-se ao domínio “dados” para disseminação da informação e conhecimento científico?

1.2 Objetivos

A presente investigação tem como objetivos:

1.2.1 Objetivo geral

O objetivo geral constitui-se em sistematizar a apresentação da Curadoria Digital na elaboração de planejamentos de dados de pesquisa e melhoria da execução de pesquisas em si pelos cientistas e pesquisadores científicos.

1.2.2 Objetivos específicos

Os objetivos específicos que sustentam o objetivo geral, foram delineados em:

- Identificar a fundamentação teórica frente à Curadoria Digital e Gestão de Dados de Pesquisa, contextualizando as temáticas na literatura;
- Descrever: 1) o conceito e aplicação da Curadoria Digital, enfatizando busca e recuperação digital; 2) o conceito e aplicação da Gestão de Dados de Pesquisa, bem como, o Plano de Gestão de Dados, o Ciclo de Vida dos Dados (CVD) conjuntamente ao Ciclo de Vida da Pesquisa Científica (CVPC), tipos de dados de pesquisa, a importância dos metadados para armazenar e preservar dados científicos e, por fim;

- Mapear os conceitos de Curadoria Digital e Gestão de Dados de Pesquisa para: 1) Análise de conteúdo (análise de dados qualitativos) dos conceitos e 2) Análise quantitativa dos autores, anos de publicação, palavras-chave, periódicos, bases de dados, instituição ou localização geográfica e área do conhecimento com maior impacto e relevância frente às temáticas contempladas pelo trabalho.

1.3 Justificativa

Como intuito principal, a temática da proposta justifica-se pela contribuição científica na área de Biblioteconomia e Ciência da Informação (BCI), bem como, seus objetos de estudo sobre dados de pesquisa/científicos, ciclo de vida dos dados, curadoria digital, gestão, planejamento e compartilhamento de dados, de modo a contribuir para a área de Ciência da Informação por meio das temáticas apresentadas.

A motivação pessoal para a temática proposta deu-se igualmente, a partir da identificação e dos conhecimentos adquiridos no Curso Bacharelado em Biblioteconomia e Ciência da Informação pela Universidade Federal de São Carlos (UFSCar), a integração da aluna ao Grupo de Pesquisa Dados e Metadados (GP-DM) da Universidade Federal de São Carlos (UFSCar), coordenado pela Prof.^a Dr.^a Ana Carolina Simionato Arakaki e aos conhecimentos agregados em atividade de extensão, *Curadoria de acervos: manutenção, organização e representação de documentos (2ª Oferta)*, coordenado pela Prof.^a Dr.^a Ana Carolina Simionato Arakaki por meio da Pró-Reitoria de Extensão (ProEx) da Universidade Federal de São Carlos (UFSCar).

Outro aspecto de motivação pessoal deu-se pelo desenvolvimento do projeto de pesquisa, *Estudo da Gestão de Dados de Pesquisa*, pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pelo desenvolvimento do projeto de pesquisa, *Curadoria Digital: a importância para a Gestão de Dados de Pesquisa*, pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), trazendo maior aprofundamento e estima para as temáticas de Curadoria Digital e Gestão de Dados de Pesquisa.

Neste contexto, tais conhecimentos e experiências agregadas contribuíram ao trabalho, estudo e análise do domínio dados e como este se apresenta

conjuntamente aos seus processos de gestão de dados e curadoria digital, para melhor aprimoramento em pesquisa na área, de valor pessoal e coletivo a fim de agregar conhecimento em âmbito geral e posterior desenvolvimento de projetos para pesquisas futuras.

A importância das duas temáticas abordadas faz-se necessária para a soma de literatura de caráter científico com tais abordagens no Brasil, além de contribuição para o desenvolvimento de pesquisas nesses escopos mencionados junto ao gerenciamento de dados, além da contribuição para a gestão informacional gerada pelos conhecimentos científicos e projetos de pesquisa.

Ademais, a temática de Gestão de Dados, vinculada ao plano de gerenciamento de dados, vem tornando-se parte relevante nos processos de pesquisa e reconhecida como essencial para análises, métodos e promoção de novas pesquisas. Desse modo, a Curadoria Digital pode ser compreendida como um processo conjunto na gestão de dados em que ambas auxiliam no gerenciamento da massa exponencial de dados, informação e conhecimento emergentes.

A presente proposta justifica-se pela reflexão de como as duas abordagens se correlacionam e em sua somatória, vinculam-se ao domínio “dados” para disseminação da informação e conhecimento científico, de modo a contribuir não apenas para a Ciência da Informação, como também, para a Ciência como um todo, em parâmetros de gestão de dados pelos pesquisadores e comunidade científica.

Pretende-se não apenas abranger diversos conhecimentos científicos, mas sim, contribuir para a própria maneira de disposição e apresentação de tais conhecimentos na elaboração de planejamentos de dados de pesquisa e melhoria da execução de pesquisas em si pelos pesquisadores na gestão de seus dados científicos.

1.4 Procedimentos Metodológicos

De caráter investigativo, analítico e reflexivo, a presente metodologia de pesquisa em consonância aos seus objetivos, caracterizou-se como investigação qualitativa, sendo também classificada como pesquisa de abordagem exploratória, utilizando-se de revisão de literatura e proporcionando maior familiaridade ao tema, elucidação e construção de hipóteses.

Em Estado da Arte, para Mapeamento, apresentou parâmetros bibliométricos na análise quantitativa da coleta de dados realizada, visando compreensão e análise frente à importância da Curadoria Digital para a Gestão de Dados de Pesquisa, o comportamento dos dados em seu ciclo de vida, bem como, reflexão e análise da correlação entre as duas abordagens temáticas.

Quanto à sua natureza, caracterizou-se como pesquisa teórico-aplicada, visto que, embora não aborde cunho prático direto para seu desenvolvimento, a pesquisa se constituiu de conhecimentos, método/plano e base teórica para a aplicação do leitor-pesquisador em termos de Gestão de Dados e Plano de Gerenciamento de Dados, de modo a subsidiar o avanço científico e tecnológico, como também, auxiliar cientistas e pesquisadores na delimitação e gerenciamento de dados e informações em suas pesquisas e projetos científicos.

Em referência aos procedimentos técnicos e tipo de pesquisa, classificou-se como uma pesquisa bibliográfica, visto que, para embasamento e contextualização teórica, foram utilizadas fontes bibliográficas como fundamentação a fim de alcançar os resultados para reflexão de objetos de informação envolvidos contemplados nas literaturas científicas das áreas de Biblioteconomia, Tecnologia da Informação e Comunicação, Computação e Ciência da Informação com intuito de fundamentação teórica para a pesquisa relacionada às duas abordagens temáticas: Curadoria Digital e Gestão de Dados de Pesquisa e as ramificações que partem dessas esferas, com critério de pertinência aos temas escolhidos.

Em primeiro momento, foi proposto embasamento e investigação teórica acerca das temáticas abordadas e contextualização em literatura. Em segundo momento, a investigação foi direcionada à abordagem conceitual e de aplicação das duas esferas juntamente às suas ramificações: ciclo de vida dos dados; ciclo de vida da pesquisa científica; metadados; entre outros. E em terceiro momento, a problemática de pesquisa foi analisada acerca de como as duas abordagens se correlacionam e em sua somatória, vinculam-se ao domínio “dados” para disseminação da informação e do conhecimento científico.

Neste escopo, o plano de trabalho para a execução da pesquisa foi contemplado em seis etapas de desenvolvimento, conforme a seguir:

1ª etapa: Levantamento bibliográfico e seleção do material obtido – que se trata da busca e identificação das bibliografias para embasamento e investigação teórica

acerca das temáticas abordadas, bem como, contextualização em literatura pertinente às temáticas de pesquisa. Para análise e estudo, o levantamento bibliográfico foi realizado em nível nacional e internacional, considerando materiais impressos e digitais, visando abranger objetivos (geral e específico) da pesquisa. Em relação aos materiais obtidos, estes foram analisados segundo sua pertinência ao tema proposto, idioma e atualidade dos documentos.

2ª etapa: Leitura, interpretação, análise e sistematização das informações – leitura do material bibliográfico utilizado e selecionado para o desenvolvimento do embasamento teórico acerca das temáticas abordadas, visando elaboração da fundamentação teórica para identificação e posterior discussão de diferentes percepções na literatura a respeito do tema, a fim de criar alicerces para a elaboração (redação) do trabalho.

3ª etapa: Análise e estabelecimento das características fundamentais extraídas da literatura – encontro de pontos e características principais encontradas na literatura para análise referencial ao tema, contemplando a questão abordada pela pesquisa (como as duas abordagens se correlacionam e em sua somatória, vinculam-se ao domínio “dados” para disseminação da informação e conhecimento científico?) e sua elucidação.

4ª etapa: Sistematização do estudo qualitativo, exploratório de cunho teórico-aplicado – reflexão e análise de informações abrangendo o objetivo geral da presente pesquisa, que objetivou compreender a importância da Curadoria Digital para a Gestão de Dados de Pesquisa e identificação das relações existentes entre as abordagens.

5ª etapa: Elaboração e redação para o exame de Trabalho de Conclusão de Curso para BCI II (TCC II) – Apresentação realizada à banca examinadora sobre as considerações preliminares do estudo proposto.

6ª etapa: Elaboração de redação final e divulgação do Trabalho de Conclusão de Curso – após orientações e sugestões da banca examinadora, a elaboração de redação final foi desenvolvida, a partir das informações e conhecimentos adquiridos que sustentou as considerações finais do trabalho de conclusão de curso, visando

divulgação à comunidade científica dos resultados obtidos com o desenvolvimento do estudo.

A forma de análise dos resultados – posterior à Fundamentação Teórica – contemplou em estado da arte, o *Mapeamento para Análise dos Conceitos de Curadoria Digital e Gestão de Dados de Pesquisa*, a fim de realizar um estudo bibliométrico.

Nesta análise bibliométrica, objetivou-se coletar dados respondidos em formulário em consonância ao processo de revisão de literatura para posteriormente, construir indicadores de identificação da abordagem conceitual e teórica a respeito das temáticas, tal como, autores, ano de publicação, palavras-chave, periódicos, bases de dados, instituição ou localização geográfica, além da identificação da área do conhecimento com maior impacto e relevância frente às temáticas contempladas pelo trabalho.

1.5 Estrutura do Trabalho

Em capítulo de abertura – a *Introdução* – aborda aspectos iniciais relacionados a pesquisa e a contextualização da temática na Ciência da Informação, em tratamento e gestão da informação, bem como, o surgimento do domínio de dados diante das tecnologias, para comunicar a importância de atividades como a Curadoria Digital, a prática de Gestão de Dados de Pesquisa e um plano de gerenciamento de dados, interligando as duas temáticas apresentadas. O capítulo também dispõe a definição do problema de pesquisa, tal como, os objetivos (geral e específicos), a justificativa em relevância ao trabalho de investigação, os procedimentos metodológicos de pesquisa e a estrutura do trabalho.

Os demais capítulos e subcapítulos da investigação apresentam-se em seguinte sucessão:

Capítulo 2 – Fundamentação Teórica – O segundo capítulo apresenta a *Fundamentação Teórica*, contextualizando as temáticas de *Curadoria Digital e Gestão de Dados de Pesquisa*. O capítulo segue por uma perspectiva geral do estudo que tende a se afunilar, servindo de fundamento a construção da pesquisa. Inicia-se em abordagem ao *Conceito de dados*, abarcando os conceitos de dado primitivo, estruturado e não estruturado.

Em sequência, abrange os *Dados de pesquisa ou dados científicos* e explana acerca dos tipos de dados de pesquisa. Posteriormente apresenta o conceito de *Metadados*, em destaque aos *Tipos de metadados* e a importância dos *Metadados para armazenar e preservar dados científicos ou de pesquisa*. Explana-se, em sequência, acerca do *Ciclo de vida dos dados*, partindo da premissa de que estes padronizam o processo de Curadoria Digital, visto que, a curadoria abrange os processos de ciclo de vida dos dados. São apresentados alguns modelos de ciclo de vida dos dados, tais como, O Ciclo de Vida dos Dados do *Data Documentation Initiative* (DDI) – segundo Santana (2013) – e o Ciclo de vida dos dados do DataONE (2013), que propiciam que tais dados sejam tratados, processados e preservados de maneira adequada.

Faz-se importante ressaltar que no *corpus* da pesquisa, há mais um modelo enfatizado, sendo o Modelo de Ciclo de Vida da Curadoria Digital do *Digital Curation Centre* – DCC (2008), presente no subcapítulo *Mapeamento dos conceitos de Curadoria Digital* e através de elaboração própria, utilizando o software *Adobe Illustrator*, apresenta-se o Modelo do Ciclo de vida da Curadoria Digital do DCC em sua versão traduzida, objetivando facilitar a visualização direta e gráfica do ciclo em nosso idioma para posterior interpretação.

Capítulo 3 – Mapeamento para Análise dos Conceitos de Curadoria Digital e Gestão de Dados de Pesquisa – No terceiro capítulo, a investigação é conduzida a um *Mapeamento para Análise dos Conceitos de Curadoria Digital e Gestão de Dados de Pesquisa*, abarcando em subcapítulos, consecutivamente, o *Mapeamento dos conceitos de Curadoria Digital* e posterior *Mapeamento dos conceitos de Gestão de Dados de Pesquisa*, em que, discorre-se sobre o Plano de Gestão de Dados (PGD) ou *Data Management Plan* (DMP), apresentando conceitos, reflexões e sua importância em assegurar a integridade dos dados e fatores de autenticidade, validade, reprodução e replicação da pesquisa científica, contribuindo para organizações de P&D e para o progresso científico.

Apresentam-se também, os resultados alcançados, discussão e análise para interpretação dos dados atingidos. O capítulo também contém o subcapítulo *gênese*, em que se explana sobre *Estado da Arte e do Conhecimento*.

Capítulo 4 – Considerações Finais – Apresentam-se as considerações e reflexões finais sobre as temáticas em referência aos resultados atingidos durante a investigação, sendo eles qualitativos ou quantitativos. Incluindo breve retrospecto do que foi contemplado pela pesquisa, bem como, o alcance do objetivo principal, demais objetivos específicos e elucidação da problemática de pesquisa.

Referências – São apresentadas as referências consultadas que formam o *corpus* teórico do trabalho, dispondo autores pesquisadores e produções científicas que fundamentaram a investigação.

2 FUNDAMENTAÇÃO TEÓRICA

Como caracterizado por Santos e Santana (2013), faz-se necessária a percepção e compreensão de que o conceito de “dado” apresenta-se como elemento básico presente nos “fluxos informacionais” e tal conceito é variavelmente definido em consonância às distintas áreas da ciência sendo, portanto, esse conceito em específico, redirecionado à Ciência da Informação, embora expresso de acordo à Ciência da Computação.

Curty e Cervantes (2016, p. 1) mencionam que “Vivenciamos o quarto paradigma da ciência, o qual tem redefinido o *modus operandi* da práxis científica como consequência dos desafios impostos pela produção de dados em larga escala”. Os autores enfatizam também a era *big data* como a causa de demandas em maneiras estratégicas de gerenciar o grande volume e tipologias de dados em grandes organizações, cabendo mencionar aqui as agências de fomento e demais instituições.

Segundo Davis e Rush (1979 *apud* SANTOS; SANTANA, 2013, p. 201), “[...] a forma mais simples de definir dados é dizendo que eles são o resultado de mensuração ou observação [...]”, trazendo também a abstração de que, como salienta Burch et al. (1983 *apud* SANTOS; SANTANA, 2013, p. 201), “[...] dados são substitutos simbólicos, linguísticos, matemáticos ou outros, que são aceitos como atributos para representação de pessoas, objetos, eventos e conceitos [...]”.

Le Coadic (1996, p. 90-91) ao mencionar as evoluções técnicas eletrônicas de informação desde 1948, explana sobre a evolução que transpassa do analógico ao digital, sendo este o fenômeno a atribuir que “[...] Toda informação pode, de agora

em diante, ser processada, codificada, armazenada e transmitida pelos mesmos processos e nas mesmas redes".

Conceitua-se, dessa maneira, a informação também como objeto de processamento digital, integrando posteriormente e a exemplo, sistemas eletrônicos, armazenamento em bancos de dados e técnicas eletrônicas ou digitais que, em avanços e modificações posteriores, permitiriam a representação em meio digital, o armazenamento, a recuperação, o acesso e a disseminação de dados, estabelecendo relação ao processamento de dados e seu tratamento em meio digital.

Voltando para a década de 1960, marcada pelo impacto da internet¹ em ciência e tecnologia que culminou em modificações na ciência e adequação desta ao surgimento de informação em grande escala, correspondente ao volume de dados gerados em tempo mínimo de maneira exponencial e, concomitantemente, pesquisa, publicação e comunicação científica adequaram-se aos dados e explosão da informação junto à ciência e tecnologia.

Nota-se que estas, evoluíram e se consolidaram ao longo dos anos, processo evolutivo este considerado contínuo, acrescentando futuras inovações, vindas de sociedades modernas, bem como, as hoje denominadas eras da informação e conhecimento, além de informações e dados gerados sucessivamente em tempo real a transporem a sociedade contemporânea.

Conforme a disponibilização e a quantidade de dados aumentavam ao longo dos anos, o gerenciamento de dados tornava-se dificultoso para análises informacionais de avanços científicos e tecnológicos, causando impasse, sobrecarga, desarranjo e excesso de informações, trazendo a concepção de que, organização e visualização adequada dos dados são essenciais para a Ciência, pesquisadores e para o desenvolvimento pragmático científico.

¹ A internet foi inventada no final da década de 1960 no contexto da Guerra Fria através da rede de computadores criada pelo Departamento de Defesa dos EUA. Especificamente, no ano de 1969 entrou em operação a rede ARPAnet como um projeto do exército norte-americano visando transmitir informações militares sigilosas e interligar departamentos de pesquisa pelos EUA.

Para mais informações:

FESCINA, D. Quem inventou a internet? Super Interessante. 2016. Disponível em: <https://super.abril.com.br/mundo-estranho/quem-inventou-a-internet/>. Acesso em: 2 fev. 2021.

TARNOFF, B. How the internet was invented. The Guardian. 2016. Disponível em: <https://www.theguardian.com/technology/2016/jul/15/how-the-internet-was-invented-1976-arpa-kahn-cerf>. Acesso em: 28 jan. 2021.

O curto período tido como era digital atribuiu às instituições científicas, comprometimento em salvar dados científicos (SAYÃO; SALES, 2012). Nesse contexto, compreende-se dados científicos como registros realizados durante o processo de pesquisa (SIMIONATO, 2018) e dados de pesquisa como dados componentes centrais de um processo de pesquisa, sendo registros científicos que fundamentam resultados de pesquisa publicados em formato de teses, dissertações, artigos e demais trabalhos científicos. (SIBI USP, 2016).

De acordo com Hey e Trefethen (2003, p. 1820), a qualidade de como metadados descrevem os dados é considerada essencial, ao mesmo passo que a construção dos motores de busca para extração de informações significativas, construídos a partir dos metadados dispostos em documentos armazenados em formato eletrônico, passa a demonstrar que a qualidade do mecanismo de pesquisa torna-se tão satisfatório quanto os metadados que ele referencia, sendo este, portanto, o modo pelo qual os metadados se tornam vitais para armazenagem e preservação de dados científicos.

Nessa busca de gerenciar e manusear o dilúvio de dados da era de informação digital visto seu aumento em volume e velocidade, torna-se necessário o uso da Curadoria Digital que abrange a Gestão de Dados, ferramentas essas de preservação e métodos de gerenciamento, planejamento, proteção, utilização e reutilização de dados, principalmente em âmbito científico e tecnológico.

Dados produzidos a todo instante em meio à alta velocidade tecnológica, culminam rapidamente em obsolescência tecnológica e desgaste do suporte em ambiente digital, sendo esse o motivo pelo qual é necessário atender aos métodos de curadoria e gestão de dados.

Por esta perspectiva, Sayão e Sales (2012) trazem a problemática da obsolescência tecnológica no que tange a dados quando armazenados, desde que atividades de pesquisa se tornam cada vez mais vinculadas a materiais digitais e, por isso, a necessidade de cuidados específicos e uma gestão adequada de conteúdos:

[...] os dados de pesquisa [...] adormecerão armazenados em computadores e mídias pessoais que inexoravelmente serão tragados pela obsolescência tecnológica, pela fragilidade das mídias e, sobretudo, pela falta de intencionalidade de preservá-los adequadamente de forma que sirvam de ponto de partida para novas pesquisas. [...] Para que haja avanço do conhecimento científico com

um nível mais aceitável de duplicação de esforços, é necessário o estabelecimento de metodologias e compromissos de longo prazo que garantam a capacidade dos dados em formatos digitais, que estão sendo gerados agora, de serem acessados, interpretados e reutilizados com a tecnologia corrente à época de acesso. (SAYÃO; SALES, 2012, p. 180).

Isto posto, intensifica-se por Sayão e Sales (2012, p. 180) o quanto “[...] o arquivamento persistente, a preservação digital e o estabelecimento de modelos de informação para preservação de registros científicos estão se tornando questões-chave para as áreas da pesquisa”.

A Carta sobre a Preservação do Patrimônio Digital – *Charter on the Preservation of Digital Heritage* – de 15 de outubro de 2003, ao reconhecer que recursos de informação estão sendo cada vez mais produzidos, distribuídos, acessados, mantidos em formato digital, e criando um novo legado considerado herança digital, adota princípios de preservação, acessibilidade e políticas de salvaguarda do patrimônio digital.

A Carta reúne princípios de acesso ao patrimônio digital, proteção contra a perda de patrimônio, continuidade digital, medidas para o desenvolvimento de estratégias e políticas de preservação do patrimônio digital. Também contempla disseminação de resultados de pesquisa e melhores práticas de democratização do acesso a técnicas de preservação digital. (UNESCO, 2003).

Desse modo, o acesso a herança digital poderá oferecer oportunidades de criação, comunicação e compartilhamento de conhecimento, sendo necessária sua preservação para o benefício de gerações presentes e futuras, enfatizada também como um tópico urgente de preocupação mundial.

Cabe assim relacionar os princípios da Carta de Preservação às medidas de armazenagem, acesso e preservação de dados científicos ou de pesquisa, utilizando-se da curadoria digital como um processo de preservação a longo prazo dos dados digitais a abranger o gerenciamento de dados de pesquisa, salvaguardando-os em perpetuidade.

Sucedese que, para melhoria de planejamentos de gestão de políticas de dados, faz-se necessário compreender que os dados se enquadram às atribuições de processos de um ciclo de vida de pesquisa, relacionando ao contexto apresentado a questão intrínseca a desenvolver-se nesta pesquisa em: como as

duas abordagens se correlacionam e em sua somatória, vinculam-se ao domínio “dados” para disseminação da informação e conhecimento científico?

Em resumo, enfatiza-se a importância que a ciência traz aos dados e sua gestão, como os Planos de Gerenciamento de Dados (PGD) e compartilhamento de dados coletados, visto a importância da gestão de dados de pesquisa desde o acesso, curadoria, gestão, qualidade, segurança e resultados positivos do gerenciamento de dados e curadoria digital.

Além do elo entre o planejamento, preservação e melhor disposição de dados de pesquisa e publicações científico-acadêmicas quando bem orientadas, levando ao “[...] surgimento de concepções surpreendentes de documentos mais apropriados ao paradigma da ciência computacional e orientada por dados.” (SAYÃO; SALES, 2012, p. 182).

Cabe salientar que um eficiente plano de gestão de dados de pesquisa que envolva manipulação, armazenamento, acesso, agregação de valor, entre outros processos anteriormente citados viabilizam desenvolvimento ao campo de pesquisa, dada a natureza dos dados de pesquisa (algoritmos, imagens, números, etc.), bem como, sua classificação (experimental, observacional, computacional, etc).

O plano de gestão busca estabelecer como esses dados serão coletados, processados e gerados durante um desenvolvimento de projeto de pesquisa, tratado por meio de todo o processo, do início ao fim da curadoria digital de dados de pesquisa, descrevendo padrões e metodologias a serem utilizados e posterior compartilhamento e disponibilização dos dados, prevendo preservação e reuso. (SAYÃO; SALES, 2015).

Neste cenário, pretendeu-se no desenvolvimento da pesquisa, recorrer às produções relacionadas, de modo geral, a Curadoria Digital e Gestão de Dados pelos conceitos de diversos pesquisadores, entre eles, Abbott (2008); Ball (2012); Bell, Hey e Szalay (2009); Bohle (2013); Hey e Trefethen (2003); Le Coadic (1996); Santana (2013), acerca do ciclo de vida dos dados; Santos e Santana (2013); Santos e Vidotti (2009); Sayão e Sales (2012; 2013; 2015; 2016), em destaque aos trabalhos com temas relacionados à curadoria digital, preservação de dados digitais de pesquisa, curadoria de dados e guia de gestão de dados de pesquisa e; Simionato (2018), referente a gestão de dados de pesquisa.

A partir do retrospecto descrito para a condução da pesquisa, as seções organizadas a seguir apresentam maior aprofundamento teórico, de modo a

destrinchar em seu desenvolvimento os tópicos relacionados, facilitando a compreensão do tema e tornando-o melhor elucidado junto à problemática de pesquisa.

2.1 Conceito de dados

A principiar o conceito de dados, a unanimidade que se encontra frente ao termo “dado” ou “dados” compele: elemento, informação ou conjunto de informações, conhecimento, caracteres e símbolos, a base para um ponto de partida ou construção de um conhecimento.

Para tanto, define Setzer (1999): “[...] *dado* como uma sequência de símbolos quantificados ou quantificáveis.”, exemplificando o autor, que determinado texto é um dado, em que as letras são os símbolos quantificados e que a constituição do alfabeto é por uma base numérica, como os dados são imagens, sons, eventualmente símbolos quantificáveis ou quantificados que constituem uma sequência de dados e, como símbolos quantificáveis, estes podem ser armazenados por máquinas e processados pela mesma.

Bertin, Visoli e Drucker (2017, p. 38) trazem que “Embora não exista uma definição consensual na literatura, pode-se afirmar que dados compreendem o elemento tangível, ou o registro do qual derivam a informação e o conhecimento.”

Nota-se que as múltiplas definições referidas à terminologia cabem ao seu contexto de uso e em seu sentido, passam por alterações de acordo com a área do conhecimento que o aborda, interpretado tal como uma entidade matemática, textual, administrativa, informacional, entre outros contextos do conhecimento humano a traduzi-lo de acordo com suas possibilidades interpretativas, tornando-o um vocábulo amplamente utilizado de maneiras distintas.

A Ciência da Informação, sendo de natureza interdisciplinar a abranger a informação desde sua origem aos processos de tratamento de dados e transformação desses em conhecimento, preocupando-se com a coleta, análise, tratamento, armazenamento, maneiras de recuperação e disseminação da informação, traz na abordagem aos dados o acesso à informação e apropriação do conhecimento envolvendo a manipulação de dados, como bem destacam Santos e Santana (2013).

Dessa maneira, falar de Ciência da Informação compreende falar de fluxos informacionais aos quais os dados são elementos básicos vinculados ao processo de geração da informação e que, retratada por Santos e Santana (2013, p. 200) “A Ciência da Informação busca soluções para responder à necessidade informacional da sociedade enquanto ciência social aplicada [...]”.

Le Coadic (1996) interpreta dados como a representação codificada da informação que permite ser processada em meio eletrônico, contemplando também em seu conceito, o sistema de gerenciamento de bases de dados que, por sua vez, contém grande quantidade de arquivos de dados e as relações entre eles. Dessa maneira, coloca-se que uma base de dados, uma vez gerenciada por um sistema, visa descrever e manipular dados, mantendo também a integridade dos mesmos.

Baca (2016 *apud* SIMIONATO, 2017) explana que dados se caracterizam por uma estrutura mínima constituída, podendo multidimensionar possibilidades referentes a acesso, uso e reuso que, na área de Ciência da Informação, vem se tornando um campo de estudo em constante progresso. Ademais, relaciona-se que a representação, manutenção e organização informacional também ganha espaço neste estudo, uma vez que busca-se por tratamento da informação e padronização adequados e hábeis em termos de organização da informação.

Fry e Sibley (1976, p. 14) compreendem dados como um conjunto de fatos, em que, o objetivo do processo de estruturação de dados envolve a coleta de dados, sendo estes, por sua vez, um conjunto de dados que podem ser registrados, em sentido de armazenamento destes quando se dialoga sobre bases de dados, uma vez que, a definição dos autores acerca de dados pauta-se sobre entidades e atributos em bases de dados.

Tal como consideram Santos e Santana (2013, p. 208), tratando da área de Ciência da Informação e enfatizados aqui como uma das mais detalhadas abordagens conceituais sobre dados, “[...] o tratamento da informação, a representação de recursos, a recuperação e a disseminação de informação se tornam áreas vinculadas à descrição, ao armazenamento, à preservação, ao acesso e à gestão de dados.”

Observa-se que a noção de dados pode variar consideravelmente entre pesquisadores e, ainda mais, entre áreas do conhecimento, tornando visível que, sendo os dados gerados por finalidades distintas, diferentes pesquisadores e comunidades científicas e acadêmicas, além da compreensão de que os dados são

também gerados por meio de diferentes processos, denota a multiplicidade existente no mesmo, que o leva a existência de diferentes tipologias para seus devidos fins, de acordo com sua criação, tipo e diferentes interpretações que possui também para cada pesquisa.

Enfatiza-se os dados científicos ou de pesquisa como preocupação emergente para o progresso científico, visto que, os dados foram sempre componentes essenciais de validade, disseminação e reprodução da pesquisa.

No contexto de “dados” é possível que se encontre este em sua gênese: o dado bruto ou dado primitivo/primário, como aquele que não passou por nenhum tipo de tratamento ou alteração, permanecendo em seu estado original. A exemplo, Semidão (2014, p. 71) traz em sua acepção que “Dado [...] figura como ‘fato’ ou ‘átomo’, um elemento bruto, desprovido de significação imediata. Algo perceptível e de baixo teor semântico”.

Para Piwowar, Day e Fridsma (2007), além de serem usados para confirmar os resultados originais, os dados brutos podem ser usados para explorar novas hipóteses, bem como, hipóteses relacionadas ou derivadas, principalmente quando combinadas com outros conjuntos de dados disponíveis ao público. Tais dados são indispensáveis ao investigar e desenvolver métodos de estudo, técnicas de análise e implementações de software.

Isso traz benefícios também para a comunidade científica, visto que, o compartilhamento de dados incentiva perspectivas diversas; auxilia na identificação de erros, desencorajando a fraude e torna-se útil para o treinamento de novos pesquisadores, aumentando o uso eficiente de recursos financeiros e evitando a coleta duplicada de dados.

Desse modo tem-se que dados primários carecem de processamentos adequados para posterior interpretação, configurando a necessidade de métodos para lidar com a carga exponencial de dados gerados ou coletados que em primeiro momento, possuem baixo teor semântico necessitando tratamento e filtragem.

Santos e Santana (2013), irão refletir acerca de dados estruturados e não estruturados, trazendo a relevância de metadados para a estrutura semântica de determinado dado, uma vez que, por meio da estrutura (ou semântica posicional), pode-se viabilizar o uso e reuso de dados. Os autores exemplificam a diferenciação entre dados estruturados e não estruturados analisando dois exemplos de dados,

sendo o primeiro um livro e o segundo, um registro bibliográfico, enfatizando ambos em formato digital,

[...] tem-se a percepção de que o livro é não estruturado e que o registro bibliográfico é estruturado. O que configura a diferenciação é que, no registro bibliográfico, sua estrutura interna está explícita, ou seja, seu conteúdo é fixado por um código, e no livro a identificação de sua estrutura requer um interpretador externo, o que o configura como um dado não estruturado. O mesmo poderia ser aplicado a uma música ou vídeo disponíveis em ambientes digitais. (SANTOS; SANTANA, 2013, p. 206).

Na concepção dos autores, compreende-se o livro como um dado não estruturado, visto que, para a identificação de seu conteúdo precisa-se de um interpretador externo, uma vez que a identificação de sua estrutura não está explícita e o registro bibliográfico, um dado estruturado, pelo seu conteúdo (estrutura interna) estar explícito por meio de um código (metadados).

Os autores explanam que dados em geral e principalmente, dados não estruturados – para que sejam interpretados – requerem o uso de metadados (simplificando: dados sobre dados), enquanto que, já existindo no conteúdo interno do dado estruturado uma semântica já explicitada, o conteúdo desse dado já está passível de interpretação de maneira autônoma por máquina.

A perspectiva de que dados não estruturados exigem dos sistemas um esforço maior, é bem colocada por Lima (2015) quando este denota que principalmente pelo aumento de dispositivos móveis conectados, tal como, usuários da web, perpetua-se que a produção de dados não estruturados cresça de forma exponencial e, por conseguinte, exige-se dos sistemas maior atenção para os processos de coleta, recuperação e análise.

Isto é, enquanto a informação de “dados organizados em um modelo conceitual” pode ser melhor recuperada, visto que os dados estruturados contêm sua organização determinada para serem recuperados, fazendo com que a informação seja recuperada por consultas em linguagem padrão; os dados não estruturados, a exemplo de um texto, possui conteúdo ambíguo e ruídos, tal como, sentimentos expressos, contendo informações “carregadas de desafios para uma máquina” – como bem salienta Lima (2015, p. 431) – poder extrair informação e conhecimento

Entende-se, dessa maneira, que um dado estruturado permite que seu conteúdo seja interpretado de maneira autônoma por máquinas, enquanto que a interpretação de dados não estruturados carece do uso de metadados (representação de dados), enfatizando que metadados são também considerados dados estruturados que permitem que “dados” possam ser identificados e interpretados numa percepção de que metadados registram, interpretam, explicam, descrevem, localizam e viabilizam a recuperação, por meio da representação e da estrutura padronizada, tornando-os melhor preservados e acessíveis.

2.2 Dados de pesquisa ou dados científicos

Os autores Sayão e Sales (2015, p. 7) trazem em seu *Guia de Gestão de Dados de Pesquisa* (2015) que o relatório *Princípios e diretrizes para o acesso a dados de pesquisa financiados por fundos públicos* do ano de 2007 da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), descreve dados de pesquisa como “registros factuais” ou registros reais que como fontes primárias, próximas da ideia em estudo, são utilizados para a pesquisa científica, sendo considerados por pesquisadores e aceitos como necessários para a validação de resultados de trabalho científico.

Como bem evidencia a denominação, dados científicos são provenientes de pesquisas científicas, as quais geram-se dados científicos de acordo com a finalidade de cada pesquisa ou comunidade acadêmica. Qin et al. (2010, p. 128) retrata que o termo "metadado científico" é comumente utilizado ao que se refere a dados que descrevem conjuntos de dados coletados e que são gerados por meio de pesquisas científicas.

Dados de pesquisa, como dados provenientes da produção científica pautam-se como alicerces para a produção e desenvolvimento científico e acadêmico em diversas áreas do conhecimento.

[...] estudos têm apontado que dados de pesquisa bem organizados, documentados, preservados, acessíveis e verificados quanto à sua acurácia e validade são mais facilmente compartilháveis e reutilizáveis, ocasionando: a ampliação do impacto, da visibilidade e da credibilidade do pesquisador [...]. (BERTIN; VISOLI; DRUCKER, 2017, p. 37-38).

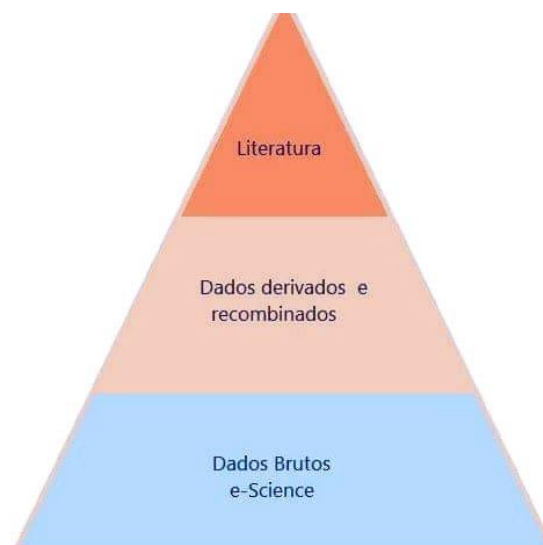
Ressaltando também que, não são todos os dados derivados de pesquisas que são publicados em artigos científicos, tal como, mesmo que tais dados tenham sido gerados em determinada conjuntura e finalidade de pesquisa, estes podem ser utilizados em outros tipos de projetos de pesquisa até mesmo com vertentes distintas e por outra comunidade acadêmica ou pesquisadores. Faz-se a importância em gerenciar, compartilhar e preservar os dados de pesquisa, visto que sua reutilização pode servir para a recriação de novas ciências. (BERTIN; VISOLI; DRUCKER, 2017, p. 38).

Nessa conjuntura, dados de pesquisa contemplam todo o tipo de registro produzido, gerado, compilado, utilizado ou reutilizado ao longo do desenvolvimento da pesquisa, em que, o que é entendível por sendo dado de pesquisa, em prática tende a variar conforme o contexto, disciplina, área do conhecimento e finalidade, como já destacados.

Bertin, Visoli e Drucker (2017, p. 38) retratam que para que os dados de pesquisa sejam verdadeiramente úteis, estes devem dispor de estrutura e organização e, nesse sentido, configura-se relevância e necessidade de desenvolvimento de mecanismos, ferramentas e práticas para o gerenciamento de dados de pesquisa em âmbito científico, acadêmico e organizações de Pesquisa e Desenvolvimento (P&D).

A figura 1 a seguir, retrata a estrutura dos dados científicos, de acordo com Costa e Cunha (2014) fundamentada em Gray (2007):

Figura 1 - Estrutura dos dados científicos



Fonte: Elaboração própria adaptada de Costa e Cunha (2014, p. 194) – elaborada pelos autores fundamentada em Gray (2007).

A figura 1 denota que, segundo Gray (2007 *apud* COSTA; CUNHA, 2014, p. 194), a estrutura dos dados científicos contempla:

- Dados Brutos; *e-Science*: provenientes do processo de coleta de dados através de controle laboratorial, por meio de sensores especializados, telescópios, satélites, ensaios de laboratórios, etc.
- Dados derivados e recombinaados: provenientes do processo inicial de análise de dados coletados que é tratado pelos bibliotecários, além de armazenado em repositórios, em alguns casos.
- Literatura: procedente do processo de comunicação científica e revisão por pares, tal como, tratado pelos bibliotecários e armazenados em bibliotecas.

Dessa maneira, como bem destacam Sayão e Sales (2015), faz-se importante considerar que dados científicos ou de pesquisa portam tempo de vida maior que a dos próprios projetos de pesquisa que os geraram e, por isso, é necessário que o uso e reuso desses dados possam ser viabilizados para que se reinicie um novo ciclo de vida dos dados de pesquisa.

Nesse processo, revelam-se os dados em sua forma bruta, desprovidos de significado em primeiro momento até que seu tratamento e interpretação sejam realizados para posterior uso, reuso e acesso, uma vez que, pode-se considerar dados científicos como a premissa pela qual cientistas-pesquisadores produzem, replicam e validam novos conhecimentos científicos e, portanto, carecem de técnicas que promovam a gestão dos dados, deste sua criação até seu acesso, disseminação e preservação.

A interpretação relacionada a dados varia entre áreas do conhecimento e seus pesquisadores, a um montante pelo qual estes também são gerados por diferentes processos, objetivos, pesquisadores, diferentes constatações e principalmente, enquadrando-se em diferentes contextos trazendo ao vocábulo sua multiplicidade interpretativa.

Nesta conjuntura de multiplicidade encontram-se as diferentes tipologias de dados, desde a visão de dados como sendo elementos, conjunto de informações, caracteres e símbolos, como também, números, vídeos, imagens, textos, algoritmos, equações, *software*, entre outros exemplos, como bem salientam Sayão e Sales (2015).

Os autores ainda destacam que tais tipos de dados possuem valor duradouro e imediato, enquanto outros adquirem valor apenas ao longo do tempo, ao mesmo passo que enquanto capturam-se alguns dados em momentos específicos e irrecuperáveis, outros dados são passíveis de reprodução:

Essa heterogeneidade intrínseca aos dados de pesquisa implica que é necessário formular estratégias de gestão de amplo espectro que englobem os vários tipos de dados. O reconhecimento dessas diferenças torna-se crucial quando se estabelecem as opções gerenciais e tecnológicas para o arquivamento persistente e para a curadoria digital das coleções de dados de pesquisa. (SAYÃO; SALES, 2015, p. 7).

Dessa maneira, ao referir-se em “heterogeneidade dos dados”, discorre-se que a diversidade de dados exige diferentes tipos de dados para distintas finalidades, tornando necessária uma gestão de dados em que se conheça sua tipologia e o estabeleça em armazenamento adequado, bem como processamento e acesso.

Sayão e Sales (2015, p. 8) explanam que dados de pesquisa podem ser classificados quanto sua natureza, origem ou "status no fluxo de trabalho da pesquisa" e, por isso, são caracterizados de diferentes maneiras.

Os tipos de dados de pesquisa classificam-se de acordo com sua origem como:

- **Dados Observacionais:** dados obtidos por meio de observações diretas; que, como descrevem Sayão e Sales (2015, p. 7), tais dados observacionais que possuem natureza instantânea, são qualificados como registros históricos. Sendo instantâneos, não são coletados uma segunda vez e, dessa maneira, necessitam de processos de curadoria para preservação contínua. Exemplos: lugares ou tempos específicos, tal como uma fotografia de um evento em determinado tempo, em determinada data específica, entre outros.
- **Dados Computacionais:** esse tipo de dado provém de resultados de execução de modelos computacionais ou simulações. Tal como salientam Sayão e Sales (2015, p. 8), estes dados podem ser replicados e, por isso, sua preservação a longo prazo não se faz necessária. Porém, os autores destacam que, sendo esses dados de ambientes virtuais, a futura réplica do modelo ou simulação pode exigir maior quantidade de informações, como

descrição de dependências de *software*, *hardware*, dados de entrada, bem como, dependências técnicas, e em alguns momentos faz-se necessário preservar somente os dados de saída. Exemplo: criação de ambientes virtuais (culturais ou educacionais).

- Dados Experimentais: são tipos de dados obtidos em laboratórios, através de experimentação ou por meio de situações controladas. Exemplo: medidas de reações químicas.

Sendo a classificação de dados de pesquisa considerando sua natureza:

- Números;
- Imagens;
- Vídeos ou Áudio;
- *Software*;
- Algoritmos;
- Equações;
- Animações ou Modelos;
- Simulações.

Por fim, os tipos de dados de pesquisa, segundo sua fase de pesquisa:

- Dados Brutos, Crus ou Preliminares: dados provenientes diretamente de instrumentos científicos;
- Dados Derivados: dados provenientes de resultados obtidos por meio de processamento ou combinação de dados brutos ou outros dados;
- Dados Canônicos ou Dados Referenciais: coleções de dados consolidados e arquivados, como bem destacam Sayão e Sales (2015, p. 9), comumente são arquivados em grandes centros de dados. Exemplos: estrutura química; sequência genética; entre outros.

Como dito anteriormente e o que também objetiva esta pesquisa, para que haja um efetivo gerenciamento e posterior compartilhamento desses dados, de maneira a impulsionar o avanço científico e tecnológico auxiliando pesquisas e pesquisadores a progredirem cientificamente, faz-se necessário a compreensão do processo integral dos dados, origem, suas tipologias e finalidade até que estes se tornem informação.

2.3 Metadados

A princípio, o termo metadados foi utilizado em contextos de sistemas de bancos de dados com a finalidade de descrever e controlar a gestão, bem como, uso dos dados. A criação de metadados também foi direcionada a facilitar descobertas e recuperação de informações relevantes a determinado contexto, seja em ambiente biblioteca ou web. (SAYÃO, 2010). Nesse contexto e, sendo uma ferramenta que viabiliza descrição, os metadados quando bem gerenciados implicam na descoberta de recursos e identificação de informações precisas a que se busca. É dessa maneira que, como considera Simionato (2017), “Os metadados são conceituados na Ciência da Informação, como os valores e atributos referentes aos recursos informacionais”.

Smiraglia (2005 *apud* SANTOS; SIMIONATO; ARAKAKI, 2014, p. 150) narra que na década de 1960, o presidente da *The Metadata Corporation*, Jack E. Myers, evidenciou metadados como aquele que descreve conjuntos de dados a serem utilizados em representação, organização e localização de recursos, podendo trabalhá-los de diferentes maneiras, de maneira a aplica-los no gerenciamento de recursos informacionais, constituindo-se em unidades menores do que os dados, de fato, representam.

Os metadados explicitam os diferentes aspectos do recurso que descreve: sua estrutura, conteúdo, qualidade, contexto, origem, propriedade e condição. E auxiliam na organização, favorecem a interatividade, validam as identificações e asseguram a preservação e principalmente, otimizam o fluxo informacional melhorando o acesso aos dados e a localização dos recursos informacionais. (SANTOS; SIMIONATO; ARAKAKI, 2014, p. 150-151).

Conforme enfatiza Dziekaniak (2007), utilizar metadados proporciona melhorias em busca e avaliação de informações, tal como melhor desenvoltura em pesquisas e manutenção desses dados. Nota-se metadados como dados que estruturam um dado principal.

Os metadados são a descrição de dados de maneira organizada e representada, sendo eles a informação estruturada que visa descrever, localizar, interpretar, explicar e também possibilitar que determinado recurso informacional seja recuperado, facilitando o seu gerenciamento e, nesse sentido, metadados sendo dados estruturados, no contexto de dados de pesquisa, viabilizarão a qualidade, o uso e reuso desses dados, descrevendo e tornando o conteúdo

explicitado através de uma forma estruturada aos conjuntos de dados gerados ou coletados por uma pesquisa.

Em adendo cabe destacar que, utilizando-se de padrões de metadados, configuram-se definição e descrição de metadados de modo a possibilitar, busca e recuperação de dados mais eficazes, facilitando manutenção da pesquisa por meio da descrição dos dados e, por conseguinte, também a manutenção desses dados. Dessa forma utilizando-se de representações adequadas, proporciona-se aos dados a padronização, remanejando-os para uma melhor qualidade, portanto, proporcionando melhor gerenciamento destes.

Dziekaniak (2007, p. 235) apresenta que pela ação de compartilhar os dados por meio de padrões de metadados, o tempo no desenvolvimento de pesquisas é minimizado, reduzindo também o tempo que se investe ao tratamento das informações geradas pela mesma e, dessa forma, a partir do uso e aproveitamento de um padrão de metadados, as comunidades científicas passam a agregar valor ao processo de tratamento de suas bases de dados, através da atividade de importar e exportar dados entre si, de maneira a homogeneizar a tradução desses dados por meio de uma estrutura genérica.

É dessa maneira que, metadados situam-se como principal fator ao que se objetiva acesso, interpretação, compartilhamento, disseminação, além dos fatores de preservação aos conjuntos de dados. Tornando-os passível de compreensão e interpretação, estes auxiliarão em processos de localização, recuperação, uso e reutilização.

Vale salientar que, utilizar padrões ou esquemas de metadados para gerenciamento de dados e representação dos mesmos, tornará mais eficaz a via que oferta melhor armazenamento de dados, localização e recuperação por meio da padronização, organização e descrição possibilitada pelos padrões de metadados.

2.3.1 Tipos de metadados

São existentes diversas categorias de metadados, como: metadados administrativos, metadados descritivos, metadados de preservação, metadados técnicos ou estruturais e metadados de uso, como bem ilustrado e explorado em maior profundidade por Gilliland-Swetland (2002, p. 5 *apud* LOURENÇO, 2007, p. 68), o quadro 1 a seguir, retrata os tipos de metadados quanto à sua função:

Quadro 1 - Tipos de metadados quanto à sua função

Tipo	Definição	Exemplos
Administrativo	Metadado utilizado na administração de recursos de informação	<ul style="list-style-type: none"> • Aquisição de informação • Direitos de reprodução • Critérios de seleção para digitalização etc.
Descritivo	Metadado para descrição de recursos de informação	<ul style="list-style-type: none"> • Catalogação de registros • Índices especializados etc.
De Preservação	Metadado utilizado para preservação de recursos de informação	<ul style="list-style-type: none"> • Documentação das condições físicas dos recursos etc.
Técnico	Metadado utilizado para conhecer as funções de um sistema ou o comportamento dos metadados	<ul style="list-style-type: none"> • Hardware e software • Dados de segurança • Documentação etc.
De Uso	Metadado relativo ao nível e tipo de uso de um recurso de informação	<ul style="list-style-type: none"> • Registros de exibição • Sumário de reuso e de versões etc.

Fonte: Elaboração própria adaptada de Gilliland-Swetland (2002, p. 5 *apud* LOURENÇO, 2007, p. 68).

O quadro 1 ilustra os cinco níveis do contexto metadados em seus diferentes aspectos, desde o que descreve estrutura, conteúdo, propriedades, entre outros, de recursos informacionais, viabilizando também acesso e recuperação dos mesmos, a se exemplificar por seus tipos e funções evidenciadas por Gilliland-Swetland (2002, p. 5 *apud* LOURENÇO, 2007, p. 68), seja na administração, descrição, preservação de recursos da informação ou a nível de se conhecer o comportamento dos metadados.

Sendo os metadados administrativos responsáveis pela gestão de recursos informacionais; metadados descritivos, responsáveis pela descrição que viabiliza a identificação do recurso informacional; metadados de preservação, responsáveis

pela conservação e preservação do recurso informacional; metadados técnicos, como explicitado no quadro, responsáveis pelo funcionamento do sistema e comportamento dos metadados e por fim, os metadados de uso, a abrangerem o nível e tipo de uso de um recurso informacional.

2.3.2 Metadados para armazenar e preservar dados científicos ou de pesquisa

Considera-se que, por meio dos metadados representa-se de maneira a descrever, interpretar, identificar e recuperar conteúdo e recursos informacionais. Este processo de recuperação da informação desencadeia resultados de buscas precisos quando tais dados científicos são bem gerenciados e armazenados.

Nesta perspectiva, em suas diferentes tipologias, metadados tornam-se essenciais para interpretação e acesso de recursos informacionais digitais e processos de gestão relacionados a esses recursos ao que tange também dados de pesquisa.

Na implementação de metadados, deve ser realizada uma análise de domínio, ou seja, devem ser criadas estruturas que garantam a utilização de padrões de descrição e que favoreçam a interpretação de cada metadado. Sendo que esses metadados atuarão como possíveis pontos de acesso tanto ao recurso isoladamente, quanto aos seus relacionamentos com outros recursos informacionais. (SANTOS; SIMIONATO; ARAKAKI, 2014, p. 151).

Dessa maneira, a partir do efetivo gerenciamento de dados de pesquisa e, por meio do uso de metadados para sua representação, uma vez que a representação interfere na identificação eficaz de determinado registro informacional, garante-se o armazenamento efetivo de dados, propiciando juntamente consistência na descrição dos dados científicos, de modo que este registro possa ser conservado, utilizado, reutilizado em outros contextos interpretativos perpetuando sua preservação.

Tal como salienta Sayão (2010, p. 3), a preservação da informação digital a longo prazo envolve um conjunto de variáveis, desde planejamento, à própria tecnologia e a ideia de disponibilizar patrimônios informacionais para futuras gerações.

Esta preservação será consolidada, bem como, todas as variáveis que a cercam serão resolvidas “[...] a partir da identificação de um conjunto de dados e

informações, expressos na forma de metadados, que ancorem os processos de gestão da preservação digital.”

Abrange-se metadados de preservação num contexto em que estes possuem a função de “[...] instruir e documentar os processos de preservação digital de longo prazo, garantindo que os conteúdos digitais possam ser acessados e interpretados no futuro” (SAYÃO, 2010) e os metadados como um conjunto de dados que podem abrigar informações também eficazes para o gerenciamento da preservação desses dados, além da representação dos dados em si, pode contribuir para a preservação dos dados científicos ou de pesquisa, desde que bem gerenciados, descritos e armazenados para uso e reuso de futuras gerações.

Consolida-se a importância dos metadados para armazenar ou preservar dados científicos ou de pesquisa, uma vez que estes, contemplando fins de acesso, interpretação, compartilhamento e disseminação, serão fator preponderante para a preservação desses conjuntos de dados e registros científicos coletados e gerados pelo processo de pesquisa, tornando-os passíveis de interpretação e compreensão para pesquisas científico-acadêmicas futuras.

O tópico central a sua relevância está em como metadados representam dados provenientes de pesquisa, responsabilizando-se justamente pela estruturação, organização, descrição e disponibilização desses dados que, por sua vez, não se apresenta como tarefa fácil em gerenciamento e estruturação, uma vez que, o processo considera não apenas coletá-los e sim, estudar maneiras eficazes de como coletar, como descrever, onde armazenar e como preservar a longo prazo, de modo a assegurar sua qualidade, salvaguarda e acesso, possibilitando juntamente seu reuso e integração a novos dados.

2.4 Ciclo de vida dos dados

Partindo da premissa de que a “[...] curadoria de dados de pesquisa permite que os dados possam ser tratados, arquivados em ambientes digitais confiáveis, preservados e reconfigurados de forma que possam ser aplicados em novos contextos científicos [...]” (SAYÃO; SALES, 2013, p. 4), tem-se que os ciclos de vida dos dados padronizam o processo de curadoria digital. (BALL, 2010).

Sayão e Sales (2016) explanam sobre a Curadoria Digital que abrange os processos de ciclo de vida dos dados:

O uso e a geração intensiva de dados pelas atividades acadêmicas e de pesquisa criam a necessidade urgente de infraestruturas gerenciais e tecnológicas que tratem de forma dinâmica o ciclo de vida dos dados – do seu planejamento até a seu arquivamento confiável – e insira esses ativos na infraestrutura mundial de informação para a pesquisa. Neste ponto que os procedimentos práticos e teóricos, que coletivamente chamamos de curadoria digital, se inserem como essencial para a captura, tratamento, preservação, arquivamento e acesso, pelo tempo que for preciso, dos dados de pesquisa considerados de valor contínuo. (SAYÃO; SALES, 2016).

Considerando a curadoria digital como um conjunto de ações de planejamento e gestão de dados até processos de preservação dos dados digitais por meio de seu ciclo de vida, torna-se atividade essencial para o tratamento abrangente aos dados de pesquisa, de valor contínuo e a longo prazo, contribuindo para gerações futuras e comunidades científicas.

Santana (2013, p. 1) retrata que faz-se necessário que se estude e proponha melhorias em todas as fases do processo que vão desde o planejamento até a visualização dos dados, exigindo uma proposta de um modelo de identificação e descrição de tais fases que, por conseguinte, propiciará o alinhamento de esforços, recursos e resultados diante do Ciclo de Vida dos Dados.

Cabe enfatizar a existência de diversos métodos para a gestão de ciclo de vida dos dados com algumas variáveis presentes tanto nos modelos, quanto nas etapas e práticas do ciclo. Estas variáveis se encontram em consonância a comunidade em que foi criada, para qual finalidade e a qual domínio abrangente.

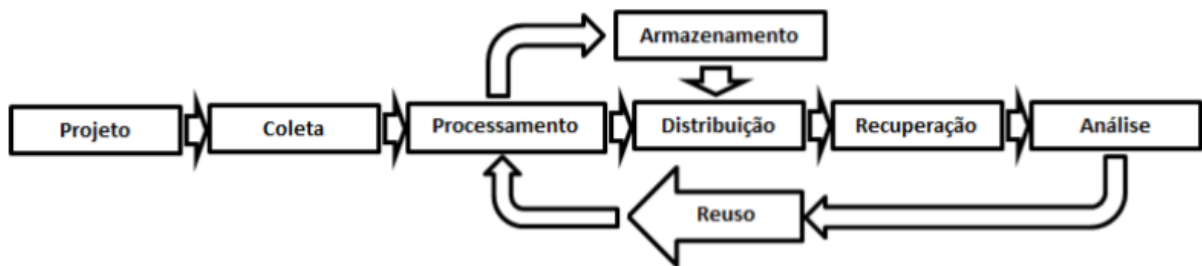
Diante disso, e pela multiplicidade de modelos existentes em referência ao ciclo de vida dos dados, destacam-se alguns modelos como, por exemplo, o Ciclo de Vida dos Dados do *Material Data Management Consortium* (MDMC), Ciclo de Vida de Dados de Pesquisa da *University College London* (UCL), o Ciclo de Vida de Dados do *Data Documentation Initiative* (DDI), Ciclo de Vida de Dados do DataONE, entre outros.

Como bem explicitado, os modelos de ciclo de vida dos dados propiciam que os dados sejam tratados, processados e preservados de maneira adequada, objetivando conduzir a fase de planejamento para a realização eficiente da gestão de dados. Um dos modelos de ciclo de vida de dados digitais anteriormente apresentado, refere-se ao Modelo de Ciclo de Vida da Curadoria Digital do DCC, *Curation Lifecycle Model* (2008).

O *Data Documentation Initiative* (DDI), segundo Santana (2013, p. 1) refere-se a um modelo que fundamentou outros, bem como, traz características em iniciativas de gestão de dados de biblioteca como referência mundial, tendo sido adotado pelo sistema de bibliotecas do *Massachusetts Institute of Technology* (MIT).

A figura 2, a seguir, apresenta o modelo de Ciclo de vida dos dados do DDI:

Figura 2 - Ciclo de Vida dos Dados do DDI

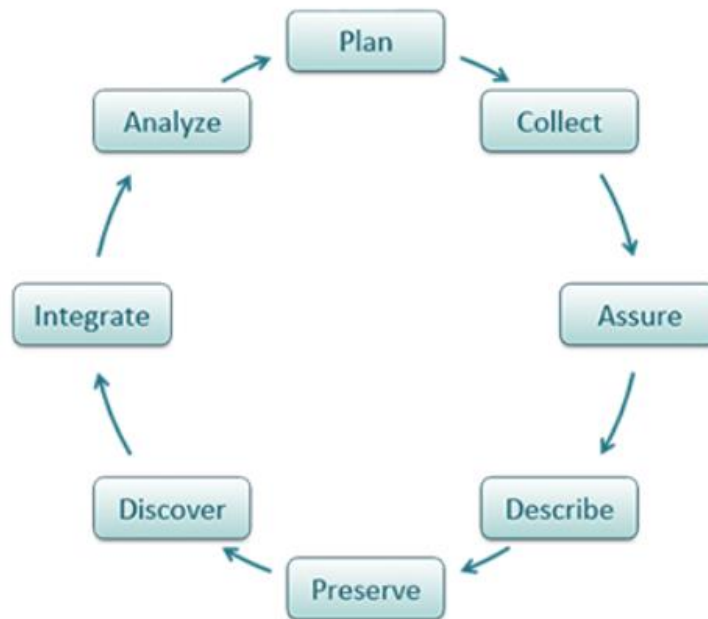


Fonte: Santana (2013, p. 3) adaptado de DDI (2004).

A figura 2 apresenta as fases do Ciclo de Vida dos Dados do DDI, compostas por: Projeto, Coleta, Processamento, Armazenamento/Distribuição, Recuperação, Análise, Reuso, Processamento, Armazenamento e Distribuição em processos contínuos.

Outro modelo de ciclo de vida de dados digitais a considerar-se, refere-se ao Modelo do *Data Observation Network for Earth* (DataONE) – Rede de Observação de Dados para a Terra, como apresenta a figura 3 a seguir:

Figura 3 - Ciclo de vida dos dados – DataONE



Fonte: DataONE² (2013).

A figura 3 retrata que, segundo Sayão e Sales (2015), o ciclo de vida do DataONE compreende oito etapas consistentes em:

1) Planejar, que compreende a descrição de dados a serem compilados e como estes serão gerenciados e posteriormente disponibilizados para acesso no decorrer de seu tempo de vida e, portanto, nesta primeira etapa, cria-se o Plano de Gestão de Dados – PGD (*Data Management Plan – DMP*);

2) Coletar, em que, “dados são colocados em formas digitais” por meio de observações realizadas;

3) Assegurar a qualidade, utilizando controles e inspeção de modo a assegurar a qualidade dos dados;

4) Descrever, utilizando padrões de metadados adequados para descrição dos dados;

5) Preservar, com a finalidade de preservação a longo prazo, em que, os dados são devidamente submetidos para arquivos, centro de dados apropriados, entre outros;

² Fonte da figura disponível no website do DataONE: <https://www.dataone.org/best-practices>

6) Descobrir, em que, dados úteis são descobertos e acessados e, junto, informações relevantes sobre os dados representados (metadados);

7) Integrar, em que, forma-se um conjunto de dados a partir de dados provenientes de fontes diversas a serem combinados e este conjunto de dados homogêneo torna-se prontificado para ser analisado, e;

8) Analisar, como fase em que realiza-se a análise dos dados.

Sayão e Sales (2015) explanam que a gestão dos CVDs propicia aos cientistas-pesquisadores, professores e demais profissionais, a continuidade ou reuso de dados de pesquisas antes utilizados, coletados e gerados até mesmo em seus próprios projetos de pesquisa, visto que, os dados e conjunto de dados científicos possuem tempo de vida maior que os projetos de pesquisas que os criaram.

Dessa maneira, proporcionam-se reuso futuro de objetos digitais e a representação destes em preservação e acesso. Por esta perspectiva, segundo Ball (2010), a curadoria envolve gerenciar e promover o uso de dados partindo de seu ponto de criação, em que planeja-se cuidadosamente a maneira pela qual os dados serão produzidos e processados.

Nesse sentido, tem-se o ciclo de vida dos dados como processo contínuo no gerenciamento dos dados objetivando uso, reuso, preservação e acesso.

Nota-se que a Ciência da Informação como campo de estudo responsável por todos os aspectos presentes na informação, desde análise, coleta, classificação, recuperação e disseminação da informação, deve conduzir e pode contribuir para os processos de Curadoria Digital a abranger os modelos de ciclo de vida dos dados e, por sua vez, os ciclos de vida da pesquisa científica, para que as fases de gerenciamento de dados tenham melhorias constantes, estudos de armazenamento de maneira contínua, aprimorando a gestão de dados, seu uso, reuso, acesso e disponibilização a longo prazo.

3 MAPEAMENTO PARA ANÁLISE DOS CONCEITOS DE CURADORIA DIGITAL E GESTÃO DE DADOS DE PESQUISA

3.1 Estado da Arte e do Conhecimento

Para fins desse estudo e a respeito do termo abranger diferentes denominações, entende-se mapeamento como estado da arte e do conhecimento. Para Ferreira (2002, p. 257) nos últimos anos "[...] no Brasil e em outros países, tem se produzido um conjunto significativo de pesquisas conhecidas pela denominação 'estado da arte' ou 'estado do conhecimento' [...]."

Estas pesquisas de caráter bibliográfico mapeiam e discutem produções acadêmicas em diferentes campos do conhecimento, trazendo aspectos a serem destacados não apenas em diferentes épocas, como em diferentes lugares por meio da metodologia de construção de levantamento e avaliação do conhecimento sobre determinado tema. (FERREIRA, 2002).

A realização do Mapeamento para Análise dos Conceitos de Curadoria Digital e Gestão de Dados de Pesquisa objetiva um estudo bibliométrico que para análise, visa coletar dados respondidos em formulário, em conformidade ao processo de revisão de literatura para, posteriormente, construir indicadores de identificação da abordagem conceitual e teórica a respeito das temáticas, bem como, autores, ano de publicação, periódicos, instituição ou localização geográfica e, se possível, identificar a área do conhecimento com maior impacto e relevância frente às temáticas contempladas pela pesquisa, que auxiliarão na sistematização de informações abordadas pelo projeto *Curadoria Digital: a importância para a Gestão de Dados de Pesquisa*.

Para melhor visualização dos tópicos relacionados, o quadro 2 a seguir, apresenta os itens inseridos em formulário para preenchimento:

Quadro 2 - Tópicos para Mapeamento dos Conceitos de Gestão de Dados de Pesquisa e Curadoria Digital

Tópicos para Mapeamento dos Conceitos de Gestão de Dados de Pesquisa e Curadoria Digital

- Título (de artigo; texto em site; etc.);
- Autoria(s);

- Ano de Publicação;
- Palavras-chave;
- Periódico;
- Base de Dados (ex. Brapci, SciELO). Considere aqui também: Portal de Periódicos; Biblioteca Virtual/Digital; Repositório Digital;
- Instituição ou localização geográfica;
- Conceito(s) teórico(s);
- Área do Conhecimento.

Fonte: Elaboração própria.

Dessa maneira, como apresenta o quadro 2, o escopo do mapeamento realizado é a práxis em estado da arte em tecer isoladamente os conceitos de Curadoria Digital e Gestão de Dados de Pesquisa para posterior análise da abrangência de diferentes produções científicas na tentativa de integrar aspectos conceituais enfatizados por diferentes autores de diferentes periódicos e países, tal como, questionar se ao longo dos anos houveram maior relevância do tema ou modificações conceituais e se sim, qual a modificação presente nesses conceitos, seguindo pela metodologia de caráter descritivo.

Cabe enfatizar que, posto em prática o mapeamento das produções acadêmicas e pesquisas na área de Ciência da Informação em abrangência a Curadoria Digital e Gestão de Dados de Pesquisa, os conceitos mapeados apresentam uma amostra acerca das temáticas retratadas e não uma totalidade em estudos e pesquisas de ambas as áreas.

Em vista disso, havendo dois mapeamentos a serem realizados envolvendo duas temáticas distintas, mesmo que interligadas entre si, optou-se em estudo, colher dados de um conjunto de 10 artigos para cada temática, totalizando 20 artigos referenciais como quantidade considerável para o desenvolvimento deste capítulo sendo, portanto, 10 artigos utilizados referente à Curadoria Digital e 10 artigos referente a Gestão de Dados de Pesquisa, levando em consideração os seguintes tópicos:

1) ambas as temáticas serem relativamente novas em domínio científico, logo, ainda estarem sob avanços conceituais, e;

2) diante de possíveis pluralidades de enfoques – quando abordam-se conceitos em maior quantidade de autores – sobre ambas as temáticas, a proposta de mapeamento dos conceitos constitui uma parte da pesquisa em apenas um capítulo (Capítulo 3) e não o foco de pesquisa. Considerando também que demais produções científico-acadêmicas apresentam maior aprofundamento de conceitos quando retratado o tema como escopo principal de sua produção científica.

Apresenta-se as análises qualitativas e quantitativas da coleta de dados realizada por meio de formulário utilizado, bem como, a discussão desses resultados.

Cabe salientar que as análises estão dispostas em consonância ao segundo objetivo específico delineado para esta pesquisa, isto é, descrever o conceito e aplicação da Curadoria Digital e descrever o conceito e aplicação da Gestão de Dados de Pesquisa que será realizado através de mapeamento a abranger os conceitos dos termos de Curadoria Digital e Gestão de Dados de Pesquisa.

Destaca-se juntamente em complemento ao objetivo principal do estado da arte proposto e em amostra, autores que sobressaíram quando a perspectiva é a definição dos termos, além de tomar como referência do estudo os dados gerados e integrados à produção científica desses autores.

Isto posto, o capítulo segue pela:

- Análise de conteúdo (análise de dados qualitativos) dos conceitos, pela qual ressalta-se que o mapeamento dos conceitos foi elaborado a partir de citações diretas e indiretas retiradas dos artigos englobados pela pesquisa;
- Análise quantitativa dos autores, utilizando o gráfico quantificado pelo Formulários Google (*Google Forms*), por meio dos dados gerados;
- Análise quantitativa dos anos de publicação e palavras-chave, em que, a quantificação dos dados (análise bibliométrica) foi inicialmente realizada em contagem manual sendo, posteriormente, os dados de contagem armazenados em planilha-tabela do Excel, permitindo a criação dos gráficos.
- Por fim, para a análise referente aos periódicos, bem como bases de dados, instituição ou localização geográfica e área (s) do conhecimento

apontada (s) em amostra, a quantificação dos dados foi elaborada através de contagem manual (bibliometria manual).

3.2 Mapeamento dos conceitos de Curadoria Digital

Em consonância aos dados coletados, segundo Sayão e Sales (2012, p. 184), os conhecimentos acumulados nas últimas décadas, no que se concerne às práticas de preservação e acesso aos recursos digitais, levaram a execução de atividades, conjunto de estratégias e abordagens tecnológicas denominadas “curadoria digital”. Sendo ainda um conceito em evolução, apresenta-se certo que esta, envolve gestão e metodologias de preservação de recursos digitais em seu ciclo de vida, processos de arquivamento e preservação digital, tal como, criação e gestão de dados.

Entretanto, como bem destacam Siebra; Borba e Miranda (2016, p. 21), “O termo Curadoria Digital ainda gera dúvidas sobre seu significado devido a sua natureza polissêmica [...]” e por emergir como um campo de prática e pesquisa interdisciplinar que, tal como salienta Weber (2017, p. 131), “[...] reflete uma abordagem holística para o gerenciamento do objeto digital.” Convergiendo com Rosenbaum (2011 *apud* MOLINA; SANTOS, 2019, p. 90) em que:

[...] Nesse contexto, a atuação da “Curadoria Digital” é um termo que engloba distintas terminologias e níveis de desempenho como: “curadoria de informação”, “curadoria de conteúdo”, “curadoria de conhecimento” e “curadoria de dados”, que na maioria dos casos põe como protagonista os seres humanos, os quais têm a capacidade de filtrar informações e reorganizá-las para uma vasta quantidade de usuários. (ROSENBAUM, 2011 *apud* MOLINA; SANTOS, 2019, p. 90).

Enfatizando ainda a dificuldade em estabelecer um conceito único para o termo, Weber (2017, p. 144) retoma a ideia a partir da definição de que o conceito de curadoria digital está associado, em estado atual, a termos como “curadoria de dados”, “curadoria de dados digitais” e vinculado ao reuso de dados científicos, fornecimento de metadados de qualidade, compartilhamento de dados, tal como, ao ciclo da preservação digital e melhoria da ciência.

Ainda por esta perspectiva, Sales e Sayão (2012) enfocam a compreensão de que a curadoria digital deve ser vista para além de um controle de repositório que irá arquivar e armazenar recursos. Sendo ela quem irá assegurar a sustentabilidade dos dados científicos para o futuro.

Sayão e Sales (2012, p. 184) segue pela mesma perspectiva, ao complementarem que a curadoria digital assegura que os conjuntos de dados possam ser pesquisados e continuem capazes de leitura e interpretação contínua, visto que a curadoria digital inclui juntamente a gestão de conjuntos de dados em grande volume para uso diário e, “[...] a ideia de curadoria digital estende-se além do controle do repositório que arquiva os recursos e envolve a atenção do criador do conteúdo e dos usuários futuros.”

Em conformidade, Lee e Tibbo (2007) compilam diversas especificidades ao retratar que a curadoria digital envolve seleção e avaliação por criadores e arquivistas; evoluindo na oferta de acesso intelectual; armazenamento; transformações de dados; e, para alguns materiais, um compromisso com a preservação a longo prazo. Faz-se da curadoria digital, um conjunto de atividades a fornecer reprodutibilidade e reutilização de dados digitais autênticos, bem como outros ativos digitais.

Envolve também o desenvolvimento de repositórios digitais confiáveis e duráveis; tal como, princípios de boa criação e captura de metadados; usos de padrões abertos para formatos de arquivo e codificação de dados; além da promoção da alfabetização em gerenciamento de informações essenciais para a longevidade dos recursos digitais.

Convergindo com o *Digital Curation Centre* (DCC), a curadoria digital é o gerenciamento, preservação de dados e informações digitais a longo prazo. Em sua atividade, envolve-se manutenção, preservação e agregação de valor aos dados digitais de pesquisa durante todo o seu ciclo de vida.

O gerenciamento ativo dos dados de pesquisa reduz ameaças ao seu valor de pesquisa a longo prazo, reduzindo risco de obsolescência digital. Além disso, a curadoria digital e a preservação de dados são processos contínuos que exigem investimentos de tempo e recursos adequados, para que se realizem ações em promover a curadoria e a preservação de forma perpétua ao ciclo de vida dos dados.

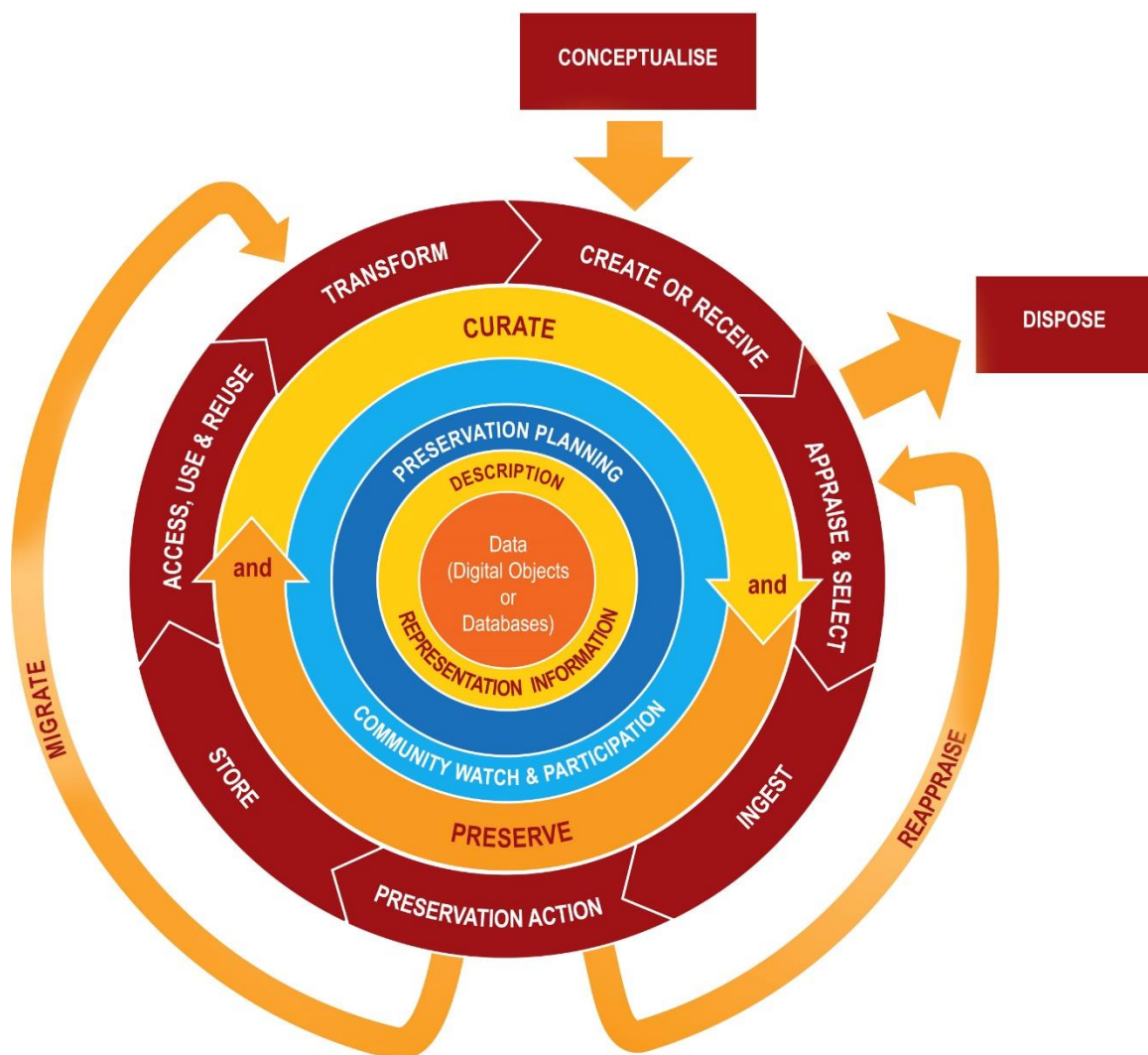
Tal como conceituam Molina e Santos (2019, p. 82), quando definem a curadoria digital como, “[...] o gerenciamento do objeto digital contendo atividades que abarcam todo o ciclo de vida desse objeto, com o intuito de que ele continue acessível e se consiga recuperar por quem dele precise. [...]”.

Este ponto também é destacado por Siebra, Borba e Miranda (2016, p. 21) em relação aos desafios associados à curadoria digital que, antes de serem

técnicos, são principalmente de ordem humana quando, em seu processo curatorial, presume a sustentabilidade de dados digitais a longo prazo.

Faz-se importante trazer, em adendo, que o *Digital Curation Centre* (DCC) desenvolveu um modelo do Ciclo de vida da Curadoria Digital que possibilita melhor interpretação a partir da representação visual de cada estágio do ciclo, conforme ilustra a figura 4 a seguir:

Figura 4 - Curation Lifecycle Model – DCC

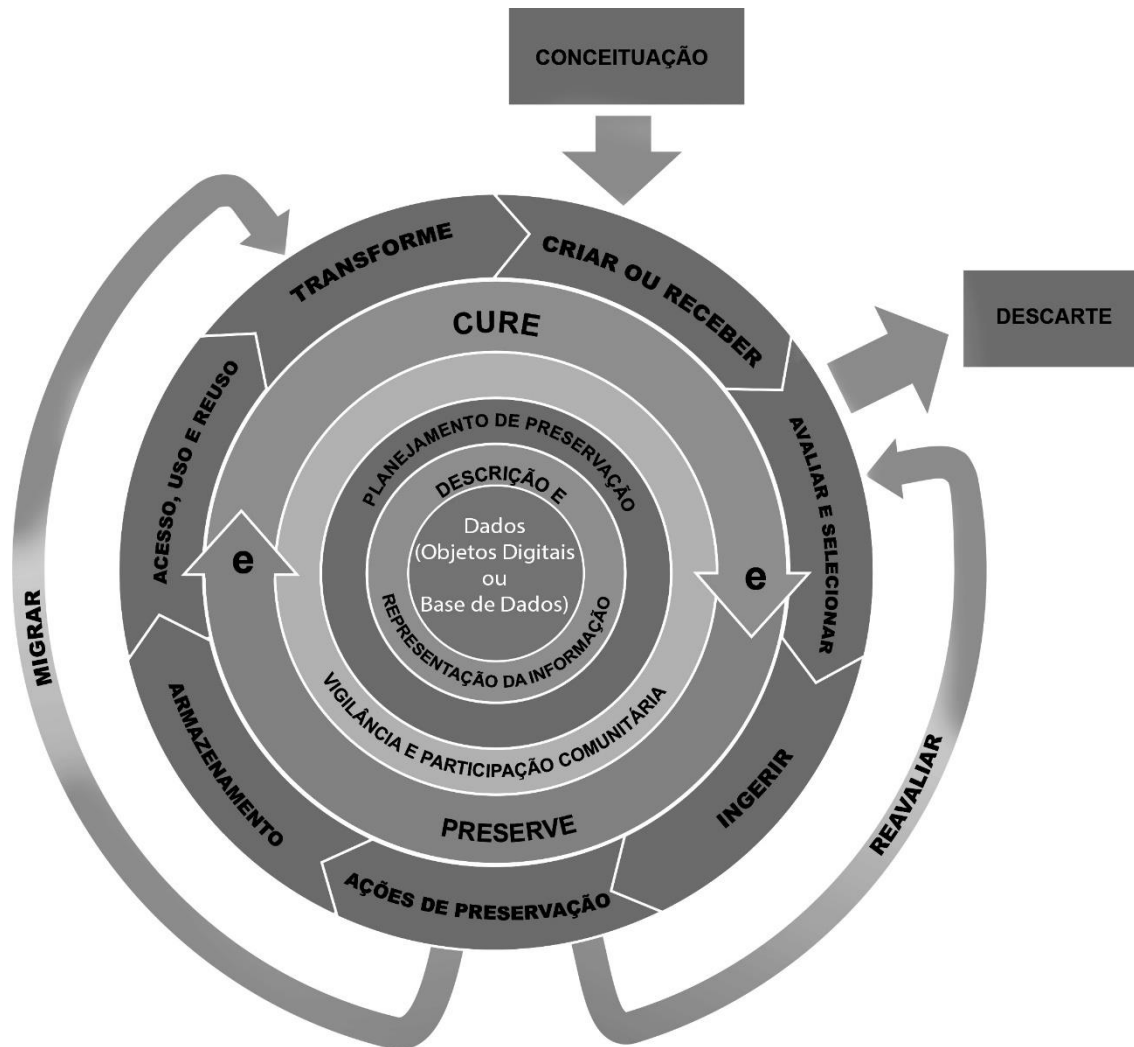


Fonte: *The DCC Curation Lifecycle Model* (2008).

Pautando-se pela figura 4 apresentada acima, optou-se por, através de elaboração própria utilizando o *software Adobe Illustrator*, traduzir a representação gráfica do ciclo de vida da Curadoria Digital de modo a facilitar sua visualização

direta em proporção e interpretação em nosso idioma, como dispõe a figura 5 a seguir:

Figura 5 - Ciclo de vida da Curadoria Digital do DCC



Fonte: Elaboração própria.

Com maior enfoque ao entendimento do ciclo e, tomando por base a análise de Segundo e Araújo (2019), as figuras 4 e 5 retratam que o modelo opera por meio de camadas do ciclo que, por sua vez, contém estágios iniciados a partir do núcleo em que os dados (*Data - Digital Objects or Databases*) estão dispostos no primeiro estágio, sendo, Descrição e Representação da Informação (*Description Representation Information*). Nesse estágio atribuem-se metadados administrativos, descritivos, técnicos, estruturais e de preservação.

O segundo estágio, Planejamento de Preservação (*Preservation Planning*), consiste em executar planejamento e preservação, pelos quais elaboram-se planos estratégicos visando gerenciar todas as etapas do ciclo de vida. Seguindo pelo terceiro estágio, Vigilância e Participação Comunitária (*Community Watch & Participation*) que consiste em informar a gestão da curadoria acerca de mudanças ocorrentes na comunidade geradora dos dados, além de fomentar o desenvolvimento de padrões, ferramentas e softwares adequados que possam acompanhar tais modificações. Posteriormente, tem-se o quarto estágio denominado Cure e Preserve (*Curate and Preserve*), contemplando ações que favoreçam a curadoria digital e preservação a longo prazo durante o ciclo de vida.

Ao redor dos quatro estágios descritos e do núcleo, localizam-se as Ações Sequenciais que se iniciam com as etapas descritas pelo DCC, em que, o ciclo tem início na Conceituação (*Conceptualise*) que, de acordo com o DCC (2008) é a etapa que concebe e planeja a criação de objetos digitais, incluindo métodos de captura de dados e opções de armazenamento.

O ciclo segue pela etapa, Criar ou Receber (*Create or Receive*), que consiste em produzir objetos digitais e atribuir metadados de arquivos administrativos, descritivos, estruturais e técnicos, ou seja, geram-se dados, seus metadados, bem como, pode-se receber dados de outro repositório.

Posteriormente os objetos digitais são avaliados e selecionados na etapa, Avaliar e Selecionar (*Appraise & Select*), em que, dados são avaliados de modo a ser selecionado aqueles que requerem curadoria e preservação a longo prazo.

Cabe destacar que nessa etapa, de acordo com a relevância dos dados, alguns serão descartados ou transferidos para outros curadores ou repositórios digitais, como elucida a próxima etapa denominada, Ingerir (*Ingest*), que consiste em transferir objetos digitais para um arquivo, repositório digital confiável, entre outros.

Depois de avaliar, selecionar/descartar ou transferir os dados, segue-se pelas Ações de Preservação (*Preservation Action*), em que, realizam-se ações que garantam a preservação a longo prazo dos objetos digitais e, promovendo confiabilidade, autenticidade e usabilidade dos dados que, como bem salientam Segundo e Araújo (2019), tais ações incluem limpeza, validação de dados, bem como, atribuição de metadados de preservação, e informações de representação, estrutura e compatibilização aos formatos adequados.

Continua-se pela etapa de Armazenamento (*Store*), que consiste em manter os dados armazenados de forma segura, para que o processo siga pela etapa de Acesso, Uso e Reuso ou reutilização (*Access, use & reuse*), que consiste em garantir facilidade de acesso aos usuários. Salienta-se que alguns dados podem estar disponíveis ao público, enquanto outros podem estar protegidos e não disponíveis e, por isso, pede-se para que se verifique se os dados estão acessíveis para uso e reuso.

Posteriormente, segue-se para a última ação sequencial denominada Transforme (*Transform*), que consiste em criar novos dados, considerados dados derivados, a partir dos dados originais, migrando para um formato diferente quando, por exemplo, um formato (como hardware ou software) torna-se obsoleto, fazendo-se necessário migrar tais dados para um novo formato diferente do anterior, finalizando o ciclo a ser novamente iniciado e contínuo.

Em fechamento ao adendo anterior e retorno aos conceitos, Luz (2018, p. 100) retoma o mesmo princípio do DCC, ao trazer que a curadoria digital possui a missão de garantir a sustentabilidade dos dados a longo prazo, vinculando-se a manutenção, preservação e agregação de valor aos dados em toda sua vida útil, tornando-a, por conseguinte, uma forma de suporte ao profissional da informação ao que se concerne a realização de medidas de preservação, responsável pela garantia de contínuo acesso à informação a longo prazo.

Em aspecto de longevidade, preservação e memória dos dados, Molina e Santos (2019, p. 96) apontam que, “A Curadoria Digital pode ser considerada como um fator fundamental para a constituição da memória, visto que pressupõe oferecer diferencial, agregar valor, aprimorar relacionamentos, rememorar histórias [...]”

A via que interliga curadoria digital a longevidade pode ser também retomada por Abbott (2008) que complementa sua abordagem, retratando a curadoria digital como o gerenciamento e a preservação de dados digitais a longo prazo envolvendo, portanto, todas as atividades referentes ao gerenciamento de dados, desde o planejamento de sua criação até as melhores práticas de digitalização e documentação, bem como, a garantia de disponibilidade e adequação para descoberta e reuso futuro.

Portanto, como o quadro 3 a seguir sugere, a curadoria digital tende a propiciar:

Quadro 3 - Benefícios da Curadoria Digital

1.	Acesso persistente a dados digitais confiáveis;
2.	Melhor qualidade dos dados em si e seu contexto de pesquisa;
3.	O uso de padrões comuns em diferentes conjuntos de dados que, por sua vez, direciona a mais oportunidades de <i>cross-searching</i> – “pesquisa cruzada” – e colaborações;
4.	Verificações de autenticidade, melhorando a confiabilidade dos dados;
5.	Garantir que os dados sejam válidos como um registro formal – quando apropriado – o que significa que podem funcionar no futuro como evidência legal;
6.	Explorar o investimento inicial, garantindo que os dados estejam disponíveis para uso, reuso e protegendo o valor financeiro das informações;
7.	Maior velocidade e variedade de acesso, oportunidades de compartilhamento e análise de dados, além de outros benefícios de pesquisa.

Fonte: Elaboração própria adaptado de Abbott (2008) em *"What is Digital Curation?"*, DCC Briefing Papers: Introduction to Curation, Edinburgh: Digital Curation Centre.

Por meio do quadro 3, nota-se que a autora ainda traça tópicos de exemplo ao que a Curadoria Digital proporciona em termos de benefícios a curto prazo e valor a longo prazo, visto que, embora a mesma garanta sustentabilidade dos dados a longo prazo, ela também possui valor imediato para os criadores de dados e usuários. (ABBOTT, 2008).

A autora destaca juntamente tópicos de exemplos de que a curadoria digital garante valor a longo prazo, como no quadro 4 a seguir, por:

Quadro 4 - Valor a longo prazo da Curadoria Digital

1.	Preservar dados e protegê-los contra perda e obsolescência (particularmente crucial quando os dados não são reproduzíveis ou são extremamente valiosos);
2.	Permitir acesso contínuo aos dados, apesar do financiamento de curto prazo ou de mudanças institucionais;
3.	Incentivar a reutilização de dados;
4.	Maximizar a exploração de materiais digitais acumulados ao longo do tempo;
5.	Fornecer informações sobre o contexto e a proveniência dos dados;

6.	Pelo uso de ferramentas e serviços para migrar dados, metadados e outras estruturas de representação para novos formatos, garantindo que permaneçam significativos para os usuários;
7.	E por uma infraestrutura de gerenciamento para preservação e disseminação perpétua de dados.

Fonte: Elaboração própria adaptado de Abbott (2008) em *"What is Digital Curation?"*, DCC Briefing Papers: Introduction to Curation, Edinburgh: Digital Curation Centre.

No quadro 4, Abbott (2008) traça as garantias de valor a longo prazo da Curadoria Digital no que se concerne à preservação, de maneira a proteger e salvaguardar os dados, permitindo acesso contínuo, incentivando seu uso e reuso, além de fornecer informações a respeito do contexto e proveniência desses dados, bem como, pelo uso de estruturas de representação, como metadados, viabilizando um melhor gerenciamento dos dados, acesso, preservação e recuperação, enfatizando a aplicação da Curadoria Digital em busca e recuperação digital.

Outro aspecto relevante na definição é destacado por Siebra, Borba e Miranda (2016, p. 27) quando as autoras verificam, por fim, que a curadoria digital em referência à área da Ciência da Informação, vem sendo utilizada em destaque a três abordagens mencionadas por elas sendo, “[...] na filtragem e seleção de dados na Web (Curadoria de Conteúdo ou de Informações), na agregação de valor, gestão ativa e preservação de dados digitais (Curadoria Digital) e na curadoria de dados de pesquisa (*e-science*).”

Percebe-se que, embora o termo Curadoria Digital ainda não seja consolidado de maneira absoluta, ainda é o termo abrangente mais utilizado em relação a outros termos relativamente frequentes como “curadoria dos dados”, “curadoria de conteúdo” ou “curadoria de informação”; e também mais usual que o termo “preservação digital”, embora este seja utilizado de forma frequente, porém em menor escala ou em quantidade próxima ao termo de “curadoria digital”.

A análise conceitual proporcionada por este mapeamento, faz considerar que as definições não divergem tanto quanto a construção de outros termos que são abrangidos em grande escala por diversas áreas do conhecimento, a exemplo do conceito “dado”, fortemente presente em áreas como Matemática, Ciência da Computação, Informática, Ciência da Informação, acompanhado de uma multiplicidade de conceitos fincados em mesmo termo.

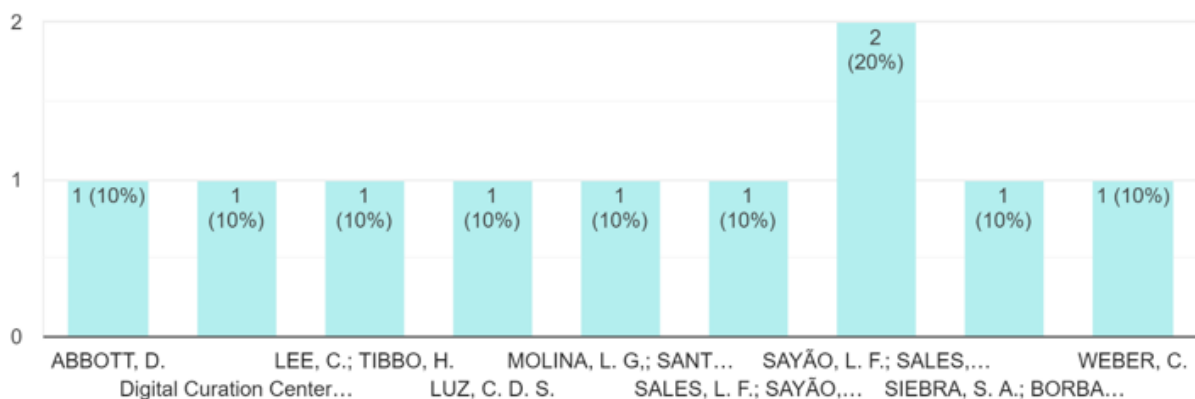
Nota-se que os conceitos dialogam entre si e complementam-se, trazendo a perspectiva de que o conceito de “curadoria digital” vem sendo construído de forma convergente e de maneira conjunta no que tange dados publicados de conhecimento científico referente a temática.

Essa convergência também é estabelecida por meio de produções científico-acadêmicas em validade, reprodução e replicação de pesquisas, junto a iniciativas, como a fundação do *Digital Curation Centre* (DCC), que além de seu propósito central, promove também os conhecimentos acerca da Curadoria Digital, levando-o a comunidades científicas, cientistas-pesquisadores, agências de fomento e demais interessados, considerando que até mesmo em sua nomenclatura, torna-se o termo maior consolidado frente a outros diversos ainda emergentes em diferentes produções científicas.

Como exposto, para a elaboração do tópico de análise e discussão dos resultados e, anteriormente realizada a análise e mapeamento dos conceitos, os apontamentos dos resultados foram dispostos em complemento ao objetivo principal do mapeamento dos conceitos de Curadoria Digital proposto e, em amostra, verificam-se os resultados obtidos e apresentados a seguir.

O gráfico 1 retrata as autorias coletadas por meio da compilação de artigos acerca de Curadoria Digital, como exposto a seguir:

Gráfico 1 - Autorias coletadas referente à Curadoria Digital



Fonte: Elaboração própria.

Desse modo, o gráfico 1 permite analisar em amostra que entre as autorias coletadas, os autores Sayão e Sales (2012 e 2016) lideram as publicações

referentes a temática de Curadoria Digital, pelo menos no Brasil, revelando um cenário anteriormente perceptível, quando pretendia-se para o desenvolvimento da pesquisa recorrer às produções científico-acadêmicas relacionadas aos temas da pesquisa.

Trata-se aqui, em especial, a temática de Curadoria Digital que logo em compilação inicial, notou-se que os trabalhos de Sayão e Sales destacaram-se acerca dos temas de curadoria digital, preservação de dados digitais de pesquisa, curadoria de dados e guia de gestão de dados de pesquisa, em que, além de comporem um dos autores mais relevantes do tema, estes trabalham também em redes de coautoria e com frequência produzem artigos relacionados ao tema, indicando forte colaboração e certa frequência no engajamento crescente da literatura.

Pode-se considerar Sayão e Sales como os autores ou coautores precursores do desenvolvimento do tema de Curadoria Digital na área de Ciência da Informação no Brasil, desenvolvendo crescentes pesquisas em parceria nas temáticas de Gestão de Dados de Pesquisa, Curadoria Digital de Dados de Pesquisa e Preservação Digital.

Dessa maneira, a relevância de análise das autorias coletadas, permite compreender autores que trabalharam em suas pesquisas de maneira individual ou conjunta, contribuindo para a produção no domínio ao longo dos anos ou inserção deste a comunidade científica, permitindo também encontrar autores que mais desenvolveram pesquisas na área.

Outro aspecto relevante a pontuar, diz respeito ao tópico referente ao ano de publicação. A quantificação dos dados (análise bibliométrica) foi inicialmente realizada em contagem manual sendo, posteriormente, armazenada em planilha-tabela do Excel, como apresenta a tabela 1 a seguir:

Tabela 1 - Número de publicações por ano

Ano	Npub
2004 a 2020 ³	1

³ 2004 a 2020: não foram encontrados referenciais de ano de publicação no site do *Digital Curation Centre*. Constando apenas "*Digital Curation Centre 2004-2020*".

2007	1
2008	1
2012	2
2016	2
2017	1
2018	1
2019	1

Fonte: Elaboração própria.

A tabela 1 retrata o número de publicações (Npub), totalizando em 10 publicações referente aos 10 artigos utilizados (referente ao ano).

Vale salientar que no primeiro item localizado na coluna Ano, há a numeração de “2004 a 2020”, sendo esta em referência ao site do *Digital Curation Centre* (DCC), em que, não foram encontrados referenciais de ano de publicação além da informação de o material ser protegido por direitos autorais, seguido pelo símbolo de *copyright* e "*Digital Curation Centre 2004-2020*". Para encontrar esta informação, basta acessar ao site⁴ do DCC e seguir por: *Home > About > "What is digital curation?"*.

Após a realização da bibliometria manual e inserção dos dados em Excel, elaborou-se o gráfico 2, também referente ao número de publicações por ano acerca do tema de Curadoria Digital apresentado a seguir:

Gráfico 2 - Número de publicações por ano



Fonte: Elaboração própria.

⁴ Website do *Digital Curation Centre* (DCC): <https://www.dcc.ac.uk/>

Conforme analisa-se pelo gráfico 2, indica-se que houve uma crescente de publicações acerca da temática de Curadoria Digital nos anos de 2012 e 2016, considerando que nesses dois anos ou entre o período destacado pode-se ter dado maior relevância ao tema ou sendo ele relativamente atual, deu-se início a discussão do tema de Curadoria Digital nesses períodos.

Ainda relativamente discreta até 2008, as produções e publicações sobre Curadoria Digital, segundo a amostra, tomaram proporção no ano de 2012 que, ao consta, parece representar período de crescimento nas produções sobre o tema, com aumento relativamente perceptível em relação ao ano anterior.

No tocante a tênue queda na quantidade de artigos ou publicações a partir de 2017, pode-se considerar que o ano de 2017 parece marcar um período de estabilização da pesquisa e produções sobre o tema, divergindo do crescimento dos anos de 2012 e 2016 anteriores.

Outro tópico relacionado ao mapeamento, diz respeito às palavras-chave encontradas e número de coocorrência dos termos em palavras-chave. Desse modo, pelo mesmo método realizado para o tópico referente ao ano de publicação, a quantificação dos dados (análise bibliométrica) das palavras-chave foi inicialmente realizada em contagem manual sendo, posteriormente, armazenada em planilha-tabela do Excel, como apresenta o quadro 5 a seguir:

Quadro 5 - Número de palavras-chave

Palavras-chave	Npc
Curadoria Digital (<i>Digital Curation</i>)	7
Preservação Digital (<i>Digital Preservation</i>)	4
<i>e-Science</i>	3
Ciência da Informação	2
Repositório Digital Confiável (<i>Trusted Repository</i>)	2
Curadoria de Conteúdo	1
Dados de Pesquisa	1
Arquivo Digital	1
Custódia Arquivística	1
Entorno Digital	1
Biblioteconomia	1
Reuso de Dado Científico	1
Comunicação Científica	1
Curadoria de Dados	1
Curador	1
Ciência Aberta	1
Repositórios Científicos	1
Audit	1
Certification	1
Documento Digital	1
Documento Eletrônico	1
Memória	1

Fonte: Elaboração própria.

O quadro 5 retrata o número de palavras-chave (Npc)⁵ encontradas nos 10 artigos utilizados. O gráfico elenca os termos de maior ocorrência sendo: Curadoria Digital (*Digital Curation*), Preservação Digital (*Digital Preservation*), *e-Science*, Ciência da Informação, Repositório Digital Confiável (*Trusted Repository*), como os que apareceram mais vezes em diferentes artigos e prossegue pelas palavras-chave de apenas uma ocorrência.

Após a realização da bibliometria manual e inserção dos dados em Excel, elaborou-se o gráfico 3, indicando coocorrências de palavras-chave, visto que, a relação de coocorrência entre duas ou mais palavras-chave é determinada pela quantidade de artigos, em que ambos termos ocorrem conjuntamente no campo de palavras-chave.

⁵ Para maior esclarecimento, o número de palavras-chave (Npc) indica a frequência de palavras-chave – frequência de ocorrências das palavras-chave elencadas.

Desse modo, o gráfico 3, também referente ao número de palavras-chave acerca do tema de Curadoria Digital, restringiu a apresentação gráfica elencando as palavras-chaves de maior ocorrência, ilustrando também quatro termos de apenas uma ocorrência, como retratado a seguir:

Gráfico 3 - Número de palavras-chave



Fonte: Elaboração própria.

Conforme o gráfico 3 apresenta, verifica-se que dos 10 artigos de amostra, 7 apresentaram a ocorrência do termo "Curadoria Digital" no campo das palavras-chave, seguido de 4 ocorrências do termo "Preservação Digital" em palavras-chave, 3 ocorrências do termo *e-Science*, 2 ocorrências do termo "Ciência da Informação" e 2 ocorrências do termo "Repositório Digital".

Dessa forma, evidencia-se o termo Curadoria Digital como o mais usual em busca e recuperação em bases de dados, visto que, ao utilizar o termo em campo de busca, permite-se alcançar maiores quantidades de resultados de pesquisa por meio dos buscadores, de modo a sintetizar o conteúdo presente em bases de dados e relacioná-los com os termos buscados.

A ocorrência de palavras-chave mais usuais, faz também com que artigos relacionados caminhem de encontro a pesquisa do usuário, otimizando sua experiência e engajando a produção de conteúdos relacionados. Considera-se que o impacto das palavras-chave relaciona-se com via direta ao ponto de necessidade

de busca do usuário, via esta também de produção de resultados de alto impacto para atração de usuários até determinado artigo e eficiência na recuperação da informação.

Cabe enfatizar que a alta relevância do uso de palavras-chave na representação e recuperação da informação, contribui para uma maior visibilidade e conduz a possibilidades de aumento no índice de citação e maior impacto da produção científica por autores, como da própria instituição, localização geográfica, periódicos, entre outros indicadores.

Nota-se que as palavras-chave representam uma fonte de acesso proficiente e relevante aos artigos científicos, enfatizando a necessidade de que sejam com atenção elegidas pelos autores para a representação de seus artigos indexados em bases de dados.

Por fim, para a análise referente aos periódicos, bases de dados, instituição ou localização geográfica, bem como, área (s) do conhecimento apontada (s) em amostra, a quantificação dos dados foi elaborada através de contagem manual (bibliometria manual), como apresenta-se a seguir:

- **Periódicos**

Dos 10 artigos compilados, não foram apresentados periódicos com maior ocorrência. No entanto, a partir da contagem dos países vinculados a determinados periódicos, fez-se possível encontrar ocorrência de repetição de países pelos quais os periódicos foram criados e são atualizados. A exemplo dos periódicos 1) *Informação & Sociedade: Estudos (I&S)*; 2) a *Revista Eletrônica AtoZ: Novas Práticas em Informação e Conhecimento*; 3) *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*; 4) *Informação & Tecnologia (ITEC)* e 5) *Revista P2P & Inovação*; ambas revistas/periódicos referentes ao Brasil.

Tal como, a primeira ocorrência de periódico encontrado foi no 1) repositório digital de pesquisa produzida na Universidade de Edimburgo, Escócia, Reino Unido, o ERA (*Edinburg Research Archive*) – Arquivo de Pesquisa de Edimburgo e redirecionado ao *Edinburgh: Digital Curation Centre: DCC Briefing Papers*. Segunda ocorrência no 2) periódico *Journal of Digital Information* que é uma revista do Reino Unido e terceira ocorrência sendo o, 3) o site do *Digital Curation Centre (DCC)* que possui sede em Edimburgo no Reino Unido.

Além dos periódicos, 1) *Prisma.com – Revista de Ciências e Tecnologias da Informação* e 2) *Comunicação e Páginas A&B, Arquivos e Bibliotecas*, ambos periódicos de Portugal. Com isso, nota-se que além das coocorrências de países como, em primeiro lugar, Brasil, em segundo, Reino Unido e terceiro, Portugal, os periódicos mostram-se receptivos a pesquisas que abrangem o contexto de Curadoria Digital em avanço crescente.

- **Bases de dados**

A base de dados que mais se sobressaiu para o encontro dos artigos foi a Bapci, apresentando-se em um indicador de 5 vezes, relacionadas a 5 coocorrências (de 10) artigos encontrados através da base. Isto relaciona-se pelo escopo da base atentar-se a Ciência da Informação que, a seguir, será apresentada como a área do conhecimento mais contemplada no mapeamento realizado, em amostra.

- **Instituição ou localização geográfica**

Referente a instituição ou localização geográfica com mais de uma ocorrência nos artigos coletados, tem-se que, instituições no país e, portanto, no Brasil foram apresentados indicadores de maior coocorrência (5), seguido por instituições e país, Reino Unido, elencado em segundo lugar, apresentando indicador de coocorrência (3) e Portugal, em terceiro lugar, apresentando em instituição e país, indicador de coocorrência (2).

- **Área do conhecimento**

Nota-se que, embora em diferentes periódicos e bases de dados, os 10 artigos abrangem a área do conhecimento sendo contemplada pela Ciência da Informação, o que denota a não ou quase não dispersão da produção científica dessa área do conhecimento para outras áreas ao longo dos anos.

Justificável pelo tema tratar justamente o campo de atuação de bibliotecários e cientistas da informação que podem auxiliar pesquisadores a como serem curadores ou gerenciadores de seus dados digitais de pesquisa.

É possível identificar alguns elementos que podem ser mais comuns às tentativas de conceituar Curadoria Digital e que a dinâmica deste conceito está presente com maior afinco na área de Ciência da Informação e suas subáreas sendo, Gestão da Informação, Biblioteconomia, entre outros, mas que, em amostra,

não permeia tanto por outras áreas do conhecimento, embora possa ser mencionada e estudada por elas, visto que o tema abrange a melhoria de curadoria digital em atividades de gerenciamento de dados científicos e de pesquisa presentes em todas as áreas do conhecimento.

Desde o mapeamento redirecionado aos conceitos até as análises e discussão de resultados gerados, tem-se que, estas análises sugerem que há interesse crescente diante da comunidade acadêmica em relação a Curadoria Digital, ainda que esteja como um conceito em evolução, manifestando-se aos poucos pela propagação de pesquisas, contribuindo para o aprimoramento científico. Salienta-se que sendo uma amostra, não abrange a totalidade do domínio, com elementos traçados sendo inerentes a esta pesquisa.

Nesse sentido, seria interessante comparar resultados de amostra a uma análise mais abrangente da temática, visto que os estados da arte abrangem maiores quantidades de artigos como escopo central de pesquisa e, por fim, vale enfatizar que algumas limitações ou restrições presentes neste estudo bibliométrico em amostra, pode vir a servir como tópico de partida para futuros projetos de pesquisa.

3.3 Mapeamento dos conceitos de Gestão de Dados de Pesquisa

Em consonância aos dados coletados, segundo Sayão e Sales (2015, p. 80), a gestão de dados de pesquisa define-se por “Conjunto de práticas de gestão voltadas para o tratamento de dados de pesquisa durante o seu ciclo de vida; inclui todos os aspectos de manutenção, compartilhamento, segurança e preservação.”

Por outro prisma, Gattelli (2015) abrangendo “dados de pesquisa” como “dados de investigação” e o motivo pelo qual a “gestão de dados de investigação” vem se tornando preocupação das comunidades científicas, retrata que dados de pesquisa “[...] tem sido um tema frequente nas agendas de vários grupos envolvidos com investigação, sejam universidades, agências financiadoras, editores de literatura científica, governos ou a própria sociedade.”

Convergindo com Sayão e Sales (2012, p. 181) ao explanarem que este fato coloca em cheque para a ciência a nova problemática sendo a gestão de dados de pesquisa em mundo digital, também interligada por redes de computadores que, por

sua vez, compele grande fluxo de dados, advindos de diferentes fontes em que são gerados, processados e compartilhados.

Em complemento, Bertin, Visoli e Drucker (2017) conceituam a GDP como,

[...] uma expressão abrangente que envolve tanto os aspectos rotineiros de planejamento, aquisição, organização, estruturação, definição de fluxos analíticos e ferramenta computacional apropriada para o armazenamento de dados, quanto às questões relativas à preservação, à organização, ao compartilhamento, à proteção e à confidencialidade destes para a instituição que possui o direito sobre tais dados, bem como o acesso e disponibilização para a sociedade. (BERTIN; VISOLI; DRUCKER, 2017).

Neste aspecto, Borgman et al. (2015) relata que gerenciar dados de pesquisa pode ser visto como um desafio e, portanto, pauta-se sobre uma tarefa complexa, uma vez que, tornar dados de pesquisa úteis para outros desconhecidos para fins imprevistos, abrange um processo ainda mais complexo. À medida que os pesquisadores abordam os limites de gerenciamento de dados das ferramentas e recursos disponíveis, eles atingem o problema de dimensionamento.

O autor enfatiza que ter mais dados requer não apenas ferramentas maiores ou mais rápidas, mas ferramentas diversificadas, bem como, diferentes modos de consulta, por um viés de que, o problema de escala está ocorrendo de maneira diferente em cada campo, laboratório, projeto e site de pesquisa e, somente comparando vários casos por longos períodos de tempo, faz-se possível identificar o conjunto de desafios que a gestão de dados propõe.

Sayão e Sales (2015, p. 27) retomam que uma das parcelas frente a importância na Gestão de Dados de Pesquisa se faz justamente em assegurar que os dados possam ser compreendidos e interpretados, como eles salientam, por qualquer usuário.

Exige-se qualidade em descrição detalhada dos dados, tal como informações adicionais ou contextuais que viabilizem que tais dados venham a transmitir informações e conhecimentos no tempo e no espaço, como destacam os autores, esclarecendo que isso é possível através da documentação que deve acompanhar os dados, sendo esta denominada como a Documentação dos Dados.

Outra vinculação presente nas definições parte de Dudziak (2016), ao abordar que, a gestão de dados de pesquisa (*Research Data Management* – RDM) abarca serviços, ferramentas, bem como, infraestruturas que compreendem o ciclo de vida

da pesquisa científica como um todo. Dudziak (2016) aponta que os vários aspectos da GDP distribuem-se em diferentes departamentos acadêmicos, tal como, serviços de apoio, a exemplo: bibliotecas, serviços de TI e escritórios de pesquisa.

A autora ainda acrescenta que os cientistas-pesquisadores carecem de apoio nas etapas de planejamento, organização, segurança, documentação, bem como, na preparação dos conjuntos de dados para depósito e preservação, e também em questões que abrangem propriedade intelectual, direitos autorais e licenciamento.

A autora finaliza trazendo que é por este sentido que a importância e papel dos bibliotecários se faz na “[...] identificação e localização de dados, apoio às condições de acesso e reutilização de dados, suporte à citação e referenciação, até a correta organização e preservação.”

Silva et al. (2019, p. 312) retoma a ideia do papel do bibliotecário, quando salienta que sua formação sendo vasta em modalidades de organização, tratamento, armazenamento, disseminação e recuperação de conteúdos, além de ferramentas e funcionalidades que adicionam valor aos conteúdos digitais, colocam os profissionais da informação, bibliotecários e cientistas da informação como colaboradores no suporte e no gerenciamento de dados.

Destaca ainda que, “[...] A falta de interesse dos pesquisadores é um dos principais motivos para o não compartilhamento. No entanto, muitos também não o fazem porque não sabem como fazer ou porque simplesmente desconhecem os benefícios da gestão de dados. [...]”. (SILVA et al., 2019, p. 306).

Entretanto, conforme sugerem Sayão e Sales (2016), é necessário observar que uma parcela das coleções de dados advém de pesquisas que precisam que seus produtos finais de pesquisa sejam arquivados, tal como, acessados de maneira restrita.

Os autores explanam que isso ocorre por alguns fatores, entre eles, segurança, dados sensíveis que carecem de tratamentos específicos sendo um desses a anonimização, tal como, possibilidade de patentes e interesses comerciais.

Ocorre que, as próprias políticas de gestão de dados têm como previsão um intervalo de tempo para que tais dados sejam utilizados por seus autores de maneira privilegiada, sendo este denominado como tempo de embargo e que, mesmo que não compartilhados em primeiro momento, é preciso manter os dados preservados e bem gerenciados para reutilização futura, e a seleção do que será ou não

preservado e compartilhado já é parte de uma das etapas da Curadoria Digital de Dados de Pesquisa.

Borgman et al. (2015) retoma o que se concerne às agências de fomento alegando que a gestão de dados tornou-se uma prioridade muito maior no processo de pesquisa devido aos requisitos das agências de financiamento e dos periódicos para liberar dados de pesquisa no momento da publicação do artigo.

Como destaca, desafios significativos para a implementação da gestão de dados surgem das complexidades da colaboração científica moderna como, noções concorrentes sobre o que conta como dados, visões díspares de pesquisa e inovação, incentivos diferentes para compartilhamento e liberação de dados, além de debate sobre questões econômicas e de propriedade intelectual relacionadas a produtos do conhecimento e políticas públicas.

O autor enfatiza que, se o “potencial da ciência” com uso intensivo de dados for realizado, sistemas, serviços, ferramentas, conteúdo, políticas, práticas e recursos humanos adequados serão necessários para descobrir e explorar os dados como produtos de pesquisa.

Piwovar, Day e Fridsma (2007), apontam que o compartilhamento de dados de pesquisa fornece benefícios à comunidade científica em geral e, por isso, a gestão de dados de pesquisa assume relevância crescente dentro do cenário científico nacional e internacional, visto que tais dados, se não forem bem gerenciados, poderão encontrar-se desorganizados e perdidos.

Associa-se, além disso, o compartilhamento de dados detalhados ao aumento da taxa de citação e reuso de dados arquivados, porém, o benefício é menos óbvio para o investigador que disponibiliza seus dados, dessa maneira, a correlação entre dados disponíveis ao público e maior impacto na literatura pode motivar os pesquisadores a compartilhar seus dados detalhados de pesquisa.

Por este prisma, Sayão e Sales (2018, p. 80) trazem que, “O requisito essencial para uma boa gestão de dados de pesquisa é que ela se realize por meio de plataformas que assegurem infraestrutura tecnológica e gerencial, sustentabilidade econômica e política de longo prazo [...]”, enfatizando o conceito e aplicação da Gestão de Dados de Pesquisa.

Os autores complementam que os serviços oferecidos por plataformas colaborativas de gestão de dados de pesquisa devem abarcar todos os processos dos “[...] ciclos de vida das coleções de dados – que são específicos, variam em

muitas direções e podem perdurar por tempo indeterminado.” (SAYÃO; SALES, 2018, p. 88).

Destaca-se que a oferta de serviços para a gestão de dados é dependente de variáveis como usuários, tipos de dados, tecnologias disponíveis, além do conhecimento por parte dos profissionais envolvidos em relação aos dados de pesquisa no que tange geração, coleta, etapas de processamento e tratamento pelas quais passaram os dados, bem como, ferramentas e metodologias para aplicação. (SAYÃO; SALES, 2018, p. 88).

Silva et al. (2019, p. 307) apresenta por meio da figura 6 a seguir, as fases da gestão de dados de pesquisa:



Fonte: SILVA et al. (2019, p. 307).

A figura 6 retrata as fases de gestão de dados de pesquisa em três momentos:

1. Antes da pesquisa começar: que contempla a etapa de planejamento, utilizando-se de projetos de pesquisas e Plano de Gestão de Dados (PGD);

2. Durante a pesquisa: que contempla a coleta e/ou geração de dados, bem como, processamento, controle de qualidade, metadados disciplinares, armazenamento seguro, *backups* e análise. E, por fim;
3. Pesquisa finalizada: que abrange avaliação, catalogação, contextualização, questões éticas e legais, publicação, preservação a longo prazo, bem como, compartilhamento, acesso e reuso. ,

Silva et al. (2019, p. 307) também atribui que a figura 6 retrata o papel do profissional bibliotecário a respeito da preservação de dados digitais a longo prazo e dessa maneira, faz-se considerar a curadoria dos dados científicos como serviço a ser ofertado por bibliotecas e demais unidades de informação, de modo a garantir acesso contínuo e reuso de dados em novas produções científicas.

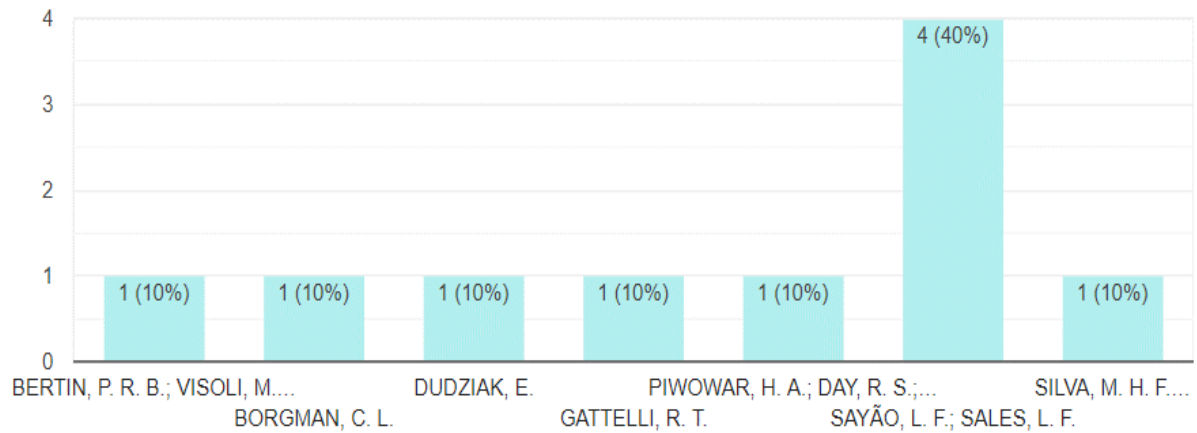
Dentre as definições de Gestão de Dados de Pesquisa, destaca-se a forte vinculação com a atuação a ser desempenhada pelo bibliotecário como de alta relevância no que tange à oferta de serviços, utilização de ferramentas, além de planos e métodos em como executar a gestão de dados.

Silva et al. (2019, p. 303) sugere o melhoramento das competências e “[...] habilidades dos bibliotecários no apoio aos pesquisadores, na implantação de infraestruturas de serviços e de letramento, no desenvolvimento de metodologia para o compartilhamento e reuso de dados [...]”.

Faz-se necessário que o conceito de Gestão de Dados de Pesquisa (GDP) seja entendido e percebido como elemento propulsor do avanço científico e tecnológico, especialmente em um momento em que tanto se discute sobre o acesso a dados, o favorecimento de sua visibilidade, bem como seu uso e reuso.

Para a elaboração do tópico de análise e discussão dos resultados e, anteriormente realizada a análise e mapeamento dos conceitos, os apontamentos dos resultados foram dispostos em complemento ao objetivo principal do mapeamento dos conceitos de Gestão de Dados de Pesquisa proposto e, em amostra, verificam-se os resultados obtidos e apresentados a seguir.

O gráfico 4 retrata as autorias coletadas por meio da compilação de artigos acerca de Gestão de Dados de Pesquisa, como exposto abaixo:

Gráfico 4 - Autorias coletadas referente a Gestão de Dados de Pesquisa

Fonte: Elaboração própria.

O gráfico 4, permite analisar em amostra que entre as autorias coletadas, como a temática de Curadoria Digital, os autores Sayão e Sales (2012, 2015, 2016 e 2018) lideram as publicações referentes a temática de Gestão de Dados de Pesquisa, pelo menos no Brasil, em que, como anteriormente explicitado, além de comporem um dos autores mais relevantes do tema, estes trabalham também em redes de coautoria e com frequência produzem artigos relacionados ao tema, indicando engajamento crescente da literatura científica, desenvolvendo crescentes pesquisas em parceria nas temáticas de Gestão de Dados de Pesquisa, Curadoria Digital de Dados de Pesquisa e Preservação Digital.

Como antes colocado, a relevância de análise das autorias coletadas, permite compreender autores que trabalharam em suas pesquisas de maneira individual ou conjunta, contribuindo para a produção no domínio ao longo dos anos ou inserção deste a comunidade científica, permitindo também encontrar autores que mais desenvolveram pesquisas na área.

Outro aspecto relevante a pontuar, diz respeito ao tópico referente ao ano de publicação. A quantificação dos dados (análise bibliométrica) foi inicialmente realizada em contagem manual sendo, posteriormente, armazenada em planilha-tabela do Excel, como apresenta a tabela 2 a seguir:

Tabela 2 - Número de publicações por ano

Ano	Npub
2007	1
2012	1
2015	3
2016	2
2017	1
2018	1
2019	1

Fonte: Elaboração própria.

A tabela 2 retrata o número de publicações (Npub) – totalizando em 10 publicações referente aos 10 artigos utilizados – por ano.

Após a realização da bibliometria manual e inserção dos dados em Excel, elaborou-se o gráfico 5, também referente ao número de publicações por ano acerca do tema de Gestão de Dados de Pesquisa apresentado a seguir:

Gráfico 5 - Número de publicações por ano

Fonte: Elaboração própria.

Conforme analisa-se pelo gráfico 5, indica-se que houve uma produção crescente de publicações acerca da temática de Gestão de Dados de Pesquisa nos anos de 2015 e 2016, considerando que nesses dois anos ou entre o período

destacado pode-se ter dado maior relevância ao tema ou sendo ele relativamente atual, deu-se início a discussão do tema de Gestão de Dados de Pesquisa nesses períodos, embora já estivesse sendo contemplado em temáticas de Curadoria Digital provenientes dos anos anteriores.

Ainda relativamente discreta até 2012, as produções e publicações sobre Gestão de Dados de Pesquisa, segundo a amostra, tomaram proporção no ano de 2015 que, ao consta, parece representar período de crescimento nas produções sobre o tema, com aumento relativamente perceptível em relação ao ano anterior, sendo 2012. No tocante a tênue queda na quantidade de artigos ou publicações a partir de 2017, pode-se considerar que 2017 parece marcar um período de estabilização da pesquisa e produções sobre o tema, divergindo do crescimento dos anos de 2015 e 2016 anteriores.

Outro tópico relacionado ao mapeamento, diz respeito às palavras-chave encontradas e número de coocorrência dos termos em palavras-chave. Desse modo, pelo mesmo método realizado para o tópico referente ao ano de publicação, a quantificação dos dados (análise bibliométrica) das palavras-chave foi inicialmente realizada em contagem manual sendo, posteriormente, armazenada em planilha-tabela do Excel, como apresenta o quadro 6 a seguir:

Quadro 6 - Número de palavras-chave

Palavras-chave	Npc
Curadoria Digital	3
Dado (s) de Pesquisa	3
Ciência da Informação	2
Preservação Digital	1
Novo Documento	1
Comunicação Científica	1
Plataforma de Gestão de Dado	1
Avaliação	1
Gestão da Informação	1
Gestão de Dado de Pesquisa	1
Plano de Gestão de Dado	1
Ciclo de Vida	1
Coleta	1
Validação	1
Descrição	1
Descoberta	1
Preservação Digital	1
Integração	1
Análise	1
Dados de investigação	1
Gestão de dados de investigação	1
Metadados	1
Perfil de Aplicação	1
Ontologia	1
Oceanografia Biológica	1
e-Science	1
Ciência Aberta	1
Repositórios Científicos	1
Gestão de dados	1
Bibliotecário de dados	1
Competência dos Bibliotecários	1

Fonte: Elaboração própria.

O quadro 6 retrata o número de ocorrências das palavras-chave, ou seja, número de palavras-chave (Npc) encontradas nos artigos utilizados.

Cabe salientar que de 10 artigos utilizados, 6 continham o campo de palavras-chave, sendo que em 4 artigos por alguma razão as palavras-chave não foram disponibilizadas. O gráfico elenca os termos de maior ocorrência sendo: Curadoria Digital (*Digital Curation*) – como detalhe curioso, visto que, no mapeamento dos conceitos de Curadoria Digital, o termo também foi elencado como de maior ocorrência – seguido por, Dado (s) de Pesquisa e Ciência da Informação como os

termos que apareceram mais vezes em diferentes artigos, prosseguindo pelas palavras-chave de apenas uma ocorrência.

Após a realização da bibliometria manual e inserção dos dados em Excel, elaborou-se o gráfico 6, indicando coocorrências de palavras-chave, visto que, a relação de coocorrência entre duas ou mais palavras-chave é determinada pela quantidade de artigos, em que, ambos termos ocorrem conjuntamente no campo de palavras-chave.

O gráfico 6, também referente ao número de palavras-chave acerca do tema de Gestão de Dados de Pesquisa, restringiu a apresentação gráfica elencando as palavras-chaves de maior ocorrência, ilustrando também seis termos de apenas uma ocorrência, como retratado a seguir:

Gráfico 6 - Número de palavras-chave



Fonte: Elaboração própria.

Conforme o gráfico 6 apresenta, verifica-se que dos 10 artigos de amostra, 3 apresentaram a ocorrência do termo "Curadoria Digital" no campo das palavras-chave, sendo mesmo valor atribuído às 3 ocorrências do termo "Preservação Digital" e 2 ocorrências do termo "Ciência da Informação".

Dessa forma, evidencia-se o termo Curadoria Digital como o mais usual em busca e recuperação em bases de dados. Como elucidado anteriormente, a ocorrência de palavras-chave mais usual, faz com que artigos relacionados caminhem de encontro a pesquisa do usuário, otimizando sua experiência e engajando a produção de conteúdos relacionados.

Considera-se que o impacto das palavras-chave relaciona-se com via direta ao ponto de necessidade de busca do usuário, via esta também de produção de resultados de alto impacto para atração de usuários até determinado artigo e eficiência na recuperação da informação.

Enfatizando novamente que a alta relevância do uso de palavras-chave na representação e recuperação da informação, contribui para uma maior visibilidade e conduz a possibilidades de aumento no índice de citação e maior impacto da produção científica por autores, como da própria instituição, localização geográfica, periódicos, entre outros indicadores.

Por fim, para a análise referente aos periódicos, bases de dados, instituição ou localização geográfica, bem como, área (s) do conhecimento apontada (s) em amostra, a quantificação dos dados foi elaborada através de contagem manual (bibliometria manual), como apresenta-se a seguir:

- **Periódicos**

Dos 10 artigos compilados, apresentaram-se a coocorrência de dois periódicos em dois artigos diferentes, sendo o periódico *Ponto de Acesso – Revista do Instituto de Ciência da Informação da UFBA* (Universidade Federal da Bahia), influenciando também na coocorrência de mesma instituição e localização geográfica a ser apresentada a seguir.

- **Bases de dados**

A base de dados que mais se sobressaiu para o encontro dos artigos novamente foi a Brapci, apresentando-se em um indicador de 4 vezes, relacionadas a 4 coocorrências (de 10) artigos encontrados através da base. Isto relaciona-se pelo escopo da base atentar-se a Ciência da Informação que, a seguir, será apresentada como a área do conhecimento mais contemplada no mapeamento realizado, em amostra.

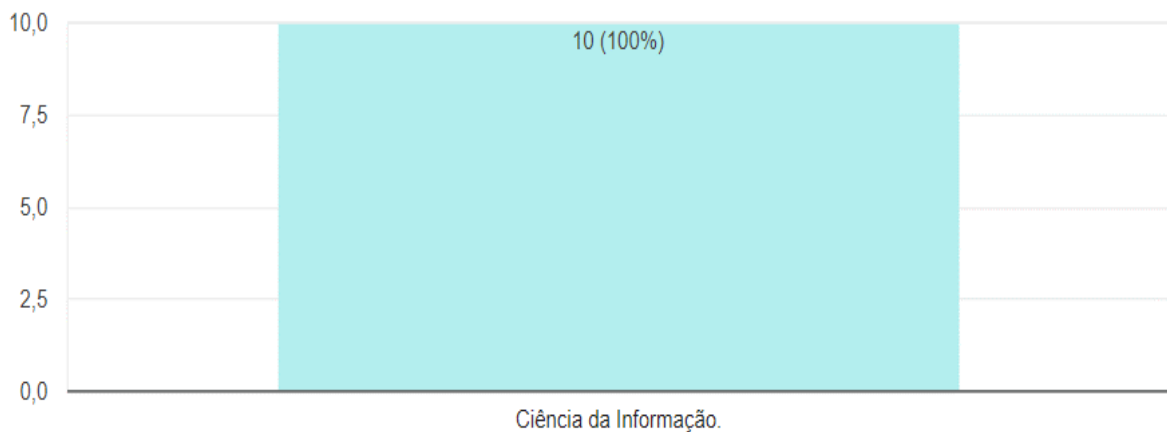
- **Instituição ou localização geográfica**

Referente a intuição ou localização geográfica com mais de uma ocorrência nos artigos coletados, tem-se que, instituições no país e, portanto, no Brasil foram apresentados indicadores de maior coocorrência (7), seguido por instituições e país, EUA, elencado em segundo lugar, apresentando indicador de coocorrência (2) e Portugal, em terceiro lugar, apresentando em intuição e país, indicador de coocorrência (1).

- **Área do conhecimento**

Constata-se que como para Curadoria Digital, os dados gerais da amostra denotam que não houve dispersão do conhecimento do campo de Ciência da Informação para outros campos do conhecimento acerca de Gestão de Dados de Pesquisa, indicando ser um tema atual, ainda em avanço e conhecimento para publicações de cientistas-pesquisadores até que futuramente se confirme seu processo de consolidação. Este indicador pode ser bem observado no gráfico 7, a seguir:

Gráfico 7 - Área do Conhecimento coletada referente a Gestão de Dados de Pesquisa



Fonte: Elaboração própria.

Embora os dados tenham sido coletados em diferentes artigos, de diferentes periódicos e bases de dados, tem-se que os 10 artigos abrangem a área do conhecimento sendo contemplada pela Ciência da Informação em sua totalidade, o que denota nesta amostra, a não dispersão da produção científica dessa área do conhecimento para outras áreas.

Como anteriormente exposto, o indicador é justificável pelo tema tratar justamente o campo de atuação de bibliotecários e cientistas da informação que podem auxiliar pesquisadores a como serem curadores ou gerenciadores de seus dados de pesquisa.

Cabe enfatizar também que, desde o mapeamento redirecionado aos conceitos até as análises e discussão de resultados gerados, tem-se que, estas análises sugerem que há interesse crescente diante da comunidade acadêmica não só em relação ao tema de Curadoria Digital, como também Gestão de Dados de Pesquisa, manifestando-se aos poucos pela propagação de pesquisas e, contribuindo para o aprimoramento científico. Salienta-se que sendo uma amostra, não abrange a totalidade do domínio, com elementos traçados sendo inerentes a esta pesquisa.

Como antes explanado, seria interessante comparar resultados de amostra também referente a Gestão de Dados de Pesquisa a uma análise mais abrangente da temática, visto que os estados da arte abrangem maiores quantidades de artigos como escopo central de pesquisa e, por fim, destaca-se que limitações ou restrições presentes neste estudo bibliométrico em amostra, podem servir como tópico de partida para futuros projetos de pesquisa.

O Plano de Gestão de Dados (PGD) ou *Data Management Plan* (DMP), responsabiliza-se pela definição dos tipos de dados a serem produzidos, métodos de preservação, tratamento e manutenção dos dados, bem como, políticas de acesso e compartilhamento de dados, além de ações de arquivamento e medidas de salvaguarda, segurança e integridade dos mesmos, dessa forma, elaborar um plano de gestão de dados economiza esforços, tempo e planejamento a como processar e tratar os dados.

De acordo com Corti et al. (2014), a gestão de dados de pesquisa contempla todas as práticas, manipulações, melhoras e processos que possibilitam assegurar a qualidade dos dados, bem como, garantir que estejam organizados, documentados, preservados, sustentáveis, acessíveis e reutilizáveis.

Tal como explanam Sayão e Sales (2018, p. 83), conforme os dados de pesquisa tornam-se de modo acelerado parte importante do fluxo de trabalho em laboratórios de pesquisa, instituições e centros de pesquisa passam a voltar sua atenção para uma plano de gerenciamento de dados e começam a tomar ações

para definir políticas, desenvolver soluções, estabelecer serviços de interesse referente a suas comunidades, enquanto que,

[...] na outra direção, os pesquisadores precisam identificar repositórios adequados para a publicação dos seus dados – uma exigência cada vez mais presente das agências de fomento, dos editores científicos e da comunidade onde está inserido - e para descobrir coleções de dados de qualidade que possam ser reutilizadas para prosseguimento de seus empreendimentos científicos. Para tal, é necessário refletir sobre quais são os parâmetros – dentre as inúmeras variáveis presentes - que podem dimensionar os requisitos de uma plataforma de gestão de dados. (SAYÃO; SALES, 2018, p. 83).

Verifica-se que, como bem salientam Sayão e Sales (2015), a prática de uma gestão de qualidade de dados, tal como, gestão do ciclo de vida dos dados de pesquisa se dá por meio do uso efetivo desses dados quando gerenciados, tratados, armazenados e preservados como um meio para o progresso científico, uma vez que, a relevância de dados gerados, quando devidamente tratados, serve de fonte de recurso informacional para a pesquisa científica de alta qualidade, corroborando para a evolução da ciência, bem como, aprendizagem e ensino científico e, portanto, esses dados habilitam pesquisadores a inovar métodos científicos, questionamentos e hipóteses.

Porém, em unanimidade, cientistas e pesquisadores apresentam dúvidas referentes a realização de PGD e sua elaboração, uma vez que não estejam devidamente informados sobre aspectos de preservação, armazenamento, depósito, representação e organização de dados científicos.

Diante disso, para a resolução do impasse, tomar medidas de apoio informacional aos pesquisadores, seria um dos passos iniciais, através de ofertas de *workshops* acerca de como desenvolver planejamentos de gestão de dados, consultorias individuais para a elaboração dos planos de gestão de dados, utilizando-se de entrevistas a respeito de qual conteúdo ou área de pesquisa estão presentes e irão desenvolver projetos de pesquisa, alertando-os também a possíveis problemas de preservação, nesse sentido, os PGDs podem contribuir de maneira significativa quando bem conduzidos e por meio de processos colaborativos entre profissionais da informação, bibliotecários, pesquisadores e repositórios. (SILVA, 2016).

Bertin, Visoli e Drucker (2017, p. 39), consideram que o "Planejamento" pode ser destacado como um dos princípios para a Gestão de Dados de Pesquisa. Dessa maneira, a etapa situa refletir sobre o gerenciamento de dados ainda como concepção e ideia antecedente a execução do projeto em si. Portanto, os autores destacam que para o planejamento da GDP, devem considerar-se tópicos de:

- Tipos, formatos, bem como conjuntos de dados existentes;
- Métodos existentes para a coleta de dados;
- Questões legais, éticas, de direitos autorais e propriedade intelectual;
- Níveis de acesso;
- Maneiras de compartilhamento e reutilização dos dados;
- Gestão, curadoria de dados e armazenamento de curto prazo;
- Depósito (arquivo) e, como salientam os autores, preservação em longo prazo.

Ainda por esta perspectiva, os autores mencionam a necessidade do estabelecimento de um programa institucional de Gestão de Dados de Pesquisa, de modo a incorporar a gestão de dados a estratégias organizacionais, sendo que, como objetivo principal, um programa institucional de GDP ofertaria aos cientistas-pesquisadores, ferramentas, treinamentos, apoio e orientação para uma gestão apropriada dos dados científicos que compele a todo seu ciclo de vida.

Bertin, Visoli e Drucker (2017, p. 45) apresentam questões chave a serem respondidas de modo a se compreender a conjuntura dos dados científicos na instituição, como uma espécie de questionário para levantamento do gerenciamento de dados pelos pesquisadores:

1. Com que tipos de dados de pesquisa a organização lida?
2. Como esses dados são produzidos e validados?
3. Onde são armazenados os dados e os conjuntos de dados de pesquisa, e por quanto tempo esses dados podem ser úteis?
4. Como esses dados serão preservados, de modo que possam ser acessíveis daqui há alguns anos?
5. Qual a qualidade dos conjuntos de dados da organização?
6. Quem são os autores dos dados?

7. Quais dados de pesquisa provêm bases para inovação tecnológica e, por isso, necessitam ser protegidos? E quais dados podem ser disponibilizados abertamente?
8. Quem pode reutilizar os dados de pesquisa produzidos pela organização? E de que modo os dados de pesquisa podem ser reutilizados?

Para a execução apropriada de um PGD de qualidade, faz-se necessário considerar a responsabilidade do pesquisador na identificação de melhores práticas e ações para a gestão de dados em sua área de pesquisa e atuação, enquanto que, por outro lado, cabe a instituição reforçar benefícios da prática de gestão de dados, organização, preservação e compartilhamento responsável de dados de pesquisa, de modo a oferecer subsídios, ferramentas e serviços essenciais para uso, reuso, acesso e preservação de dados, garantindo também que autores pesquisadores recebem os devidos créditos por seus dados citados. (BERTIN; VISOLI; DRUCKER, 2017, p. 46).

Sayão e Sales (2015, p. 16), em seu *Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores* (2015), elencam alguns motivos para se criar um Plano de Gestão de Dados de Pesquisa, em que, além da gestão de dados ser uma das áreas primordialmente responsável pela condução de pesquisas em ambientes científicos atuais, auxilia pesquisadores a considerar em fase de planejamento, como os dados serão gerenciados durante a pesquisa e, posteriormente, preservados e compartilhados para a comunidade científica.

Entre os motivos estão: assegurar a integridade da pesquisa e o seu potencial de replicação, assegurar que os dados e demais objetos de pesquisa sejam autênticos e confiáveis, aumentar a eficiência do pesquisador, evitar riscos de perda, duplicação de esforços em "recoleta" ou "regeração" de dados, de modo a possibilitar que demais pesquisadores utilizem-se de seus dados para outros contextos e novas perspectivas, bem como, aumentar a segurança dos dados, visibilidade da pesquisa e ajustar medidas de preservação e arquivamento.

Uma vez que os PGDs descrevem quais dados serão coletados e gerados e quais metodologias utilizar para seu processo de tratamento e gerenciamento, além de como torná-los acessíveis para reuso, pensando em medidas de preservação a longo prazo, influenciando também na descrição de conjuntos de dados a serem posteriormente localizados e recuperados de maneira mais econômica e eficaz.

A importância pauta-se em assegurar a integridade dos dados, contribuindo para fatores como autenticidade, estabelecendo validade, reprodução e replicação da pesquisa científica, de modo a atuar diretamente em instituições, centros de pesquisa, agências de fomento, comunidades científicas e organizações de P&D a contribuírem para o progresso científico.

4 CONSIDERAÇÕES FINAIS

Evidenciou-se a princípio que, o conceito de dados passa por consideráveis variações entre pesquisadores e áreas do conhecimento, visto que, os dados são gerados por finalidades distintas, diferentes pesquisadores, comunidades científicas e acadêmicas, além de diferentes processos, denotando sua multiplicidade que também o leva a existência de diferentes tipologias e interpretações de acordo com a finalidade de cada pesquisa.

Ainda no aspecto dados, a noção de dados estruturados conceitua-se por metadados, visto que estes possibilitam que dados sejam representados, encontrados, identificados e interpretados, permitindo que a leitura de seu conteúdo seja realizada de forma autônoma por máquinas tornando mais fácil sua recuperação, enquanto que, dados não estruturados exigem interpretador externo, tendo em vista que esse tipo de dado requer de sistemas um esforço maior aos processos de coleta, análise e recuperação.

Verificou-se que dados de pesquisa contemplam todo o tipo de registro produzido, gerado, utilizado ou reutilizado ao longo do desenvolvimento da pesquisa e dados científicos são provenientes de pesquisas científicas, sendo gerados em consonância a finalidade de cada pesquisa ou comunidade científico-acadêmica.

A partir de uma gestão de qualidade dos dados de pesquisa, utilizando-se de metadados para sua representação, garante-se o armazenamento efetivo de dados, sendo este, portanto, o modo pelo qual os metadados tornam-se essenciais para armazenagem e preservação de dados de pesquisa, uma vez que, utilizar padrões ou esquemas de metadados para gerenciamento de dados culmina em melhor armazenamento, localização e recuperação por meio da padronização e representação dos dados.

Em desafio e necessidade de se preservar o conjunto de dados coletados ou gerados, surge o conceito de Curadoria Digital de dados de pesquisa ou dados científicos, a abranger a gestão de dados e visando, sobretudo, salvaguardar os dados de pesquisa, promovendo sua longevidade e mantendo-os acessíveis para utilização futura de comunidades científico-acadêmicas.

Dessa maneira, em abordagem mais densa, objetivou-se aprofundar os conceitos de Curadoria Digital e Gestão de Dados de Pesquisa por meio de formulário com perguntas estratégicas para mapeamento, visando tabular os

conceitos dos autores e complementos por meio de dados oriundos dos artigos científicos, sendo gerados e posteriormente interpretados por contagem manual e análise bibliométrica.

Nesse sentido, inferiu-se que o conjunto de conceitos identificados ao longo do mapeamento apresenta que os conceitos dialogam entre si e complementam-se, bem como, dentre as definições de Gestão de Dados de Pesquisa, destaca-se a forte vinculação com a atuação do bibliotecário como de alta relevância no que tange à oferta de serviços, utilização de ferramentas, além de planos e métodos em execução de planos para a gestão de dados.

Nota-se a Curadoria Digital como um ciclo contínuo tanto em processo interno – ciclo da curadoria digital – quanto em processo externo, considerando políticas externas que conduzam a curadoria digital a um processo continuado, uma vez que este é um processo a longo prazo que envolve um conjunto de atividades que visam garantir preservação, acesso, recuperação e disseminação de dados digitais em perpetuidade.

Portanto, para que não se perca em obsolescência tecnológica, faz-se necessário que práticas e ferramentas, como padrões de metadados, plano de gestão de dados sejam sempre atualizados de maneira a salvaguardar os dados digitais.

Ademais, no mapeamento redirecionado aos conceitos, bem como, na discussão de resultados gerados, observou-se que estas análises sugerem que há interesse crescente diante da comunidade acadêmica em relação a Curadoria Digital e Gestão de Dados de Pesquisa, ainda que estejam como um conceito em evolução, manifestando-se aos poucos pela propagação de pesquisas.

Tem-se o Ciclo de Vida dos Dados como processo contínuo no gerenciamento dos dados objetivando uso, reuso, preservação e acesso, além da concepção de que os modelos de Ciclo de Vida dos Dados padronizam o processo de Curadoria Digital, contribuindo para resultados diante de etapas iniciais e finais em abrangência aos dados em todo seu ciclo.

Os Planos de Gestão de Dados descrevem quais dados serão coletados e gerados e, quais metodologias serão utilizadas para seu processo de tratamento, gerenciamento, acesso, uso e reuso, contemplando medidas de preservação a longo prazo, de modo a assegurar a integridade dos dados e contribuir para fatores de autenticidade, validade, reprodução e replicação da pesquisa científica.

Verificou-se que a comunidade científica tem dado importância à temática de Gestão de Dados de Pesquisa evidenciando impasses de cientistas-pesquisadores em gestão de dados científicos, bem como, uso, reuso e metodologias de preservação, armazenamento e salvaguarda.

Nessa constante, aprimoram-se as metodologias presentes no processo de Curadoria Digital que abrange a Gestão de Dados de Pesquisa e um conjunto de métodos a fim de se preservar dados de pesquisa a longo prazo, sendo este um processo complexo, porém significativo ao progresso científico.

Em retrospecto, a execução da investigação apresentou o objetivo principal contemplado e os demais objetivos específicos alcançados. Os objetivos específicos atingidos foram divididos em três tópicos, sendo alcançado o primeiro objetivo específico na identificação da fundamentação teórica frente à Curadoria Digital e Gestão de dados de Pesquisa, contextualizando as temáticas na literatura.

Bem como, o segundo objetivo específico sobre descrever o conceito e aplicação da Curadoria Digital, enfatizando busca e recuperação digital e o conceito e aplicação da Gestão de Dados de Pesquisa, bem como, o Plano de Gestão de Dados, o Ciclo de vida dos dados conjuntamente ao Ciclo de vida da pesquisa científica.

E o terceiro objetivo específico, em mapear os conceitos de Curadoria Digital e Gestão de Dados de Pesquisa para análise de conteúdo (análise de dados qualitativos) dos conceitos e análise quantitativa dos autores, anos de publicação, palavras-chave, periódicos, bases de dados, instituição ou localização geográfica e área do conhecimento com maior impacto e relevância frente às temáticas contempladas pelo trabalho.

Nesse sentido, em estado da arte, a pesquisa contemplou o Mapeamento dos Conceitos de Curadoria Digital e Gestão de Dados de Pesquisa, com conceitos melhor aprofundados para análise e discussão, em diferentes campos do conhecimento e produções científicas na tentativa de integrar aspectos conceituais enfatizados por diferentes autores, periódicos, entre outros tópicos abordados.

Foram notadas características representacionais dos dados relacionando-os ao símbolo, havendo também relações conceituais em sua definição, tornando-o abrangente e distinto entre as áreas da ciência.

Por um lado isto denota uma definição por vezes “distinta” e não fixa e, por outro, traz para a área de Ciência da Informação que lida diretamente com isso,

desafios em definições significativas de dados que abarquem a conjuntura informacional discutida pela Ciência da Informação a envolver dados desde sua gênese à transformação deste em conhecimento e informação por meio do tratamento de dados.

Para que haja um gerenciamento efetivo, armazenamento, preservação e posterior compartilhamento de dados de pesquisa, a compreensão do processo integral dos dados até que os mesmos se tornem informação se faz necessária, o que implica desde o processo em que dados encontram-se em sua forma bruta até o tratamento e interpretação desses dados, adentrando em posterior representação, recuperação, disseminação, uso, reuso e acesso, impulsionando também estudos científicos por pesquisadores em seus respectivos campos do conhecimento.

Os domínios de Curadoria Digital e Gestão de Dados de Pesquisa vem gerando e podem gerar ainda mais avanços significativos não apenas para a Ciência da Informação como também para a comunidade científico-acadêmica, em progresso à ciência e atividade de pesquisa que possui resultados difundidos para a sociedade e comunidade científica.

A elucidação da problemática de pesquisa sobre como as duas abordagens se correlacionam e em sua somatória, vinculam-se ao domínio “dados” para disseminação da informação e do conhecimento científico, fez-se presente na importância que não apenas a comunidade científica, como também, a gestão do conhecimento científico em si, tem redirecionado aos dados de pesquisa e a preocupação em como gerir esses dados no tocante à compreenderem melhorias em processos de compartilhamento, acesso, disseminação, armazenamento, salvaguarda, uso, reuso e legitimação dos dados coletados ou gerados em suas pesquisas.

Enfatizou-se a relevância da Curadoria Digital e Gestão de Dados de pesquisa quando vinculam-se ao domínio de dados, visando não apenas disseminação da informação e conhecimento científico, como também, métodos de acesso, gestão de qualidade, segurança, a longevidade a que se redireciona a gestão, estabelecendo elo entre planejamento, preservação e melhor disposição de dados de pesquisa e publicações científico-acadêmicas quando bem conduzidas.

A Ciência da Informação como campo de estudo responsável pelos aspectos presentes na informação, desde análise, coleta, classificação, recuperação e disseminação da informação, pode contribuir para os processos de Curadoria Digital

a abranger os modelos de ciclo de vida dos dados e, por sua vez, os ciclos de vida da pesquisa científica, conduzindo para que etapas de gerenciamento de dados tenham melhorias constantes, atualizações e estudos para o armazenamento de maneira contínua, de modo a aprimorar a gestão de dados, seu uso, reuso, acesso e disponibilização a longo prazo.

Visto que um dos objetivos da presente pesquisa seria contribuir para a área de Ciência da Informação, espera-se que este trabalho tenha contribuído com os estudos na área e possa incentivar pesquisadores a aplicar os conhecimentos elencados.

Procurou-se delinear conceito e aplicação da Curadoria Digital e Gestão de Dados de Pesquisa, abarcando informações capazes de auxiliar, mesmo que preliminarmente, cientistas-pesquisadores em como aplicar o processo de Curadoria Digital, ações de gerenciamento de dados de pesquisa, modelos de ciclo de vida dos dados e Planos de Gestão de Dados de pesquisa.

Faz-se perceptível a conscientização de agências de fomento, comunidades científicas, instituições e organizações de P&D referente a necessidade em preservar, salvaguardar e tornar acessível os dados digitais, provenientes de pesquisas e domínios científicos.

Torna-se constante o reconhecimento da Curadoria Digital e sua importância para a Gestão de Dados de Pesquisa que consiste ao fato de que os métodos utilizados pelo processo de curadoria digital, como etapas de arquivamento e preservação digital, incluem também ao processo, etapas que visam dar atenção a criação e efetiva gestão de dados de pesquisa, tornando dados passíveis de interpretação, réplica, uso, reuso e acesso ao conhecimento científico de maneira a gerar novas pesquisas e perspectivas.

Por fim, acredita-se que junto à gama de conhecimento promovida pelo dilúvio de dados, era *big data*, bem como, a globalização da informação, sobretudo, potencializada pelas redes e Revolução Digital em crescimento exponencial de dados, a curadoria digital e o gerenciamento de dados digitais vem se tornando um dos grandes desafios da ciência no século XXI, cabendo aprofundar e aplicar conhecimentos em gestão de dados de pesquisa como elemento propulsor para o avanço científico.

REFERÊNCIAS

- ABBOTT, D. What is digital curation? **DCC Briefing Papers: Introduction to Curation. Edinburgh: Digital Curation Centre**, 2008. Disponível em: <https://www.dcc.ac.uk/guidance/briefing-papers/introduction-curation/what-digital-curation>. Acesso em: 16 maio 2020.
- BALL, A. A review of data management lifecycle models. **Bath, UK: University of Bath**, 2012. Disponível em: <http://opus.bath.ac.uk/28587/1/redm1rep120110ab10.pdf>. Acesso em: 20 out. 2018.
- BALL, A. Review of the state of the art of the digital curation of research data. **Bath, UK: University of Bath**, 2010. Disponível em: <https://purehost.bath.ac.uk/ws/portalfiles/portal/293012/erim1rep091103ab12.pdf>. Acesso em: 25 maio 2020.
- _____. Charter on the Preservation of Digital Heritage. **United Nations Educational, Scientific and Cultural Organization (UNESCO)**, 2003. Disponível em: http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html. Acesso em: 30 abr. 2020.
- BELL, G.; HEY, T.; SZALAY, A. Computer Science: Beyond the Data Deluge. **Science**, v. 323, n. 5919, p. 1297–1298, 2009. Disponível em: <http://www.uvm.edu/pdodds/files/papers/others/2009/bell2009a.pdf>. Acesso em: 28 out. 2018. DOI:10.1126/science.1170411. ISSN 0036-8075.
- BERTIN, P. R. B.; VISOLI, M. C.; DRUCKER, D. P. 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, p. 34-48, 2017. Disponível em: <http://www.brapci.inf.br/index.php/res/v/81812>. Acesso em: 16 fev. 2020.
- BOHLE, SHANNON. What is E-science and How Should it Be Managed? **Scilogos. Scientific and Medical Libraries. Nature and Spektrum der Wissenschaft**, 2013. Disponível em: http://www.scilogos.com/scientific_and_medical_libraries/what-is-e-science-and-how-should-it-be-managed/. Acesso em: 25 out. 2018.
- BORGMAN, C.L. et al. Knowledge infrastructures in science: data, diversity, and digital libraries. **International Journal on Digital Libraries**, v. 16, p. 207–227, 2015. Disponível em: <https://escholarship.org/uc/item/32q2z1c9>. Acesso em: 7 maio 2020.
- CORTI, L. et al. **Managing and sharing research data: a guide to good practice**. Los Angeles: SAGE, 2014. ISBN: 978-1-4462-6726-4. Disponível em: https://books.google.com.br/books?id=-gg6AAQBAJ&printsec=frontcover&hl=pt-BR&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false. Acesso em: 22 maio 2020.
- COSTA, M. M.; CUNHA, M. B. O bibliotecário no tratamento de dados oriundos da e-science: considerações iniciais. **Perspectivas em Ciência da Informação**, v. 19, n. 3, p. 189-206, 2014. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/37440>. Acesso em: 28 nov. 2019.

CURTY, R. G.; CERVANTES, B. M. N. Data Science: ciência orientada a dados. **Informação & Informação**, Londrina, v. 21, n. 2, p. 1-4, maio/ago., 2016. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/27929/0>. Acesso em: 5 maio 2020.

DATAONE Data Observation Network for Earth. **Best Practices**. Disponível em: <http://www.dataone.org/best-practices>. Acesso em: 27 maio 2020.

DATAONE Data Observation Network for Earth. **Data Management Guide for Public Participation in Scientific Research**. 2013. Disponível em: <https://www.dataone.org/sites/all/documents/DataONE-PPSR-DataManagementGuide.pdf>. Acesso em: 27 maio 2020.

DIGITAL CURATION CENTRE. **What is digital curation?** Disponível em: <https://www.dcc.ac.uk/about/digital-curation>. Acesso em: 31 mar. 2020.

DIGITAL CURATION CENTRE. **The DCC Curation Lifecycle Model**. 2008. Disponível em: <https://www.dcc.ac.uk/sites/default/files/documents/publications/DCCLifecycle.pdf>. Acesso em: 31 mar. 2020.

DUDZIAK, E. A. **Competências do Bibliotecário na Gestão de Dados de Pesquisa, Comunicação Científica e Acesso Aberto**. São Paulo, 2016. Disponível em: <https://www.aguia.usp.br/noticias/competencias-gestao-dados-pesquisa/>. Acesso em: 28 abr. 2020.

DZIEKANIAK, G. V. Mapeamento do uso de metadados por comunidades científicas. **BIBLOS - Revista do Instituto de Ciências Humanas e da Informação**, v. 20, p. 229-243, 2007. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/22802>. Acesso em: 2 jan. 2020.

FERREIRA, N. S. de A. As pesquisas denominadas "estado da arte". **Educação & Sociedade**, v. 23, n. 79, p. 257-272, 2002. Disponível em: <https://www.scielo.br/pdf/es/v23n79/10857.pdf>. Acesso em: 27 maio 2020.

FRY, J.P.; SIBLEY, E.H. Evolution of Data-Base Management Systems. **ACM Computing Surveys**, v.8, n.1, p. 7-42, march 1976. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.520.808&rep=rep1&type=pdf>. Acesso em: 28 maio 2020.

GATTELLI, R. T. **Gestão de dados de investigação no domínio da Oceanografia Biológica: criação e avaliação de um perfil de aplicação baseado em ontologia**. Dissertação (Ciência da Informação) - Faculdade de Engenharia e Faculdade de Letras, Universidade do Porto, Portugal, 2015. Disponível em: <https://repositorio-aberto.up.pt/bitstream/10216/79336/2/35451.pdf>. Acesso em: 28 abr. 2020.

HEY, T.; TREFETHEN, A. e-Science and its implications. In: Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences. **The Royal Society**, UK, v. 361, p. 1809-1825, 2003. Disponível em:

https://eprints.soton.ac.uk/257964/1/e_Science_and_its_Implications.pdf. Acesso em: 30 out. 2018.

LE COADIC, Yves-François. **A Ciência da Informação**. Brasília, DF: Briquet Lemos Livros, 1996.

LEE, C.; TIBBO, H. Digital Curation and Trusted Repositories: Steps Toward Success. **Journal of Digital Information**, v. 8, n. 2, 2007. Disponível em: <http://journals.tdl.org/jodi/article/viewArticle/229/183>. Acesso em: 22 fev. 2020.

LIMA, P. O. Revisão sistemática sobre uso de ontologia para análise de sentimento em conteúdo da web. **Ciência da Informação**, v. 44, n. 3, 2015. Disponível em: <http://www.brapci.inf.br/index.php/res/v/17938>. Acesso em: 3 jan. 2020.

LOURENÇO, C. A. Metadados: o grande desafio na organização da Web. **Informação & Sociedade: Estudos (I & S)**, João Pessoa, v.17, n.1, p.65-72, jan./abr., 2007. Disponível em: http://www.brapci.inf.br/_repositorio/2010/11/pdf_59d6ff3ea6_0012781.pdf. Acesso em: 30 nov. 2019.

LUZ, C. D. S. Curadoria digital, custódia arquivística e preservação digital: relações possíveis. **Páginas A&B, Arquivos e Bibliotecas (Portugal)**, n. 10, p. 92-103, 2018. Disponível em: <<http://hdl.handle.net/20.500.11959/brapci/109277>>. Acesso em: 20 maio 2020.

MARKOFF, J. A Deluge of Data Shapes a New Era in Computing. **The New York Times**. 2009. Disponível em: <https://www.nytimes.com/2009/12/15/science/15books.html>. Acesso em: 27 jan. 2021.

MOLINA, L. G.; SANTOS, J. C. Curadoria digital: novos suportes documentais e a preservação da memória. **Prisma.com (Portugal)**, n. 38, p. 82-101, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/112256>. Acesso em: 13 mar. 2020.

PIWOWAR, H. A.; DAY, R. S.; FRIDSMA, D. B. Sharing detailed research data is associated with increased citation rate. **Plos One**, v. 2, n. 3, 2007. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0000308>. Acesso em: 21 fev. 2020.

QIN, J. et al. Linking entities in scientific metadata. **International Conference on Dublin Core and Metadata Applications**, 2010. Disponível em: <https://dcpapers.dublincore.org/pubs/article/view/1013/980>. Acesso em: 21 out. 2019.

SALES, L. F.; SAYÃO, L. F. Há Futuro para as Bibliotecas de Pesquisa no Ambiente de eScience?. **Informação & Tecnologia (ITEC)**, Marília/João Pessoa, v. 2, n. 1, p. 30-52, jan./jul., 2015. Disponível em: <http://www.periodicos.ufpb.br/ojs/index.php/itec/article/view/26029/14677>. Acesso em: 15 out. 2018.

SALES, L. F.; SAYÃO, L. F. O impacto da curadoria digital dos dados de pesquisa na comunicação científica. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 17, n. esp., p. 118-135, 2012. Disponível em: <https://www.brapci.inf.br/index.php/res/v/36022>. Acesso em: 16 abr. 2020.

SANTANA, R. C. G. Ciclo de Vida dos Dados e o papel da Ciência da Informação. **Encontro Nacional de Pesquisa em Ciência da Informação**, v. 4, 2013. Disponível em: <http://www.brapci.inf.br/index.php/article/view/0000013772/fac7d2fec2c83f1e9208204b2fef10>. Acesso em: 19 maio 2020.

SANTOS, P. L. V. A. DA C.; SANTANA, R. C. G. Dado e Granularidade na perspectiva da Informação e Tecnologia: uma interpretação pela Ciência da Informação. **Ciência da Informação**, Brasília, v. 42, n. 2, p. 199–209, maio/ago. 2013. Disponível em: <http://revista.ibict.br/ciinf/article/view/1382>. Acesso em: 2 nov. 2019.

SANTOS, P. L. V. A. DA C.; SIMIONATO, A. C.; ARAKAKI, F. A. Definição de metadados para recursos informacionais: apresentação da metodologia BEAM. **Informação & Informação**, v. 19, n. 1, p. 146-163, jan./abr. 2014. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/15251>. Acesso em: 12 out. 2019.

SANTOS, P. L. V. A. DA C.; VIDOTTI, S. A. B. G. Perspectivismo e tecnologias de informação e comunicação: acréscimos à Ciência da Informação? **DataGramZero: revista de Ciência da Informação**, v. 10, n. 3, 2009. Disponível em: <http://www.brapci.inf.br/index.php/article/view/0000007096/b876db99ad9c3921c7293614a7d53a2f/>. Acesso em: 16 out. 2018.

SAYÃO, L. F. Uma outra face dos metadados: informações para a gestão da preservação digital. **Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, v. 15, n. 30, p.1-31, 2010. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/viewFile/1518-2924.2010v15n30p1/19527>. Acesso em: 4 jan. 2020.

SAYÃO, L. F.; SALES, L. F. Curadoria digital e dados de pesquisa. **AtoZ: novas práticas em informação e conhecimento**, v. 5, n. 2, p. 67–71, 2016. Disponível em: <https://revistas.ufpr.br/atoz/article/view/49708/30179>. Acesso em: 29 abr. 2020.

SAYÃO, L. F.; SALES, L. F. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade: Estudos (I &S)**, João Pessoa, v. 22, n. 3, p. 179-191, set./dez., 2012. Disponível em: <http://periodicos.ufpb.br/index.php/ies/article/view/12224>. Acesso em: 28 maio 2020.

SAYÃO, L. F.; SALES, L. F. Dados de pesquisa: contribuições para o estabelecimento de um modelo de curadoria de dados para o país. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 6, n. 1, 2013. Disponível em: <http://www.brapci.inf.br/index.php/article/view/0000014157/0f3127bfd6475f45e6950e1e0da20371>. Acesso em: 24 out. 2018.

SAYÃO, L. F.; SALES, L. F. Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores. **CNEN**, Rio de Janeiro, 2015. Disponível em: <http://www.cnen.gov.br/component/content/article/75-cin/material-didatico-cnen/160-guia-de-gestao-de-dados-de-pesquisa>. Acesso em: 16 fev. 2020.

SAYÃO, L. F.; SALES, L. F. Subsídios para a construção de um modelo de avaliação de sistemas de gestão de dados de pesquisa. **Ponto de Acesso**, v. 12, n. 3, p. 80-108, 2018. Disponível em: <https://brapci.inf.br/index.php/res/v/119905>. Acesso em: 21 fev. 2020.

SEGUNDO, S. J. da S.; ARAÚJO, W. J. de. Curadoria e 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, p. 113-152.

SEMIDÃO, R. A. M. Dados, Informação e Conhecimento enquanto Elementos de Compreensão do Universo Conceitual da Ciência da Informação: Contribuições Teóricas. 198 p. **Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação – Universidade Estadual Paulista ‘Júlio de Mesquita Filho’ (UNESP)**, Marília, 2014. Disponível em: https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/semidao_ram_me_mar.pdf. Acesso em: 5 jan. 2020.

SETZER, V. W. Dado, informação, conhecimento e competência. **DataGramZero**, v. 0, n. 0, 1999. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/7327>. Acesso em: 18 nov. 2019.

SIBI USP. Apoio ao pesquisador: dados de pesquisa. **SIBI USP**, São Paulo, 2016. Disponível em: <http://www.sibi.usp.br/apoio-pesquisador/dados-pesquisa/>. Acesso em: 30 out. 2018.

SIEBRA, S. A.; BORBA, V. R.; MIRANDA, M. K. F. O. Curadoria digital: um termo interdisciplinar. **Informação & Tecnologia**, v. 3, n. 2, p. 21-38, jul./dez. 2016. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/41848>. Acesso em: 26 fev. 2020.

SILVA, M. H. F. X. et al. Competências dos bibliotecários na gestão dos dados de pesquisa. **Ciência da Informação**, v. 48, n. 3, 2019. Disponível em: <https://brapci.inf.br/index.php/res/v/136517>. Acesso em: 8 maio 2020.

SILVA, F. C. C. O papel dos bibliotecários na gestão de dados científicos. **Revista Digital de Biblioteconomia & Ciência da Informação**, v. 14, n. 3, p. 387-406, 2016. DOI: 10.20396/rdbci.v14i3.8646333. Acesso em: 23 maio 2020.

SIMIONATO, A. C. **Gestão de dados de pesquisa**. Ago. 2018. 37 slides. Série de Seminários de Pesquisa do Departamento de Computação da UFSCar. Disponível em: <https://www.slideshare.net/karusimionato/gesto-de-dados-de-pesquisa>. Acesso em: 29 out. 2018.

SIMIONATO, A. C. Mapeamento dos Metadados para Dados Científicos. **Encontro Nacional de Pesquisa em Ciência da Informação**, 2017. Disponível em:

<http://enancib.marilia.unesp.br/index.php/xviiienancib/ENANCIB/paper/view/563>. Acesso em: 14 out. 2019.

USP Universidade de São Paulo. **O 4º Paradigma da Pesquisa Científica: como vamos enfrentar a tsunami de dados?**, 2007. 28 slides. Disponível em: https://edisciplinas.usp.br/pluginfile.php/3117041/mod_resource/content/3/1_quarto_paradigma.pdf#:~:text=O%204o%20paradigma%20%C3%A9%20concebido,perdido%20no%20mar%20em%202007. Acesso em: 27 jan. 2021.

VASCONCELOS, Y. O avanço da eScience impacta o modo tradicional de fazer ciência. **Biblioteca Central UFRGS**, Rio Grande do Sul, 2014. Disponível em: <https://www.ufrgs.br/blogdabc/o-avanco-da-escience-impacta-o-modo/>. Acesso em: 29 out. 2018.

WEBER, C. Curadoria digital de dados científicos: pelo viés de um periódico. **Revista P2P & Inovação**, Rio de Janeiro, RJ, v. 3 n. 1, p. 130-147, set./mar. 2017. Disponível em: <https://www.brapci.inf.br/index.php/res/v/9049>. Acesso em: 29 maio 2020.