

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO
CURSO DE ARQUIVOLOGIA

Maria Luiza da Fonseca

A CONTRIBUIÇÃO DA ARQUIVOLOGIA PARA A GESTÃO DE DADOS DE PESQUISA

Porto Alegre

2018

Maria Luiza da Fonseca

A CONTRIBUIÇÃO DA ARQUIVOLOGIA PARA A GESTÃO DE DADOS DE PESQUISA

Trabalho de conclusão de curso de graduação apresentado como requisito para a obtenção do título de Bacharela em Arquivologia da Universidade Federal do Rio Grande do Sul.

Orientador: Prof^o Dr. Rafael Port da Rocha

PORTO ALEGRE

2018

CIP - Catalogação na Publicação

Fonseca, Maria Luiza da
A contribuição da Arquivologia para a gestão de
dados de pesquisa / Maria Luiza da Fonseca. -- 2018.
68 f.
Orientador: Rafael Port da Rocha.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Faculdade
de Biblioteconomia e Comunicação, Curso de
Arquivologia, Porto Alegre, BR-RS, 2018.

1. Arquivologia. 2. Gestão documental. 3. Gestão de
dados de pesquisa. I. Rocha, Rafael Port da, orient.
II. Título.

Maria Luiza da Fonseca

A contribuição da Arquivologia para a gestão de dados de pesquisa

Trabalho de conclusão de curso de graduação apresentado como requisito para a obtenção do título de Bacharela em Arquivologia da Universidade Federal do Rio Grande do Sul.

Aprovado em: ____ de _____ de ____.

BANCA EXAMINADORA

Orientador: Prof. Dr. Rafael Port da Rocha
Universidade Federal do Rio Grande do Sul

Prof. Jorge Eduardo Enríquez Vivar
Universidade Federal do Rio Grande do Sul

Prof.^a Ma. Rita de Cássia Portela da Silva
Universidade Federal do Rio Grande do Sul

AGRADECIMENTOS

A todos que me motivaram a fazer esta graduação, meus pais e minha irmã, colegas e amigos.

À UFRGS e aos professores da Arquivologia, em especial àqueles que me conduziram à realização deste trabalho: professor Rafael e professora Rita.

RESUMO

A presente pesquisa busca verificar a aplicação da gestão documental, conceitos e instrumentos, na gestão de dados da pesquisa, atividade ainda pouco evidente no Brasil, mas bastante desenvolvida no exterior. Para isso foram analisados os principais aspectos que compõem ambas as tipologias de gestão, procurando estabelecer as concepções, diferenças e semelhanças entre elas. Num segundo momento, passa a identificar os benefícios da gestão documental e verificar a possibilidade de obtê-los também no ambiente de dados de pesquisa. Os conceitos da Arquivologia são desenvolvidos sob a perspectiva da gestão de dados, promovendo percepções sobre a possível influência e contribuição desta ciência sobre os dados de pesquisa. De forma geral, busca apresentar a Arquivologia e o profissional arquivista como potenciais colaboradores para a gestão de dados. Como embasamento foram consultados referências nacionais e internacionais do campo da Arquivologia e da gestão de dados de pesquisa.

Palavras chave: Arquivologia, gestão de documentos, gestão de dados de pesquisa.

ABSTRACT

This research seeks to verify records management's application, its concepts and instruments, in the research data management, an activity still not very evident in Brazil, but rather developed abroad. For that, the main aspects that make up both types of management were analyzed, trying to establish the conceptions, differences and similarities between them. In a second moment, it starts to identify the records management's benefits and verify the possibility of obtaining them also in the research data environment. The Archivology's concepts are developed from the data management's perspective, promoting perceptions about the possible influence and contribution of this science for the research data. In general, it seeks to present Archivology and the archivist as potential collaborators for data management. As background, national and international references were consulted in the field of Archivology and research data management.

Keywords: Archivology, records management, research data management.

SUMÁRIO

1 INTRODUÇÃO.....	9
2 DADOS DE PESQUISA.....	12
2.1 GESTÃO DOS DADOS DE PESQUISA.....	14
2.2 CICLO DE VIDA DO DADO DA PESQUISA.....	14
2.3 PLANO DE GESTÃO DE DADOS.....	18
2.4 DOCUMENTAÇÃO DOS DADOS.....	19
2.5 A PROVENIÊNCIA DOS DADOS DE PESQUISA.....	20
2.6 PRESERVAÇÃO DOS DADOS.....	21
2.7 COMPARTILHAMENTO DE DADOS.....	23
3 GESTÃO DOCUMENTAL.....	25
3.1 DOCUMENTO ARQUIVÍSTICO.....	25
3.2 OS PRINCÍPIOS ARQUIVÍSTICOS.....	27
3.2.1 Princípio do ciclo de vida.....	27
3.2.2 Princípios da Proveniência e de Respeito à Ordem Original.....	29
3.2.3 Princípio da Territorialidade.....	31
3.3 GESTÃO ARQUIVÍSTICA DE DOCUMENTOS.....	32
4 METODOLOGIA.....	36
5 A GESTÃO DOCUMENTAL ARQUIVÍSTICA COMO SUPORTE À GESTÃO DE DADOS DE PESQUISA.....	38
5.1 OS PRINCÍPIOS ARQUIVÍSTICOS NA GESTÃO DE DADOS DA PESQUISA.....	38
5.2 IDENTIFICAÇÃO E CONTEXTO.....	48
5.3 AS FUNÇÕES DA GESTÃO DOCUMENTAL E AS POSSÍVEIS RELAÇÕES COM A GESTÃO DE DADOS DE PESQUISA.....	52
5.4 O PAPEL DA DOCUMENTAÇÃO PARA A GESTÃO DE DADOS DE PESQUISA. .	57
6 CONCLUSÃO.....	61
REFERÊNCIAS.....	63

1 INTRODUÇÃO

É crescente no Brasil a preocupação com a produção, armazenamento e disponibilização dos dados gerados a partir do desenvolvimento de pesquisas científicas, realidade já vivenciada em outros países, especialmente na América do Norte, Europa e Austrália. Nestas regiões a gestão dos dados passou a ser mais uma exigência para o pesquisador e saber como administrá-los é fator determinante para alcançar os resultados desejados, assegurar a validação da pesquisa e disponibilidade a outros pesquisadores e, além disso, para a obtenção de financiadores para o fomento de seus trabalhos.

Pesquisa é um processo de investigação sobre as relações existentes entre fatos, fenômenos, situações ou objetos com o objetivo de oferecer respostas a um problema proposto. Para ser considerada “científica”, a pesquisa deve ser desenvolvida de maneira organizada e sistemática, seguindo um planejamento estabelecido pelo pesquisador. É a tipologia da pesquisa que determina a variedade e quantidade de dados e os documentos produzidos (UDESC, s.d)

A partir desta necessidade de gestão dos dados, com forte tendência de se estabelecer no Brasil, nota-se que os pesquisadores possuem à disposição guias e orientações, mas que são compilações das melhores práticas e muitas vezes focam apenas em partes do processo, carecendo de orientações a partir de uma visão integrada de todo o ciclo de vida do dado, dos processos e dos documentos produzidos. Como mesmo afirma a *Inter-university Consortium for Political and Social Research (ICPSR)* a respeito do seu Guia para Preparação e Arquivamento de Dados das Ciências Sociais, “é uma compilação das melhores práticas obtidas das experiências de muitos arquivistas e investigadores. [...] não pretende estabelecer políticas e procedimentos específicos para determinados arquivos, uma vez que sofrem variações” (ICPSR, 2012, p. 9, tradução nossa).

Apesar de hoje ser realizada principalmente no âmbito da Ciência da Informação e Biblioteconomia, existe uma possibilidade de a gestão de dados ser conduzida sob ótica da Arquivologia e desenvolvida com apoio dos arquivistas através das técnicas da gestão documental. Arquivologia é uma ciência que estuda a gestão documental em uma perspectiva orgânica, ou seja, seu objeto de estudo são as informações únicas vinculadas aos processos institucionais e para isso se vale de princípios e instrumentos instituídos. Da mesma forma como ocorre com os documentos arquivísticos, pode partir da análise dos contextos que influenciam a produção e trajetória dos dados de pesquisa, e por isso os conceitos e princípios

base da Arquivologia são capazes de contribuir para a construção de métodos aplicáveis.

Tomando como exemplo uma Universidade, a ação de gerir dados de pesquisa pode ser analisada como parte integrante de uma das suas atividades-fim: a pesquisa científica e a produção de conhecimento. Os dados registrados neste contexto podem ser vistos pela Arquivologia como documentos, pois são informações em um suporte resultantes de uma atividade/ação no âmbito da missão da Universidade. Sendo assim, considera-se aqui todo o decurso da pesquisa científica como um processo institucional gerador de informações e que também por esta razão carece de uma gestão eficiente.

A gestão documental, definida pelo Dicionário Brasileiro de Terminologia Arquivística como um “conjunto de procedimentos e operações técnicas referentes à produção, tramitação, uso, avaliação e arquivamento de documentos” (ARQUIVO NACIONAL, 2005, p. 100), pode ser vista como um instrumento possível para se operacionalizar a gestão de dados da pesquisa. Não apenas os resultados de pesquisa em si (trabalho final) podem ser objetos de estudo (documento) da Arquivologia, como também o dado durante seu ciclo de vida pode ser gerenciado arquivisticamente.

Com o aumento do nível de exigência em relação a apresentação do plano de gestão de dados da pesquisa, os arquivistas presentes nas Universidades e demais fomentadores da produção de conhecimento poderão se deparar com a missão de atuar neste nicho de gestão. Além disso, pode-se também avaliar um outro aspecto: os arquivistas como pesquisadores. O uso do conhecimento destes profissionais para gestão dos seus próprios dados de pesquisa pode ser um fator de contribuição no desenvolvimento mais célere e eficaz da Arquivologia.

Problema de pesquisa:

Princípios e métodos da Arquivologia para a gestão de documentos podem contribuir para o aprimoramento da gestão de dados da pesquisa?

Objetivo Geral:

Este estudo tem por objetivo verificar contribuições da gestão arquivística de documentos para o desenvolvimento da gestão de dados de pesquisa.

Objetivos Específicos:

- Identificar e contextualizar a gestão de dados de pesquisa;
- Definir e contextualizar a gestão documental;
- Identificar os princípios da Arquivologia no contexto da gestão de dados de pesquisa;
- Analisar o uso da gestão documental para promover o desenvolvimento da gestão de dados de pesquisa.

Para perceber a gestão de dados na perspectiva da gestão documental praticada pela Arquivologia, partiremos do reconhecimento e contextualização destas modalidades de gestão, compreendendo o que são e quais suas finalidades, em quais realidades estão inseridas, quais são seus princípios e atividades executadas, além dos benefícios que promovem em relação a produção de informação e conhecimento ao pesquisador, à ciência e à sociedade como um todo. Será no contexto da pesquisa científica, através da identificação dos princípios arquivísticos e das características dos documentos arquivísticos digitais que analisaremos a possibilidade do desenvolvimento da gestão de dados por meio das ações de gestão documental.

2 DADOS DE PESQUISA

Dados de pesquisa são elementos registrados, como fatos e números, utilizados como matéria prima para reflexões, discussões ou cálculos na construção de informação e conhecimento durante o desenvolvimento de uma pesquisa científica. Mais precisamente

são definidos como registros factuais (pontuações numéricas, registros textuais, imagens e sons) usados como fontes primárias para pesquisa científica, e que são comumente aceitos na comunidade científica como necessários para validar os resultados da pesquisa. (OECD, 2007, p. 13, tradução nossa)

Alguns exemplos são os dados estatísticos, resultados de experimentos, medições e observações de trabalho de campo e podem apresentar-se em diversos formatos, como áudio, imagem, vídeo ou animação. Neste universo plural, dados são classificados conforme sua origem (observacionais, computacionais, experimentais), natureza (números, imagens, vídeos ou áudio, software, algoritmos, equações, animações ou modelos e simulações) e fase da pesquisa (*raw data* ou dados brutos, crus ou preliminares, derivados, canônicos ou referenciais). Acima de tudo estes registros promovem a validação da pesquisa, pois demonstram a origem e comprovam os resultados obtidos. Como justificativa para o compartilhamento de dados, Borgman faz referência em relação às ciências sociais e humanas:

[...] (1) para reproduzir ou verificar pesquisas, (2) para disponibilizar ao público os resultados de pesquisas financiadas com fundos públicos, (3) permitir que outros façam novas perguntas sobre dados existentes e (4) para avançar no estado de pesquisa e inovação. (2012, p. 1.059, tradução nossa)

Como já visto, dados podem ser registrados nos mais diversos suportes. Porém, diante da realidade da sociedade em rede, amparada pela defesa do compartilhamento, da economia e da eficiência, “dado” neste trabalho se refere aquele em formato digital, objeto de técnicas de preparação e armazenamento diferenciadas, visto as particularidades que apresenta.

A produção de dados pode ser bastante intensa durante o processo de pesquisa. A diversidade permite que dados coletados e não publicados para o objetivo inicialmente proposto sigam a ser trabalhados mesmo após o término do financiamento, que os conjuntos de dados sejam reforçados ou sejam reaproveitados por outros pesquisadores e por outras áreas, resultando em otimização do tempo e da qualidade dos novos estudos.

Entre as diversas motivações para a divulgação e compartilhamento dos dados de pesquisa destacam-se ainda o incentivo à investigação científica e ao debate, a oportunidade de novas colaborações entre usuários e criadores de dados, maior eficiência ao evitar o retrabalho, maximização da transparência e responsabilidade, aumento do impacto e visibilidade da pesquisa, possibilidade de fornecer o crédito direto ao pesquisador e a difusão de recursos importantes para educação e treinamento.

O pesquisador possui à sua disposição diferentes maneiras de compartilhamento dos dados. Ele pode optar pelo depósito em um repositório institucional, centro de dados especializado, arquivo ou banco de dados, pela disponibilização on-line através de um projeto ou site institucional, pela apresentação a um periódico científico ou fazê-lo informalmente entre os pesquisadores.

Para auxiliar a comunidade científica na formulação e administração de seus dados, alguns princípios orientam a promoção da reutilização deste material. Os chamados princípios *FAIR* “ênfaticamente o aprimoramento da capacidade das máquinas de encontrar e usar os dados automaticamente, além de apoiar sua reutilização por indivíduos” (WILKINSON *et al.*, 2016, p. 1, tradução nossa). De acordo com a Associação de Bibliotecas Europeias de Investigação (LIBER, 2017), *FAIR* significa que os dados devem ser:

- localizáveis (*findable*): Dados e materiais complementares possuem metadados suficientemente ricos e um identificador único e persistente;
- acessíveis (*accessible*): Metadados e dados são compreensíveis para os seres humanos e máquinas e são depositados em um repositório confiável;
- interoperáveis (*interoperable*): Metadados usam uma linguagem formal, acessível, amplamente compreendida e aplicável para a representação do conhecimento;
- reutilizáveis (*reusable*): Dados e coleções possuem licenças de uso e fornecem informações precisas sobre suas proveniências.

É possível notar que estas características são consumadas principalmente através da formulação de metadados de qualidade durante todo o ciclo de vida do dado e poderão ser percebidas ao longo da análise das ações necessárias em cada fase do ciclo proposta neste trabalho.

2.1 GESTÃO DOS DADOS DE PESQUISA

A gestão de dados pode ser definida como um conjunto de ações pensado para se obter o máximo de controle, segurança e qualidade em todas as etapas da pesquisa científica. Para isso, considera um ciclo de vida do dado com demandas específicas em cada fase.

O aproveitamento de dados começou a ser discutido na década de 80 com publicações mais minuciosas a respeito das razões para aprimorar a proteção e o compartilhamento. Já é de conhecimento das agências de fomento que diante da intensa produção de informação e de toda a tecnologia disponível e em permanente evolução, não é mais cabível o desperdício de dados de qualidade. Com o objetivo de valorizar ainda mais a pesquisa científica a gestão de dados passou a ser pensada e exigida por estas agências assim como pelos seus financiadores.

Para a compreensão do que é a gestão de dados de pesquisa, passamos a definir o ciclo de vida, as operações executadas e os elementos necessários que delineiam cada etapa, a qualidade deste gerenciamento e, por consequência, a qualidade da própria pesquisa.

2.2 CICLO DE VIDA DO DADO DA PESQUISA

Diversos conceitos remetem a palavra ciclo à continuidade, renovação constante, evolução, o tempo entre acontecimentos. De modo geral, os dados de pesquisa possuem vida útil maior do que os projetos que os originam, justificando o uso do termo. Como explica Sant'Ana, a remissão ao ciclo se deve a necessidade de existir uma estrutura básica que auxilie na contextualização dos “momentos, características e requisitos bem como fatores que permeiam estes momentos” no processo de acesso e uso dos dados (2016, p. 117).

Modelos de ciclo de vida do dado são subconjuntos de modelos de ciclo de vida do processo geral de pesquisa, pois representam um aspecto separado deste processo. Frequentemente os modelos de ciclo de vida do dado fazem parte de todo o processo de pesquisa, mas destacam os aspectos ou atributos específicos dos dados (CARLSON, 2014).

A concepção de um ciclo de vida do dado contribui para o estabelecimento de ações específicas concernentes as necessidades de coleta, processamento, análise, compartilhamento, recuperação, segurança, preservação, armazenamento e usos, ou seja, contribui para o gerenciamento do dado gerado ao longo de sua existência.

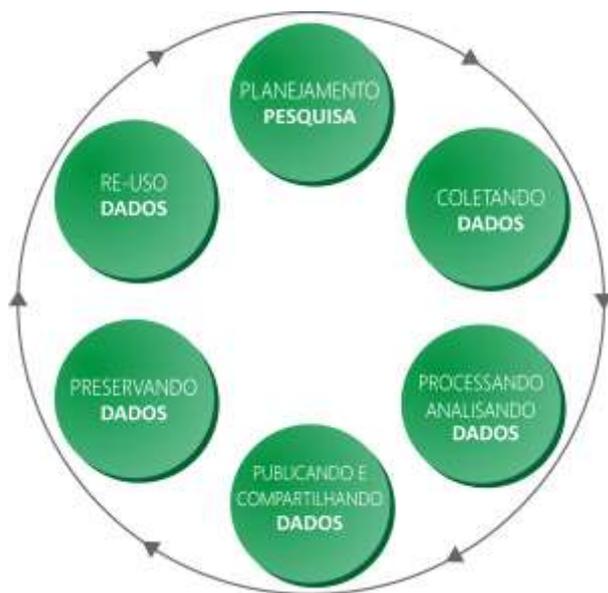
Conforme Carlson, é uma demanda proveniente da complexidade do processo de

pesquisa, pois envolve a “coordenação entre pessoas, agências, recursos e eventos”. O autor também esclarece a respeito dos modelos de ciclo de vida:

ajudam a definir e ilustrar visualmente esses processos complexos, facilitando a identificação das partes componentes ou etapas distintas dos dados da pesquisa. Ao quebrar o processo de como os dados surgem e como eles crescem e evoluem à medida que são aplicados para cumprir seu propósito em estágios inter-relacionados, as necessidades específicas do pesquisador são mais prontamente identificadas. Dividir o processo de dados em estágios de componentes também ajuda a identificar e contabilizar os papéis, responsabilidades, marcos e outros componentes-chave na oferta de serviços de dados. A definição dessas necessidades e componentes informa diretamente a articulação de quais serviços de dados podem ser desenvolvidos para atender a essas necessidades e quando esses serviços podem ser oferecidos para serem mais úteis. (2014, p. 65, tradução nossa)

Um exemplo de ciclo de vida é o concebido pelo *UK Data Service*, representado na figura 1:

FIGURA 1 – Tradução do modelo de Ciclo de Vida do *UK Data Service*



FONTE: Dados de Pesquisa Abertos.
Disponível em: dadosdepesquisa.rnp.br/category/dados-de-pesquisa/

Este modelo compreende as seguintes etapas:

- Coletar: função que representa a produção de dados pelo pesquisador e engloba a captura e a aquisição de dados de terceiros.
- Processar: indica as ações de introduzir, digitalizar, transcrever e traduzir dados; conferir, validar, limpar, anonimizar, definir a procedência; descrever e documentar os

dados, realizar o gerenciamento.

- Analisar: abrange a exploração e interpretação dos dados, verificação de possibilidades de pesquisa e as fontes de dados.
- Compartilhar: estabelecer direito autoral, criar a documentação para os usuários e os metadados das descobertas, definir o acesso mais adequado à tipologia dos dados e publicar e promover a sua difusão.
- Armazenar: envolve ações de preservação, como a definição dos melhores formatos e mídias, cópias de segurança (*back up*), curadoria e a documentação destas ações.
- Reusar: significa a condução de análises secundárias e revisões da pesquisa e das publicações, realização de pesquisas de acompanhamento, exame dos resultados, criação de estudos semelhantes e o uso dos dados por outros e para diferentes objetivos, como, por exemplo, para a educação.

Cabe ressaltar que não existe atualmente uma definição de ciclo de vida padrão, variações podem ocorrer entre as instituições que adotam a gestão de dados e conforme a matéria pesquisada e material produzido. Assim, as variações também ocorrem entre as representações de ciclo. A figura 2 demonstra o ciclo definido pelo *DataOne* representado pelas operações sobre os dados de forma mais fragmentada em relação ao anterior e somente até a etapa de análise.

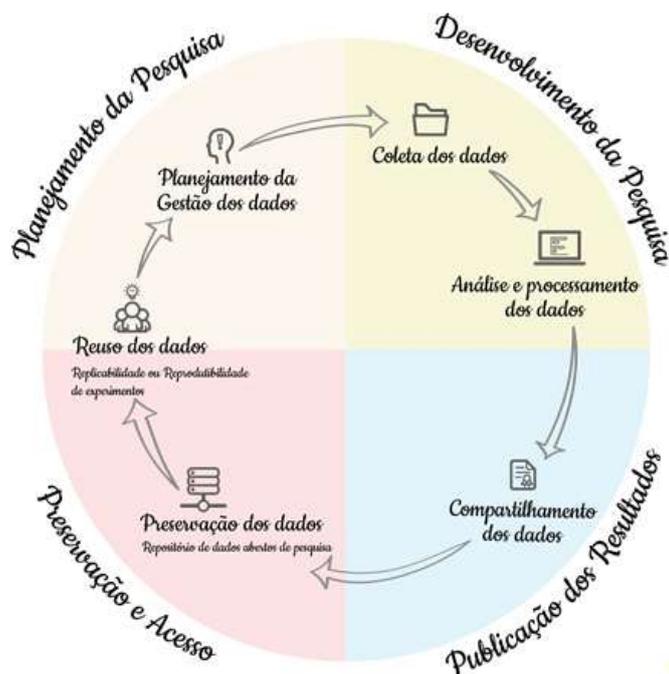
FIGURA 2 – Tradução do modelo de Ciclo de Vida do *DataOne*



FONTE: Sayão e Sales (2015, p. 12)

Já o ciclo elaborado pelo CEDAP, como apresentado na figura 3, demonstra as etapas da pesquisa e a relação com o ciclo dos dados. Verifica-se como o estabelecimento de etapas para pesquisa orienta as operações necessárias em relação aos dados.

FIGURA 3 – Relação entre os ciclos de vida da pesquisa e dos dados



FONTE: CEDAP - Centro de Documentação e Acervo Digital da Pesquisa da Universidade Federal do RS (UFRGS)

Portanto, diante das possibilidades de definição de ciclo de vida e de adaptação para representar as etapas e funções, contribuem Sayão e Sales:

A importância desses modelos é que eles oferecem uma estrutura que representa as muitas operações que precisarão ser realizadas sobre os registros de dados durante a sua vida, garantindo que eles possam ter a sua usabilidade otimizada e estendida (2015, p. 11).

O ciclo de vida se mostra portanto como uma ferramenta para o pesquisador na organização e articulação das tarefas que necessita cumprir para gerir seus dados e receber os melhores resultados sobre seus esforços e investimentos.

2.3 PLANO DE GESTÃO DE DADOS

A importância dada à fase inicial de planejamento é representada pelo quesito de elaboração de um plano de gestão de dados representado por meio de um documento formal a “ser tratado como uma carta de intenções que considere o que realmente é necessário para a preservação, compartilhamento e reuso dos dados” (SAYÃO e SALES, 2015, pg. 15). Segundo Schiermeier:

Um plano de gerenciamento de dados explica como pesquisadores vão lidar com seus dados durante e depois de um projeto, e abrange criar, compartilhar e preservar dados de pesquisa de qualquer tipo, incluindo texto, planilhas, imagens, gravações, modelos, algoritmos e *software*. (2018, p. 403, tradução nossa)

O objetivo deste plano é programar as metodologias, operações e documentos que devem ser formalizados para a efetiva gestão dos dados, minimizando o risco de perda e fornecendo orientação para outros pesquisadores sobre como usar os dados no futuro. O planejamento auxilia ainda a identificar e ao mesmo tempo amenizar possíveis barreiras ao compartilhamento dos dados, como os acordos de confidencialidade, e às oportunidades de comercialização da pesquisa (DAVIDSON *et al.*, 2014). Melhorias nos serviços de gerenciamento de dados de pesquisa também podem ser implementadas a partir dos resultados do planejamento.

Especialmente nos países de vanguarda da pesquisa científica, como por exemplo Estados Unidos, Austrália, Reino Unido e Holanda, órgãos de financiamento passaram a exigir dos pesquisadores a elaboração do plano de gestão de dados para a obtenção de financiamento de seus trabalhos. No Brasil, para a FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo – o plano de gestão de dados se tornou um componente obrigatório na fase de submissão de um projeto ao entender que a gestão adequada racionaliza recursos, facilita a reprodutibilidade e a promoção de outras análises por novas gerações de pesquisadores¹.

Igualmente no Brasil, o Guia De Gestão de Dados de Pesquisa, iniciativa da Comissão

¹ Exemplos sobre a exigência da apresentação do plano de gestão de dados nos países citados podem ser observados através dos *links*:

<www.nsf.gov/bfa/dias/policy/dmp.jsp>

<libguides.library.curtin.edu.au/c.php?g=202401&p=1333108>

<wellcome.ac.uk/>

<rdm.uva.nl/en/planning/funder-requirements/funder-requirements.html>

<www.fapesp.br/gestaodedados/>

Nacional de Energia Nuclear – CNEN descreve os elementos mais frequentemente recomendados para constituir o plano e que possibilitam conhecer e definir:

- As características qualitativas e quantitativas dos dados e as principais ações futuras de gestão;
- Quais metadados serão necessários, como se dará a criação e captura e qual o padrão escolhido (quais informações serão apresentadas);
- A política de acesso, compartilhamento e reuso;
- Preservação digital dos dados para arquivamento de longo prazo;
- Orçamento e custos envolvidos na gestão dos dados.

Muito comumente as instituições já fornecem aos pesquisadores modelos para a elaboração do plano de gestão de dados. Além disso, o pesquisador pode contar com ferramentas *online* que o auxiliam na elaboração, revisão e compartilhamento do seu plano de gestão de dados, como o *Data Management Planning Tool - DMP Tool*² e o *DMP Online*³. Oferecendo modelos de requisitos, textos auxiliares, orientações e recursos para as melhores práticas, estas ferramentas funcionam como assistentes para que o plano esteja adequado às exigências das instituições e dos financiadores.

2.4 DOCUMENTAÇÃO DOS DADOS

Ação fundamental é a descrição completa dos dados, ou seja, a operação de documentação. Números e caracteres transmitem pouco ou até mesmo nenhum significado, a menos que sejam documentadas de alguma forma.

A documentação pode se apresentar sob a forma de metadados ou de documentos (texto, vídeos, planilhas, entre outros formatos) e é um processo que deve andar em paralelo com todas as práticas de gestão, pois desta forma permite que sejam “compreendidos e interpretados por qualquer usuário agora e no futuro” e “explica como estes recursos foram coletados ou gerados, o que os dados significam, qual é o seu conteúdo e estrutura, quais foram as manipulações a que eles foram submetidos” (SAYÃO e SALES, 2015, p. 27).

Metadados, ou o conjunto de informações sobre o dado, fornecem detalhes sobre a

² Disponível em: dmptool.org

³ Disponível em: dmponline.dcc.ac.uk

coleta, processamento, interpretação e permitem que os dados e arquivos sejam descobertos, usados e citados corretamente. A função dos metadados vai além:

Metadados incluem descrições de como os dados e arquivos são nomeados, fisicamente estruturados e armazenados, bem como detalhes sobre os experimentos, métodos analíticos e contexto de pesquisa. Geralmente, a utilidade e a longevidade dos dados se relacionam diretamente com quão completos e compreensíveis são os metadados. (MICHENER, 2015, tradução nossa)

Devido a todo este potencial de referência que possuem, os metadados são classificados em quatro tipos:

- Descritivos: para identificar o dado, descrevem informações como autor, criação, assunto, etc;
- Administrativos: apresentam informações a respeito do gerenciamento do dado, como proveniência, preservação digital, direitos de acesso, aquisição, etc;
- Técnicos: apresentam aspectos técnicos como resolução, tipo de arquivo, formato, etc;
- Estruturais: descrevem como se dá a estrutura de todos estes metadados que formam o documento e a relação entre eles.

Nota-se portanto que o objetivo dos metadados não é apenas a descrição do conteúdo. Cada tipologia vem a contribuir para o potencial de registro e a documentação, quanto mais completa, melhor garante a credibilidade dos significados que o dado se propõe a transmitir.

2.5 A PROVENIÊNCIA DOS DADOS DE PESQUISA

As informações de proveniência sobre um recurso fornecem informações a respeito de sua origem, apontando o produtor e aspectos sobre criação e modificação. Em diversas áreas a compreensão do significado de proveniência vem sendo estudado e adotado como princípio fundamental e vital para avaliar a qualidade e a confiabilidade das informações.

O Serviço de Dados Australiano contribui com a definição de que “a proveniência está associada não apenas ao histórico de um produto de dados no tempo, mas também às relações entre um produto de dados e outras entidades que permitem a criação dos dados” (ANDS, 2018, p. s.n), ou seja, sendo um registro que descreve entidades e processos envolvidos ou com o poder de influência na produção e entrega dos recursos, a proveniência fornece uma

base crítica para avaliar a autenticidade, promovendo a confiança e permitindo inclusive a reprodutibilidade dos dados.

A documentação (ou registro) da origem dos dados de pesquisa se dá com o fornecimento de metadados de proveniência e torna-se item importante para determinar também a qualidade da informação, a integração de diversas fontes e como creditar os originadores na reutilização dos dados.

Cabe considerar a diferença entre metadados descritivos e de proveniência:

Enquanto os metadados descrevem um recurso de informação, os metadados de proveniência são considerados uma forma de metadados administrativos e se diferem daqueles que descrevem o conteúdo intelectual. Por exemplo, o título, autor e resumo de um trabalho são considerados descritivos. Por outro lado, as informações que descrevem a fixidez de arquivos, como a Impressão Digital Numérica Universal (UNF) ou o *hash* criptográfico, ou descrevem a versão e o histórico de alterações, como a data de criação, seriam metadados de proveniência, porque se referem à cadeia de controle e modificação do objeto. (ALTMAN apud CODATA-ICSTI, 2013, p. 13, tradução nossa)

Isto demonstra que as informações sobre proveniência, uma vez consideradas metadados administrativos, se relacionam mais significativamente com a gestão dos dados do que com a caracterização. Estes metadados proporcionam o conhecimento de diversos aspectos que influenciam ou até mesmo alteram o dado ao longo do seu ciclo de vida. Devido a esta multiplicidade de ações que afetam os dados, é importante conhecê-las e documentá-las, seja uma alteração, aplicação de criptografia ou evolução de formato, pois assim a manutenção da integridade pode ser assegurada.

2.6 PRESERVAÇÃO DOS DADOS

Tão fácil quanto produzir dado é a sua possibilidade de perda, corrompimento ou a falta de controle em relação ao seu uso. Item de imprescindível importância para a gestão dos dados é a segurança, aspecto que se obtém e aprimora com o “*backup*, arquivamento, preservação e proteção física, criptografia e ainda as leis que governam a proteção dos dados” (SAYÃO e SALES, 2015, p. 39). O planejamento da segurança deve considerar todo o ciclo de vida dos dados e abarcar aspectos desde a realização de cópias de segurança e locais de armazenamento, controle de acesso, *firewall* e antivírus, registro de eventos, manutenção dos equipamentos utilizados no processamento e arquivamento, envio e destruição dos dados, até

a proteção de informações pessoais que podem compor um dado.

As iniciativas de preservação também devem ser contínuas em todo o ciclo de vida do dado e incluem a definição de um local de arquivamento a longo prazo apto a executar a denominada curadoria de dados de pesquisa. Como explicam Sayão e Sales (2012, p. 180) o principal desafio da curadoria é “preservar não somente o conjunto de dados, mas de preservar, sobretudo, a capacidade que ele possui de transmitir conhecimento para uso futuro das comunidades interessadas”. Em resumo, a curadoria nada mais é do que a gestão ativa e a preservação dos recursos digitais durante todo o ciclo de vida. Corrobora ainda Arellano sobre a visão do conceito de “gerenciamento digital”⁴ que considera a preservação digital como “componente de um conjunto de serviços, políticas e especialistas que constituem o contexto do ciclo de vida da informação digital” (2008, p. 42).

É através da gestão do objeto digital⁵ que os repositórios, além de prover armazenamento e preservação segura, fomentam a visibilidade dos dados, as possibilidades de reuso e compartilhamento, a interação entre os pesquisadores e a administração de licenças. Como exemplificam Sayão e Sales (2012, p. 69) é “o estabelecimento de *links* entre dados e artigos de periódicos que possibilita a formulação de novas concepções de publicações científicas e de um alto grau de contextualização para os dados, artigos e projetos” e de fato, mais do que uma infraestrutura de preservação, os repositórios funcionam como uma intersecção entre produtores e usuários de dados, informações e documentos. Porém, para serem considerados confiáveis e eficazes, devem possuir os requisitos de proteção dos dados contra a obsolescência tecnológica e a fragilidade das mídias (SAYÃO, 2010). Além disso, a validade legal de documentos eletrônicos depende de sistemas que garantam a autenticidade e integridade das informações neles contidos (ARELLANO, 2008).

A gestão no repositório acompanha o fluxo dos dados desde a ingestão ou entrada, passando pela gestão dos metadados, acesso até a administração e planejamento da preservação e armazenamento. Conforme Sayão (2010) o repositório garante a confiabilidade dos dados quando, entre algumas das responsabilidades:

- Oferece um sistema organizacional que apoie sua viabilidade de longo prazo e a informação digital da qual ele tem responsabilidade;
- Demonstra responsabilidade fiscal e sustentabilidade;

4 LAVOIE; DEMPSEY apud ARELLANO, 2008, p. 42.

5 Arellano (2008, ps. 47, 48) define objeto digital como “um componente digital necessário para reproduzir um documento eletrônico” composto pelos diversos tipos de metadados, arquivos de preservação e arquivos de disseminação.

- É projetado de acordo com convenções e padrões aceitos no sentido de assegurar a gestão, o acesso e a segurança contínua dos materiais depositados;
- Estabelece metodologias para avaliação dos sistemas;
- Adota políticas, práticas e desempenho que possam ser auditáveis e mensuráveis;

Outro fator que agrega segurança na manipulação de dados são as certificações de repositórios confiáveis que podem se dar em diversos padrões ou níveis de auditoria, como a ISO 16363, Nestor, *Data Seal Approval* e *Core trusted Seal*.

Para cumprir os requisitos de preservação, um ferramental importante na seleção de uma base que atenda às especificidades do trabalho desenvolvido são os diretórios de repositórios de dados, como o Re3data⁶. Apresentando as opções disponíveis a nível mundial, através deles o pesquisador pode verificar as tecnologias empregadas, métodos de gestão e submissão, licenças, fluxos, custos, níveis de certificação e estratégias de preservação mais adequadas. Destaca-se ainda a iniciativa da USP – Universidade de São Paulo, junto a outras Universidades brasileiras, que desenvolveu um metabuscador⁷ de dados científicos visando compilar em um único lugar os dados científicos disponibilizados por pesquisadores das instituições envolvidas.

2.7 COMPARTILHAMENTO DE DADOS

O compartilhamento dos dados da mesma forma deve ser adequadamente inserido no plano de gestão, afinal este é um dos principais objetivos da definição de um ciclo de vida: habilita o acesso para a validação da pesquisa e para o reuso dos dados. O compartilhamento depende da descrição precisa do dado, seu armazenamento confiável e da preservação adequada de formatos e suportes. Aqui o pesquisador deve demonstrar que atua de acordo com os princípios éticos aos quais se submeteu e neste quesito os direitos e licença para o uso atribuídos também garantem segurança e qualidade aos dados compartilhados.

É especialmente nesta etapa que a adoção dos princípios *FAIR* se reforça, pois direcionam à promoção do máximo aproveitamento de dados de pesquisa e fornecem orientação para a sua administração, sendo relevantes para todas as partes interessadas no

⁶ Disponível em: www.re3data.org

⁷ Disponível em: metabuscador.sc.usp.br

atual “ecossistema digital” (LIBER, 2017, p. 1)

Com estas considerações é possível perceber como a definição de um ciclo de vida pode otimizar as ações de gestão dos dados, pois, devido sua natureza progressiva e circular, cada fase é uma evolução da sua precedente, as ações estão interligadas, dependendo e contribuindo para as demais. São inúmeros fatores que condicionam a trajetória do dado e o pesquisador pode planejar cada etapa individualmente, mas sem a necessidade de ser um projeto estanque diante da dinamicidade da pesquisa.

Corroborando o *UK Data Service* (2018, tradução nossa): “Dados bem organizados, bem documentados, preservados e compartilhados são inestimáveis para o avanço da investigação científica e para intensificar as oportunidades de aprendizado e inovação”.

3 GESTÃO DOCUMENTAL

A gestão de documentos na Arquivologia é melhor definida ao se conhecer as especificidades que a diferenciam e sua trajetória na história que também posiciona e reforça a relação com os princípios e as necessidades de controle da massa documental.

Nesta seção são apresentados os elementos que fundamentam a atividade de gestão documental arquivística. São eles: o documento arquivístico e os princípios norteadores da teoria arquivística que correspondem ao princípio do ciclo de vida, de proveniência, respeito a ordem original e ao princípio da territorialidade. Em seguida são expostos os conceitos e instrumentos de gestão documental.

3.1 DOCUMENTO ARQUIVÍSTICO

Segundo Santos (2015, p. 106) “documento é um conjunto de dados estruturados, apresentados em uma forma fixa, representando um conteúdo estável, criado com o objetivo de registrar, transmitir e/ou preservar informações”. No escopo da Arquivologia busca-se o vínculo com o objetivo da produção do documento e não com o local de armazenamento, sendo assim diferenciado o “documento arquivístico” do “documento de arquivo”. Somado ao conceito anterior, Santos (2015, p. 113) define o documento arquivístico como aquele “[...] produzido ou recebido por pessoa física ou jurídica (pública ou privada), no exercício de uma atividade, observando os requisitos normativos da atividade à qual está relacionado, e preservado como evidência da realização dessa atividade”.

Esta definição permite compreender a gestão documental executada pela Arquivologia, pois não é uma atividade realizada exclusivamente por ela, outras áreas também possuem seus métodos. São as particularidades dos documentos arquivísticos que distinguem a gestão arquivística de documentos da simples gestão documental. Para evidenciá-las, servem como base as características do documento arquivístico elencadas por Santos (2015), visto sua construção alicerçada em ampla pesquisa a especialistas de todo o mundo:

- **Fixidez:** nos documentos, a estabilidade do conteúdo e da forma são inerentes à manutenção da percepção desejada pelo seu criador em relação às informações registradas. É observada tanto nos documentos físicos quanto nos digitais.

- **Organicidade:** relação que um documento possui com os demais vinculados a mesma atividade e com o fundo ao qual pertence. Sua contextualização se dá a partir da relação com os demais.
- **Naturalidade:** os documentos arquivísticos são originados de um processo natural, em conformidade com os interesses do produtor e a legislação, e assim também são acumulados.
- **Unicidade:** cada via ou até mesmo cada cópia do documento exerce uma função única vinculada à sua produção e uso.
- **Autenticidade:** significa a existência de um controle do processo de criação, manutenção e custódia do documento arquivístico para prevenir as alterações não autorizadas e para permitir a identificação das modificações. O documento é o que parece ser.
- **Imparcialidade:** o documento é criado para atender as demandas relacionadas à execução da atividade que exige sua produção e dessa forma são mais confiáveis porque demonstram fidelidade aos fatos e ações.

Em relação aos documentos digitais, a completude, confiabilidade e integridade também definem o documento arquivístico:

- **Completude:** presença de todos os elementos intrínsecos e extrínsecos exigidos pelo produtor e pelo sistema jurídico-administrativo a que pertence (CONARQ, 2011, p. 126).
- **Confiabilidade:** sinônimo de fidedignidade, a capacidade de representar os fatos tratados; “conferida aos registros pelos controles exercidos sobre a criação e pela integridade/completude de sua forma.” (DURANTI apud CONARQ, 2011, p. 22)
- **Integridade:** refere-se ao documento completo que não sofreu nenhum tipo de corrupção ou alteração não autorizada nem documentada (CONARQ, 2011, p. 129).

Cabe salientar que estas características não são fixas, ou seja, são diversas as interpretações a respeito das nomenclaturas e da presença nos documentos arquivísticos. No entanto, estas características acabam por reforçar o significado de documento arquivístico uma vez que em todas elas é constante a presença da relação produtor-atividade-documento.

3.2 OS PRINCÍPIOS ARQUIVÍSTICOS

3.2.1 Princípio do ciclo de vida

O Ciclo Vital dos documentos é definido como “sucessivas fases por que passam os documentos de um arquivo, da sua produção à guarda permanente ou eliminação” (ARQUIVO NACIONAL, 2005). Porém, a origem do termo não é unânime. Para Barros (2010) no ideário promulgado pelo arquivista britânico Hilary Jenkinson na década de 20, os documentos passavam por uma série de períodos de vida e, baseados nisso, os arquivistas Schellenberg e Posner criaram a metáfora do ciclo vital dos documentos. Outros autores apontam a proposição do termo pelo arquivista americano Philip C. Brooks em 1940, apesar de constatações que apontam “a existência de várias práticas relacionadas ao conceito antes de seu surgimento ‘oficial’” (FILHO e SOUSA, 2016, p.190).

Como as etapas até a destinação final não são “distintas, engessadas e sem ligação” como entenderam Rosseau e Couture (MEDEIROS e AMARAL, 2010, p. 301), neste trabalho busca-se apresentar uma noção geral a respeito do princípio, sem desconsiderar as demais interpretações desenvolvidas por autores e profissionais de diferentes realidades.

Uma anterior concepção sobre o campo de atuação da Arquivologia ditava que os arquivos eram locais de custódia e responsáveis principalmente pela guarda da memória. Graças a sua característica dinâmica, a Arquivologia teve a habilidade de adaptação às diferentes realidades e necessidades da sociedade diante da produção documental e informacional ao longo dos séculos. Atualmente a arquivística se depara com a necessidade de dominar a trajetória do documento até mesmo antes de ser produzido, ou seja, as ações de gestão devem ser pensadas pela entidade produtora em conjunto com todos os envolvidos e interessados (interna e externamente) desde o momento da concepção do documento, passando pela tramitação, organização, uso e destinação final. Isso se deve a enorme e veloz produção informacional estimulada também pelas tecnologias e que conduz os arquivos pela chamada Arquivística integral na Sociedade do Conhecimento e os considera como sistemas de informação (SANTOS, 2015).

Sob a perspectiva do paradigma pós-moderno, com foco agora nas informações vinculadas aos processos institucionais (SANTOS, 2015), há uma significativa ampliação de escopo e a multiplicidade dos objetos de estudo, incluindo um “macro” e “intelectual” que é “a informação ou, mais precisamente, os dados que possibilitam a informação” (BELLOTTO

apud SANTOS, 2015, p. 91).

Mas é voltando-se à origem dos termos que se compreende a associação do estabelecimento de um ciclo de vida dos documentos com a gestão documental. Esta noção de ‘transcurso do tempo’ surgiu nos Estados Unidos na década de 30 e foi pensada como uma nova alternativa para as necessidades dos arquivos da Administração Federal daquele país diante da massiva produção documental, mudando o conceito dos arquivos para locais com “potencial valor informativo para pesquisadores” (DINGWALL, 2017, P. 209). Nasce então o ‘modelo do ciclo vital’ caracterizado como uma representação dos estágios da existência do documento, começando com a criação e terminando com a sua destruição ou preservação permanente.

Derivada do ciclo vital, a Teoria das Três Idades surge da necessidade de se classificar o documento no contexto da atividade que lhe deu origem, especialmente em relação a frequência de uso, e de se preservar por motivos superiores ao da sua criação (como prova ou memória, por exemplo). A três idades são assim caracterizadas:

- 1ª idade ou corrente: os documentos representam as atividades ainda em curso ou são objetos de consultas frequentes;
- 2ª idade ou intermediária: os documentos já não são de uso corrente, há alguma consulta, porém, por interesse administrativo, aguardam a destinação final normalmente em arquivo diverso dos documentos correntes.
- 3ª idade ou permanente: os documentos são mantidos pelo valor histórico, probatório e informativo e são definitivamente preservados.

Lousada também contribui em relação às fases:

na fase corrente e intermediária, os usuários são, em sua maioria, os próprios produtores dos documentos; trata-se de uma consulta imediata a fim de atender objetivos específicos, ao passo que, na fase permanente, o uso está relacionado a questões de cultura, como a memória, por exemplo – o que não impede que haja consulta por razões ligadas à administração (2015, ps. 51 e 52).

A sequência de estágios pela qual passa o documento inclui a criação, manutenção e uso, destinação final, armazenamento e disponibilidade no arquivo permanente ou outra instituição arquivística, conforme o caso (SANTOS, 2015). Não há um consenso ou uma definição precisa sobre estas etapas, pois a caracterização varia entre os autores e profissionais da área.

O início do século XX foi igualmente marcado pelo desenvolvimento das Teorias da Administração e a aplicação dos conceitos por elas concebidos passaram da mesma forma a ser aplicados na área dos arquivos surgindo o termo *records management* ou gestão dos documentos. Através de técnicas e procedimentos orientados para a produção, organização dos documentos durante o uso pelo produtor até o recolhimento ou eliminação, a gestão documental deu à atividade arquivística um novo papel, pois passou a atuar em etapas e participar de soluções antes não contempladas.

Como cita Mundet, a finalidade da gestão documental e sua relação com o ciclo de vida podem ser assim compreendidos:

[...] a gestão se estende ao ciclo completo dos documentos, desde a produção até a eliminação ou conservação permanente. O objetivo é assegurar uma documentação adequada, evitar o não essencial, simplificar os sistemas de criação e uso dos papéis, melhorar a forma de organizar e recuperar os documentos, proporcionar o cuidado adequado e o depósito a baixo custo dos documentos nos arquivos intermediários e assegurar o expurgo adequado dos documentos desnecessários. (RICKS apud MUNDET, 2012, p. 91, tradução nossa)

Importante também a interpretação de que: “[...] a concepção do ciclo vital implica o reconhecimento de que os arquivos não devem mais gerenciar apenas um produto, mas um processo” (JARDIM apud LOUSADA, 2015, p. 52).

3.2.2 Princípios da Proveniência e de Respeito à Ordem Original

Também chamado de princípio de respeito aos fundos, é considerado a base do posterior desenvolvimento científico da arquivística, uma vez que proporcionou suas próprias perspectivas e características, diferenciando-a das demais áreas às quais esteve ligada ao longo da sua história. Para compreendê-lo, cabe ressaltar os significados e a relação entre fundo e organicidade:

- Fundo: “conjunto orgânico de documentos produzidos e/ou recebidos por uma entidade pública ou privada no decurso da sua atividade” (UNIVERSIDADE DO PORTO, 2018).
- Organicidade: relação natural entre os documentos do arquivo (fundo) em decorrência das atividades da sua entidade produtora (ARQUIVO NACIONAL, 2005).

Como explica Santos, em 1841 Natalis de Wally, chefe da Seção Administrativa dos

Arquivos Departamentais do Ministério do Interior francês, postulou “que todos os documentos provenientes de uma instituição deveriam ser mantidos juntos quando enviados aos grandes depósitos arquivísticos, não sendo misturados aqueles acumulados por outras instituições” (2015, p. 148). Foram os arquivistas prussianos (Alemanha), os responsáveis por complementá-lo em relação a organização interna dos arquivos, recomendando que os documentos de cada fundo fossem mantidos na ordem dada pelo seu produtor, derivando daí outro conceito da área, o de respeito à ordem original. Este conceito é aplicável aos documentos de valor permanente, recolhidos por seu valor probatório ou de memória, e seu objetivo é manter ou recuperar a organização que possuíam nas suas fases iniciais.

Schellenberg considera o princípio da proveniência pela sua importância para a teoria arquivística e para a classificação de documentos de arquivo, que deve respeitar a proveniência dos documentos. “O princípio serve para tornar conhecidas a natureza e significação dos documentos” (SCHELLENBERG, 1963, p. 110), além de manter a sua integridade. Desse modo, o autor relaciona a importância da aplicação do princípio da proveniência com a razão de ser do documento (SOUZA, 2012).

Com o intuito de colaborar com a compreensão das seções seguintes deste trabalho, no quadro 1 são evidenciados os conceitos de proveniência e ordem original com base nas definições do Dicionário Brasileiro de Terminologia Arquivística (ARQUIVO NACIONAL, 2005):

Quadro 1 – Os conceitos de proveniência e dos princípios da proveniência e do respeito à ordem original

Proveniência	Princípio da proveniência	Princípio do respeito à ordem original:
Indica a entidade coletiva, pessoa ou família produtora de documentos.	O conjunto de documentos produzido por uma entidade coletiva, pessoa ou família não deve ser misturado aos de outras entidades produtoras. Também chamado princípio de respeito aos fundos.	O arquivo deveria conservar o arranjo dado pela entidade coletiva, pessoa ou família que o produziu.

Segundo Lousada (2015) a formação das correntes teóricas na Arquivologia é delineada pelo paradigma inicial de características custodial, patrimonialista, historicista e

tecnicista e, em um segundo momento e ainda em consolidação, pelo paradigma informacional, científico e pós-custodial quando, a partir da segunda metade do século XX, a tecnologia passa a influenciar a dinâmica informacional. A autora ainda contextualiza que na década de 80 novas publicações recuperaram o princípio de proveniência baseado agora na informação, no estudo da função e do contexto de criação dos registros. De físico e estático, passou a ser dinâmico e funcional se adaptando especialmente aos documentos eletrônicos. Surge também a defesa de uma “proveniência social” ao perceber a análise sociocultural como mais um fator de influência na produção e nos processos documentais (TOGNOLI apud LOUSADA, 2015, p. 78).

A proveniência tem se mostrado bastante flexível em relação às frequentes mudanças pelas quais passam o documento e seus usuários e se mantém como uma ferramenta essencial ao arquivista na compreensão de “como as organizações criam, usam e descartam as informações na era digital” (LOUSADA, 2015, p. 79). Sendo assim, um fundo entendido como o produto das atividades de uma entidade, leva automaticamente à necessidade de se conhecer quem o produziu, sua natureza, estrutura, evolução, área e forma de atuação, resultando em classificação correta e adequada ao uso dos documentos e na transformação das informações em possibilidades de conhecimento (COOK, 2012).

3.2.3 Princípio da Territorialidade

Este princípio é definido por diversos autores como derivado do princípio da proveniência, instituindo que os documentos devem ser conservados nos serviços de arquivo do seu território de origem. Segundo Corrêa e Marques (2016, p. 18) “tem raízes em questões ligadas ao direito de propriedade. Ele surge, inicialmente, para solucionar questões de sucessão de documentos, contenciosos arquivísticos, em caso de conflitos [...]”.

Atualmente é utilizado como caracterizador da “jurisdição a que pertence cada documento, de acordo com a área territorial, a esfera de poder e o âmbito administrativo onde foi produzido e recebido” (CAMARGO; BELLOTTO apud SANTOS, 2015, p. 156)

Corrêa e Marques (2016) também enumeram as colaborações deste princípio:

- Para as práticas arquivísticas, na resolução de conflitos jurídicos;
- Para a teoria arquivística, ao complementar o princípio da proveniência, especificando-o em relação aos conflitos que envolvem a sucessão de documentos;

- Para a constituição de uma disciplina científica.

A partir destas observações pode-se perceber que a maneira como os documentos foram geridos desde o início da aplicação da arquivística se transformou inúmeras vezes, conforme os acontecimentos que induziram a humanidade a se adequar às necessidades surgentes a partir deles. Segundo a abordagem pós-moderna desenvolvida no Canadá:

o discurso teórico arquivístico está mudando de produto para processo, de estrutura para função, de arquivos para arquivamento, de registro para contexto de registro, de resíduo “natural” ou subproduto passivo da atividade administrativa para memória social [...] (COOK apud LOUSADA, 2015, p. 83).

Apesar dos mais diversos e até controversos entendimentos sobre os princípios arquivísticos, os objetos de estudo e a terminologia, pode-se inferir que se adaptaram às mudanças e que por isso mantêm a Arquivologia viva, disponível para novos debates e ativa na produção de conhecimento.

3.3 GESTÃO ARQUIVÍSTICA DE DOCUMENTOS

A gestão arquivística de documentos é um processo que busca principalmente o controle da massa documental e os benefícios resultantes; permite a aplicação da teoria das três idades no ciclo vital dos documentos:

“busca intervir nas fases que compõem o ciclo vital dos documentos determinando o tempo necessário de permanência nas fases de arquivamento (corrente e intermediária) até que seja dada a destinação final (eliminação ou guarda permanente) [...]”. (MEDEIROS e AMARAL, 2010, p. 301)

Quando uma instituição elabora seu programa de gestão documental são definidas as normas e procedimentos técnicos que configurarão a produção, tramitação, classificação, avaliação, uso e destinação dos documentos durante todo os respectivos ciclos de vida, com a definição de seus prazos de guarda e destinação final, requisitos estes necessários inclusive para o desenvolvimento de sistemas informatizados de gestão de informações (BERNARDES e DELATORRE, 2008).

Ainda com base nos conceitos elaborados por Bernardes e Delatorre (2008), cada atividade da gestão de documentos pode ser assim compreendida:

- Produção: elaboração padronizada de tipos e séries documentais, também envolve a implantação de sistemas de organização da informação e de novas tecnologias.
- Tramitação: engloba as ações relacionadas à padronização e controle do fluxo documental;
- Utilização: representada por todas as atividades de Protocolo (recebimento, classificação, registro, distribuição, tramitação e expedição), todas as atividades de Arquivo (organização e arquivamento, reprodução, acesso à documentação e recuperação de informações) e a gestão de sistemas de protocolo e arquivo manuais ou informatizados.
- Classificação: recupera o contexto de produção documental, isto é, a função e a atividade que a determinou e identifica os tipos e séries através do agrupamento dos documentos do fundo, de forma hierárquica respeitando o princípio da proveniência e o da ordem original.
- Avaliação: estabelece o prazo de guarda e destinação final dos documentos, garantindo a preservação daqueles de guarda permanente e a eliminação segura dos desprovidos de valor probatório e informativo.

Os principais produtos para execução destas atividades são o Plano de Classificação e a Tabela de Temporalidade de Documentos de Arquivo, instrumentos que organizam e controlam a massa documental produzida sempre com o objetivo intrínseco de garantir o acesso à informação e o direito à memória.

O Plano de Classificação representa o agrupamento dos documentos, elaborado de acordo com o órgão produtor, a função, subfunção e atividade responsável pela produção ou acumulação. O Dicionário Brasileiro de Terminologia Arquivística (2005) o define como um “esquema de distribuição de documentos em classes, de acordo com métodos de arquivamento específicos, elaborado a partir do estudo das estruturas e funções de uma instituição e da análise do arquivo por ela produzido”. O e-Arq Brasil (2011) também contribui: “por meio dele, são estabelecidas a hierarquia e a relação orgânica dos documentos, devidamente demonstradas na forma como eles são organizados em unidades de arquivamento”.

Também segundo o Dicionário Brasileiro de Terminologia Arquivística (2005) a Tabela de Temporalidade é o “instrumento de destinação, aprovado por autoridade competente, que determina prazos e condições de guarda tendo em vista a transferência,

recolhimento, descarte ou eliminação de documentos”. Ela então define os prazos de guarda dos documentos em razão de seus valores administrativo, fiscal, jurídico-legal, técnico e histórico, além de autorizar a eliminação ou determinar a guarda permanente.

Para a implantação de um sistema de gestão é essencial o conhecimento profundo da instituição produtora dos documentos através da análise de todos os contextos nos quais está inserida e, da mesma forma, contribui para este sistema a metodologia denominada *identificação*.

Os contextos que influenciam a produção documental são classificados em jurídico-administrativo, de proveniência, de procedimentos, tecnológico e documental. Assim, a análise resulta no reconhecimento das razões e objetivos de produção, da origem, das exigências e normas norteadoras, dos meios e ferramentas e das necessidades de registro.

De forma mais detalhada, o diagnóstico de cada contexto aponta a relação com alguns elementos, conforme proposto pelo e-arq Brasil (2011):

- Jurídico-administrativo: Leis e normas externas à instituição produtora de documentos que controlam a condução de suas atividades.
- De proveniência: organogramas, regimentos e regulamentos internos que identificam a instituição produtora de documentos, o que ela faz.
- De procedimentos: normas internas que regulam a produção, tramitação, uso e arquivamento dos documentos da instituição, como é feito.
- Tecnológico: ambiente tecnológico (hardware, software e padrões) que envolve o documento.
- Documental: código de classificação, guias, índices e outros instrumentos que situam o documento dentro do conjunto a que pertence, ou seja, no fundo de arquivo; ações e serviços que visam o arquivo; fornece subsídio para entender o contexto documental.

Na Arquivologia a *identificação*, tarefa de levantamento sobre o documento de arquivo e o seu órgão produtor, é definida como “uma metodologia de pesquisa que se desenvolve, nos parâmetros do rigor científico, como tarefa preliminar e necessária às funções arquivísticas de classificação, avaliação, descrição e também para o planejamento da produção documental” (RODRIGUES, 2012, p. 197).

Com o objetivo de obter o “[...] conhecimento exaustivo da instituição que produziu os documentos [...]” (LA TORRE MERINO *et al.* apud RODRIGUES, 2012, p. 204), a *identificação* é considerada um trabalho de pesquisa e de crítica sobre a gênese documental

através da identificação do organismo produtor, das competências, funções e atividades, do tipo documental (procedimentos administrativos e tramitação) e através da delimitação da série documental (RODRIGUES, 2012). Sousa (2003) ainda defende a relação entre a gênese e a ordem original, esta servindo como “proteção” para a ligação orgânica entre os documentos que representam a missão, funções, atividades e tarefas definidos e executados pelo produtor.

4 METODOLOGIA

Este trabalho busca identificar as necessidades da gestão de dados de pesquisa, as características da gestão documental e em quais aspectos a Arquivologia pode agregar ao pesquisador na produção, preservação e disponibilização de seus dados.

Com base nos objetivos propostos, caracteriza-se como uma pesquisa exploratória, pois busca definir elementos que possibilitam a identificação de princípios e métodos da gestão arquivística de documentos capazes de contribuir para o aprimoramento da gestão de dados da pesquisa..

De natureza básica, a intenção é abordar fundamentos, porém trazendo novas discussões para a realidade que se apresenta para a Arquivologia e para a produção de informação e conhecimento. Com abordagem qualitativa, busca compreender os cenários que compõem e influenciam a gestão documental e de dados de pesquisa.

Habilitam este estudo como procedimento metodológico a pesquisa bibliográfica e a consulta aos *sites* de instituições que desenvolvem a temática. Para identificar e contextualizar a gestão de dados da pesquisa serão consultadas publicações de instituições nacionais e internacionais detentoras de maior conhecimento teórico e prático sobre o tema. Os autores mais consagrados e a produção de profissionais da arquivística e de áreas afins serão fonte para a definição e contextualização da gestão documental assim como para a identificação dos princípios arquivísticos.

O quadro 2 define com maior clareza a relação de cada objetivo específico com o método de pesquisa escolhido e as fontes selecionadas.

Quadro 2 – Apresentação dos métodos e fontes utilizados para o alcance dos objetivos específicos

Objetivos específicos	Método	Fontes
1) Identificar e contextualizar a gestão de dados de pesquisa	Pesquisa bibliográfica e em materiais produzidos por instituições da área	Publicações sobre a temática dados de pesquisa e gestão de dados de pesquisa, <i>sites</i> e cursos organizados pelas instituições da área
2) Definir e contextualizar a gestão documental	Pesquisa bibliográfica	Publicações sobre o histórico da Arquivologia e gestão documental

3) Identificar princípios da Arquivologia no contexto da gestão de dados de pesquisa	Pesquisa bibliográfica	Publicações sobre a Arquivologia, seus princípios e sobre a gestão de dados de pesquisa; Estudos desenvolvidos nos objetivos 1 e 2
4) Analisar o uso da gestão documental para promover o desenvolvimento da gestão de dados de pesquisa	Pesquisa bibliográfica	Publicações em arquivologia e gestão de dados (guias, cursos e vídeos); Estudos desenvolvidos nos objetivos 1, 2 e 3

5 A GESTÃO DOCUMENTAL ARQUIVÍSTICA COMO SUPORTE À GESTÃO DE DADOS DE PESQUISA

Após a conceituação e a compreensão da evolução e papel de cada elemento que define a Arquivologia e a pesquisa científica, a análise realizada por esta investigação inicia com a intenção de evidenciar as relações e contribuições existentes e ainda possíveis entre estes dois universos.

De acordo com a abordagem apresentada nas seções anteriores, a gestão arquivística se define ao considerar estes principais aspectos:

- **Organicidade:** relação do documento com os demais vinculados a mesma atividade e com o contexto funcional da sua produção.
- **Princípios arquivísticos:** proveniência, respeito à ordem original, territorialidade e ciclo de vida.
- **Contextos da produção documental:** jurídico-administrativo, de proveniência, procedimental, tecnológico e documental.

Neste momento a análise realizada por este trabalho recai sobre os conceitos e funções da gestão documental que podem complementar a gestão de dados de pesquisa. Serão explorados:

- A percepção da atuação dos princípios arquivísticos na gestão de dados de pesquisa;
- A importância de se conhecer os contextos que atuam sobre a produção de dados;
- As funções da gestão documental e da gestão de dados e as possíveis relações;
- A produção documental a partir das funções de gestão de dados de pesquisa.

5.1 OS PRINCÍPIOS ARQUIVÍSTICOS NA GESTÃO DE DADOS DA PESQUISA

Nesta seção os princípios arquivísticos passam a ser analisados no ambiente da gestão de dados de pesquisa. Apesar do enfoque diferenciado, conceitos similares são desenvolvidos e considerados neste ambiente, permitindo que se faça a relação com os conceitos estabelecidos pela Arquivologia.

A bagagem teórica da Arquivologia em relação a integridade e autenticidade de materiais digitais e a recuperação da informação contribuem para a gestão de dados de

pesquisa. Os princípios e propriedades arquivísticas, como proveniência, organicidade e confiabilidade podem servir como uma base segura para a guarda e uso dos dados.

Para a Arquivologia, o conceito de proveniência, concebido na época dos documentos físicos, além da necessidade de se conhecer a procedência para que os documentos sejam devidamente classificados e retratem a realidade do produtor, trata-se de evitar a mistura de materiais de origens diversas, respeitando a delimitação em fundos. Isso fica evidente na própria definição do princípio de proveniência constante no Dicionário Brasileiro de Terminologia Arquivística (2005): “o conjunto de documentos produzido por uma entidade coletiva, pessoa ou família não deve ser misturado aos de outras entidades produtoras. Também chamado princípio do respeito aos fundos.”

As transformações do comportamento da sociedade em relação a necessidade de informação ao longo das décadas influenciaram o princípio arquivístico de proveniência. Também devido às mudanças provocadas pela era digital, seu significado é rediscutido na Arquivologia e o conceito de proveniência social é um indicativo desta nova realidade. Retomando este conceito, também chamado de pós-custódia, pós-moderna ou funcional, ambiental ou virtual, Cook contribui para sua definição:

Proveniência, em suma, está ligada virtualmente, através de múltiplas relações, fluídas e dinâmicas, entre a criação ou atividades de criação e o documento resultante e suas diversas audiências. Proveniência liga documentos a estas relações multifuncionais e de atividade, em vez de ao seu lugar estrutural de origem. (2012, p. 147).

A proveniência do dado de pesquisa foi definida por Groth e Moreau (2013, s.n, tradução nossa) como “informação sobre entidades, atividades e pessoas envolvidas na produção de um dado ou objeto, que pode ser usada para avaliar sua qualidade, confiabilidade ou fidedignidade”.

Partindo deste conceito, de maneira geral a gestão de dados entende a referência de proveniência como um meio que torna possível e seguro o compartilhamento e o reuso dos dados. Fornece informações aos pesquisadores para o uso dos dados de maneira confiável, isto é, demonstram suas origens, quem os produziu e de que forma ocorreu. A proveniência evidencia as relações entre os dados e entre as outras entidades e demais participantes do processo de pesquisa que os geraram.

As definições de proveniência pela ótica da gestão de dados e pela ótica da Arquivologia se aproximam. A Arquivologia observa a proveniência como a necessidade de

identificar as relações existentes entre o documento e seu processo de criação. Isso é ampliado na proveniência social em que passam a ser considerados todos os elementos que influenciam a produção de informação, desde leis, intervenções naturais, sociais e culturais, relações, métodos, ferramentas e normativos, tudo o que de certa forma colabora para o alcance dos meios até o produto final. Estes elementos também são observáveis no ambiente da pesquisa científica para a identificação das informações de proveniência de seus dados.

Na Arquivologia, conforme evidenciado na seção 3.2.2, o princípio arquivístico do respeito à ordem original dita que “o arquivo deveria conservar o arranjo dado pela entidade coletiva, pessoa ou família que o produziu” (ARQUIVO NACIONAL, 2005, p. 137). Entendido por diversos autores como sendo uma consequência ou aprofundamento do princípio da proveniência, o respeito à ordem original “indica que o sistema de classificação deve estar baseado na organização original do organismo produtor, ou seja, os documentos devem ser ordenados da mesma forma que a instituição produtora classificou” (LOUSADA, 2015, p. 38). É certo que a interpretação deste princípio pode ser bastante controversa, pois há autores que defendem uma definição única para o princípio da ordem original e para o princípio da proveniência, ou entendem o respeito à ordem original como uma segunda etapa do respeito à procedência. (SOUZA, 2012, p. 59). Mas independente da perspectiva adotada, para a maioria dos autores o respeito à ordem original é considerado fundamental para a análise e organização documental.

As próprias concepções pós-modernas em relação aos documentos arquivísticos e o princípio da ordem original indicam a passagem do objeto físico para um objeto abstrato formado por dados e metadados, sem um significado físico de armazenamento. Cook acrescenta que “virtualmente combina conteúdo, contexto e estrutura para fornecer evidência” (LOUSADA, 2015, p. 79). Ainda segundo Cook, este objeto abstrato pode ser ordenado de várias formas e “as ordens refletem usos múltiplos nos processos de trabalho, em vez de arranjo físico de objetos registrados” (LOUSADA, 2015, p. 81).

Percebendo a ordem original relacionada à entidade produtora, se faz necessário definir quem é o produtor dos dados sob a ótica da Arquivologia e do documento arquivístico. Neste estudo primeiramente investiga-se dados da pesquisa como documento arquivístico, para então identificar o produtor neste contexto. Para tanto, alguns conceitos essenciais são recapitulados:

1. Como descrito na seção 3.1, documento arquivístico é aquele produzido e/ou recebido por uma pessoa física ou jurídica, no decorrer das suas atividades, independente do

suporte, e dotado de organicidade.

2. Já dados de pesquisa, conforme a seção 2, são registros coletados, observados ou criados, considerados fonte primária para o desenvolvimento da pesquisa científica e necessários para validar seus resultados.

3. E finalmente a pesquisa, definida como um processo de investigação que, a partir da análise de dados obtidos, busca soluções para um problema.

Assim, destes conceitos depreende-se que dados de pesquisa podem ser considerados documentos produzidos e/ou recebidos no decorrer das atividades da pesquisa, à quem é atribuído o papel de produtor/recebedor de dados.

Considerando uma pesquisa como órgão produtor ou consumidor de dados, o princípio da ordem original deveria estar relacionado à ordem de produção e recebimento destes dados e este arranjo deveria ser registrado junto ao fundo gerado. Pensando na aplicação deste princípio na gestão de dados de pesquisa, conservar o arranjo construído pelo pesquisador em um repositório de dados é uma tarefa impraticável neste ambiente virtual. São os metadados que de certa forma cumprem este papel, ou seja, quanto mais abundantes e completos, melhor retratam estas relações que possibilitarão a percepção de uma ordem entre os dados obtidos. Usando as informações de metadados para remissão, outros instrumentos seriam necessários para a representação e visualização do fundo e do arranjo dado à pesquisa.

Apesar da possibilidade de explorar a aplicação do princípio da territorialidade sobre os dados, compará-la com a dos documentos demanda reflexão. Este viés se apresenta devido à complexidade do meio digital e das possibilidades de compartilhamento que extrapolam as barreiras territoriais consideradas pelo princípio arquivístico.

O momento histórico pode ser um potencial influenciador na origem de um conceito. É o que ocorre com o princípio da territorialidade que, em circunstância de conflitos por território, provém da noção de defesa do direito de propriedade. Conforme Corrêa e Marques (2016), os documentos relativos às terras conquistadas funcionavam como certificados das vitórias e por isso eram exigidos pelos novos “proprietários”. Outros aspectos que determinaram a definição deste princípio foram a divisão e a transmissão de territórios, em época que reinos, países ou cidades eram dominados ou tornavam-se independentes.

De acordo com o Dicionário de Terminologia Arquivística que também o denomina de “proveniência territorial”, o princípio de territorialidade sustenta que os “arquivos deveriam ser conservados em serviços de arquivo do território no qual foram produzidos, excetuados os documentos elaborados pelas representações diplomáticas ou resultantes de operações

militares” (2005, p. 141). Somado à explicação de Souza (2012, p. 59) de que “sua essência volta-se para a manutenção dos documentos no país de origem (fundo), sem mesclá-los com os documentos de outros países”, estes conceitos demonstram a perspectiva jurisdicional deste princípio. Um dos objetivos neste caso é defender a manutenção da relação do documento com um dos aspectos da sua proveniência: o local de produção/origem. Também é uma ferramenta de defesa quando se discute a respeito do interesse, direito e responsabilidades sobre alguma documentação.

Para a gestão de dados de pesquisa, de história muito mais recente e influenciada por outros aspectos, as características do princípio arquivístico da territorialidade podem colaborar em relação à garantia de propriedade e de autoria dos dados. É possível associar esta função do princípio com o direito de propriedade intelectual, o *copyright*. Conforme Sayão e Sales, “a maioria dos produtos de pesquisa – incluindo planilhas, publicações, relatórios e programas de computador – se enquadra como obras literárias e é, portanto, protegida pelas leis de *copyright*” (2015, p. 73). O direito sobre os dados é mantido pelo detentor do *copyright* que autoriza o arquivamento, o compartilhamento e a reutilização. E para estabelecer os direitos que deseja manter ou renunciar pode atribuir licenças de uso para seus dados, como por exemplo a *Open Data Commons*⁸.

Ainda conforme os autores, a política de atribuição de *copyright* adotada pela instituição na qual atua o pesquisador indica diferentes situações sobre o direito de propriedade. Geralmente os pesquisadores que criam os dados detêm os direitos associados a eles, mas

no ambiente de uma instituição acadêmica, em tese, o empregador é o proprietário do *copyright* de uma obra realizada durante o período em que o pesquisador está trabalhando na instituição. Entretanto, muitas instituições atribuem o *copyright* dos materiais de pesquisa, dados e publicações – ou seja, resultados de pesquisa – aos pesquisadores que os criaram. (2015, p. 73)

Existem também as situações de fontes variadas, como as pesquisas colaborativas e o uso de dados derivados onde o *copyright* é atribuído em conjunto a vários pesquisadores ou instituições provenientes de diversos locais (Sayão e Sales, 2015).

Passando para a análise do “território” como o local geográfico de origem dos dados e onde se encontram, com a realidade globalizada e digital a identificação da proveniência se mostra mais complexa. No decorrer da execução de uma pesquisa científica dados podem ser

⁸ Disponível em: opendatacommons.org

produzidos por pesquisadores sediados em locais distintos, referindo-se a fenômenos e momentos em determinado tempo, tornando-se complexa a definição da jurisdição.

Assim, o princípio da territorialidade aplicado na gestão de dados de pesquisa pode ser analisado sob o viés de alguns aspectos que apontam para a transposição de barreiras territoriais, considerando o território virtual de produção de dados:

1. A realidade da rede: a informática e a internet transformaram a capacidade da pesquisa e de seus dados de se tornarem disponíveis nos mais diversos locais e para um número maior de pessoas. A realidade do compartilhamento de informações significa para o pesquisador a ampliação do campo de obtenção de elementos para a pesquisa que desenvolve; já para os usuários de dados significa usufruir de recursos muitas vezes indisponíveis no local em que se encontram.

2. Como abordado, a definição da jurisdição: quando do intercâmbio na produção de dados que em muitos casos é decorrente de projetos envolvendo grupos de vários locais e países.

3. A disponibilidade e possibilidade de escolha entre diferentes tipos de repositórios e centros de dados para recebê-los, gerenciá-los e disseminá-los. Além disso, podem estar localizados em país diferente ao do pesquisador e da origem da pesquisa e produção dos dados, por exemplo: o pesquisador brasileiro define um servidor americano para abrigar seu material.

Casos atualmente discutidos envolvendo dados pessoais colaboram para interpretar a complexa ação de definir território e suas características de direito de propriedade e jurisdição. A título de exemplo são os conflitos de interesse quando uma empresa estrangeira se instala no Brasil sob as normas brasileiras e armazena os dados obtidos no seu país de origem, considerando a sua legislação em relação a sigilo, jurisdição, propriedade e etc. Exemplificam estes casos as disputas da justiça brasileira com a Google Brasil⁹ e a Microsoft versus o governo americano, o que resultou inclusive na recente “Lei da Nuvem”¹⁰ nos Estados Unidos.

Estas reflexões vêm reforçar que o direito de propriedade é item extremamente importante para a gestão de dados de pesquisa, mas que o aspecto “local de conservação” do princípio arquivístico da territorialidade não se adapta tão facilmente à realidade dos dados.

9 “STJ impõe soberania de dados na nuvem ao Google Brasil” - Convergência Digital, 06/06/2013. Disponível em: www.convergenciadigital.com.br

10 “Governo dos EUA cria Lei da Nuvem e quer acesso a dados em qualquer lugar” - Luís Osvaldo Grossmann, 27/03/2018. Disponível em: www.convergenciadigital.com.br

Na Arquivologia, independente do uso deste princípio, cabe o entendimento de Corrêa e Marques (2016) quando interpretam a territorialidade como “ações de influência dos arquivos”. Ainda segundo os autores, estas ações refletem todos os processos sobre os conjuntos documentais (produção, acumulação, organização, acesso e uso), validados pela organicidade. Da mesma forma, apesar de influências diferentes, os dados de pesquisa também sofrem ações e cabe a reflexão diante da complexidade do tema.

Como definido em seção anterior, na Arquivologia o ciclo vital representa as fases pelas quais passam os documentos desde quando produzidos até a destinação final. Neste trabalho apontamos como fases a criação, manutenção e uso, armazenamento, eliminação e disponibilidade no arquivo permanente ou em outra instituição arquivística, conforme considerou Santos (2015.)

A respeito do uso da metáfora

é uma das que é mais frequentemente aplicada em conjuntos de objetos que possuem uma existência temporal definida pela ligação entre sua criação e extinção (nascimento e morte) e que experimentam uma mudança na natureza de suas propriedades (crescimento, amadurecimento e declínio) à medida que progridem pelas várias etapas entre essas fronteiras (DINGWALL apud FILHO E SOUSA, 2016, p.195)

Ao longo do ciclo de vida são diversas as ações que levam o documento a uma progressão, ou seja, as mudanças que passa por influência das necessidades de uso, dos locais de armazenamento e do que significam para o produtor e usuários.

Segundo o Dicionário Brasileiro de Terminologia Arquivística, para a teoria das três idades os arquivos são considerados correntes, intermediários ou permanentes, de acordo com a frequência de uso pelo produtor e a identificação de seus valores primário e secundário (ARQUIVO NACIONAL, 2005).

Conforme análises de Filho e Sousa (2016) a relação entre as fases arquivísticas e as idades se dá especialmente em razão da consideração dos seguintes aspectos:

- Temporal: momento da ação à qual o documento vincula-se, recebendo o arquivo as denominações corrente, intermediário e permanente;
- Localização do documento: arquivo setorial, de depósito ou central e arquivo público;
- Frequência de uso do documento: com base no interesse do produtor é denominado arquivo ativo, semiativo, inativo; e com base no interesse do usuário denomina-se arquivo de gestão, pré arquivo, histórico.

Ainda conforme os autores, foi o arquivista francês Yves Perótin na década de 60 o responsável pela associação dessas idades (corrente, intermediária e permanente) ao conceito do ciclo vital dos documentos.

O ciclo de vida pode ser compreendido como “uma metáfora ou analogia para os processos que integram a gestão de documentos — a criação, captura, manutenção, uso e disposição dos documentos” (BROTHMAN apud FILHO e SOUSA, 2016, p.195). Ocorre então “a aplicação prática do ciclo de vida em concordância com os procedimentos e operações técnicas que integram a gestão de documentos” (FILHO e SOUSA, 2016, p.196).

O estabelecimento de um ciclo de vida no contexto da gestão de dados é plenamente defendido e já de maneira bastante similar ao princípio da Arquivologia.

A adoção de um modelo de Ciclo Vital dos dados tem como consequência os melhores procedimentos de guarda e a formulação das políticas necessárias a uma gestão responsável de informações, além de contribuir com o pesquisador na organização do seu fluxo de trabalho. Visto como uma ferramenta que garante a usabilidade dos dados da pesquisa, destaca-se a inter-relação entre as etapas e a importância de se pensar nos seus requisitos desde o início do processo:

a manutenção de dados é um processo, não um estado final onde os dados são simplesmente entregues a um repositório na conclusão de um estudo. Em vez disso, deve começar no início de um projeto e incorporar um cronograma para depositar produtos ao longo do seu ciclo de vida e a criação e preservação de metadados precisos [...] (Jacobs e Humphrey apud ICPSR, 2012, p. 7, tradução nossa)

Visto as contribuições propostas à pesquisa e ao pesquisador, da mesma forma que ocorre com os documentos arquivísticos, a adoção de um ciclo de vida para os dados orienta as atividades e funções da gestão destes dados, relação que passa a ser analisada também sob a perspectiva das idades arquivísticas.

As atividades da gestão de dados representam as ações executadas e suas funções sobre eles em cada fase do ciclo de vida. Com base no modelo elaborado pelo *UK Data Service* demonstrado na seção 2.2, o ciclo compreende: coleta, processamento, análise, compartilhamento, armazenamento e reuso.

Na pesquisa científica, ao investigar o ambiente, identificar a proveniência, analisar as características referentes à territorialidade e definir o ciclo de vida dos dados é possível perceber a ação da gestão arquivística. Em cada fase do ciclo considerar estes aspectos conduz o pesquisador e a pesquisa a parâmetros de maior confiabilidade. A obtenção de dados através

da produção ou coleta é a atividade que demanda atenção a estes aspectos, pois serão determinantes para as demais atividades e para o resultado final da pesquisa e seu reuso.

Durante a execução de uma pesquisa são sistematizadas as fases de coleta, processamento e análise que consistem no recebimento dos dados conforme as políticas definidas, na atribuição de metadados apropriados, na avaliação dos dados e seleção para a curadoria e preservação a longo prazo. Posicionando estes estágios na conceituação de ciclo de vida da Arquivologia, é possível considerá-las nas fases corrente e intermediária, uma vez que indicam o maior uso dos documentos e da mesma forma um período de tratamento mais intenso dos dados.

A partir do conhecimento obtido em relação a origem, contextos e relações faz-se o processamento dos dados obtidos de forma mais segura. Mais precisamente o processamento equivale às ações que visam organizar os dados para a análise, como: transcrever e traduzir, validar, definir a procedência, descrever e documentar e realizar o gerenciamento. Com os dados prontos para o uso inicia a atividade de análise, onde a possibilidade de aplicar o material obtido é definida. A análise se operacionaliza por meio de “ações e métodos que ajudam a garantir a qualidade dos dados ao descrever feitos, detectar padrões, desenvolver explicações e provar hipóteses” (CORREA, 2016, p. 98, tradução nossa).

A difusão dos dados através das ações que permitem o compartilhamento seguro e eficaz também se diferencia com a consideração dos aspectos relacionados à origem. São estas noções que embasam o pesquisador na definição da documentação para os usuários, do acesso mais adequado à tipologia dos dados e da publicação. Nesta fase, a essência da técnica de avaliação arquivística pode ser considerada, servindo como um instrumento de verificação do valor dos dados para outros possíveis usos e em outras pesquisas. Segundo Ieda Pimenta Bernardes (2012)¹¹, para a Arquivologia a avaliação pode servir como um instrumento de preservação e de eliminação e é considerado o único recurso técnico legítimo para a escolha de documentos históricos e o único possível para construir o patrimônio documental (capacidade de expressar a memória de um grupo social). Sendo assim, será o ambiente, as necessidades e seus significados para o produtor e usuários que conduzirão a avaliação sobre os dados para determinar a guarda e disponibilização.

O armazenamento leva em consideração as características das fases precedentes, um diferencial para o seu propósito principal de preservação dos dados que demanda do

¹¹ Disponível em: www.youtube.com/watch?v=FhuSR29ME44

pesquisador identificar os dados que possui e seus objetivos de compartilhamento. Indica a transferência dos dados para um arquivo, repositório ou base de dados com a definição das ações de preservação para garantir que os dados permaneçam autênticos, confiáveis, utilizáveis e íntegros. Estas ações apontam semelhanças com a preparação para a fase permanente do ciclo de vida do documento, necessárias desde o princípio.

Nas fases de acesso, uso e reuso os dados devem estar disponíveis para os usuários de forma livre ou com os devidos controles de acesso e autenticação. Estas fases também se assemelham com a fase permanente dos documentos arquivísticos e especialmente com as ações de difusão que objetivam a promoção do acesso e uso.

A característica de continuidade no ciclo de vida é percebida no reuso que consiste na transformação ou no aproveitamento dos dados para a criação de novos dados e relações com outros materiais. Isso significa que os dados farão parte do estado corrente de uma nova pesquisa e é a credibilidade conferida pelo conhecimento da origem e de todo o ciclo de vida que garante a ampliação e renovação da aplicação dos dados e a diversidade de suas finalidades.

Esta verificação também demonstra que as funções são dependentes umas das outras assim como se complementam. Agrega na compreensão da interligação entre as funções, assim como na ideia de ciclo, a afirmação de Sant'Ana:

O eventual acesso futuro, e, portanto, a **disseminação** destes dados, **deve ser considerada já desde a fase de coleta** de tal forma que a viabilidade de uma maior encontrabilidade e acesso seja possível, **exigindo que informações (atributos)** que mesmo que não estejam ligados diretamente a necessidade atual **sejam incluídas no planejamento da estrutura de obtenção para que seja possível**, por exemplo, **identificar elementos contextuais dos dados que possam favorecer sua localização e interpretação na fase de recuperação**. Isso pode gerar, inclusive, a ampliação do volume de dados obtidos e de **aporte de informações sobre a própria coleta** como parte das estruturas a serem preenchidas com os dados coletados. (2016, p. 126, grifo nosso)

A definição de um modelo de ciclo de vida é capaz de fornecer uma visão geral dos estágios exigidos para uma curadoria bem-sucedida e a preservação dos dados. A partir da definição de funções e responsabilidades no modelo, aprimoramentos também são possíveis através da implementação de novos padrões e tecnologias.

5.2 IDENTIFICAÇÃO E CONTEXTO

Ao longo da análise sobre a gestão documental percebe-se como o conhecimento sobre os contextos é fator importante para a aplicação dos princípios arquivísticos, assim como para compreender a produção documental e fornecer “informações subsidiárias para a criação de um sistema classificatório” de um fundo ou conjunto de documentos (PENHA e RODRIGUES, 2013, p. 11). Estes contextos abrangem uma diversidade de elementos internos e externos que influenciam o produtor assim como suas atividades geradoras de documentos e os documentos em si. Dessa forma, são classificados em cinco tipos, como descritos na seção 3.3: jurídico-administrativo, de proveniência, de procedimentos, tecnológico e documental (E-ARQ BRASIL, 2011). Analisar a maneira como os diversos elementos afetam a produção leva ao conhecimento mais profundo a respeito do produtor, de forma a conduzir a implantação de um sistema de gestão mais adequado e garantindo a produção de documentos arquivísticos.

Tendo como uma das bases para o alcance de seus objetivos a obtenção de informações sobre contexto, nasceu a metodologia denominada *identificação*. A origem do termo provém da realidade dos países ibero-americanos de acumulação irregular de documentos na década de 80 e foi definida pelo Dicionário Brasileiro de Terminologia Arquivística (2005) como um “processo de reconhecimento, sistematização e registro de informações sobre arquivos, com vistas ao seu controle físico e/ou intelectual”.

Para Penha e Rodrigues, a identificação

é realizada com base no princípio da proveniência, com o objetivo de gênese documental, a partir do reconhecimento das funções e atividades da entidade produtora que ficam registrados nos documentos de arquivo. Pode-se dizer que a identificação torna-se base para delimitação de qualquer fundo de arquivo, sendo fundamental para integrar os conceitos de função e atividade administrativa à produção documental e, além desses usos, também fundamentam a delimitação de séries de documentos dentro do contexto institucional. (2013, p. 14)

Segundo Rodrigues, “no Brasil, os processos de *identificação* incidiram sobre órgãos produtores, documentos e arquivos” e neste último “para fins de definição de estratégias para implantação de sistemas de arquivos” (2012, p. 202). Entendendo os sistemas de arquivos como a integração de instituições ou serviços com as finalidades de custódia, processamento técnico, conservação e acesso (consulta) a documentos (ARQUIVO NACIONAL, 2005), percebe-se como a *identificação* pode colaborar desde os fundamentos de toda uma ampla

estrutura de gestão.

A fim de estabelecer a aproximação desta técnica e a gestão de dados de pesquisa, pontua-se este conceito de *identificação* no campo da arquivística definido por Rodrigues:

identificação vem a ser o ato de determinar a identidade do documento de arquivo, de caracterizar os elementos próprios e exclusivos que conferem essa identidade. Significa determinar estes elementos que o individualizam e o distinguem em seu conjunto. O processo de produção deste conhecimento implica em reunir informações sobre o documento em seu contexto de produção e descrever estes elementos que formam sua identidade, que revelam o seu vínculo arquivístico (RODRIGUES, 2012, p. 200)

É portanto no ato de reunir informações no contexto de produção que vislumbra-se a colaboração da metodologia de *identificação* para a gestão dos dados. O olhar arquivístico sobre o ambiente de dados da pesquisa paira primeiramente na necessidade de conhecer o cenário no qual o pesquisador e a pesquisa que desenvolve, na figura de produtor dos dados, estão inseridos. Através do reconhecimento das funções e atividades da entidade produtora de dados, cujo papel atribuímos à pesquisa, estas passam a explicitar sua relação e seus significados em relação ao produtor. Assim, destes conceitos depreende-se que dados de pesquisa são documentos produzidos e/ou recebidos no decorrer das atividades da pesquisa, à quem é atribuído o papel de produtor/recebedor de dados. Para esta averiguação pode buscar subsídios através das informações de contexto que influenciam o desenvolvimento da pesquisa, como legislações, regulamentações, organogramas, ferramentas e suportes.

A análise realizada por este trabalho busca identificar que configurações assumem os contextos arquivísticos no caso da pesquisa e seus dados. Para tal, tem como fonte principal as informações contextuais presentes nas estruturas propostas para a elaboração de planos de gestão de dados. Estas estruturas foram percebidas através da análise dos guias de gestão de dados consultados, em especial os de iniciativa do CNEN (Brasil), do ICPSR (EUA) e do *UK Data Archive* (Reino Unido).

O quadro 3 foi elaborado com o intuito de tangibilizar a relação entre os contextos arquivísticos e os elementos de um plano de gestão de dados. A primeira coluna elenca os cinco contextos considerados pela Arquivologia e sob quais formas podem se fazer presentes. Na segunda coluna, uma seleção de elementos baseada na análise do conteúdo mais frequentemente citado pelas instituições e guias relacionados à gestão de dados como o mínimo necessário para a composição de um plano. Além dos guias mencionados, contribui para esta seleção a análise das informações disponibilizadas através de vídeos e cursos

produzidos pelas instituições e iniciativas que investem fortemente na gestão de dados, como a Universidade de Amsterdã (Holanda), Edina e *Data Tree* (Reino Unido), *ANDS* (Austrália), entre outros.

É válido reforçar que, em linhas gerais, o conteúdo do plano é o que deve ser executado pelo pesquisador para a gestão e preservação dos dados ao longo da pesquisa e após seu término, ou seja, orientar a exploração dos contextos oferece ao pesquisador uma base mais sólida para a elaboração do plano e conseqüentemente para a gestão dos seus dados.

Quadro 3 – Os contextos identificados através dos elementos recomendados para o plano de gestão de dados

Contextos	Elementos para o plano de gestão de dados	Exemplos
Jurídico-administrativo: Leis e normas externas à instituição produtora de documentos que controlam a condução de suas atividades	Definição das questões éticas e legais, restrições ao compartilhamento de dados, direitos autorais e direitos de propriedade intelectual	Brasil. Lei 9610/98, Regula os direitos autorais e dá outras providências (1)
Proveniência: Organogramas, regimentos e regulamentos internos que identificam a instituição produtora de documentos, o que ela faz	Funções e responsabilidades de gerenciamento de dados	Estatuto da Universidade Federal do RS (UFRGS): determina a pesquisa como um de seus fins (2)
Procedimentos: Normas internas que regulam a produção, tramitação, uso e arquivamento dos documentos da instituição, como é feito	Normas e medidas de garantia de qualidade, custos ou recursos necessários	Circular PRP-STI 001/2016 de 28 de junho de 2016 da Universidade de São Paulo: Utilização do Sistema Nuvens Interconectadas (interNuvem USP) para pesquisa (3)
Tecnológico: Ambiente tecnológico (hardware, software e padrões) que envolve o documento; o suporte	Escolha do repositório para o armazenamento de dados e medidas de backup	Construídas a partir de software livre, a Universidade de São Paulo (USP) disponibiliza as plataformas: <i>CKAN</i> , <i>Dataverse</i> e <i>Dspace</i> (4)

<p>Documental: Código de classificação, guias, índices e outros instrumentos que situam o documento dentro do conjunto a que pertence (no fundo de arquivo); ações e serviços que visam o arquivo; fornece subsídio para entender o contexto</p>	<p>Definição de quais dados serão gerados durante a pesquisa, planos para compartilhar dados; metadados</p>	<p>Plano de gestão de dados. Exemplo: Programa de pesquisa de genoma vegetal (<i>PGRP</i>) (5) <i>National Science Foundation (NSF), Biology Directorate</i></p>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------

FONTE: elaborado pela autora

(1) Disponível em: www.planalto.gov.br/ccivil_03/Leis/L9610.htm

(2) Disponível em: www.ufrgs.br/ufrgs/a-ufrgs/estatuto-e-regimento

(3) Disponível em: pedido.intervuem.usp.br/resources/docs/circular-PRP-STI-001-2016-28jun2016.pdf

(4) Disponível em: www.sti.usp.br/gestao-de-dados-cientificos/

(5) Disponível em: stanford.box.com/shared/static/2jo7mtlgutjyf5z8yz39.pdf

A análise da tabela leva a percepção de que a atividade de identificação dos contextos aplicada pela Arquivologia pode colaborar principalmente no momento do planejamento da pesquisa científica, mas também para uma proposta de classificação dos dados.

Nota-se a presença dos contextos na estruturação do plano de gestão de dados. Assim, o plano se torna uma representação dos contextos a serem considerados pela pesquisa ou aqueles nos quais a gestão de seus dados está inserida. Neste sentido a Arquivologia colabora a medida que estuda e desenvolve interpretações e técnicas que auxiliam a compreensão e a identificação dos elementos que representam os contextos da produção documental. Por isso uma prévia análise dos contextos, através da perspectiva arquivística, contribui para a construção do plano de gestão de dados.

A abordagem sobre os contextos é capaz de auxiliar o pesquisador na preparação de suas atividades, pois fornecem um panorama completo do que influencia ou afeta os dados e os resultados da pesquisa. Assim, antecipa-se em relação às exigências normativa e ferramental, aos itens de maior complexidade a serem providenciados assim como aos benefícios que esta análise aporta ao andamento da pesquisa. Dentre estes benefícios destacam-se: a organização adquirida através da definição de métodos e instrumentos promovendo agilidade na busca e produção de resultados, segurança em relação as questões éticas e direitos e aos embasamentos legais característicos de cada tipo de pesquisa, segurança

também em relação a manipulação e armazenamento dos dados, economia e a abertura dos dados e da pesquisa, entre outros.

Atuando na gestão de dados de pesquisa, a Arquivologia pode ainda contribuir na adaptação de suas definições de contextos. Esta reestruturação de conceitos colaboraria para que o pesquisador compreendesse e o profissional arquivista aplicasse com mais precisão esta técnica arquivística.

É possível depreender que através desta percepção sobre o ambiente a gestão se torna mais clara e organizada, há o domínio do que é essencialmente necessário e o controle do que é produzido fica evidenciado. Consta-se também que a Arquivologia já contribui para com este entendimento e possui meios para adaptar suas técnicas à análise dos contextos que envolvem os dados de pesquisa.

5.3 AS FUNÇÕES DA GESTÃO DOCUMENTAL E AS POSSÍVEIS RELAÇÕES COM A GESTÃO DE DADOS DE PESQUISA

A gestão de documentos, retomando a definição do Dicionário Brasileiro de Terminologia Arquivística (ARQUIVO NACIONAL, 2005), consiste em um “conjunto de procedimentos e operações técnicas referentes à produção, tramitação, uso, avaliação e arquivamento de documentos em fase corrente e intermediária, visando sua eliminação ou recolhimento”, ou seja, envolve uma série de atividades coordenadas abrangendo três fases: produção, utilização e destinação dos documentos (UNB, 2015).

Em relação aos benefícios da gestão documental, o Manual de gestão de documentos da Universidade de Brasília (UNB) elenca aqueles obtidos pela própria instituição, mas que são universais e se estendem a outros produtores. Define que, ao controlar e acompanhar a documentação arquivística produzida ou recebida, garante a uniformização de procedimentos, a eficiência na recuperação da informação e na criação e manutenção dos documentos. O resultado é a maior transparência, desburocratização da instituição, legalidade dos atos administrativos e a destinação correta dos documentos (UNB, 2015).

Como descritas na seção 3.3, as atividades de gestão documental são definidas em:

- Produção
- Tramitação
- Classificação

- Utilização
- Avaliação

A fase de produção de documentos se dá em razão da existência do organismo produtor através da execução de suas atividades. O principal aspecto considerado nesta fase da gestão é a necessidade de racionalização, priorizando o essencial. Assim, a criação deve ser normatizada, delimitando por exemplo os tipos de suporte, as formas de registro da informação, os responsáveis pela produção, entre outros. O controle da produção garante a economia de recursos investidos e visa especialmente à “preservação da integridade física e intelectual do documento e, conseqüentemente, ao acesso à informação” (CRUZ, 2013, p. 21). Para isso, percebe-se como conhecer a gênese documental é fundamental para que todos estes itens de controle possam ser definidos conforme as necessidades e possibilidades do produtor.

A fase de utilização significa todo o “fluxo percorrido pelos documentos para o cumprimento de sua função administrativa, assim como de sua guarda, após cessar o trâmite” (ARQUIVO NACIONAL, 2011, p. 22). Para conhecer este fluxo são necessários métodos de controle e “técnicas específicas de classificação, organização e elaboração de instrumentos de recuperação da informação” e a implantação de sistemas de arquivo (IDEM). Esta fase engloba portanto as atividades de tramitação e classificação.

Como instrumentos da fase de tramitação constam o protocolo, definido como um conjunto de operações (recebimento, registro, distribuição e movimentação) que asseguram a localização e recuperação mais eficiente dos documentos, e o *workflow*, associado à definição do fluxo de documentos eletrônicos (CRUZ, 2013).

Considerada uma atividade fundamental do processo de gestão de documentos de arquivo, a classificação exige o conhecimento dos contextos de produção e das necessidades de uso da documentação. Indica a organização dos documentos realizada pela “análise e identificação do conteúdo dos documentos, com a seleção da categoria de assuntos sob a qual serão classificados, arquivados e acessados” (ARQUIVO NACIONAL, 2011, p. 25).

O produto desta atividade é o plano de classificação, assim sintetizado por Cruz:

indica a organização dos documentos, desde sua origem, com base em esquema sistematizado das funções e atividades desenvolvidas pela administração que os gerou, e padroniza a forma de arquivamento, propiciando a recuperação dos documentos de forma eficaz. (2013, p. 21)

Os resultados provenientes da ação de classificar ainda refletem nas etapas seguintes

do ciclo documental: a avaliação, seleção, eliminação, transferência, recolhimento e acesso, pois a organização já garante o domínio dos tipos documentais e seus conteúdos.

Terceira e última fase, a destinação é representada pelos procedimentos de análise, seleção e mais precisamente de avaliação que consiste em

uma atividade essencial para a racionalização do ciclo de vida dos documentos, na medida em que o processo de análise e seleção de documentos visa estabelecer prazos de guarda nas fases corrente e intermediária, definindo o momento em que poderão ser eliminados ou destinados ao arquivo permanente, de acordo com o valor e o potencial de uso que apresentam para a administração que os gerou e para a sociedade. (ARQUIVO NACIONAL, 2011, p. 25)

Determinar o valor dos documentos habilita a fixação dos prazos de guarda. A valoração decorre da verificação da frequência de uso das informações contidas nos documentos, da existência de leis ou decretos que regulem a prescrição legal, da existência de outras fontes com as mesmas informações e da necessidade de guarda dos documentos por precaução (CRUZ, 2013).

Os resultados da avaliação são representados na tabela de temporalidade, considerada um instrumento de destinação vinculado ao plano de classificação que

facilita não só os procedimentos de transferência, recolhimento e eliminação, mas, também, a identificação de documentos de valor probatório e informativo, destinados à guarda permanente, distinguindo-os daqueles destituídos de valor secundário e, portanto, destinados à eliminação. (CRUZ, 2013, p. 25)

A construção de critérios de valor é realizada com a participação de profissionais de diversas áreas dentro da instituição produtora, diante da necessidade de se identificar a variedade de utilidades das informações contidas nos documentos. Assim, para a tarefa de avaliar, devem-se definir as Comissões Permanentes de Avaliação de Documentos de Arquivo (CPADs), podendo ser integradas pelo arquivista ou responsável pela guarda da documentação, servidores das unidades às quais se referem os documentos, profissional da área jurídica responsável pela definição do valor legal dos documentos e dos prazos de prescrição, profissionais ligados ao campo do conhecimento de que trata o acervo objeto da avaliação (historiador, economista, sociólogo, engenheiro, médico, etc), entre outros (CRUZ, 2013).

Passando para a análise dos dados de pesquisa, percebe-se que atividades de gestão de documentos podem contribuir para o aperfeiçoamento da gestão de dados. Sua gestão indica a

condução de atividades desde o planejamento de todo o processo de existência dos dados, da produção ao compartilhamento e reuso.

O planejamento requer a capacidade de prever os requisitos para a administração dos dados sob os mais diversos aspectos e para isso o pesquisador precisa definir previamente as necessidades em cada fase do ciclo de vida. Nesta primeira etapa é importante para o pesquisador que obtenha e documente o maior número de informações, desde a origem da pesquisa, incluindo as instituições envolvidas, seus objetivos, qual a necessidade de dados e de recursos (financeiro, tecnológico, humano, físico, etc) para a obtenção, manutenção e compartilhamento destes, quais os normativos a serem obedecidos, requisitos de segurança exigidos, local de armazenamento, entre outros.

A disponibilização destas informações é de certa forma materializada pelo plano de gestão de dados ao formalizar as ações e compromissos que serão estabelecidos em relação a eles. Sayão e Sales sintetizam o conteúdo do plano de gestão de dados:

descreve, de uma forma geral, que dados serão processados, coletados ou gerados; quais as metodologias e padrões que serão utilizados nesses processos; se, como e sob que condições esses dados serão compartilhados e/ou tornados abertos para a comunidade de pesquisa; e como eles serão curados e preservados. (2015, p. 15)

Este documento, além de proporcionar segurança para os futuros usuários, certamente colabora com o próprio pesquisador na execução daquilo que foi planejado durante o andamento da pesquisa. De certa forma as informações obtidas no planejamento acabam por demonstrar a própria gênese dos dados, fundamentando a produção mais consciente. Significa uma análise dos contextos da pesquisa que orienta a produção e obtenção de dados mais organizada e voltada para as etapas seguintes de arquivamento e reuso.

A ênfase dada pela Arquivologia para o controle de produção e tramitação e para a classificação dos documentos pode ser considerada pela pesquisa científica. Como acontece no controle da produção documental, a preservação da integridade física e intelectual dos dados pode ser obtida através dos mecanismos para a gerência destas atividades. Um exemplo é a possibilidade do plano de gestão direcionar os procedimentos de controle da criação e tramitação a partir do estabelecimento prévio das pessoas envolvidas e dos seus acessos e autorizações. Desta forma, além das ações sobre os dados serem conhecidas e gerenciadas, dominam-se as inúmeras versões geradas ao longo do ciclo de vida.

Ainda no contexto de uso o pesquisador pode estabelecer uma organização do material processado, definindo uma modalidade de classificação com o objetivo de manter a relação

orgânica entre os dados. A partir das noções estabelecidas no planejamento sobre os contextos de produção e os objetivos de usos dos dados, esta organização pode ser prevista com vistas ao processo de análise mais ágil e enriquecido, assim como o compartilhamento e reuso, ao controle do ciclo de vida dos dados e ainda à contribuição para a atribuição de graus de sigilo (NEGREIROS, 2008).

Diante do volume de dados obtidos, é possível pensar sobre a conveniência da abertura da totalidade dos materiais produzidos e a guarda em repositório por tempo indeterminado, visto que não é tudo que se torna útil ao compartilhamento ou ao reuso. Segundo Sayão e Sales nem todos os dados precisam ser preservados, deve-se selecionar os dados que passarão por processo de gestão de longo prazo. Ainda afirmam que

em geral todos os dados brutos devem ser mantidos; todo produto que se configure como dado de pesquisa que tenha exigido tempo e muitos recursos para ser obtido deve ser preservado. Qualquer dado que não pode ser facilmente substituído deve ser preservado. (2015, p. 24)

Visando a qualidade do retorno na busca por dados e ponderando a capacidade de armazenamento dos repositórios, considera-se que os dados possam passar por um processo de avaliação que, além de contribuir para a qualidade do compartilhamento, poderia estabelecer prazos de guarda e até a eliminação dos dados antes de seguirem para o armazenamento. Os critérios e/ou procedimentos de avaliação poderiam ser estabelecidos no plano de gestão, pois uma vez conhecidos os contextos de produção, a posterior seleção dos dados para o armazenamento seria facilitada.

A Universidade de Amsterdã, através das experiências de pesquisadores vinculados¹² a ela, apresenta as principais características dos dados que determinam a preservação de longo prazo: valores científicos e históricos, unicidade, potencial de reuso, impossibilidade de reprodução e custos. São portanto valores a considerar desde o planejamento da gestão dos dados para a escolha mais adequada do repositório e demais questões que envolvem o armazenamento, como o tempo de guarda, sigilo, etc.

Todos os aspectos verificados para o conhecimento e controle dos dados produzidos refletirão na fase do reuso. A organização estabelecida a partir do conhecimento sobre proveniência, somado a valoração dos dados proporcionam facilidades aos novos pesquisadores e acima de tudo ampliam a segurança e credibilidade da pesquisa.

¹² Disponível em: libguides.vu.nl/researchdata/data-archiving

5.4 O PAPEL DA DOCUMENTAÇÃO PARA A GESTÃO DE DADOS DE PESQUISA

A importância do ato de registrar as ações sobre a gestão dos dados ao longo da pesquisa pode ser percebida sob alguns pontos de vista. O “registro” equivale ao ato de documentar, ou seja, de gerar a documentação respectiva a cada ação.

Sayão e Sales fazem referência a um primeiro aspecto, a efetiva transmissão de conhecimento:

Os dados para transmitirem conhecimento precisam de tratamentos específicos, de catalogação e de documentação que inclui metadados descritivos e disciplinares e documentos que garantam sua interpretação, como cadernos de laboratório, roteiro de entrevistas, manuais, dicionário de dados etc. [...] (2016, p. 68)

Um outro aspecto observado durante a realização de uma pesquisa científica é a documentação servindo como fonte de memória para o pesquisador, auxiliando a lembrá-lo a respeito das motivações pela escolha de procedimentos, quais os propósitos definidos, assim como para revisá-los. Além disso, as diversas análises e reanálises sem o devido registro podem levar o pesquisador a perder o foco da pesquisa, pois os resultados muitas vezes provém das revisões que, sem o controle da proveniência e organização dos dados, prejudica a correção de erros e a reutilização do material (MACINNES, 2018).

Sendo assim, após considerar o papel da proveniência e as funções do ciclo de vida dos dados, além das possibilidades de integrar os outros princípios e técnicas arquivísticas à gestão dos dados, percebe-se que a documentação se fortalece como um ponto essencial para a elaboração, desenvolvimento e uso adequado do material gerado pela pesquisa, pois fornece as informações fundamentais para garantir a sua qualidade e autenticidade.

A documentação pode ser analisado sob as formas de:

- Metadados: se referem aos dados; descrevem formalmente os vários atributos de cada elemento ou conjunto de elementos de dados;
- Documentos que registram as operações pelas quais passam os dados no ambiente da pesquisa: fazem referência ao uso específico dos dados (CORREA, 2016).

Para a Arquivologia um metadado é considerado uma anotação, compondo a forma intelectual do documento arquivístico e “seu objetivo é identificar o documento individualmente e estabelecer a sua relação com os demais documentos integrantes do dossiê”

(RONDINELLI, 2005, p. 61). Estas anotações também são consideradas na avaliação da confiabilidade e autenticidade dos documentos arquivísticos.

Da mesma forma esses elementos complementam os dados de pesquisa. Sayão e Sales contribuem para a definição dos metadados como parte da documentação dos dados:

Metadado é um subconjunto padronizado e estruturado da documentação dos dados, formado por elementos de informação bem definidos – por exemplo, “título”, “autor”, “resumo”, “fonte” - que ajudam a conferir contexto e informar a proveniência dos seus dados, ou seja, a procedência e o histórico desses dados para pessoas e sistemas. Nessa direção, os metadados informam sobre a origem, propósito, tempo de referência, localização geográfica, autor, condições de acesso e termos de uso das coleções de dados e muito mais. (2015, p. 27)

Conforme citado, a relação entre o dado e contexto no qual foi coletado e criado é representada pelos metadados, responsáveis por manter as informações sobre a origem e processos ao longo da existência do dado, além de “frequentemente serem a única forma de comunicação entre o analista secundário e o produtor de dados” (ICPSR, 2012, p. 11, tradução nossa). Controlar as ações sofridas e fases pelas quais passam os dados ao longo do ciclo de vida também se torna um diferencial de segurança para o trabalho do pesquisador e posteriormente para a reutilização por parte de outros usuários.

A descrição do dado pelos metadados é importante para a sua localização, compreensão e interpretação. Como visto, o significado dos dados muitas vezes não existe sem uma documentação que mantenha as características e relações registradas, podendo resultar em perda do material e de todos os recursos investidos. Uma outra percepção sobre o benefício de investimento em metadados de qualidade é quando podem representar um diferencial para a difusão e reutilização por outros usuários. Isso porque tanto o conteúdo quanto as ações sobre ele ficam evidentes para os novos usuários, resultando em maior reaproveitamento e de maneira mais qualificada e ágil. Considerando também a possibilidade de classificar os dados na fase de processamento, os metadados descritivos auxiliariam nesta atividade ao trazer informação a respeito da composição do dado, posicionando-o na estrutura da pesquisa.

Reforçam Eynden *et al.*: “sempre que dados são usados, informação contextual mínima é necessária para fazer sentido a estes dados” (2011, pg. 9, tradução nossa). Os autores também sintetizam as informações que uma adequada documentação de dados deve incluir, ratificando a importância de se conhecer os contextos e de produzir os metadados:

- O contexto da coleta de dados: história do projeto, intenções, objetivos e hipóteses;

- Métodos e processos de coleta de dados, instrumentos utilizados assim como hardware e software, situação temporal e geográfica e fontes de dados secundárias utilizadas;
- Estrutura de conjunto de dados, estudos de caso, relações entre arquivos;
- Validação dos dados, verificação, revisão, limpeza e procedimentos de garantia de qualidade realizados;
- Alterações feitas nos dados ao longo do tempo e identificação de diferentes versões de arquivos de dados;
- Informações sobre condições de acesso e uso ou confidencialidade dos dados.

Segundo Sayão e Sales, a outra forma possível consiste em “documentos em formatos variados – texto, vídeos, planilhas, etc. – que apoiam o entendimento sobre como os dados foram coletados, gerados, processados e de como estão estruturados, organizados, nomeados” (2015, p. 29). São portanto documentos que auxiliam o uso dos dados, podendo ser considerados componentes do processo de pesquisa. As tipologias inclusive podem ser previstas já durante o planejamento, com vistas à construção de modelos mais eficazes e à facilidade do processo de registro.

Cada função da gestão de dados de pesquisa gera documentos que servem como ferramenta e prova das atividades e tarefas executadas. A documentação gerada a partir das operações desenvolvidas em cada fase do ciclo de vida do dado é um sinalizador de que há gestão, afinal o registro destas ações organiza e norteia a produção de dados para o produtor e garante a segurança do processo e das informações a serem compartilhadas. Além disso, sob o ponto de vista do esforço financeiro afirma-se: “em certas áreas da ciência, a falta de planejamento e execução da documentação e arquivamento apropriados de conjuntos de dados é um dos principais impedimentos para a obtenção do valor máximo do investimento em dados de pesquisa” (OECD, 2007, p. 19, tradução nossa).

O documento que pode parecer mais representativo de todo o processo é o plano de gestão de dados que guia o pesquisador na organização do seu trabalho diante da diversidade de aspectos a serem considerados, além de servir como demonstrativo para os investidores da pesquisa do que se pretende e do que será necessário em relação a custos e estruturas.

As fases subsequentes da pesquisa geram ou são representadas por documentação diversa, conforme a modalidade e a área de estudo e as necessidades do seu processo. Das referências consultadas para este trabalho, revelaram-se conteúdos cujo registro é recomendado ou tipos de documentos que já são consequência natural das atividades da

gestão dos dados. Eis alguns exemplos:

- Instrumentos de coleta e análise, como questionários, cadernos de laboratório, informações de proveniência sobre fontes de dados derivados ou digitalizados;
- Instrumentos resultantes da fase de análise, dicionários de dados;
- Informações sobre configurações de equipamentos e instrumentos para o compartilhamento;
- Relatórios técnicos sobre coleta e processamento dos dados;
- Relatórios de metodologia para a preservação a longo prazo e esquema de banco de dados;
- Documentação sobre a forma como o crédito institucional e pessoal deve ser reconhecido;
- Documentação sobre a organização adotada para os arquivos relacionados a um conjunto de dados;
- Documentação para estabelecer a responsabilidade pelo acesso e gerenciamento de dados, como descrição das tarefas formais das instituições, pedidos de subvenção, contratos para a pesquisa, acordos de publicação e licenças.

Estes exemplos levam à percepção de que o uso da tecnologia não necessariamente implica na diminuição da produção de outros documentos, os complementares ou de suporte seguem sendo necessários. Isto ocorre porque as facilidades de produção resultam em grande volume de dados que se distribuem por diversos “arquivos” (computadores pessoais e outros instrumentos) e acabam por exigir mais registros para o domínio da produção e das demais funções sobre os dados. Além disso, documentar significa demonstrar a origem e todo o ciclo de vida para a abertura dos dados e para prolongar seus usos no futuro.

Por fim, boas práticas em documentação de dados colaboram para um melhor padrão de reprodutibilidade da pesquisa, considerado um princípio fundamental do método científico (EDINA, 2018), e para a recuperação da informação, além da promoção da confiabilidade que “é condição *sine qua non* para que um dado seja útil” (SANT’ANA, 2016, p. 125).

6 CONCLUSÃO

A temática sobre dados abrange diversos fenômenos atualmente discutidos, como a *e-Science*, *Big Data*, dados abertos, curadoria digital e a cultura da colaboração. Optou-se aqui por uma análise em relação à possível atuação da Arquivologia para que todos estes fenômenos possam ser futuramente discutidos também sob a sua ótica e demonstrar o potencial dos arquivistas como possíveis profissionais de dados.

A Arquivologia pode contribuir com as pesquisas científicas ao aportar sua expertise de gestão documental à gestão de dados. As ações de gestão documental assim como suas finalidades podem sofrer variações entre os países conforme as necessidades diante da realidade da produção documental e as técnicas e conceitos empregados. Ainda assim, os princípios que norteiam a gestão documental arquivística, somados às técnicas e demais ferramentas estão aptos de tradução no ambiente da pesquisa, com as devidas adaptações em relação a realidade dos dados.

A Arquivologia também se mostra capaz de atuar diante da cultura da colaboração que requer clareza e credibilidade em relação às informações e suas fontes. Isto se deve porque suas ações e princípios se voltam para a documentação, ou seja, seu objeto de estudo é, além da disponibilização da informação resultado das atividades do produtor, a informação gerada para se conhecer e definir os processos que envolvem a produção e orientam a gestão.

A gestão de dados já estabelece algumas necessidades similares às da gestão documental, como o conhecimento da proveniência dos dados por exemplo. A Arquivologia, tendo como princípio base a proveniência, vem colaborar com o conhecimento que já possui e que segue aprimorando e reciclando a cada realidade vivenciada. A influência do princípio da proveniência pôde ser verificada na caracterização dos outros princípios arquivísticos abordados: o da territorialidade, de respeito à ordem original e a teoria das três idades.

Da mesma forma o princípio da proveniência se mostra de grande valia para a gestão de dados de pesquisa. Isso porque pode garantir aos usuários, considerando os envolvidos na pesquisa e a comunidade em geral, a idoneidade da origem, presando pela manutenção da relação dos dados com o produtor, local e seus métodos de obtenção.

A retomada de conceitos e a percepção das similaridades entre a gestão documental e a gestão de dados lançou um olhar sobre a possibilidade de considerar dados de pesquisa como documentos arquivísticos. Além disso, tanto a classificação como a avaliação e seus instrumentos se mostram capazes de apropriação pela gestão de dados, sendo necessário ainda

um aprofundamento sobre como poderiam se adequar à pesquisa.

O trabalho consistiu principalmente em uma análise dos conceitos que compõem a Arquivologia de forma a visualizá-los no ambiente dos dados de pesquisa. Mesmo não sendo possível a “transferência” do uso, foi possível refletir sobre os mais diversos aspectos que se assemelham ou se diferenciam e que são impeditivos ou passíveis de adoção e adaptação.

REFERÊNCIAS

- ANDS. **Data provenance.** Disponível em: <<https://www.ands.org.au/working-with-data/publishing-and-reusing-data/data-provenance>> Acesso em: 08 Out. 2018.
- ARELLANO, M. A. M.. **Critérios para a preservação digital da informação científica.** 356 f. Tese (Doutorado em Ciência da Informação) - Universidade de Brasília, Brasília, 2008. Disponível em: <<http://repositorio.unb.br/handle/10482/1518>> Acesso em 05 Jul. 2018.
- ARQUIVO NACIONAL (Brasil). **Dicionário Brasileiro de Terminologia Arquivística.** Rio de Janeiro: Arquivo Nacional, 2005. Disponível em: <http://conarq.arquivonacional.gov.br/images/publicacoes_textos/dicionario_de_terminologia_arquivistica.pdf> Acesso em: 02 Jul. 2018.
- ARQUIVO NACIONAL (Brasil). **Gestão de documentos: curso de capacitação para os integrantes do Sistema de Gestão de Documentos de Arquivo – SIGA, da administração pública federal.** Rio de Janeiro: Arquivo Nacional, 2011. Disponível em: <http://www.siga.arquivonacional.gov.br/images/cursos_capitacao/Apostila_gestao_documentos_2015.pdf> Acesso em: 26 Out. 2018
- BARROS, Thiago H. B. **A construção discursiva em arquivística: uma análise do percurso histórico e conceitual da disciplina por meio dos conceitos de classificação e descrição.** 2010. 132 f. Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2010. Disponível em: <<http://hdl.handle.net/11449/93667>>. Acesso em: 08 Out. 2018.
- BERNARDES, Ieda Pimenta; DELATORRE, Hilda. **Gestão Documental Aplicada.** São Paulo: Arquivo Público do Estado de São Paulo, 2008. Disponível em: <http://www.arquivoestado.sp.gov.br/site/assets/publicacao/anexo/gestao_documental_aplicada.pdf> Acesso em 08 Out. 2018.
- BERTIN, Patrícia Rocha Bello; VISOLI, Marcos Cezar; DRUCKER, Debora Pignatari. **A gestão de dados de pesquisa no contexto da e-science: benefícios, desafios e oportunidades para organizações de p&d.** Ponto de Acesso, Salvador, v.11, n.2, p. 34-48, ago. 2017. Disponível em: <<http://dx.doi.org/10.9771/rpa.v11i2.21449>> Acesso em: 05 Jul. 2018.
- BORGMAN, Christine L. **The conundrum of sharing research data.** Journal of the American Society for Information Science and Technology, vol. 63, n.6, 1059 a 1078 (2012). Disponível em: <<https://onlinelibrary.wiley.com/doi/epdf/10.1002/asi.22634>> Acesso em: 05 Jul. 2018.
- BUSCHLE, Cristala A. **Metodologia Científica e da Pesquisa.** Joinville: UDESC, 2014. Disponível em: <<http://www.joinville.udesc.br/portal/professores/cristala/materiais/Unidade3aPesquisaCientifica.pdf>> Acesso em: 08 Out. 2018.

CARLSON, Jake. The use of life cycle models in developing and supporting data services. In: RAY, Joyce M. **Research data management: practical strategies for information professionals**. Indiana: Purdue University Press, 2014. Disponível em: <<https://books.google.com.br/books?id=qZStAQAAQBAJ>> Acesso em 05 Jul. 2018.

CODATA-ICSTI Task Group on Data Citation Standards and Practices (2013). **Out of Cite, Out of Mind: The Current State of Practice, Policy, and Technology for the Citation of Data**. Data Science Journal. 12. Disponível em: <<http://doi.org/10.2481/dsj.OSOM13-043>> Acesso em: 08 Out. 2018.

COOK, T. (2012). Entrevista - Terry Cook (PORTUGUÊS). *InCID: Revista De Ciência Da Informação E Documentação*, 3(2), 142-156. Disponível em: <<https://doi.org/10.11606/issn.2178-2075.v3i2p142-156>> Acesso em: 08 Out. 2018.

CORREA, F. C. **Gestión de datos de investigación**. Barcelona: Editorial UOC, 2016.

CORRÊA, Fernando Gabriel, MARQUES, Angelica Alves da Cunha. **Princípio da territorialidade: trajetórias e definições**. In: ÁGORA, ISSN 0103-3557, Florianópolis, v. 26, n. 52, p. 05-29, jan./jun., 2016). Disponível em: <<https://agora.emnuvens.com.br/ra/article/view/582>> Acesso em 05 Jul. 2018.

CRUZ, Emília Barroso. **Manual de gestão de documentos**. Belo Horizonte: Secretaria de Estado de Cultura de Minas Gerais, Arquivo Público Mineiro, 2013. Disponível em: <http://www.siaapm.cultura.mg.gov.br/acervo/acervo_gestao/Manual_Gestao.pdf> Acesso em: 26 Out. 2018

DATA TREE. Disponível em: <<https://datatree.org.uk/?redirect=0>> Acesso em: 08 Out. 2018.

DAVIDSON, Joy; JONES, Sarah.; MOLLOY, Laura. **Big data: the potential role of research data management and research data registries**. In: IFLA World Library and Information Congress 80th IFLA General Conference and Assembly, Lyon, França, 16-22 ago 2014. Disponível em: <<http://eprints.gla.ac.uk/102160/1/102160.pdf>> Acesso em 05 Jul. 2018.

DINGWALL, Glenn. Modelo do ciclo vital e modelo continuum. In: EASTWOOD, Terry; MACNEIL, Heather. **Correntes atuais do pensamento arquivístico**. Belo Horizonte: Editora UFMG, 2017

E-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos / CONARQ - Câmara Técnica de Documentos Eletrônicos. 1.1. versão. - Rio de Janeiro: Arquivo Nacional, 2011. Disponível em: <<http://www.siga.arquivonacional.gov.br/images/publicacoes/e-arq.pdf>>. Acesso em: 08 Out. 2018

EUROPEAN COMMISSION DIRECTORATE-GENERAL FOR RESEARCH & INNOVATION. **Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020**. 3.2.ed. 2017. Disponível em: <http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf> Acesso em: 02 Jul. 2018.

EYNDEN, Veerle Van den, CORTI, Louise; WOOLLARD, Matthew; BISHOP, Libby; HORTON, Laurence. **Managing and Sharing Data: Best Practice for Researchers**. 3.ed. Colchester: UK Data Archive, 2011. Disponível em: <<http://www.data-archive.ac.uk/media/2894/managingsharing.pdf>> Acesso em 05 Jul. 2018.

FILHO, C. M. A. C.; SOUSA, R. T. B. **O ciclo vital dos documentos no âmbito da arquivologia: surgimento, disseminação e interpretações**. *Pesq. Bras. em Ci. da Inf. e Bib.*, João Pessoa, v. 11, n. 1, p. 187-202, 2016. Disponível em: <<http://www.periodicos.ufpb.br/ojs/index.php/pcbic/article/view/28664/15790>> Acesso em: 28 Out. 2018.

FREIRE, Juliana; SILVA, Claudio T.; CALLAHAN, Steven P.; SANTOS, Emanuele; SCHEIDEGGER, Carlos E.; VO, Huy T. **Managing Rapidly-Evolving Scientific Workflows**. University of Utah: IPAW 2006. LNCS 4145, pp. 10–18, 2006. Disponível em: <https://link.springer.com/content/pdf/10.1007/11890850_2.pdf> Acesso em 08 Out. 2018.

GERHARDT, Tatiana E. (org); SILVEIRA, Denise T. (org). **Métodos de pesquisa**. Porto Alegre: UFRGS, 2009. Disponível em: <<http://www.ufrgs.br/cursopgdr/downloadsSerie/derad005.pdf>> Acesso em: 08 Out. 2018.

GROTH, P.; MOREAU, L. **PROV-Overview - An Overview of the PROV Family of Documents**. W3C Working Group, Note 30 (2013). Disponível em: <<https://www.w3.org/TR/prov-overview/>> Acesso em 08 Out. 2018.

INDOLFO, Ana Celeste. **O uso de normas arquivísticas no estado brasileiro: uma análise do poder executivo federal**. 237 f. Dissertação (Mestrado em Ciência da Informação) – Convênio Universidade Federal Fluminense e Instituto Brasileiro de Informação em Ciência e Tecnologia, Niterói; Rio de Janeiro, 2008. Disponível em: <<http://ridi.ibict.br/handle/123456789/929>> Acesso em: 06 Jul. 2018.

INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH (ICPSR). **Guide to Social Science Data Preparation and Archiving: Best Practice Throughout the Data Life Cycle**. 5.ed. Ann Arbor: ICPSR, 2012. Disponível em: <<https://www.icpsr.umich.edu/files/ICPSR/access/dataprep.pdf>> Acesso em: 02 Jul. 2018.

LIBER. **Implementing FAIR Data Principles: The Role of Libraries**. Disponível em: <<https://libereurope.eu/wp-content/uploads/2017/12/LIBER-FAIR-Data.pdf>> Acesso em: 08 Out. 2018.

LOUSADA, Mariana. **A mediação da informação na Teoria Arquivística**. 135 f. Tese (doutorado) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências, 2015. Disponível em: <<http://hdl.handle.net/11449/124379>> Acesso em: 05 Jul. 2018.

MACINNES, John. **MANTRA: documentation, metadata, citation**. Edinburgh: EDINA, 2018. Disponível em: <https://mantra.edina.ac.uk/documentation_metadata_citation/> Acesso em: 08 Out. 2018.

MEDEIROS, Nilcéia L.; AMARAL, Cléia Márcia G. **A Representação do ciclo vital dos documentos: uma discussão sob a ótica da gestão de documentos.** Em *Questão*, Porto Alegre, v. 16, n. 2, p. 297 - 310, jul./dez. 2010. Disponível em: <<http://seer.ufrgs.br/EmQuestao/article/viewFile/15108/10436>> Acesso em: 08 Out. 2018.

MICHENER, William K. **Ten Simple Rules for Creating a Good Data Management Plan.** *PLoS Comput Biol* 11(10): e1004525, 2015. Disponível em: <<https://doi.org/10.1371/journal.pcbi.1004525>> Acesso em: 02 Jul. 2018.

MUNDET, José Ramón Cruz. **Archivística: gestión de documentos y administración de archivos.** Madrid: Alianza, 2012.

NEGREIROS, Leandro R. **Uma compreensão dos fundos arquivísticos sob o enfoque da análise de domínio.** *Informação & Informação*, [S.l.], v. 13, n. 2, p. 79-95, dez. 2008. ISSN 1981-8920. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/1818/1678>>. Acesso em: 10 Nov. 2018.

OBSERVATÓRIO DE CIÊNCIA DA INFORMAÇÃO DA UNIVERSIDADE DO PORTO (Portugal). **DeltCI - Dicionário Eletrónico de Terminologia em Ciência da Informação.** Disponível em: <<https://paginas.fe.up.pt/~lci/index.php/1668>> Acesso em: 02 Jul. 2018.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **OECD Principles and Guidelines for Access to Research Data from Public Funding.** OECD, 2007. Disponível em: <<http://www.oecd.org/science/sci-tech/38500813.pdf>> Acesso em: 05 Jul. 2018.

PENHA, N. A.; RODRIGUES, A. C. **A noção de identificação arquivística na Espanha e no Brasil nas décadas de 1980-1990.** *PRISMA.COM* n.º 21 (2013). Disponível em: <<http://ojs.letras.up.pt/index.php/prismacom/article/view/1936/3186>> Acesso em: 09 Nov. 2018

RODRIGUES, A. C. Identificação: uma Metodologia de Pesquisa para a Arquivística. In: VALENTIM, M. L. P. **Estudos avançados em Arquivologia.** Marília: Oficina Universitária, 2012. p. 197 - 215. Disponível em <http://www.marilia.unesp.br/Home/Publicacoes/estudos_avancados_arquivologia.pdf> Acesso em 08 Out. 2018.

RONDINELLI, R. C. **Gerenciamento arquivístico de documentos eletrônicos: uma abordagem teórica da diplomática arquivística contemporânea.** 4ª ed. Rio de Janeiro: Editora FGV, 2005.

SANT'ANA, R. C. G. **Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação.** *Informação e informação*, Londrina, v. 21, n. 2, p. 116-142, maio/ago. 2016. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/27940/20124>> Acesso em 08 Out. 2018.

SANTOS, Vanderlei Batista dos. **A Arquivística como disciplina científica: princípios, objetivos e objetos**. Salvador: 9Bravos, 2015.

SAYÃO, Luis Fernando; SALES, Luana Farias. **Curadoria digital e dados de pesquisa. AtoZ: novas práticas em informação e conhecimento**, 5(2), 67 – 71, 2016. Disponível em: <<http://dx.doi.org/10.5380/atoz.v5i2.49708>> Acesso em 05 Jul. 2018.

SAYÃO, Luis Fernando; SALES, Luana Farias. **Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa**. Inf. & Soc.:Est., João Pessoa, v.22, n.3, p. 179-191, set./dez. 2012. Disponível em: <<http://www.periodicos.ufpb.br/ojs/index.php/ies/article/view/12224>> Acesso em 05 Jul. 2018.

SAYÃO, Luis Fernando; SALES, Luana Farias. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**. Rio de Janeiro: CNEN/IEN, 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 09 Abr. 2018.

SAYÃO, Luis Fernando. **Repositórios digitais confiáveis para a preservação de periódicos eletrônicos científicos**. Ponto de Acesso, Salvador, V.4, n.3, p. 68-94, dez 2010. Disponível em: <<http://dx.doi.org/10.9771/1981-6766rpa.v4i3.4709>> Acesso em 05 Jul. 2018.

SCHELLENBERG, T. R. **Arquivos modernos: princípios e técnicas**. 6ª ed. Rio de Janeiro: Fundação Getúlio Vargas, 2006.

SCHELLENBERG, T. R. **Documentos públicos e privados: arranjo e descrição**. 2ª ed. Rio de Janeiro: Fundação Getúlio Vargas, 1980.

SCHIERMEIER, Quirin. **Data management made simple**. Nature 555, 403-405 (2018). Disponível em: <<https://www.nature.com/magazine-assets/d41586-018-03071-1/d41586-018-03071-1.pdf>> Acesso em: 08 Out. 2018.

SOUSA, R. T. B. **Os princípios arquivísticos e o conceito de classificação**. In: RODRIGUES, Georgete Medleg; LOPES, Ilza Leite (Org.). Organização e representação do conhecimento na perspectiva da Ciência da Informação. Brasília: Thesaurus, 2003, v. 2, p. 240-269. Disponível em: <<http://repositorio.unb.br/handle/10482/1439>> Acesso em 08 Out. 2018.

SOUZA, C. L. Princípios arquivísticos na literatura internacional e nacional: mapeamento do princípio da proveniência. **Revista Ibero-Americana de Ciência da Informação**, v. 5, n. 2, 2012. Disponível em: <<http://www.brapci.inf.br/v/a/13073>>. Acesso em: 08 Out. 2018.

UK DATA SERVICE. **Research data lifecycle**. Disponível em: <www.data-archive.ac.uk/create-manage/life-cycle> Acesso em: 08 Out. 2018.

UNB. **Manual de Gestão de Documentos de arquivo da Universidade de Brasília**. Brasília: Cebraspe, 2015. Disponível em: <http://www.arquivocentral.unb.br/images/documentos/Manuais/Manual_gestao_documentos.pdf> Acesso em: 09 Nov. 2018.

WILKINSON, M. D. *et al.* **The FAIR Guiding Principles for scientific data management and stewardship.** *Sci. Data* 3:160018 doi: 10.1038/sdata.2016.18 (2016). Disponível em: <<https://www.nature.com/articles/sdata201618.pdf>> Acesso em 05 Jul. 2018.