

## Perspectivas arquivísticas na Gestão de Dados de Pesquisa: uma análise a partir da Arquivística Integrada

**José Gustavo Corrêa**

IBICT, Programa de Pós-graduação em Ciência da Informação, Rio de Janeiro, RJ, Brasil

Procuradoria Geral do Estado do Rio de Janeiro, Rio de Janeiro, RJ, Brasil

[jose.gustavo@outlook.com](mailto:jose.gustavo@outlook.com)

**Jamille Abreu Passalini de Sousa**

IBICT, Programa de Pós-graduação em Ciência da Informação, Rio de Janeiro, RJ, Brasil

[j.passalini@gmail.com](mailto:j.passalini@gmail.com)

DOI: <https://doi.org/10.26512/rici.v15.n2.2022.40762>

Recebido/Recibido/Received: 2021-11-12

Aceitado/Aceptado/Accepted: 2022-07-08

### Resumo

Apresenta uma proposta de relação entre a Arquivística Integrada e a Gestão de Dados a partir do desenvolvimento das funções arquivísticas preconizadas pelos canadenses Jean-Yves Rousseau e Carol Couture. Aborda a Gestão de Dados de Pesquisa como uma atividade multidisciplinar e desenvolve o papel da Arquivologia na Gestão de Dados de Pesquisa. Demonstra as potencialidades da integração das duas disciplinas.

**Palavras-chave:** Arquivologia. Arquivística Integrada. Gestão de Dados de Pesquisa. Função arquivística.

### Archival perspectives in research data management: an analysis from Integrated Archival

#### Abstract

It presents a proposal for a relationship between Integrated Archival Science and Research Data Management based on the development of archival functions recommended by Canadian authors Jean-Yves Rousseau and Carol Couture. It approaches Research Data Management as a multidisciplinary activity and develops the role of Archival Science in Research Data Management. It demonstrates the potential of integrating the two disciplines.

**Keywords:** Archival Science. Integrated archival Science. Research data management. Archival function.

### Perspectivas de archivo en la gestión de datos de investigación: un análisis del archivo integrado

**Resumen:** Presenta una propuesta para una relación entre la Archivística Integrada y la Gestión de Datos basada en el desarrollo de funciones archivísticas preconizadas por los canadienses Jean-Yves Rousseau y Carol Couture. Aborda la gestión de datos de investigación como una actividad multidisciplinaria y desarrolla el papel de la archivología en la gestión de datos de investigación. Demuestra el potencial de integrar las dos disciplinas.

**Palabras clave:** Archivología; Archivística Integrada; Gestión de datos de investigación; Función archivística.

O pressuposto basilar da Arquivística Integrada é promover o ciclo documental completo, ou seja, promover um tratamento arquivístico que englobe o documento desde sua produção até sua destinação final (preservação permanente ou eliminação). No âmbito das ciências documentárias, entendemos a Arquivologia como uma disciplina multi e interdisciplinar. Nesse sentido, a associação entre a Arquivística Integrada e a Gestão de Dados de Pesquisa surge naturalmente como uma atividade a ser realizada.

De imediato, é possível fazer uma associação direta com a função arquivística de preservação da informação — consideramos isso porque entendemos que, na Gestão de Dados, as atividades arquivísticas concentram-se na etapa de arquivamento e preservação. Contudo, se partirmos do pressuposto da arquivística integrada, podemos associar outras funções.

Desta forma, foi realizada uma pesquisa do tipo exploratória buscando responder à seguinte questão: De que forma os pressupostos da Arquivística Integrada podem ser apropriados pela Gestão de Dados? Nosso objetivo foi investigar as relações entre Arquivística Integrada e a Gestão de Dados de Pesquisa.

A metodologia adotada foi o levantamento bibliográfico, que lastreou a realização de uma comparação entre as funções arquivísticas e os processos desenvolvidos na gestão de dados de pesquisa. Foram apresentadas e analisadas a arquivística integrada, as funções arquivísticas, bem como a gestão de dados de pesquisa e seu caráter multidisciplinar. Este processo levou à construção de um quadro comparativo cuja análise foi realizada contrapondo-se cada uma das áreas e destrinchando suas relações de aproximação.

Tendo em vista o caráter interdisciplinar da Ciência da Informação, ao elucidar as convergências entre a Arquivística Integrada e a Gestão de Dados, pretendemos também salientar o papel do profissional arquivista nesta área e, conseqüentemente, melhorar o processo de gerenciamento de dados de pesquisa na prática diária de profissionais da informação.

## **2. A Arquivística Integrada**

O fazer científico é pautado por transformações e é dessa forma que ocorrem as mudanças de paradigmas que movimentam a ciência. A institucionalização da Arquivologia como disciplina ocorreu a partir da Revolução Francesa, pois foi o momento em que os acervos documentais começaram a ser entendidos como instrumentos de poder e voltou-se um olhar mais atento para sua organização (SANTOS, 2010). Contudo, se nesse momento o foco era no documento, outras abordagens surgiram transformando a disciplina arquivística. Uma dessas abordagens é a Arquivística Integrada, cunhada pelos canadenses Jean-Yves Rousseau e Carol Couture durante a década de 1980.

Na Arquivística Integrada, a informação orgânica assume o lugar do documento de arquivo como objeto da disciplina arquivística. Trata-se de uma informação contextualizada, ou seja, aquela que surge em função das atividades do seu produtor; nessa abordagem, o contexto de produção é fundamental para se entender a concepção do acervo arquivístico.

Ao transformar o objeto da Arquivologia em informação orgânica, os canadenses inserem a disciplina no contexto da Sociedade da Informação (TOGNOLI, 2010). Nesse sentido, as atividades desenvolvidas pela disciplina Arquivística são pautadas por uma nova dinâmica de produção de documentos em que o excesso de informação exige das profissões que lidam com o fenômeno procedimentos integrados e dinâmicos. Dessa forma, a abordagem canadense busca integrar dois fazeres arquivísticos comuns na Europa e na América do Norte: os *record managers* e os *archivists* (TOGNOLI, 2010).

A diferença entre as duas atividades surgiu no âmbito da administração pública norte-americana, onde os *record managers* eram responsáveis pela documentação corrente e os *archivists* pela documentação histórica e permanente. Porém, no final dos anos de 1980 o Arquivo Nacional do Québec enfrentava certa dificuldade em lidar com essa dinâmica. Os autores Natália Tognoli e José Augusto Guimarães (2011) expõem que é a partir dessa problemática específica do Arquivo Nacional do Québec em lidar com documentos correntes, intermediários e permanentes que uma abordagem integradora foi proposta.

Assim sendo, podemos definir a Arquivística Integrada como uma disciplina que insere o fazer arquivístico no âmbito da gestão da informação, uma vez que considera a informação orgânica como objeto da disciplina. A Arquivística Integrada prevê uma política global de gerenciamento de arquivos cujos objetivos são:

- garantir a unidade e a continuidade das intervenções do arquivista nos documentos de um organismo e permitir assim uma perspectiva do princípio das três idades e das noções de valor primário e secundário;
- permitir a articulação e a estruturação das actividades arquivísticas numa política de organização de arquivos;
- integrar o valor primário e o valor secundário numa definição alargada de arquivo (ROUSSEAU; COUTURE, 1998, p. 70).

Aqui, é importante mencionar as características dos documentos de arquivo que pautam o fazer arquivístico. A divisão entre as fases corrente, intermediária e permanente define o ciclo vital dos documentos de arquivo, conforme a abordagem do arquivista americano Theodore R. Schellenberg (2006). A fase corrente reúne documentos frequentemente consultados e que estão próximos ao produtor; a fase intermediária reúne documentos conservados para fins administrativos e que estão aguardando sua destinação final (eliminação ou guarda

permanente); já a fase permanente reúne documentos que serão preservados permanentemente.

As fases do ciclo de vida do documento de arquivo e suas especificidades podem ser percebidas pela imagem da Figura 1 a seguir proposta por Janice Gonçalves (1998):

**Figura 1 - Ciclo de vida dos documentos**

1a. idade ARQUIVO CORRENTE	. documentos vigentes, freqüentemente consultados
2a. idade ARQUIVO INTERMEDIÁRIO e/ou CENTRAL	. final de vigência; documentos que aguardam prazos longos de prescrição ou precaução; . raramente consultados; . aguardam a destinação final: eliminação ou guarda permanente
3a. idade ARQUIVO PERMANENTE	. documentos que perderam a vigência administrativa, porém são providos de valor secundário ou histórico-cultural

Fonte: Gonçalves (1998)

Ainda sobre as fases do ciclo de vida do documento de arquivo e conforme o quadro de Janice Gonçalves (1998), podemos comentar que elas se relacionam diretamente com a Teoria do Valor, preconizada por Theodore R. Schellenberg em seu clássico livro *Modern Archives: principles and techniques*, publicado em 1956 nos EUA e em 1973 no Brasil. A Teoria do Valor diz respeito, portanto, ao valor que o documento possui para o produtor.

### **3. Teoria do Valor**

Os valores preconizados pela Teoria do Valor são dois: valor primário e valor secundário. Documentos cujo valor é primário são aqueles relacionados à própria produção do documento e seu uso possui fins administrativos, fiscais ou legais. Já os documentos cujo valor é secundário são mantidos em função do seu potencial uso para fins de pesquisa ou referência. As fases corrente e intermediária reúnem documentos de valor primário ao passo que a fase permanente reúne documentos de valor secundário.

A Teoria do Valor é considerada a base da avaliação arquivística, pois a avaliação é um:

Trabalho interdisciplinar que consiste em identificar valores para os documentos (imediato e mediato) e analisar seu ciclo de vida, com vistas a estabelecer prazos para sua guarda ou eliminação, contribuindo para a racionalização dos arquivos e eficiência administrativa, bem como para a preservação do patrimônio documental (GONÇALVES, 1998, p. 14)

É necessário explorar esses conceitos ao falar de Arquivística Integrada, pois trata-se de uma abordagem disciplinar que tem como objetivo integrar os fazeres e transformar a disciplina. De acordo com Tognoli e Guimarães (2011, p. 28), na proposta da Arquivística Integrada, "a Arquivística e o arquivista não devem mais ser vistos como simples guardiões da memória histórica e institucional. Eles participam, agora, do momento de criação dos documentos ativos".

A partir dessa citação podemos entender que o arquivista deve atuar em todas as fases do ciclo documental, ou seja, da produção à destinação. Também nesse sentido a autora Natália Tognoli (2010) destaca algumas funções arquivísticas fundamentais para a Arquivística Integrada, são elas: produção, difusão e acesso, classificação, recuperação e a preservação da informação. É considerando as funções arquivísticas preconizadas pela Arquivística Integrada que podemos fazer a relação com a Gestão de Dados.

#### **4. A Gestão de Dados de Pesquisa: uma missão multidisciplinar**

Dados de pesquisa historicamente foram relegados durante o processo de investigação científica. Vistos apenas como um subproduto, os dados não eram considerados elementos com potencial de uso (ou reuso), fazendo com que a preocupação com o seu armazenamento ou a sua gestão fosse ínfima ou mesmo inexistente. Este cenário se altera significativamente com a mudança de *status* por que passam os dados de pesquisa nos últimos anos.

A compreensão da importância dos dados para a ciência por parte de agências de financiamento, pesquisadores e outros atores do fazer científico mudou a noção de que o único produto legítimo de pesquisa é o artigo de periódico, a dissertação ou a tese. Os potenciais de reuso dos dados produzidos durante a pesquisa científica são vistos agora como podendo "capacitar os pesquisadores a formular novos tipos de indagações, hipóteses e a usar métodos analíticos inovadores no estudo de questões críticas para a ciência e para a sociedade." (MAYERNIK *et al.*, 2012, p. 1).

Dados compartilhados e bem gerenciados trazem inúmeros benefícios aos pesquisadores: garantem a reprodutibilidade dos trabalhos publicados, criando a possibilidade de repetição de experimentos; inibem a perda e a destruição dos dados coletados nas pesquisas, mantendo a sua integridade, autenticidade e confiabilidade; bem como permitem a emergência de novos temas de pesquisa, que surgem a partir da análise dos dados por outros pesquisadores, retroalimentando o fluxo de investigação (SAYÃO; SALES, 2017).

Desta forma, é natural que a atividade de gestão desses dados seja um ponto importante para pesquisadores, profissionais ligados à gestão da informação científica e, em última análise, para a própria ciência. Sendo assim, a responsabilidade pela gestão dos dados

não deve se restringir apenas ao pesquisador, mas sim a todas as pessoas envolvidas no processo de pesquisa. Trata-se, portanto, de uma atividade multidisciplinar, que conta com a colaboração de profissionais de diferentes áreas. Sayão e Sales (2017, p. 112, grifo dos autores) classificam a Gestão de Dados como um “conjunto de atividades gerenciais e tecnológicas, apoiadas por políticas gerais e específicas destinadas a garantir: **arquivamento, curadoria, preservação e oferta de acesso contínuo aos dados de pesquisa**”.

É inevitável, no entanto, que profissionais da informação (bibliotecários, arquivistas, etc.) sejam demandados a atuar mais ativamente no processo de Gestão de Dados, haja vista, por exemplo, o papel central de bibliotecas de pesquisa que, a partir de agora, passam a ter seu papel alargado para também serem custodiantes de dados, além dos tradicionais materiais informacionais cuja guarda já possuíam.

Bibliotecários possuem *know-how* em gestão de informações, conhecimento sobre metadados, preservação digital e recuperação da informação. Podem, portanto, auxiliar o pesquisador desde o início do processo de gestão, contribuindo com a criação de um plano de Gestão de Dados de pesquisa que vai orientar a gestão por todo o ciclo de vida dos dados, indo desde o planejamento e a coleta, passando pela asseguuração, descrição, preservação, descoberta, integração até a análise dos dados.

Profissionais da área de computação e tecnologia, por outro lado, também têm papel importante na construção de uma efetiva Gestão de Dados de Pesquisa. Cientistas de dados com formação nesta área contribuem no “desenvolvimento de tecnologias de análise, manipulação, visualização, modelagem [e] algoritmos para as coleções de dados.” (SAYÃO; SALES, 2017, p. 128), atuando geralmente próximos aos pesquisadores. Já profissionais tecnologistas da informação contribuem no sentido de manter bases de dados operacionais, preocupando-se com a manutenção dos sistemas, a segurança (inclusive com a realização de backups dos dados) e o seu armazenamento.

Arquivistas, por sua vez, também são peças-chave na amalgama de profissionais necessária para a boa Gestão de Dados de Pesquisa. A atuação desses profissionais com arquivos de diferentes tipos, bem como a visão holística aplicada aos documentos de arquivo através da utilização de um ciclo de vida dos documentos arquivísticos contribui muito para a prática profissional envolvendo dados. Nesse sentido, arquivistas podem atuar no arquivamento e na preservação dos dados, garantindo sua autenticidade, integridade e confiabilidade(SAYÃO; SALES, 2017).

É justamente a contribuição da Arquivologia para a Gestão de Dados de Pesquisa que iremos explorar mais profundamente no tópico a seguir, sobretudo as relações entre a Arquivística Integrada e a Gestão de Dados. Se até aqui foi clarificada a contribuição de cada

profissional para uma boa gestão (e conseqüentemente a contribuição de cada área de formação), o mergulho nas relações e convergências entre a Arquivística Integrada e a Gestão de Dados de pesquisa pode não só esclarecer qual a real aproximação entre essas duas áreas de estudo, mas também descortinar contribuições e leituras arquivísticas até agora não exploradas para os problemas da Gestão de Dados.

## 5. As Funções Arquivísticas no Ciclo de Gestão de Dados

De forma geral, são sete as funções arquivísticas propostas pelos canadenses Jean-Yves Rousseau e Carol Couture (1998). Porém, neste artigo, focaremos naquelas destacadas por Tognoli (2010) e que consideramos pertinentes na associação entre Arquivística Integrada e Gestão de Dados. Ou seja: criação, difusão e acesso, classificação, recuperação e preservação.

Os canadenses Jean-Yves Rousseau e Carol Couture propõem um programa de gestão integrada da informação orgânica dividido em três componentes. O primeiro componente reúne as funções de criação, difusão e acesso. No momento da criação, pensa-se em formas de padronizar a informação através de formulários ou outras formas de controle. As atividades desenvolvidas nessa etapa visam controlar a produção excessiva de informação, pois:

A informação inútil, supérflua ou com dupla função é eliminada de partida [...] Além disso, uma intervenção imediata, desde a criação da informação, permite determinar seu encaminhamento, bem como seu tratamento ulterior. Devem se colocar várias perguntas: Quem tem acesso à informação? Qual o seu ciclo de vida? (ROUSSEAU; COUTURE, 1998, p. 68).

Portanto, é possível perceber que o momento de criação é essencial para o desenvolvimento das outras funções arquivísticas. A boa gestão desde o momento da criação permite que a informação seja estruturada e inteligível (ROUSSEAU; COUTURE, 1998).

Já a difusão e o acesso, outras funções desse primeiro componente do programa proposto pelos canadenses, organizam a informação no canal correto de difusão e proporcionam o acesso. É nesse momento que são consideradas as restrições de acesso e o público-alvo.

As funções pertinentes ao segundo componente proposto pelos autores canadenses são a classificação e a recuperação. A classificação é considerada pela literatura arquivística como a função matricial da gestão da informação orgânica, pois é justamente ela que fornece contexto e proporciona o vínculo arquivístico, conforme Janice Gonçalves ao afirmar que o objetivo da classificação é “dar visibilidade às funções e às atividades do organismo produtor de arquivo, deixando claras as ligações entre os documentos” (GONÇALVES, 1998, p. 12).

Também é a partir da classificação que se obtém uma recuperação eficaz. A classificação tem como pressuposto reunir em classes e grupos por semelhanças e diferenças. Essa dinâmica auxilia o processo de recuperação.

Por fim, o terceiro componente proposto pelos autores canadenses citados congrega a função de preservação. Os autores utilizam o termo "proteção e conservação" (ROUSSEAU; COUTURE, 1998) contudo, entendemos que se trata da preservação. A preservação deve fazer parte de um programa pautado por planejamento e políticas de ação, inclusive com ações de conservação. Jean-Yves Rousseau e Carol Couture (1998, p. 68) afirmam que "a informação bem protegida e conservada segundo normas técnicas e materiais precisas pode ser bem comunicada".

Por essa breve explanação das atividades desenvolvidas no âmbito das funções arquivísticas selecionadas fica evidente que todas elas se relacionam. Os três componentes, portanto, devem ocupar um só papel na gestão da informação.

A proposta desse artigo é articular a arquivística integrada e a Gestão de Dados de Pesquisa. Para isso, entendemos ser necessário fazer uma relação direta entre as funções arquivísticas aqui exploradas e a Gestão de Dados de Pesquisa.

## 6. Relação entre a arquivística integrada e a Gestão de Dados de pesquisa

As relações entre a Arquivística Integrada e a Gestão de Dados de Pesquisa podem ser sumarizadas no quadro abaixo. Cada uma das relações será mais bem detalhada em seguida.

Quadro 1 - Relações entre Arquivística Integrada e Gestão de Dados de Pesquisa

Arquivística Integrada (Funções arquivísticas)	Gestão de Dados de Pesquisa
Criação e Produção	Concepção da pesquisa Criação do Plano de Gestão de Dados
Difusão e Acesso	Criação de políticas de acesso aos dados de pesquisa Curadoria de dados Utilização de licenças de uso, embargos e anonimização de dados sensíveis
	Atribuição de metadados

Classificação e Recuperação	Contextualização dos dados
Preservação	<p>Definição de política de preservação dos dados de pesquisa</p> <p>Atribuição de metadados de preservação</p> <p>Garantia de autenticidade e integridade dos dados</p> <p>Observância dos modelos de preservação utilizados pelos repositórios custodiantes dos dados</p>

Fonte: os autores.

### 6.1. Criação e Produção

No âmbito da Gestão de Dados de Pesquisa, a etapa correspondente à função arquivística de “criação/produção” se dá na concepção da pesquisa, envolve diretamente o pesquisador (e, idealmente, profissionais da informação que o auxiliam) e tem como principal subproduto o Plano de Gestão de Dados (PGD).

É no PGD que o pesquisador vai delinear como os seus dados serão gerenciados, seja durante o momento da pesquisa, seja após o término de sua investigação, dando conta de como serão preservados e compartilhados. É nesse momento que serão padronizados os tipos de dados que serão coletados, processados ou gerados durante a pesquisa, quais padrões, códigos e metodologias serão úteis, quais dados deverão ser preservados a longo prazo, assim como em quais formatos e padrões de metadados. Sobre o PGD, Sales e Sayão (2015, p. 15, grifo nosso) dizem:

O PGD descreve o **ciclo de vida de gestão para todos os dados** que serão coletados, processados ou gerados por um projeto de pesquisa. De uma forma abreviada, ele se constitui em um documento formal que estabelece um compromisso de como esses dados serão tratados durante todo o desenvolvimento do projeto, e também após a sua conclusão.

A utilização de modelos de PGD é amplamente difundida pelas instituições, de modo que existem ferramentas para auxiliar os pesquisadores nessa tarefa. Uma das mais conhecidas é a Data Management Platform tool (DMP tool) criada pela Universidade da Califórnia. A partir dos *templates* estabelecidos pelas instituições ligadas à pesquisa sobre a qual se refere o PGD,

a plataforma garante que o pesquisador irá indicar informações consideradas essenciais pela instituição, agência de fomento etc.

Aqui se torna notória a relação entre a teoria arquivística e a Gestão de Dados. Se Couture e Rousseau (1998), pela perspectiva da Arquivística Integrada, afirmam que na fase de criação deve-se pontuar questões como “Quem tem acesso à informação?” e “Qual o seu ciclo de vida?”; Sales e Sayão (2015) respondem não só que essas mesmas perguntas também são objeto de atenção na Gestão de Dados, como também sustentam que as respostas a elas devem sempre estar presentes no PGD (vide a menção ao ciclo de vida de gestão dos dados na citação acima).

Não obstante, o PGD responde, na Gestão de Dados, a uma preocupação central da arquivística na fase de criação, qual seja a padronização e a racionalização da produção de documentos. Conforme Sousa (2013, p. 22) afirma:

No desempenho de suas funções as instituições produzem uma quantidade considerável de documentos[...] Esses documentos, por registrarem atividades que muitas vezes são diárias, necessitam ter a sua produção controlada. Pois podem formar volumes documentais imensos, tornando-se assim um problema para os setores de trabalho no que se refere ao espaço físico para armazenamento.

Ao padronizar e delimitar os dados que serão coletados, gerados ou processados pela pesquisa, o PGD evita o cotejo de dados que não serão necessários ao pesquisador, consequentemente diminuindo o volume de dados a ser gerenciado na pesquisa e aumentando a sua eficiência.

## 6.2. Difusão e Acesso

A difusão dos acervos documentais dos arquivos configura-se como uma função extremamente importante na prática arquivística. Trata-se de um processo, sobretudo nos arquivos públicos, que dá a conhecer à população o que essas instituições oferecem e possuem. Bellotto (2004) chega a dizer que a difusão é a atividade que melhor pode levar os arquivos a cumprir seus contornos sociais, fazendo com que eles se projetem na comunidade e ganhem uma dimensão cultural e popular.

A promoção do uso dos dados de pesquisa se dá de maneira bastante similar. Dar a conhecer os dados de pesquisa aos pesquisadores é uma atividade que está muito ligada à curadoria, que promove o uso do dado de pesquisa “a partir do seu ponto de criação para assegurar que ele está apto para propósitos correntes e **disponível para descoberta e reuso.**” (SALES, SAYÃO, p. 151, 2017, grifo nosso). Além disso, a curadoria pode agregar valor aos dados

depositados num repositório, fazendo com que não só eles sejam promovidos, mas também promovendo a pesquisa que lhes deu origem.

Não se pode esquecer, no entanto, que apesar da difusão e da promoção serem necessárias e muito bem-vindas, nem todos os documentos de arquivo ou dados de pesquisa podem ser acessados indiscriminadamente. A existência de documentos sensíveis e sigilosos é uma realidade enfrentada pela Arquivologia em diversas instituições governamentais.

Esta realidade também é amplamente enfrentada na Gestão de Dados de Pesquisa. Se o acesso aos dados de pesquisa tornou-se, no contexto atual, “um imperativo para a ciência com reflexos globais, considerando que os pesquisadores trabalham em cooperação internacional e os dados são criados, compartilhados e acessados em escala planetária[...]”(SALES; SAYÃO, 2014, p. 78), a Gestão de Dados precisou se debruçar sobre as maneiras que esse acesso é feito.

Dados de pesquisa podem conter informações sensíveis, como dados pessoais e dados institucionais sigilosos. Da mesma forma que nos arquivos, é preciso criar políticas de acesso a esses dados. A adoção de licenças de acesso e uso, a utilização de embargos temporários ou, ainda, de anonimização de dados pode garantir a segurança necessária e, ao mesmo tempo, o acesso adequado da comunidade acadêmica.

Não obstante, é de fundamental importância para um bom acesso aos dados a disponibilização de uma interface web e a oferta de aplicações e serviços acerca das coleções, tanto para o usuário humano, quanto para aplicações tecnológicas. (SALES; SAYÃO, 2015). De maneira geral, é notório que a preocupação com o acesso e com a difusão de acervos (de documentos ou de dados) está presente tanto na Arquivologia quanto na Gestão de Dados de Pesquisa. Ainda que existam diferenças metodológicas, os princípios gerais são os mesmos, fortalecendo a relação entre as duas disciplinas.

### **6.3. Classificação e Recuperação**

De acordo com Flores e Lampert (2013) o objetivo primeiro da classificação é deixar claras as funções e as atividades do organismo produtor dos documentos, fazendo com que as ligações entre eles fiquem evidentes. Além disso, a classificação é considerada por Couture e Rousseau (1998) como a primeira etapa que conduz à acessibilidade do acervo de documentos. Trata-se de uma função primordial, uma vez que dela depende o desempenho das outras funções arquivísticas.

A função de classificar um documento de arquivo deve ter por objetivo, portanto, deixar claro o vínculo entre o documento e seu contexto de produção para que “não haja perda de sentido e/ou de capacidade de refletir a atividade que o gerou, mesmo que o documento esteja

fisicamente separado do restante de seu conjunto” (LÓPEZ; CARVALHO, 2013, p. 273). Uma classificação malfeita pode ter, entre outros efeitos, uma recuperação ineficaz, prejudicando o usuário final dos arquivos. Ainda segundo López e Carvalho (2013, p. 273),

Um sistema de classificação efetivo e eficaz pressupõe um estudo prévio de diplomática e tipologia documental, que poderá, eventualmente, valer-se de elementos híbridos. É importante estudar a estrutura para determinar as funções exercidas por cada órgão. A análise das funções auxilia a identificação de processos e atividades, os quais se relacionam a espécies documentais específicas. A contextualização de cada espécie, por sua vez, é o que permite que os registros sirvam de prova da ocorrência das atividades que os geraram, bem como do trâmite percorrido. Tal aspecto é crucial no caso de arquivos públicos, já que lá os documentos probatórios têm, supostamente, fé pública.

Na Gestão de Dados, a ligação entre o dado de pesquisa e o seu contexto de produção/criação também é de suma importância para o pesquisador. Essa relação é clarificada através dos metadados, que irão identificar não só o contexto de produção, mas também todas as outras informações relevantes sobre os dados. Essa ligação é importante porque, tal como os documentos arquivísticos, dados de pesquisa também possuem valor probatório, uma vez que podem comprovar os resultados das pesquisas das quais derivam. Além disso, dados também podem ser classificados de acordo com sua natureza, origem e *status* no fluxo de trabalho da pesquisa. (SAYÃO; SALES, 2015).

Na Gestão de Dados, assim como na Arquivologia, uma classificação malfeita (atribuição de metadados inadequados) pode influenciar diretamente a recuperação e a compreensão dos dados, prejudicando a descoberta e dificultando (ou mesmo impedindo) o seu reuso, um dos principais objetivos da Gestão de Dados de Pesquisa. A classificação de dados se mostra, desta forma, como uma atividade da qual a eficácia das outras etapas da gestão (descoberta, recuperação e reuso) dependem, assim como acontece no contexto arquivístico.

#### **6.4. Preservação**

Comumente, no ciclo de vida dos dados, os arquivistas atuam no momento de arquivamento e preservação, porque são responsáveis por garantir a autenticidade, integridade e confiabilidade dos dados de pesquisa. Por isso, entendemos que esta seja a função cuja associação entre a Arquivística e a Gestão de Dados se faz mais naturalmente.

A função preservação envolve planejamento, pois está relacionada ao desenvolvimento de políticas e também à implementação de uma rotina de monitoramento constante. No ambiente digital, isso significa a definição de padrões de metadados e migração constante de suportes e formatos. Além disso, também é importante mencionar que os suportes digitais estão sujeitos aos mesmos riscos que os suportes físicos, tais como ameaças biológicas, ambientais ou

climáticas. A definição de uma política de preservação sistêmica é essencial nesse momento do ciclo de vida dos arquivos (e dos dados).

Tendo em vista o contexto atual — em que a produção de documentos nato digitais costuma superar a produção em suporte convencional — focaremos na preservação digital. De acordo com Miguel Ferreira:

A preservação digital consiste na capacidade de garantir que a informação digital permanece acessível e com qualidades de autenticidade suficientes para que possa ser interpretada no futuro recorrendo a uma plataforma tecnológica diferente da utilizada no momento da sua criação (FERREIRA, 2006, p. 20).

Isso significa promover métodos de preservação digital que podem ser estruturais ou operacionais. Os métodos estruturais envolvem a adoção de padrões, elaboração de normas e definição de metadados de preservação. Já os métodos operacionais relacionam-se com a migração de suporte, emulação, preservação da informação etc. (MARDERO ARELLANO, 2004).

Outra estratégia de preservação digital muito comum são os repositórios. Sobre isso, Santos e Flores (2016) afirmam que é preciso que os repositórios estejam de acordo com normas e padrões bem difundidos, tal como o modelo *Open Archival Information System* - OAIS. A utilização de um modelo como o OAIS garante ao dado preservado qualidades arquivísticas, como autenticidade e integridade.

Aqui, faz-se bem claros os aspectos arquivísticos relacionados à Gestão de Dados. A solicitação de ambientes seguros para preservação dos dados é uma constante que envolve os dados de pesquisa, como bem demonstram Sales e Sayão (2015) quando comentam sobre a gestão de longo prazo dos dados de pesquisa:

Existem várias opções que podem ser utilizadas para esta fase dos dados, elas incluem repositório institucional de sua instituição, repositórios associados aos periódicos científicos e repositórios e centros de dados que se dedicam a disciplinas específicas (SALES; SAYÃO, 2015, p. 24).

A utilização de repositórios implica no uso de metadados descritivos, estruturais e administrativos essenciais para o gerenciamento da preservação em longo prazo (SANTOS; FLORES, 2016). Em relação ao uso desses metadados, também podemos associar as duas disciplinas, principalmente utilizando o PGD como fundamento, já que o PGD reúne um conjunto de informações que caracterizam e identificam os dados. Dessa forma, entendemos que o PGD pode ser utilizado como uma ferramenta auxiliar no processo de descrição dos dados para fins de preservação.

Durante a elaboração do PGD, alguns elementos precisam ser definidos, entre eles o tipo de dado a ser capturado. Sales e Sayão (2015) afirmam que a categorização dos dados é

importante para a preservação, pois: "A distinção definida por essa categorização é de grande importância na escolha das estratégias de arquivamento e preservação" (SALES; SAYÃO, 2015, p. 11).

A categorização dos dados costuma ser feita durante o desenvolvimento do PGD, contudo é importante que esteja claro que esta categorização pode ser atualizada no decorrer da pesquisa e pode ter impacto direto no momento de preservação dos dados. Mais uma vez percebemos que as atividades desenvolvidas no ciclo de vida dos dados estão relacionadas, o que favorece uma análise integradora tal qual a prevista pela Arquivística Integrada.

## **7. Considerações Finais**

As novas iniciativas em torno dos dados de pesquisa — e de sua gestão — nos permitem potencializar as discussões e estudos acerca do tema, assim como a proposta que trouxemos nesta pesquisa. Através da perspectiva arquivística, buscamos relacionar duas disciplinas cujos enfoques são a informação e seus potenciais usos.

A análise das funções arquivísticas exploradas neste artigo demonstrou que elas são essenciais para a garantia da autenticidade e da integridade dos documentos de arquivo. Da mesma forma, a proposta inicial de relacionar tais funções arquivísticas – sob a égide da arquivística integrada – à gestão de dados de pesquisa nos permite, agora, explorar possibilidades práticas entre as duas disciplinas. Destacamos como a principal a utilização do PGD como um instrumento auxiliar no desenvolvimento da função classificação enquanto ferramenta de garantia do vínculo arquivístico, visto que é a partir deste vínculo que se garante o contexto dos documentos.

Vimos também que a classificação é a função arquivística responsável por dar visibilidade às funções institucionais e demonstrar a relação entre os documentos. É desenvolvida a partir de uma série de elementos, como o estudo das tipologias documentais, legislações, estrutura institucional etc. Entendemos que tais elementos estão presentes no PGD, conforme observado nos modelos amplamente difundidos, como aqueles reunidos pela DMP tool.

Finalmente, pudemos ainda ressaltar a natural relação entre a função arquivística de preservação e a gestão de dados. Se Rousseau e Couture (1998) entendiam que a preservação deve fazer parte de um programa planejado de políticas de ação, o arquivista pode, na gestão de dados, aplicar esses preceitos ao recomendar metadados descritivos, estruturais e administrativos para preservação em longo prazo. Ou na determinação de um repositório que se adeque ao tipo de dado que deseja preservar. Vimos ainda que o PGD pode ser utilizado como ferramenta auxiliar de descrição dos dados para fins de preservação.

Em suma, resta claro que as relações entre arquivística integrada e gestão de dados de pesquisa são múltiplas e benéficas para a preservação, acesso e difusão dos dados de pesquisa e não se limitam àquelas exploradas neste trabalho.

De forma geral, consideramos que novos estudos precisam ser desenvolvidos tendo em vista novas aplicabilidades para a Arquivologia no âmbito da Gestão de Dados, salientando o papel do profissional arquivista nesta área e, conseqüentemente, melhorando o processo de gerenciamento de dados de pesquisa na prática diária de profissionais da informação.

## Referências

BELLOTTO, Heloísa Liberalli. **Arquivos permanentes: tratamento documental**. Rio de Janeiro: FGV, 2008.

FERREIRA, Miguel. **Introdução à preservação digital: conceitos, estratégias e actuais consensos**. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006. 88 p. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/626/490>. Acesso em: 23 dez. 2019.

GONÇALVES, Janice. **Como classificar e ordenar documentos de arquivo**. São Paulo: Arquivo do Estado, 1998.

LAMPERT; Sérgio Renato; FLORES, Daniel. As funções de produção, classificação e avaliação de documentos arquivísticos no software NuxeoDocument Management. Rio de Janeiro, **Informação Arquivística**, v. 2, n. 1, jan./jun. 2013. Disponível em: <http://www.aaerj.org.br/ojs/index.php/informacaoarquivistica/article/view/16/12>. Acesso em: 03. jan. 2020.

LÓPEZ, André Porto Ancona; CARVALHO, Pedro Davi Silva. A classificação arquivística por assunto em documentos fotográficos: o exemplo do Arquivo Público do Distrito Federal. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 3, n. 2, p. 271-279, jul./dez. 2013. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/pgc/article/view/17470>. Acesso em: 03. jan. 2020.

MÁRDERO ARELLANO, Miguel Angel. Preservação de documentos digitais. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 15-27, maio/ago. 2004. Disponível em: <http://revista.ibict.br/ciinf/article/view/1043/1113>. Acesso em: 23 dez. 2019.

MAYERNIK, M. *et al.* The data conservancy instance infrastructure and organization service for research data curation. **D-Lib Magazine**, v. 18, n. 9/10, Sept./Oct.2012. Disponível em: <http://www.dlib.org/dlib/september12/mayernik/09mayernik.html>. Acesso em: 06 fev. 2020.

ROUSSEAU, J. Y.; COUTURE, C. **Fundamentos da disciplina arquivística**. Lisboa: Publicações Dom Quixote, 1998.

SANTOS, H. M.; FLORES, D. O documento digital no contexto das funções arquivísticas. **Páginas A&B, Arquivos e Bibliotecas**, Porto, n. 5, p. 165-177, 2016. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/65458>. Acesso em: 28 nov. 2019.

SANTOS, Paulo Roberto Elian dos. **Arquivística no laboratório**: história, teoria e métodos de uma disciplina. Rio de Janeiro: Teatral: FAPERJ, 2010. 215 p.

SAYÃO, Luís Fernando; SALES, Luana Farias. **Guia de gestão de dados de pesquisa para pesquisadores e bibliotecários**. Rio de Janeiro: CNEN, 2015. Disponível em: [http://www.cnen.gov.br/images/CIN/PDFs/GUIA\\_DE\\_DADOS\\_DE\\_PESQUISA.pdf](http://www.cnen.gov.br/images/CIN/PDFs/GUIA_DE_DADOS_DE_PESQUISA.pdf). Acesso em: 23 dez. 2019.

SAYÃO, Luís Fernando; SALES; Luana Farias. Dados abertos de pesquisa: ampliando o conceito de acesso livre. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, Rio de Janeiro, v. 8, n. 2, p. 76-92, 2014. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/viewFile/611/1252>. Acesso em: 03. jan. 2020.

SAYÃO, Luís Fernando; SALES; Luana Farias. **Gestão de dados de pesquisa**. Rio de Janeiro: CNEN, 2017. Disponível em: <http://www.cnen.gov.br/images/CIN/Cursos/minicurso-gestao-dados-pesquisa-sayao.pdf>. Acesso em: 03. jan. 2020.

SHELLENBERG, T. R. **Arquivos modernos: princípios e técnicas**. 6. ed. Rio de Janeiro: FGV, 2006.

SOUSA, Fábio Nascimento. **Funções arquivísticas**: contribuições para o cumprimento da Lei de Acesso à Informação. 2013. Trabalho de Conclusão de Curso (Especialização em Gestão em Arquivos). — Universidade Federal de Santa Maria, Cachoeira do Sul, RS, 2013. Disponível em: [https://repositorio.ufsm.br/bitstream/handle/1/114/Souza\\_Fábio\\_Nascimento.pdf?sequence=3&isAllowed=y](https://repositorio.ufsm.br/bitstream/handle/1/114/Souza_Fábio_Nascimento.pdf?sequence=3&isAllowed=y). Acesso em: 03. jan. 2020.

TOGNOLI, Natália Bolfarini. **A contribuição epistemológica canadense para a construção da arquivística contemporânea**. 2010. 119 f. Dissertação (mestrado). — Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2010. Disponível em: <http://hdl.handle.net/11449/93669>. Acesso em: 28 nov. 2019.

TOGNOLI, Natália Bolfarini; GUIMARÃES, José Augusto Chaves. A organização do conhecimento arquivístico: perspectivas de renovação a partir das abordagens científicas canadenses. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.16, n.1, p.21-44, jan./mar. 2011. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-99362011000100003](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362011000100003). Acesso em: 28 nov. 2019.