

## A INTEROPERABILIDADE NA PRESERVAÇÃO DA INFORMAÇÃO ARQUIVÍSTICA: OS METADADOS E A DESCRIÇÃO

Charlley Luz<sup>1</sup>

### RESUMO

Este artigo apresenta os impactos da realidade digital no acesso à informação e avalia a interoperabilidade da informação arquivística e os padrões que visam estruturá-la. Analisa a padronização da informação originada pela estrutura documental, sua relação orgânica e a Descrição Arquivística para associar com a definição e a gestão de metadados na preservação digital. Conclui que a criação de ontologias originadas por processos descritivos normatizados evidencia a informação arquivística e a padroniza, em forma de metadados, permitindo sua interoperabilidade por meio de plataformas e diferentes sistemas digitais.

**Palavras-chave:** Ontologia. Interoperabilidade. Padrões e descrição arquivística. Informação arquivística. Informação digital.

### 1 INTRODUÇÃO

A informação é representada por símbolos, portanto, são “estruturas simbolicamente significantes, codificadas de forma socialmente decodificável e registradas (para garantir permanência no tempo e portabilidade no espaço)” (SMIT; BARRETO, 2002, p 21-22). São compostos por significante (a forma do símbolo) e significado (conteúdo em si). Estes símbolos geralmente estão registrados em documentos e em dados. Estes documentos precisam de um processo de tratamento informacional e de significação em que seu conteúdo é organizado para descoberta e uso, se assim atender à necessidade do usuário.

---

<sup>1</sup> Arquivista, professor da Pós-graduação em Gestão de Documentos da Fundação Escola de Sociologia e Política de São Paulo (FESPSP) e consultor em estratégia de informações e ambientes digitais da Feed Consultoria. Autor dos livros *Arquivologia 2.0*, *A Informação Digital Humana* e *Primitivos Digitais: uma abordagem arquivística*. E-mail: charlleyluz@gmail.com.

No contexto digital e de redes, a informação passa por processos que garantem sua padronização para facilitar as operações de busca e recuperação de sistemas, e também como forma de criar links semânticos que façam sentido aos agentes computacionais. É onde se manifesta a indexação, a padronização temática e semântica, principalmente no caso de sistemas de classificação na organização de acervos informacionais.

Nessa realidade, parte da informação que transita pela rede possui características arquivísticas, pois “os equipamentos de informática são usados rotineiramente como meios de produzir, acumular e transmitir informações arquivísticas e não-arquivísticas” (LOPES, 2009, p.360). Parte dessa documentação é informação de arquivo, pois muitas dessas informações “se cristalizam como documentos informáticos, não existindo em outros suportes” (LOPES, op. cit., p.360). A informatização também trouxe para os arquivos a realidade dos sistemas de gestão documental, que se tornaram uma plataforma entre o usuário, representantes digitais de documentos e o uso de suas informações em ambientes que armazenam arquivos. Surge, assim, a necessidade de preservar em longo prazo e definitivamente as informações arquivísticas e, portanto, os documentos digitais em suas características recebidas desde sua origem.

O tema deste artigo foi definido pela oportunidade de reflexão sobre a padronização dessa informação sobre os arquivos, a informação arquivística, visando à interoperabilidade e sua preservação, no contexto de redes e em sistemas de gestão de documentos. Pretende-se, assim, trazer problematizações importantes e necessárias para a área, e mostrar caminhos para pesquisas futuras.

Por meio de uma abordagem arquivística, a fim de solucionar problemas de tratamento e representação informacional numa rede tecnológica, exercitam-se diálogos entre as áreas de Tecnologia da Informação e da Ciência da Informação. Em relação à Tecnologia da Informação, aborda-se a informação digital e a web 3.0. Em relação ao campo informacional, trata-se tanto da organização e representação (Ciência da Informação), como a abordagem aplicada e o compartilhamento da Informação Arquivística em sistemas de arquivo (dentro do campo específico da Arquivologia).

## 2 A INTEROPERABILIDADE

A existência da rede já pressupõe a interoperabilidade. De forma a atender a todos elementos de redes (os usuários, os usos e os nodos de rede), é necessária a criação de padrões de interoperabilidade. Os padrões de descrição de dados e sua estruturação alinham linguagens e facilitam as trocas de informações entre instituições. Afinal “vive-se em uma nova etapa, em que a presença da informação registrada mensurável em bytes, tem uma significação muito maior do se possa, a priori, imaginar” (LOPES, op. cit., p. 359).

Num sentido lato, segundo a Wikipédia, a interoperabilidade pode ser definida como a capacidade de um sistema (informatizado ou não) se comunicar de forma transparente (ou o mais próximo disso) com outro sistema (semelhante ou não). A NBR ISO 9126-1 (ABNT, 2003), que trata da qualidade de software, define a interoperabilidade como a capacidade do produto de software interagir com um ou mais sistemas especificados. A norma destaca que a interoperabilidade é usada no lugar de compatibilidade para evitar confusões com a subcaracterística “capacidade para substituir”.

Quanto às qualidades da interoperabilidade, Diallo (2011) realiza um levantamento bibliográfico e observa que dois conceitos são recorrentes nas definições de interoperabilidade encontrados na literatura científica. Um é o de intercâmbio de informações, em que afirma que os sistemas informatizados são caracterizados pela sua capacidade de trocar informações; o outro, a usabilidade da informação, que seria uma noção de utilidade das informações trocadas.

Isso gera necessidade de trabalho, teórico e prático, para a área da Ciência da Informação e seus profissionais. Estabelece-se a prática de tratamento da informação: a preparação da interoperabilidade, isto é, sistematicamente aplicar significado aos objetos informacionais, aos registros e recursos de rede de forma que tanto os humanos quanto os agentes computacionais entendam aquela informação ali manifestada.

A necessidade de interoperabilizar a informação é básica para os tempos de comunicação em rede. A interoperabilidade garante o uso e a encontrabilidade dos metadados estruturados dos objetos informacionais. As formas de estruturar a informação

arquivística, aquela referente aos acervos documentais disponíveis digitalmente, não é obrigatoriamente padronizada. Isso representa que cada serviço informacional, ou unidade de tratamento documental (um serviço de arquivo, por exemplo), pode definir a forma como vai estruturar este tipo de informação, gerando suas específicas políticas de interoperabilidade e de descrição, além do seu próprio padrão de metadados.

Existem iniciativas internacionais de normalização de dados, de tecnologias e normas que indicam caminhos de padronização para dados para a área dos arquivos, aplicável em documentos orgânicos e plataformas digitais. Verificou-se como o produto da descrição arquivística (que possui iniciativas internacionais de uniformização) pode ser utilizado como padrão de interoperabilidade de dados de acervos de arquivos, a chamada informação arquivística. Eles podem ser aplicados utilizando-se os perfis de metadados existentes, tal como o EAD (Encoded Archival Description).

Para a ciência arquivística, com um corpus científico único e tendo como objeto o documento arquivístico e todo processo relativo à sua custódia e uso, as tentativas de padronização de dados são recentes, remontando à década de 1980, culminando, em 1994, com a publicação da primeira norma internacional da área, a ISAD(G) – International Standard Archival Description - General. Tais iniciativas sempre buscaram organizar a informação arquivística e sua representação descritiva para acesso e uso, informações sobre os fundos, conjunto de documento, seus originadores, suas funções e custodiadores. Trata-se de informação arquivística, disponível num arquivo organizado como um sistema, seja ele informatizado ou não.

### **3 INTERFACES DA INFORMAÇÃO DIGITAL NA WEB 3.0 E INTEROPERABILIDADE**

A informação digital, para ser tratada por agentes computacionais e como forma estruturante do conhecimento, precisa ser organizada em suas principais características, como, por exemplo, seus temas de conhecimento ou identificando seus aspectos semânticos. As informações, que antes eram encontradas em livros e documentos, atualmente compõem complicadas redes dispostas em plataformas digitais, porém, o processo de organização da informação continua sendo necessário.

Para as instituições coletoras de cultura, no conceito de Smit (2012), a organização de informações passa a ter um papel decisivo na busca realizada pelos usuários, por meio de plataformas digitais e portais, interna ou externamente. Ao definir os processos comunicativos e de navegação de sistemas e linguagens, objetiva-se que, ao utilizar uma plataforma digital de sistema de informação de arquivo, isto leve o usuário a um novo conhecimento, estabelecendo um *locus* de mediação entre instituições e pessoas.

Ao acrescentar as páginas web ou as plataformas digitais neste rol de suportes, admite-se que a interface e os metadados são instrumentos para a gestão desses recursos. Os serviços de informação e seus sistemas, portanto, estão inseridos em redes, o que requer a adoção de padrões que possibilitem importar e exportar seus dados em caráter global. O Dublin Core e o MARC (Machine Readable Cataloging), por exemplo, são formatos sugeridos para esta padronização em relação a dados gerais e o EAD e EAC (Encoded Archival Context) para dados arquivísticos.

Iniciativas com dados abertos ou públicos já operam, criando redes semânticas de informações públicas, que é a proposta primordial da Web 3.0. De forma ampla, a aplicação do esforço em interoperabilizar os dados também necessitam de processos de organização específicos e de definições, como a criação de perfis de metadados e a aplicação de descrições, atividades já partilhadas no tratamento da informação em ambientes digitais, tanto por museus, bibliotecas, arquivos, instituições e pessoas.

A interoperabilidade é um processo de preparação da informação digital. Ela possui linguagens onde são estruturados seus metadados. Para compreender a interoperabilidade da informação arquivística, é necessário destacar que a preparação da informação passa pela padronização com a ISAD(G) e/ou NOBRADE (Norma Brasileira de Descrição Arquivística). Afinal, as normas “permiten el intercambio de descripciones archivísticas gracias a que la información se encuentra estructurada en campos predeterminados” (ALONSO, URBÁN E LEGUINA, 2008, p. 14).

A interoperabilidade pressupõe a pré-disposição da informação para o intercâmbio e portabilidade, isto é, necessitam de ambientes estruturados para tais operações. Assim, é possível disponibilizar a informação arquivística a ser utilizada na troca

de pacotes, fornecendo dados para *harvestings*<sup>2</sup> ou na busca, seja em motores de mercado em um próprio sistema ou em plataformas digitais. Esta é a chamada semântica computacional, pois se emprega este termo para denominar a propriedade que tem uma Descrição Arquivística em permitir aos programas de computadores entenderem a informação que contém e, assim, realizar tarefas de tratamento de dita informação ou inferidas a partir delas. (ALONSO, URBÁN E LEGUINA, 2008).

Outro aspecto que reforça a interoperabilidade é a forma como são estruturados estes dados padronizados, no modo que são lidos e interpretados. A esta propriedade se dá o nome de Interoperabilidade Semântica, pois, como estes autores destacaram, a “interoperabilidad semántica, que denota la capacidad de una información para ser interpretada, compartida e intercambiada por diferentes sistemas de tratamiento basados en el paradigma de la Web Semántica” (ALONSO, URBÁN E LEGUINA, 2008, p.15).

A interoperabilidade é um pressuposto da Web Semântica considerada uma evolução estrutural da web sintática, sendo um campo importante de pesquisa e acumulando algumas décadas de concepção, estudo e aprofundamento, pois:

A partir do final da década de 1990, começaram a formalizar-se pesquisas relacionadas ao desenvolvimento de uma nova geração da Web, com o objetivo de possibilitar a incorporação de ligações semânticas aos recursos informacionais, de modo que os próprios computadores possam “compreendê-las” de forma automatizada. Machine understandable information, com esta sucinta expressão Berners-Lee (1998) impulsionou os primeiros estudos em direção ao projeto da Web Semântica. (RAMALHO; VIDOTTI; FUJITA, 2007, p. 1).

Para os arquivistas significa, segundo Quisbert (2006), devido à sua responsabilidade de guarda e preservação, dar acesso a informações, por vezes usando diferentes gerações de software. O software é substituível e atualizado, mas a informação deve permanecer como original. No caso de arquivos, os documentos arquivísticos são preservados da mesma forma: através de diferentes gerações de software. Os registros devem permanecer como originais a fim de cumprir suas exigências como prova, por isso

---

<sup>2</sup> Trata-se de sistemas (robôs) coletores de dados interoperáveis e abertos da web, que disponibilizam estas informações em sistemas especialistas para busca e consulta.

a importância dos metadados e da interoperabilidade, que possibilita a evolução da informação através destas versões de sistemas.

Por outro lado, a informação digital alterou a relação dos indivíduos com a busca de conhecimento e acesso à informação, tornando-o um ser seletivo. A mediação é realizada diretamente pelos usuários, exigindo esquemas intuitivos de classificação e de recursos de navegação em interfaces. Além disso, a informação arquivística ainda terá uma classificação conforme a função e atividade que a gerou, ou seja, haverá também informações de seu contexto. Como afirma Ribeiro (2005, 2005, p.10),

com a internet criaram-se possibilidades de uma consulta à distância em frações de segundos sobre volumes incomensuráveis de informação, o que alterou, radicalmente, os comportamentos de pesquisa, os perfis dos utilizadores e as necessidades de informação, agora à escala planetária.

Para o modelo de ciclo de vida da informação, ela deve ser utilizada, com o objetivo de evitar que, no fim, seja apenas um conjunto de dados. Portanto, a informação é conhecimento que pode ser trocado. A intenção de um sistema de informação de arquivo, assim, não deve ser apenas o de armazenamento, mas também o de garantir a autenticidade, a perenidade da cadeia de custódia, além de estimular o uso de seu acervo por meio da difusão e disponibilização de metadados a agentes computacionais. A interoperabilidade só faz sentido se a informação for de fato utilizada, segundo Diallo et al. (2011), abordando o conceito do IEEE (Instituto de Engenheiros Eletricistas e Eletrônicos), diz que é “a capacidade de dois ou mais sistemas ou componentes para trocar informações e usar as informações que foram trocadas”. (DIALLO et al., p.84).

As interfaces de relacionamento de informações não solucionam os problemas de organização da informação digital, mas têm um papel importante nisso, ao serem a forma de mediar e dar acesso às informações organizadas. Planejar essa mediação, utilizando, por exemplo, a arquitetura de informação focada no usuário ou a estrutura de descrição de links da web 3.0, ou as abordagens semânticas e ontológicas na descrição e indexação é, portanto, uma forma de organizar a informação e disponibilizar para o usuário final, facilitando o trabalho dele na busca por conhecimento e no esclarecimento de um estado anômalo de conhecimento (CHOO, 2003).

Diallo (2011) também afirma que a interoperabilidade precisa criar interfaces, pois a determinação do que é útil é completamente dependente do sistema que recebe os dados. O sistema de envio não é capaz de determinar se a informação que é trocada é útil. Logo, precisa evidenciar quais informações está enviando e, do outro lado, uma interface que garanta a utilidade, intermediando estes dados por meio de modelos. Padrões de metadados é uma forma de criar esta ponte entre estes sistemas.

#### **4 A INTEROPERABILIDADE, A PRESERVAÇÃO E A ORGANICIDADE**

A área de atuação da Arquivística, englobando o documento e a gestão da produção, processamento e disseminação da informação, necessária e básica para a tomada de decisões na administração contemporânea, gera como produto a informação arquivística.

Nesse sentido, a organização da informação se torna uma linguagem comum entre as áreas da Arquivologia e da Ciência da Informação. Para ambas as áreas, trata-se de representação, indexação, organização e difusão, no caso específico da arquivologia, tudo que envolve a informação arquivística, ou sobre os fundos de documentos. Muitos elementos da Arquivologia são referenciados como organização da informação, resultando na informação arquivística.

Neste sentido, a interoperabilidade poderá estar aplicada à conservação ou preservação, pois incorpora os procedimentos relativos à manutenção da integridade física e (ou) lógica dos documentos ao longo do tempo, e as tecnologias e métodos que permitem o seu processamento e recuperação. Ou seja, o conjunto documental deve estar preparado para passar por submissão em diferentes sistemas. Em relação à informação digital, a descrição arquivística gera um conjunto de metadados e devemos considerar que [...] os metadados fazem parte do processo de preservação dos documentos arquivísticos digitais e que os mesmos devem ser preservados juntamente com os documentos arquivísticos digitais (INNARELLI, 2015, p. 209).

Interoperabilidade pressupõe padronização da informação que será trocada, por isso a importância da Descrição Arquivística, que aplica modelos criados para uso

internacional (a ISAD-G) e nacional (a NOBRADE). A descrição arquivística é uma função que perpassa todo o ciclo de vida do documento, num esforço integrado de descrição, devendo ter seus elementos adequados para cada uma de suas fases. Deve considerar a unidade documental a qual se refere e a sua classificação dentro do arquivo, aplicando o princípio multinível, que condicionou a criação das normas descritivas.

O acesso também é um componente da descrição arquivística, efetivado por meio de processos de estabelecimento de pontos de acesso (plano de classificação), para facilitar a recuperação dos documentos ou informação. Compreende a criação e utilização de índices e de vocabulários controlados e, em alguns casos, podem ser desenvolvidos tesouros, seguindo os princípios e normas oriundas da Documentação e da biblioteconomia.

Para cada processo da instituição existem funções vinculadas que, unidas, são os passos de um processo. Eles são responsáveis pelo desempenho das funções. Este conceito está presente em Thomassen, que afirma que “os arquivos são compostos por informação vinculada a processos, ou seja, informação gerada e estruturada por processos de trabalho” (THOMASSEN, 2006, p.06). Logo, para o autor, o que compõe o arquivo é informação gerada e estruturada por processos de trabalho funcionalmente inter-relacionados. Quanto ao documento de arquivo, o que o caracteriza e diferencia do senso comum é exatamente seu vínculo à ação realizada. Trata-se de “informação registrada, independente da forma ou do suporte, produzida ou recebida no decorrer das atividades de uma instituição ou pessoa, dotada de organicidade, que possui elementos constitutivos suficientes para servir de prova dessas atividades” (CONARQ, 2004, p. 7).

O documento arquivístico é “documento que um determinado organismo, seja ele pessoa física ou jurídica, produz no exercício de suas funções e atividades” (GONÇALVES, 1998, p.20) A organicidade, portanto, é o diferencial que dá valor arquivístico, uma vez que “documento é o resultado da atividade de uma pessoa física ou jurídica, podemos falar do caráter orgânico desse registro” (SOUSA, 2014, p.7). O aspecto orgânico é importante, pois é o que caracteriza a informação arquivística, como utilizaremos neste trabalho.

O aspecto orgânico destaca-se nas abordagens que mostram que esta é uma informação produzida ou recebida no decorrer das atividades de uma instituição ou

pessoa, ou “produção que pode significar tanto a elaboração do documento pelo próprio organismo, como a recepção e guarda” (GONÇALVES, 1998, p. 20). O documento arquivístico carrega em si, ainda, dois importantes pressupostos: a confiabilidade, que é a capacidade de um documento arquivístico “sustentar os fatos que se refere”, que é garantido por meio de sua completude, ou seja, tem que ser completo, ser criado pela autoridade competente e ter seus procedimentos de criação bem controlados; portanto, está relacionada ao momento da sua produção e à veracidade do seu conteúdo (CONARQ, 2006); e também a autenticidade, que é uma característica da qualidade de um documento ser de fato o que é, independente se for rascunho ou minuta, original ou cópia, e que seja livre de adulterações ou qualquer outro tipo de intervenção. Este pressuposto está relacionado com a forma de transmissão, além de suas estratégias de custódia e preservação. Logo, o documento arquivístico serve como testemunho da ação que o gerou, por isso ele precisa ser confiável e autêntico (CONARQ, 2006).

Cabe destacar, ainda o caráter orgânico do Arquivo, o vínculo da gênese documental ao próprio ciclo de produção e uso da documentação. O ciclo compreende uma entrada que são os materiais, serviços e informações. Esta entrada é processada e agregada a uma cadeia de valores, onde são executadas atividades para fim e para alimentar o meio. Durante o processo acontece o uso e a produção documental. O final de todo o ciclo também é caracterizado pela produção de materiais, serviços e informações, que alimentam uma nova cadeia de produção e geração. Este caráter orgânico é originado nas estruturas das instituições através de funções desempenhadas e atividades realizadas. Pode-se imaginar que a “organicidade é revelada pelo inter-relacionamento e pelo contexto de existência e de criação.” (SOUSA, 2014, p. 7). Parte desta organicidade é evidenciada na forma de Informação Arquivística.

O conceito de *archival bond*, ou vínculo arquivístico, engloba arquivos, documentos e organicidade. Ela é a característica primordial do documento arquivístico, isto é, fazer parte de algo que vincula o documento a uma estrutura ou instituição. Para Rocha (2011), “alguns arquivistas se referem ao conceito de *archival bond* como vínculo arquivístico. Entretanto, existe outro termo utilizado tradicionalmente pelos arquivistas brasileiros, com o mesmo significado de *archival bond*, que é o de relação orgânica”

(ROCHA, 2011, p.85). Este termo é derivado do termo organicidade, ao qual *archival bond* está relacionado. Podemos considerar a similaridade de sentido, neste ponto, entre *archival bond*, vínculo arquivístico e relação orgânica do documento, todos representados na Informação Arquivística.

Segundo a autora, o *archival bond* é relacionado à organicidade, que é característica de arquivo. Já a organicidade é consagrada na área de arquivos e extremamente referida na literatura científica da área. O diferencial do arquivo em relação a uma coleção de documentos é a organicidade, além disso, os documentos de arquivo apresentam entre si o *archival bond*. Logo, o vínculo de arquivo é parte indissociável da informação arquivística e, portanto, pertence à estrutura de um documento arquivístico digital.

Para Thomassen, o que compõe o arquivo é informação gerada e estruturada pelos processos de trabalho, funcionalmente inter-relacionados. Novamente o relacionamento entre os documentos em uma série é parte integrante do sentido de Arquivo para este autor. Para a prática do profissional de arquivo, como dito, o autor ainda destaca que “um sistema de gerenciamento arquivístico é desenvolvido para estabelecer, manter e explorar a ligação entre estes processos de trabalho e a informação que geram, a fim de otimizar os potenciais informacionais decorrentes de suas relações” (THOMASSEN 2006, p. 6).

## 5 PENSANDO NA PRESERVAÇÃO, DO SIGAD AO RDC-ARQ

Para suceder a preservação, é necessário realizar diversas operações, entre elas, mudanças de suporte e formatos, bem como atualização do ambiente tecnológico. A fragilidade do suporte digital e a obsolescência tecnológica de hardware, software e formato exigem intervenções periódicas (CONARQ, 2012). A preservação digital engloba ações visando manter a integridade e a acessibilidade dos documentos digitais ao longo do tempo. “Devem alcançar todas as características essenciais: físicas (suporte), lógicas (software e formato) e conceituais (conteúdo exibido)” (CONARQ, 2012).

Uma estratégia de preservação deve ser aplicada desde o planejamento da gestão documental e, conseqüentemente da informação arquivística, definindo-se prazos de guarda e eliminação e aplicado diretamente no objeto digital via sistemas de gerenciamento, desde sua gênese. “As ações de preservação digital têm que ser incorporadas desde o início do ciclo de vida do documento. Tradicionalmente, a preservação de documentos arquivísticos se concentra na obtenção da estabilidade do suporte da informação” (CONARQ, 2004). É necessário, portanto, o planejamento da interoperabilidade, definida como a gestão de metadados para preservação. Ela ocorre por meio dos metadados administrativos e estruturais, é definida ao entender um documento em suas características intrínsecas e extrínsecas, com a Diplomática, que ajuda a definir parâmetros para garantir a autenticidade, fidedignidade e fidelidade dos documentos.

Assim, a organicidade no documento arquivístico digital é definida por sua ligação a outros documentos pelo vínculo arquivístico e este deve ser preservado. O documento também deve ter o contexto identificável, deve-se ter acesso à informação sobre o produtor, autor, destinatário, datas. Este documento claramente participa ou apoia uma ação em uma instituição ou para uma pessoa. O contexto poderá ser representado por metadados.

O documento é feito para ser usado. O uso de Sistemas de Informação deve ser entendido como uma forma de pluralização da informação, a fim de promover o cultivo de conhecimento, não só dentro da comunidade onde atua, mas também na sociedade. A utilização dos registros como fontes de informação mantém os documentos vivos (QUISBERT, 2006). Logo, estes dados devem ser preparados tanto para migração como para difusão.

Uma visão importante no modelo brasileiro é quanto ao conjunto SIGAD (Sistema Informatizado de Gestão Arquivística de Documentos) e RDC-ARQ (Repositório Arquivístico Digital Confiável), pois somados representam as três idades documentais (corrente, intermediário e permanente). Ao migrar os documentos arquivísticos para um RDC-ARQ, define-se um ambiente de custódia que herdará pacotes de informação preparadas para a interoperabilidade, portanto, descritas utilizando padrões de descrição arquivísticos.

## 6 RECOMENDAÇÕES PARA PRESERVAÇÃO

Quanto à acessibilidade, um documento acessível é aquele que pode ser localizado, recuperado, apresentado e interpretado sempre que for necessário ou houver uma inferência. A interoperabilidade também é inerente à organização, afinal “o sistema deve garantir a transmissão de documentos para outros sistemas sem perda de informação e de funcionalidades e ser capaz de recuperar qualquer documento em qualquer tempo e apresentá-lo com a mesma forma de sua criação” (CONARQ, 2006, p.22).

Vale lembrar que a confiabilidade é a capacidade de um documento arquivístico sustentar os fatos a que se refere. Para isso, deve ser completo, isto é, criado pela autoridade competente e ter seus procedimentos de criação bem controlados. A confiabilidade está relacionada ao momento da sua produção e à veracidade do seu conteúdo. Além disso, a autenticidade caracteriza a qualidade de um documento ser de fato o que diz ser e que seja livre de adulterações ou qualquer outro tipo de corrupção. Logo, isto está relacionado com a forma de transmissão e estratégias de custódia e preservação (CONARQ, 2012).

Um SIGAD permite ter uma compreensão da organização como um todo, ao estruturar os instrumentos de mediação que propiciem a existência da Gestão de Documentos Arquivísticos, principalmente o plano de classificação e a tabela de temporalidade. É um sistema desenvolvido para produzir, receber, armazenar, dar acesso e destinar documentos arquivísticos digitais (CONARQ, 2006).

Os requisitos do e-ARQ determinam como o SIGAD deve tratar o documento arquivístico como uma unidade complexa produzida e/ou recebida por uma pessoa física ou jurídica, no decorrer das suas atividades, qualquer que seja o suporte, e dotado de organicidade. Ele sempre participa ou apoia uma ação, gerencia fluxos entre pessoas envolvidas (autor, destinatário, redator).

O SIGAD recupera todos os componentes digitais (objeto físico e lógico) que formam o documento arquivístico (objeto conceitual), atendendo a requisitos de gestão dos documentos a partir do plano de classificação para manter a relação orgânica entre os

documentos. Ele gerenciar o ciclo de vida dos documentos arquivísticos significa classificá-los de acordo com as funções e atividades da organização e aplicar os prazos de guarda específicos a cada classe.

Num SIGAD, a classificação passa por registrar e aplicar um código de classificação, além disso, a organicidade é expressa nas relações que os documentos guardam entre si ao refletirem as funções e atividades da pessoa ou organização que os produziu.

Por meio do registro de metadados dos documentos, há o controle do ciclo de vida, da autenticidade, do acesso e da preservação, além disso, a segurança no armazenamento e gestão garantem a autenticidade dos documentos e a transparência das ações (trilha de uso auditável). O SIGAD trata sistematicamente a seleção, a avaliação dos documentos arquivísticos e a sua destinação (eliminação ou guarda permanente), conforme legislação em vigor. Não há eliminação automática: é necessário realizar a exportação dos documentos para transferência e recolhimento.

Portanto, para preservação em muito tempo é necessária uma abordagem de complementação de sistemas de gerenciamento de arquivos e de preservação em longo prazo, integrando a gestão documental com os repositórios digitais confiáveis. Assim, o SIGAD incorpora procedimentos para a preservação de longo prazo dos documentos arquivísticos, e pela preservação do vínculo arquivístico e da manutenção da cadeia de custódia que serão repassados aos RDC-Arq.

Com base na resolução do CONARQ – Implementação de Repositórios Arquivísticos Digitais Confiáveis – RDC-Arq, de 2015, existem algumas considerações acerca do cenário para garantias de confiança de um sistema informático de Arquivo. Estes elementos englobam a responsabilidade pelo repositório, o processamento técnico, a independência de repositórios e as garantias de interoperabilidade. Assim, verificamos que a interoperabilidade da informação arquivística já é garantida a partir da visão de seu repositório que será capaz de preservar a sua integridade, como podemos ver no Quadro 1.

Responsabilidade pelo repositório	- compartilhada por profissionais de arquivo e de tecnologia da informação; - cumprir os requisitos tecnológicos e os procedimentos do tratamento arquivístico.
-----------------------------------	--

Tratamento arquivístico	<ul style="list-style-type: none"> <li>- capaz de organizar e recuperar os documentos, de forma a manter a relação orgânica entre eles</li> <li>- organização a partir de um plano de classificação de documentos</li> <li>- descrição multinível, de acordo com a norma internacional para descrição arquivística</li> </ul>
Independência dos repositórios	<ul style="list-style-type: none"> <li>- um repositório digital deve ter independência;</li> <li>- seu funcionamento e o acesso aos documentos não podem depender das aplicações que funcionam em conjunto com ele.</li> </ul>
Interoperabilidade	<ul style="list-style-type: none"> <li>- conformidade com as normas e padrões estabelecidos,</li> <li>- níveis de interoperabilidade com outros repositórios digitais e sistemas informatizados que tratam de documentos arquivísticos.</li> </ul>

**Quadro 1** – Características de um RDC-Arq

A mesma norma estabelece alguns princípios de preservação digital para garantir o acesso de longo prazo a documentos arquivísticos autênticos, o que implica a adoção de alguns princípios para que um RDC-Arq deva ser capaz de atender aos procedimentos arquivísticos em suas diferentes fases e aos requisitos de um repositório digital confiável.

Deve-se atender a estes requisitos pressupondo que a autenticidade dos documentos arquivísticos digitais está sob ameaça no momento da transmissão em rede (entre pessoas e sistemas) e em relação ao tempo (atualização/substituição de hardware e software usados para armazenar, processar e comunicar os documentos). Por isso, é necessário considerar que a preservação digital é um processo contínuo, que começa na concepção do documento (CONARQ, 2015).

Um ponto importante a ressaltar é que a autenticidade dos documentos arquivísticos digitais tem por base os procedimentos de gestão e preservação e a confiança tanto no repositório como no órgão responsável pela guarda desses documentos. Por isso, a elaboração de manuais e os procedimentos de preservação desempenhados pelo repositório digital apoiam a presunção de autenticidade desses documentos, assim como o registro, em metadados, das intervenções de preservação em cada documento apoia a presunção de autenticidade desses documentos (CONARQ, 2015). Esta operação pode, por exemplo, ser acompanhada de uma ontologia descritiva aplicada ao sistema, que evolui no espaço tempo (acompanha o processo de avaliação e eliminação) e registra estes tipos de alterações até definir um quadro final de arranjo que representa as funções e atividades

daquele domínio. Como se vê mais adiante, as ontologias são uma forma de representação e estruturação de conceitos.

Os repositórios confiáveis focam nos documentos arquivísticos autênticos, e não genericamente em objetos digitais. Pois é necessário arbitrar o que se considera como documento original, uma vez que a preservação digital implica a necessidade de conversão de formatos e atualização de suportes. Outro fator destacado pelo CONARQ é quanto a distinguir claramente a autenticidade e autenticação de documentos, considerando que a primeira é a qualidade de o documento ser verdadeiro, e a segunda é uma declaração dessa qualidade, feita, em um dado momento, por uma pessoa autorizada para tal. A autenticidade dos documentos digitais deve ser avaliada e presumida no momento de sua submissão ao repositório. Por fim, o repositório digital é responsável pela manutenção permanente da autenticidade dos documentos a ele submetidos.

Como vemos, os princípios almejados pelo CONARQ representam as condições mínimas para considerar uma plataforma digital confiável e de que forma devem ser tratados os registros. Alguns deles podem englobar ontologias, mas sempre se trata de uma boa gestão de metadados e registros do estado de integridade e de processamento da informação declarada como arquivística.

Para Santos e Flores (2015), a confiabilidade é também concedida pela relação de segurança e credibilidade, pois o “repositório digital deve ser o ambiente autêntico para a preservação em longo prazo, dispendo, por exemplo, de ferramentas para a implementação das estratégias de preservação e inserção de padrões de metadados.” (SANTOS; FLORES, 2015, p.205)

Os autores ainda afirmam que para este repositório todas as ações realizadas em relação aos documentos digitais (migrações, por exemplo) devem ser registradas detalhadamente, criando-se assim, um histórico de cada objeto digital armazenado, acrescentando confiabilidade aos conteúdos. Ainda destacam a interoperabilidade como ferramenta de gestão e preservação, associada às políticas e ao plano de preservação para servirem de artifícios para a criação de um repositório arquivístico digital confiável em longo prazo.

Podemos ver que os metadados têm uma importante função no quesito confiabilidade: ao analisarmos o RDC-Arq, podemos verificar seu papel ao longo do tempo nas migrações entre ambientes de custódia documental.

## 7 APONTAMENTOS FINAIS

Definidas algumas das principais funções que envolvem a organização dos objetos informacionais, quer seja identificar a existência de todos os tipos de recursos informacionais, assim que estiverem disponíveis, quer seja agrupar sistematicamente recursos informacionais em coleções (de bibliotecas, arquivos, museus, e ambientes digitais da internet), gera-se Informação Arquivística ao se produzir listas desses recursos informacionais, oferecendo metadados úteis de identificação.

Parte do processo de organização da informação digital nada mais é do que evidenciar a essência informacional de cada registro, estabelecendo o processo de descrição e de indexação por meio de metadados e assim garantindo sua preservação. Esse tratamento da informação é objeto de trabalho dos profissionais da informação. Para isso, conta-se com os metadados e os padrões, como os definidos para a descrição arquivística.

Para Diallo (2011) a maneira de alcançar a interoperabilidade é chegar a um acordo sobre quais informações devem ser trocadas em um determinado domínio e concordar sobre a forma de descrevê-lo. Como as informações devem ser trocadas, pode ser acordada com ou sem os usuários. O resultado é um modelo comum, muitas vezes padronizado, que descreve o domínio de interesse. Esta já é uma prática de tratamento da informação, a preparação da interoperabilidade de forma que tanto os humanos quanto os agentes computacionais entendam.

Cabe destacar que, no âmbito da web 3.0, e considerando a organização da informação, parte das taxonomias e dos modelos descritivos alimentam os computadores, tornando-os capazes de entender as estruturas e ontologias e, conforme aponta Robredo (2010), “a finalidade é acrescentar uma camada de significado no topo da web atual que faça dela menos um catálogo e mais um guia – e ainda forneça os fundamentos para que os sistemas possam raciocinar de forma semelhante aos humanos” (ROBREDO, 2010, p.19).

Neste cenário, as linguagens documentárias (como as taxonomias) são elementos importantes que operam combinados, sinergicamente, com outros aspectos da web 3.0 e fazendo parte de uma de suas camadas. Isso já estava previsto na formulação do conceito da Web Semântica, que foi concebido como um conjunto de tecnologias relacionadas. Em 2000, o W3C divulgou a primeira proposta de arquitetura para Web Semântica, no qual destaca que:

Com base em uma série de camadas sobrepostas, onde cada camada ou tecnologia deveria obrigatoriamente ser complementar e compatível com as camadas ao mesmo tempo em que não deveria depender das camadas superiores, possibilitando assim uma estrutura idealmente escalonável, que indicasse os passos e as tecnologias necessários para a concretização do projeto Web Semântica (RAMALHO; VIDOTTI; FUJITA, 2007, p. 2).

Faz-se imprescindível que se criem instrumentos que forneçam os sentidos lógico e semântico das informações aos computadores, sendo o uso de linguagens documentárias uma das camadas que interagem e os valores aplicados por meio de links semânticos, estes são os instrumentos, conforme Ramalho, Vidotti e Fujita (2007).

Já Smith (2008) destaca que, quando se compara as diversas ontologias de determinados campos, para a Ciência da Informação a ontologia é um dicionário de termos formulados em uma sintaxe canônica, com definições comumente aceitas, projetadas para produzir um quadro lexical ou taxonômico para o conhecimento-representação, e que pode ser compartilhado por diferentes comunidades de sistemas de informação.

Com as características sintáticas dos metadados, deve-se considerar os elementos descritivos propostos pela NOBRADE como uma forma de identificar os links e os seus respectivos registros, como experienciado por esta pesquisa no item anterior. Como visto, o resultado da descrição estaria fornecendo informação arquivística padronizada para o contexto 3.0, desde que seja:

[...] expresso em RDF, que o codifica em conjuntos de sujeito, verbo e predicado, sendo cada um desses elementos identificados com URIs (Universal Resource Identifier). Os URIs vão permitir a utilização desses sujeitos e predicados como ligações em qualquer página web e que qualquer pessoa defina um novo verbo ou conceito na Web (BERNERS-LEE; HENDLER; LASSILA, 2001, p. 2).

Em analogia, pode-se afirmar que as URIs da web 3.0 possuem elementos descritivos, suas fichas de metadados que descrevem e dão sentido a um elemento do objeto informacional (aqui usado como sinônimo de registro). Este conjunto, a que se referem Berners-Lee, Hendler e Lassila, pode ser considerado como um registro descritivo, com as relações representativas criadas pelos elementos descritivos arquivísticos, aplicados aos links.

Deste modo, essa seria a abordagem arquivística para a Web Semântica aqui detalhada, considerando a criação de ontologias comuns, originadas por processos descritivos normatizados, que evidenciam a informação arquivística e as padronizam em forma de metadados, possibilitando sua interoperabilidade por meio de plataformas e diferentes sistemas digitais (incluindo redes) ao longo do tempo. A preservação é mais facilmente implantada em sistemas que garantam que os dados sejam preparados para a interoperabilidade, pois estas independem de plataforma tecnológica e, portanto, de linguagens tecnológicas sujeitas às regras de mercado e de momento.

## **THE INTEROPERABILITY IN THE PRESERVATION OF ARCHIVAL INFORMATION: METADATA AND DESCRIPTION**

### **ABSTRACT**

It presents the impacts of digital reality on access to information and evaluates the interoperability of archival information and the standards that aim to structure it. It analyzes the information standards originated by the documentary structure, its archival bond and the Archival Description, relating to the definition and management of metadata for digital preservation. It concludes that the creation of ontologies originated by normative descriptive processes, evidences the archival information and standardizes it, in the form of metadata, allowing its interoperability through platforms and different digital systems.

**Keywords:** Ontology. Interoperability. Standards and archival description. Archival information. Digital information.

## REFERÊNCIAS

ALONSO, S. S.; URBÁN, M. Á. S.; LEGUINA, G. R. Sobre la interoperabilidad semántica en las descripciones archivísticas digitales. **Revista Española de Documentación Científica**, 2008, v. 31, n. 1, p. 18-31. Disponível em: <<http://redc.revistas.csic.es/index.php/redc/article/download/410/422>>. Acesso em: 18 fev. 2015.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. **The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities.** Scientific American, 2001. Disponível em: <[http://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American\\_%20Feature%20Article\\_%20The%20Semantic%20Web\\_%20May%202001.pdf](http://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf)>. Acesso em: 26 nov. 2015.

CHOO, C. W. **A organização do conhecimento: como as organizações usam a informação para criar conhecimento, construir conhecimento e tomar decisões.** São Paulo: SENAC, 2003.

CONSELHO NACIONAL DE ARQUIVOS (CONARQ). **Glossário de documentos arquivísticos digitais.** Câmara Técnica de Documentos Eletrônicos. Rio de Janeiro: Arquivo Nacional, 2004.

\_\_\_\_\_. **Modelo de requisitos para sistemas informatizados de gestão arquivística de documentos; e-ARQ.** Câmara Técnica de Documentos Eletrônicos. Rio de Janeiro: Arquivo Nacional, 2006.

\_\_\_\_\_. **Diretrizes para a presunção de autenticidade de documentos arquivísticos digitais.** Câmara Técnica de Documentos Eletrônicos. Rio de Janeiro: Arquivo Nacional, 2012. Disponível em: <[http://www.conarq.arquivonacional.gov.br/images/publicacoes\\_textos/conarq\\_presuncao\\_autenticidade\\_completa.pdf](http://www.conarq.arquivonacional.gov.br/images/publicacoes_textos/conarq_presuncao_autenticidade_completa.pdf)>. Acesso em: 22 mar. 2015.

\_\_\_\_\_. **Diretrizes para a implementação de repositórios arquivísticos digitais confiáveis: RDC-Arq 2015.** Câmara Técnica de Documentos Eletrônicos (CTDE). Rio de Janeiro: Arquivo Nacional, 2015. Disponível em: <[http://www.conarq.arquivonacional.gov.br/images/publicacoes\\_textos/diretrizes\\_rdc\\_arq.pdf](http://www.conarq.arquivonacional.gov.br/images/publicacoes_textos/diretrizes_rdc_arq.pdf)>. Acesso em: 09 abr. 2017.

DIALLO, S. Y. et al. **Understanding interoperability.** In: Proceedings of the Emerging M&S Applications in Industry and Academia Symposium, San Diego, CA, USA, 2011, 84–91.

GONÇALVES, J. **Como classificar e ordenar documentos de arquivos.** Associação dos Arquivistas de São Paulo; Arquivo do Estado, 1998, 38 p. (Projeto Como Fazer, 2)

Disponível em:

<[http://www.arqsp.org.br/arquivos/oficinas\\_colecao\\_como\\_fazer/cf2.pdf](http://www.arqsp.org.br/arquivos/oficinas_colecao_como_fazer/cf2.pdf)>. Acesso em: 22 mar. 2015.

INNARELLI, H. **Gestão da preservação de documentos arquivísticos digitais**: proposta de um modelo conceitual. 2015. Tese de Doutorado ECA/USP.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). NBR ISO / IEC 9126-1 **Engenharia de software - Qualidade de produto - Parte 1: Modelo de qualidade**. 2003.

LOPES, L. C. **A nova arquivística na modernização administrativa**. 2.ed. Brasília: Projecto Editorial, 2009.

QUISBERT, H. **A framework for the development of archival information systems**. Tese de Doutorado. (2006) Luleå University of Technology.

RAMALHO, R. A. S.; VIDOTTI, S. A. B. G.; FUJITA, M. S. L. **Web Semântica**: uma investigação sob o olhar da Ciência da Informação. DataGramaZero - Revista de Ciência da Informação, v. 8, n. 6, 2007. Disponível em: <[http://www.dgz.org.br/dez07/Art\\_04.htm](http://www.dgz.org.br/dez07/Art_04.htm)>. Acesso em: 26 nov. 2014.

RIBEIRO, F. **Organizar e representar informação: apenas um meio para viabilizar o acesso**. Revista da Faculdade de Letras, Ciências e Técnicas do Património, Faculdade de Letras da Universidade do Porto, v. 4, p. 83-100, 2005. Disponível em: <<http://ler.letras.up.pt/uploads/ficheiros/1239.pdf>>. Acesso em: 22 mar. 2015.

ROBREDO, Jaime. Ciência da informação e Web semântica: Linhas convergentes ou linhas paralelas? In: Jaime Robredo; Marisa Bräscher (Orgs.). **Passeios pelo bosque da informação**: estudos sobre representação e organização da informação e do conhecimento. Brasília DF: IBICT, 2010. 335 p. ISBN: 978-85-7013-072-3. Capítulo 1, p. 12-47. Edição eletrônica. (Edição comemorativa dos 10 anos do Grupo de Pesquisa EROIC). Disponível em: <<http://www.ibict.br/publicacoes/eroic.pdf>>. Acesso em: 22 mar. 2015.

ROCHA, C. L. Glossário Multilíngue do Projeto Interpares 31. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, 2011, p.76-90.

SANTOS, H; FLORES, D. **Trusted digital repositories for digital archival documents**: considerations on the preservation in long-term. Perspectivas em Ciência da Informação, 2015, v. 20, n. 2, p. 198-218.

SMIT, J. **A informação na Ciência da Informação**. InCID: R. Ci. Inf. e Doc. Ribeirão Preto, 2012, v. 3, n. 2, 84-101.

\_\_\_\_\_. W.; BARRETO, A. A. **Ciência da Informação**: base conceitual para a formação do profissional. In: Valentim, M. L. P. (Org.). Formação do profissional da informação. São Paulo: Polis, 2002.

SMITH, B. **Ontology**. The Blackwell Guide to the Philosophy of Computing and Information, 2008, pp.153-166.

SOUSA, R. T. B. **Alguns apontamentos sobre a classificação de documentos de arquivo**. Brazilian Journal of Information Science: Research Trends, 2014, v. 8, n. 1/2.

THOMASSEM, T. **Uma primeira introdução à arquivologia**. Arquivo & Administração, Rio de Janeiro, v. 5, n. 1, p. 5-16, jan./jun. 2006.

---

**Trabalho recebido em: 07 mar. 2016**

**Trabalho aceito em: 09 abr. 2017**

---