



## MODELO LÓGICO DA INFORMAÇÃO NO *OPEN ARCHIVAL INFORMATION SYSTEM*: UMA REFLEXÃO ARQUIVÍSTICA SOBRE O PACOTE DE INFORMAÇÃO PARA ARQUIVAMENTO

**Henrique Machado dos Santos**

Mestre em Patrimônio Cultural pela Universidade Federal de Santa Maria, Brasil. Arquivista da Universidade Federal do Rio Grande, Brasil.

E-mail: [henrique.hms.br@gmail.com](mailto:henrique.hms.br@gmail.com)

**Daniel Flores**

Doutor em *Metodologías y Líneas de Investigación en Biblioteconomía y Documentación* pela Universidad de Salamanca, Espanha. Professor da Universidade Federal Fluminense, Brasil.

E-mail: [dfloresbr@gmail.com](mailto:dfloresbr@gmail.com)

### Resumo

O objetivo deste estudo é analisar o modelo lógico da informação proposto pelo *Open Archival Information System* (OAIS), e identificar as suas convergências com a Arquivística. Verifica-se a estrutura do pacote de informação para avaliar a sua capacidade de preservar documentos arquivísticos digitais autênticos em longo prazo. A metodologia utilizada parte do levantamento bibliográfico de materiais previamente publicados. Os dados coletados são analisados de forma qualitativa e a discussão segue a lógica dedutiva. Assim, tem-se um artigo de revisão assistemática, que proporciona uma reflexão sobre o tema e salienta o ponto de vista dos autores. Nas conclusões, observa-se que o pacote de informação para arquivamento comporta questões pertinentes à Arquivística, como, por exemplo, garantia de autenticidade, manutenção da proveniência, organicidade, e histórico de custódia. Portanto, reafirma-se a necessidade da implementação de repositórios arquivísticos digitais confiáveis em conformidade com o OAIS.

**Palavras-chave:** Preservação digital. Documento digital. Arquivística. Repositório digital. Autenticidade.

### LOGICAL MODEL OF INFORMATION IN OPEN ARCHIVAL INFORMATION SYSTEM: AN ARCHIVAL REFLECTION ON THE ARCHIVAL INFORMATION PACKAGE

### Abstract

*The objective of this study is to analyze the logical information model proposed by the Open Archival Information System (OAIS), and to identify its convergences with Archival science. Check the structure of the information packet to evaluate its ability to preserve authentic long-term digital archival records. The methodology used is part of a bibliographical survey of previously published materials. The data collected are qualitatively analyzed and the discussion follows the deductive logic. Thus, there is an article of no systematic review, which provides a reflection on the theme and highlights the authors' point of view. In the conclusions, it is observed that the archival information package includes issues pertinent to Archival science, such as guarantee of authenticity, preservation of provenance, organicity, and custody history. Therefore, it is reaffirmed the need to implement reliable digital archival repositories in accordance with OAIS.*

**Keywords:** Digital preservation. Digital record. Archival science. Digital repository. Authenticity.

## 1 INTRODUÇÃO

O modelo *Open Archival Information System* (OAIS) é considerado a principal referência na área da preservação digital, essencialmente, possui um modelo funcional e um modelo de informação para desenvolver um repositório digital. O modelo funcional discorre sobre as entidades responsáveis pela aquisição, armazenamento, gestão de metadados, administração do repositório, planejamento da preservação e gestão do acesso à comunidade designada. Já o modelo de informação apresenta os componentes dos objetos de informação e a sua relação com os demais objetos digitais.

Dessa forma, faz-se uma distinção: o modelo funcional apresenta os fluxos de informação pelos quais perpassam os documentos que se desejam preservar; e o modelo de informação apresenta a forma com que a informação é estruturada para que possa ser preservada no longo prazo. O modelo de informação pode ainda ser dividido em duas partes: a primeira apresenta a estrutura dos tipos de informação mais comuns em um OAIS; e a segunda aborda detalhes sobre a preservação dos documentos digitais por meio de um pacote de informação.

O modelo OAIS corresponde a norma ISO 14721:2012, e já é suficientemente difundido no âmbito da comunidade de pesquisa em preservação digital. Entretanto, detalhes relacionados ao modelo funcional e ao modelo de informação ainda são pouco compreendidos. Consequentemente, tornam-se necessários estudos que aprofundem o OAIS, e descrevam a sua possível contribuição para a preservação de documentos arquivísticos digitais em longo prazo.

Diante do exposto, este estudo tem o objetivo analisar o modelo lógico da informação no OAIS, e as suas possíveis convergências com a Arquivística. A ênfase do estudo está em verificar se a estrutura do pacote de informação tem a capacidade de preservar e garantir acesso em longo prazo a documentos arquivísticos digitais autênticos.

A metodologia utilizada consiste no levantamento bibliográfico de materiais previamente publicados, que comporta: livros, publicações técnicas e artigos de periódicos científicos recuperados pela ferramenta *Google Scholar*. Os dados coletados são analisados de forma qualitativa e a discussão segue a lógica dedutiva. Como consequência, obtém-se um artigo de revisão com caráter assistemático, sua finalidade consiste em proporcionar uma reflexão sobre o tema, para salientar o ponto de vista dos autores (GIL, 2010; LUNA, 1997; SILVA; MENEZES, 2005; VOLPATO; BARRETO; UENO; VOLPATO; GIAQUINTO; FREITAS, 2013).

## 2 PACOTE DE INFORMAÇÃO

Os documentos digitais e seus respectivos metadados são, de forma lógica, entendidos como objetos separados, ainda que os metadados estejam inseridos nos documentos. Logo, tem-se uma sequência de *bits*, que se relaciona com todos os objetos necessários para tornar o documento compreensível. Desta forma, o modelo OAIS define quatro classes de objetos de informação que, em conjunto, formam o pacote de informação: informação de conteúdo, informação descritiva de preservação (*Preservation Description Information – PDI*), informação de empacotamento e descrição do pacote (SARAMAGO, 2004).

O pacote de informação é a estrutura conceitual utilizada para apoiar o processo de preservação em longo prazo. Ele é um recipiente que contém dois tipos de objetos de informação: a informações de conteúdo e a informação descritiva de preservação. Tal pacote pode ser associado com outros dois tipos de objetos de informação: a informação de empacotamento e a descrição do pacote. Sendo assim, o pacote de informação consiste em um conjunto de objetos digitais encapsulados.

O encapsulamento consiste em adicionar descrição aos objetos digitais preservados no que se referem as suas estruturas físicas ou lógicas, de modo a reunir todas as informações necessárias para sua correta interpretação no futuro (SAYÃO, 2010). Assim, os documentos podem ser encapsulados junto com a aplicação de *software* utilizada em sua produção e com a descrição do ambiente de *hardware* e *software* requerido para executar tal aplicação (MÁRDERO ARELLANO, 2004).

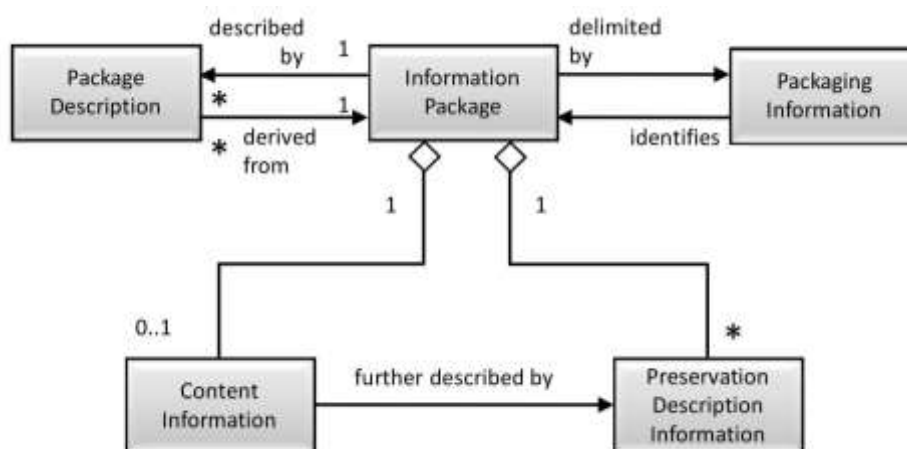
Vale ressaltar que um documento digital não é uma unidade simples, mas sim composto por um conjunto de componentes digitais que definem sua estrutura, tipo, tamanho e cores das fontes, existência de imagens etc. Alguns desses componentes (como os tipos de fonte ou imagens) estão presentes no ambiente de produção, mas não são, obrigatoriamente, carregados junto com o documento. Assim, é aconselhável o uso de formatos que encapsulem todos os componentes digitais, de modo a garantir que o documento tenha o máximo de independência do ambiente tecnológico em que foi criado (SANTOS, 2012, p. 121).

No pacote de informação, a informação de conteúdo corresponde ao documento arquivístico digital e a PDI consiste nos metadados necessários para preservar e garantir o acesso em longo prazo. Assim, o documento arquivístico e os metadados são encapsulados e identificados pela informação de empacotamento. Posteriormente, são associadas às informações descritivas do pacote, as quais correspondem aos metadados descritivos que auxiliam a localizar o pacote em um repositório digital (ROCHA, 2015).

O pacote de informação será o meio para reunir, transportar e localizar os documentos arquivísticos digitais e os seus respectivos metadados de preservação. Neste contexto, ressalta-se que os documentos arquivísticos são considerados a meta da preservação em longo prazo.

Existem diferentes tipos de pacotes de informação (*information package*) que são usados na preservação dos documentos, tais pacotes podem ter por finalidade: estruturar e armazenar os conteúdos do OAIS; transportar a informação necessária vinda do produtor até o OAIS; ou transportar a solicitação de informações entre o OAIS e os consumidores (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). A seguir a “Figura 1 – Conteúdos do pacote de informação” apresenta a relação entre seus elementos.

Figura 1 – Conteúdos do pacote de informação



Fonte: (CCSDS, 2012, p. 4-34)

Um pacote de informação poderá conter “zero” ou “um” objeto de informação de conteúdo (*content information*); “zero” ou “mais” objetos de PDI (*preservation description information*); e está associado exatamente a “uma” unidade de informação de empacotamento (*packaging information*) que o identifica e delimita. Além disso, o pacote de informação também está associado a “zero” ou “mais” descrições de pacote (*package description*) que descrevem a informação de conteúdo (*content information*) a fim de otimizar o acesso.

Assim, os documentos digitais devem ser gerenciados por um Repositório Arquivístico Digital Confiável (RDC-Arq) em conformidade com o modelo OAIS. Este irá definir pacotes de informação para reunir os documentos digitais (informação de conteúdo) e seus respectivos metadados (informação de representação) (BRASIL, 2015b).

Sendo assim, o RDC-Arq será o ambiente tecnológico complexo para armazenar e gerir os documentos digitais. Um ambiente composto por soluções informatizadas que capturam, armazenam, preservam e concedem acesso aos documentos custodiados. Logo, o RDC-Arq é formado por *hardware*, *software*, serviços, um acervo de informações digitais e os respectivos metadados associados à informação. Dessa forma, tem-se por objetivo apoiar a gestão de documentos digitais no longo prazo (ROCHA, 2015).

Caberá às organizações decidirem qual alternativa selecionar entre os aplicativos disponíveis para RDC-Arq's. Para tanto, deve considerar suas capacidades, suas limitações e, essencialmente, se o aplicativo se adapta a natureza dos documentos digitais e se comporta os requisitos arquivísticos. Alternativamente, as organizações podem optar por desenvolver o seu próprio RDC-Arq (THOMAZ, 2007).

### 3 TIPOS DE PACOTE DE INFORMAÇÃO

Um OAIS contém pacotes de informação que reúnem a informação de conteúdo, a PDI, a informação de empacotamento e a informação descritiva. Por conseguinte, define três tipos de pacotes de informação: o Pacote de Informação para Submissão (*Submission Information Package – SIP*), que é fornecido pelo produtor ao OAIS; o Pacote de Informação para Arquivamento (*Archival Information Package – AIP*), que tem por finalidade a preservação de longo prazo; e o Pacote de Informação para Disseminação (*Dissemination Information Package – DIP*), que é entregue ao consumidor mediante solicitação. Logo, as funções de alto nível do OAIS compreendem a aquisição de SIP's, o armazenamento de AIP's, a gestão de dados, a administração do sistema, e a disponibilização de DIP's para acesso (HEDSTROM, 2001).

Inicialmente, o produtor envia um pacote de informação à entidade de admissão, este pacote é denominado SIP; posteriormente, o mesmo pacote é transformado em AIP para ser arquivado no repositório; e após solicitação do consumidor, o AIP é convertido em DIP para ser entregue pela entidade de acesso ao consumidor.

A maioria dos SIP's terá alguma informação de conteúdo e algumas PDI's, mesmo assim, vários SIP's poderão ser necessários para fornecer um conjunto completo de informações de conteúdo e PDI associada. As informações de conteúdo e as PDI's possuem informação de representação associada; e observa-se ainda, que no caso de existirem vários SIP's que usem a mesma informação de representação, é provável que tais informações de representação sejam fornecidas uma única vez ao OAIS, se for aplicada a diversos SIP's. Em outro caso, a PDI será fornecida em um SIP separado, sem informação de conteúdo. Já a informação de empacotamento sempre estará presente de alguma forma (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012).

Todo SIP recebido pela entidade de admissão passará por uma verificação de qualidade e será transformado em AIP conforme os padrões de formatação de dados do repositório; além de gerar informação descritiva para auxiliar na pesquisa e na recuperação de

conteúdos. E posteriormente, a entidade de admissão irá transferir os AIP's para a entidade de armazenamento arquivístico, e as informações descritivas a entidade de gestão de dados (THOMAZ; SOARES, 2004). A entidade de admissão irá receber as informações, verificar a qualidade da transferência, gerar o AIP, adicionar informação descritiva, e posteriormente, solicitar o seu arquivamento.

Salienta-se que no ambiente OAIS, um ou mais SIP's são transformados em um ou mais AIP's para preservação. O pacote AIP possui um conjunto completo de PDI para as informações do conteúdo associadas; e esse pacote ainda pode conter um conjunto de outros AIP's (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012).

As informações de empacotamento de um AIP seguirão a conformidade com as normas internas do OAIS, e podem variar, conforme sua respectiva gestão. Caso a informação descritiva associada a um AIP seja abrangente, os consumidores podem localizar e solicitar as informações de conteúdo de interesse; o que otimiza o processo de busca e recuperação da informação.

Em resposta a uma solicitação, o OAIS irá fornecer uma parte ou até mesmo o AIP completo ao consumidor, que será entregue na forma de DIP. O pacote DIP pode incluir coleções de AIP, e conter PDI, seja ela completa ou não. Ressalta-se que as informações de empacotamento estarão presentes, de alguma forma, para que o consumidor possa distinguir claramente a informação solicitada. As informações de empacotamento podem assumir várias formas, dependendo da mídia de divulgação e das exigências dos consumidores.

Já a informação descritiva associada ao DIP, poderá ser disponibilizada em qualquer momento de sua transferência; sua finalidade é fornecer informações suficientes para que o consumidor possa reconhecer o DIP entre possíveis pacotes semelhantes, por exemplo, uma descrição textual integrando a informação de empacotamento (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012).

Ressalta-se que o RDC-Arq deve ter procedimentos para verificar, e caso necessário, adequar os documentos, para que sejam compreensíveis aos consumidores, enquanto comunidade designada (BRASIL, 2015b). Logo, entende-se que o RDC-Arq terá dupla finalidade: preservar os documentos arquivísticos digitais em longo prazo, e garantir o acesso, de modo que sejam autênticos e compreensíveis à sua comunidade designada.

Entende-se por comunidade designada o conjunto identificado de usuários potenciais com capacidade para entender um conjunto específico de informações do OAIS. A comunidade designada pode conter várias comunidades de usuários. Cabe ao arquivo definir sua comunidade designada, salienta-se que essa definição poderá ser alterada com o tempo (CCSDS, 2012; ISO 14721:2012).

No contexto da Arquivística, essa distinção de público-alvo não ocorre de forma rígida. O acesso aos documentos preservados é concedido para todos os indivíduos que tenham interesse, salvo os casos de informação sigilosa. Mesmo assim, é possível definir comunidades designadas na esfera local, regional, nacional, etc., a fim de atender as necessidades e facilitar o acesso para tais grupos.

Não se trata de aplicar uma restrição de público, mas sim, em aperfeiçoar o processo de busca para os consumidores já conhecidos. Dessa forma, a comunidade designada corresponde ao arquivo e seus usuários potenciais. Conseqüentemente, tem-se por objetivo atender quaisquer usuários interessados nas informações do RDC-Arq, e lhes proporcionar acesso a documentos que sejam facilmente compreensíveis.

#### **4 PACOTE DE INFORMAÇÃO PARA ARQUIVAMENTO**

O RDC-Arq deverá demonstrar se os SIP's foram aceitos de forma integral, em parte (justificar as lacunas) ou se foram recusados. Caso forem aceitos, será preciso descrever a

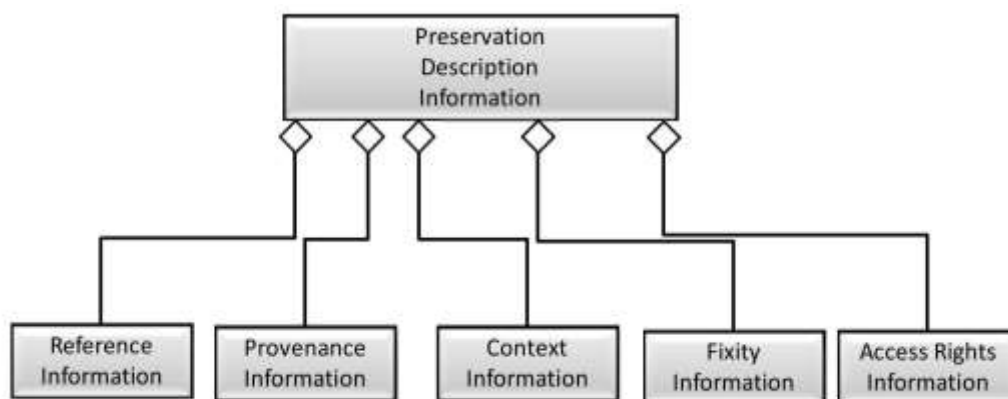
transformação dos SIP's em AIP's, verificar a completude do AIP no momento em que for gerado; além de identificar os metadados que serão adicionados como consequência desse processo. E posteriormente, devem-se adicionar identificadores com base em padrões reconhecidos amplamente (BRASIL, 2015b).

Logo, os documentos arquivísticos digitais e os metadados contidos no SIP serão reorganizados na forma de um AIP, e receberão uma informação descritiva oriunda desse processo. Tal informação é essencial para manutenção da autenticidade visto que toda a alteração nos documentos digitais deverá ser registrada para garantir que não foi realizada de forma arbitrária.

A definição do AIP fornece um conjunto de informações contendo as qualidades necessárias à preservação em longo prazo de um determinado objeto de informação. O pacote AIP, enquanto objeto de informação, constitui um recipiente de outros objetos de informação. Assim, no interior do AIP encontram-se a informação de conteúdo, considerada o alvo da preservação, e a PDI que contém informações adicionais; que são necessárias para tornar as informações de conteúdo significativas no longo prazo. Observa-se que os requisitos da PDI de um AIP são obrigatórios, de modo que todas as suas classes de informações devem estar presentes, diferentemente dos demais pacotes de informação, onde a PDI não é obrigatória (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012).

A PDI é obrigatória pelo fato do AIP ser o pacote utilizado para reunir o documento digital e as suas informações relacionadas, consideradas necessárias à preservação de longo prazo. Já no caso do SIP e do DIP, a PDI torna-se opcional, visto que estes pacotes têm finalidades diferentes do AIP, que são respectivamente: reunir a informação que se pretende preservar, e disseminar a informação desejada de forma autêntica e compreensível ao consumidor. A seguir, a “Figura 2 – Informação descritiva de preservação” apresenta os subtipos da PDI.

Figura 2 – Informação descritiva de preservação



Fonte: (CCSDS, 2012, p. 4-38)

A informação de referência (*reference information*) irá identificar e descrever os mecanismos necessários para fornecer identificadores às informações de conteúdo e aos sistemas externos, de modo que possam referenciá-las sem ambiguidades (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). Portanto, a informação de referência otimiza o processo de busca e recuperação da informação de conteúdo desejada.

A informação de proveniência (*provenance information*) irá documentar o histórico da informação de conteúdo, identificar sua origem, sua cadeia de custódia, e fornecer trilhas de auditoria para identificar quaisquer alterações, contribui assim, para elevar os níveis de confiabilidade (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). Logo, devem-se

registrar os metadados de preservação associados aos documentos, e assim, garantir a integridade, a proveniência, além de auxiliar na localização (BRASIL, 2015b). O objetivo da custódia consiste em manter o vínculo arquivístico entre os documentos e garantir a sua segurança, para assim, oferecer acesso a documentos autênticos (SILVA, 2017). Desta forma, a informação de proveniência irá registrar eventos relacionados à origem e a custódia das informações de conteúdo, e possibilitar o registro de informações pertinentes à preservação, como, por exemplo, migrações de suporte e formato.

A informação contexto (*context information*) irá documentar as relações das informações de conteúdo com o ambiente, registrar o motivo de sua produção e a sua relação com as outras informações de conteúdo do OAIS (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). A informação de contexto relaciona a informação de conteúdo de um AIP com as de outros AIP's, de modo que lhes confere organicidade, um princípio elementar à Arquivística.

A informação de fixidez (*fixity information*) irá verificar a integridade dos dados para garantir que a informação de conteúdo não tenha sido alterada sem registro. Também inclui esquemas de detecção de erro para os objetos de informação de conteúdo, no entanto não possui mecanismos para preservação da integridade, apenas especifica os requisitos mínimos (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). Assim, é pertinente dispor de ferramentas amplamente reconhecidas para monitorar os componentes digitais dos documentos (BRASIL, 2015b). Com a informação de fixidez é possível verificar a forma fixa e o conteúdo estável da informação de conteúdo, ou seja, a apresentação do documento arquivístico digital será sempre a mesma, e o seu conteúdo permanecerá inalterado.

E a informação de direitos de acesso (*access rights information*) irá identificar as restrições relacionadas à informação de conteúdo, isto inclui as licenças, o controle de acesso e a distribuição estabelecida pelo acordo de submissão. Observa-se que estas questões estão relacionadas tanto à preservação, quanto ao uso da informação de conteúdo (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). Por meio da informação de direitos de acesso será possível garantir o cumprimento com quaisquer restrições relacionadas à informação de conteúdo, o que evita futuros problemas, tanto na disseminação de conteúdos, quanto nas ações de preservação efetuadas. Por consequência, as migrações e a disseminação dos DIP's irão considerar tais restrições.

Ressalta-se que o RDC-Arq deve atender um conjunto de condições para garantir a preservação em longo prazo dos AIP's, manter a descrição multinível, a autenticidade, a organicidade dos documentos e o acesso contínuo em longo prazo. Tais condições consistem em: utilizar estratégias previamente definidas no planejamento da preservação, e registrá-las nos metadados de preservação; verificar a necessidade de migração de suporte ou conversão de formatos; preservar a informação de conteúdo do AIP, que foi originalmente admitida no repositório, assim como, aquela resultante da última migração; monitorar a integridade dos AIP's com o registro de metadados de fixidez e realizar verificações com ferramentas de *checksum*; além de manter um registro de todas as ações proferidas sobre os AIP's (BRASIL, 2015a; 2015b).

Neste ponto, observa-se a necessidade de preservar dois objetos em paralelo: o objeto submetido originalmente e o objeto de preservação. Portanto, será preferível recorrer ao objeto original caso um determinado procedimento/estratégia de preservação seja insuficiente ou proporcione resultados indesejados. Logo, torna-se possível recomeçar a linha de preservação, e optar por outras estratégias para melhor atender as necessidades da comunidade designada.

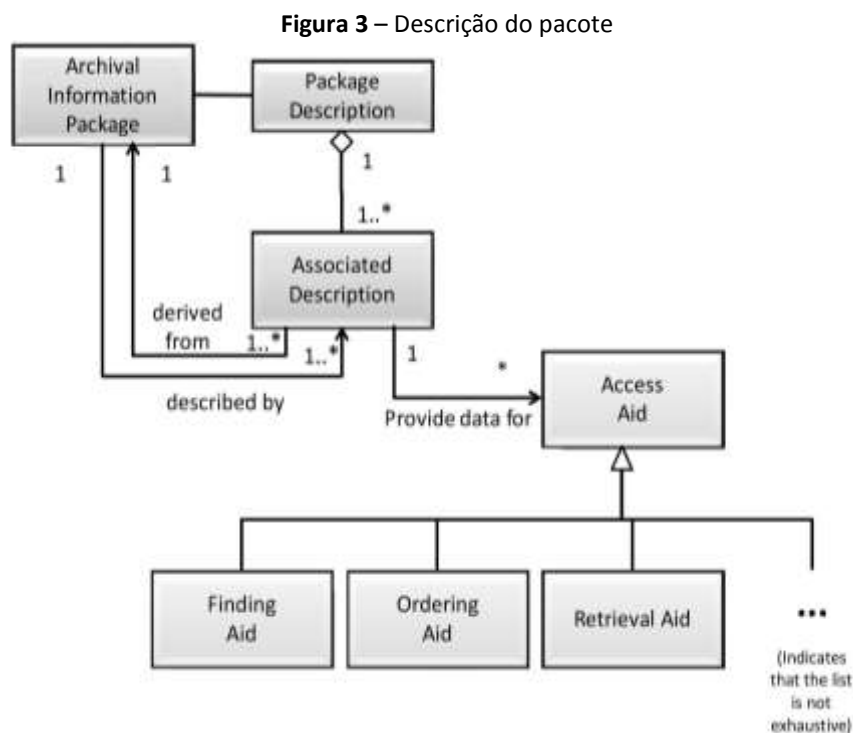
Ressalta-se que os métodos de preservação digital ainda não estão sedimentados, da mesma forma, não podem ser limitados apenas às estratégias de migração e emulação. Há documentos em diversos formatos, e a comunidade de preservação ainda está começando a

compreendê-los (CLOONAN, 2016). Como consequência, surge um processo de reformulação da preservação de documentos arquivísticos, que envolve conceitos de autenticidade, fixidez e custódia. Tal fato gera novos estudos que constituem uma bibliografia especializada sobre o tema (SANTOS; FLORES, 2018). Assim, ao considerar tal mutabilidade dos formatos de arquivo, ressalta-se a importância de se manter dois objetos digitais em paralelo: o objeto original e o objeto de preservação.

O repositório deverá descrever cada classe de informação do AIP que será preservada, assim como sua implementação. Portanto, será necessário apontar o documento arquivístico digital, sua informação de representação e a PDI; e compreender o relacionamento desses componentes (BRASIL, 2015b). Além disso, o AIP é delimitado e identificado pelas informações de empacotamento, as quais poderão estar presentes como uma estrutura na mídia que contém o AIP ou contida na entidade armazenamento arquivístico. Ressalta-se que as funções internas de delimitação e identificação devem ser precisamente definidas pelo OAIS (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012).

Cada pacote AIP está associado a uma estrutura de informações descritivas chamada de descrição do pacote, que permite ao consumidor localizar e analisar as informações de interesse e solicitar as informações desejadas. A informação necessária para um instrumento de acesso (*access aid*) é chamado de descrição associada (*associated description*).

A descrição do pacote não é requisito à preservação da informação de conteúdo, no entanto é necessária para fornecer visibilidade e acesso aos conteúdos. O conteúdo da descrição do pacote é altamente dependente da estrutura da informação de conteúdo e da PDI que ela descreve, e pode conter várias descrições associadas, conforme a quantidade de instrumentos de acesso utilizados para localizar, visualizar, recuperar ou solicitar a informação de conteúdo associada e a PDI (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). Tal esquematização é apresentada a seguir na “Figura 3 – Descrição do pacote”.



Fonte: (CCSDS, 2012, p. 4-39)



Conforme observado na figura, a descrição do pacote (*package description*) deve conter uma descrição associada (*associated description*) que fornece dados para um instrumento de recuperação (*retrieval aid*). Dessa forma, os usuários autorizados recuperam as informações de conteúdo e a PDI que são especificadas pela descrição do pacote (*package description*).

Esse instrumento de recuperação geralmente faz parte da entidade armazenamento arquivístico. Ele traduz o identificador único do AIP, então atribuído pelo OAIS, para identificar o AIP em um subconjunto de operações e nomes de arquivos de dados que são necessários para recuperar o AIP na entidade armazenamento arquivístico. Após isso, retorna ao consumidor, as informações de conteúdo e as PDI's do AIP solicitado (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012).

A descrição do pacote (*package description*) também pode conter séries de descrições associadas (*associated description*), sendo que cada uma irá conter dados para um ou mais instrumentos de acesso (*access aid*). Há ainda outros dois subtipos adicionais de instrumentos de acesso que são os instrumentos de busca (*finding aid*) e os instrumentos de solicitações (*ordering aid*) (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). Os instrumentos de busca auxiliam o consumidor a localizar a informação de interesse. Já os instrumentos de solicitação auxiliam o consumidor a descobrir o custo de solicitar um pacote AIP, além disso, o consumidor ainda pode especificar as transformações proferidas sobre o AIP antes da disseminação via pacote DIP.

Todo repositório deve gerar DIP's para a comunidade designada e oferecer opções de acesso/entrega dos documentos. É preciso implementar uma política de controle de acesso adequada às necessidades do repositório e dos produtores que ali depositaram seus documentos. Logo, haverá um nível de concessão de acesso a cada AIP, para determinados grupos de usuários, conforme o acordo estabelecido entre produtor e OAIS, que poderá requerer identificação e autenticação dos usuários. Da mesma forma, é preciso manter um registro de falhas no controle de acesso, para identificar casos de acesso negado de forma indevida, o que contribui para avaliar possíveis falhas no sistema de segurança (BRASIL, 2015b).

O controle de acesso é primordial para um RDC-Arq que segue o modelo OAIS, pois, permite verificar os níveis de segurança, bem como, se o usuário está sendo devidamente atendido. Isso requer uma política previamente definida, que considere quaisquer restrições ou condições para uso das informações de conteúdo.

O RDC-Arq deve demonstrar que o DIP gerado corresponde a solicitação do usuário, ou seja, as informações são completas e precisas, e são entregues nos formatos de arquivo desejados. Da mesma forma, deve-se demonstrar que todos os pedidos de acesso obtiveram uma resposta, seja aceitação ou rejeição. E caso aceitos, o repositório deve fornecer mecanismos para verificar a autenticidade dos DIP's, e demonstrar a transformação do AIP no DIP correspondente (BRASIL, 2015b).

## 5 ESPECIALIZAÇÃO DO AIP E DA DESCRIÇÃO DE PACOTE

O pacote AIP possui duas especializações: a Unidade de Arquivamento de Informação (*Archival Information Unit – AIU*) e a Coleção de Arquivamento de Informações (*Archive Information Collection – AIC*). Assim, a AIU é utilizada para preservar um único objeto de informação de conteúdo, que não está decomposto em outros AIP's. Já a AIC organiza um conjunto de AIP's (AIU's e outros AIC's) em uma hierarquia temática, que proporcione acesso flexível e eficiente aos consumidores.

Todos os AIP's organizados por um AIC estão contidos nas suas informações de conteúdo. A diferença básica entre AIU's e AIC's está na complexidade da sua informação de

conteúdo e de suas respectivas descrições de pacotes e informações de empacotamento associadas. O modelo OAIS considera as diferenças, entre AIU e AIC, referentes à informação de conteúdo e as funcionalidades de empacotamento e descrição associadas, as quais devem ser adequadamente complexas e interligadas para justificar a definição de classes separadas (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012).

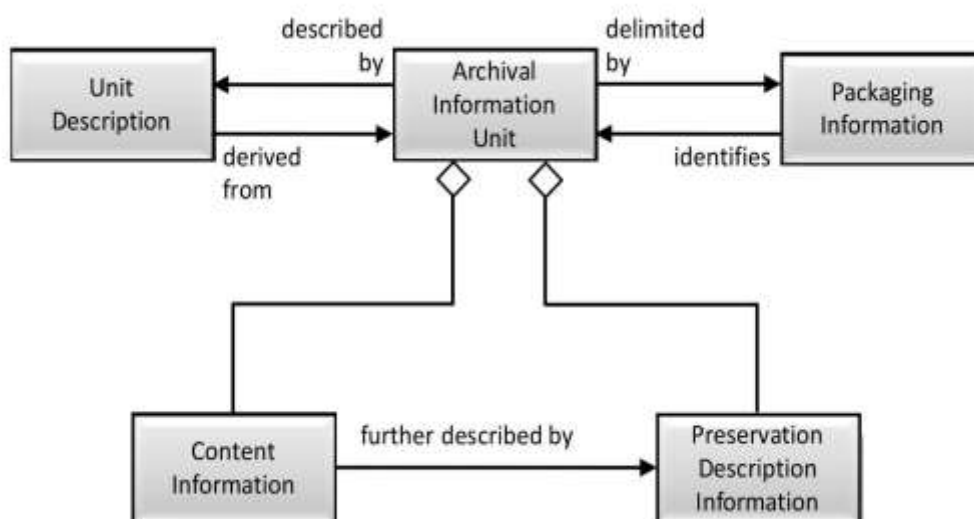
Com relação à preservação da informação, há uma clara distinção entre AIU e AIC. Observa-se que a AIU é composta por um único objeto de informação de conteúdo, que é descrito por, exatamente, um conjunto de PDI. Já a informação de conteúdo de uma AIC é composta por uma coleção de outros AIC's e AIU's, e cada um possui a sua própria PDI. Além disso, a AIC tem a sua própria PDI que descreve os critérios e os processos da coleção (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). Como consequência da especialização do AIP têm-se duas especializações da descrição do pacote: a descrição da unidade e a descrição da coleção.

### 5.1 Unidade de Informação de Arquivamento

Uma única AIU contém exatamente um objeto de informação de conteúdo (que pode consistir em múltiplos arquivos de dados) e exatamente um conjunto de PDI. Salienta-se que o OAIS é livre para decidir como construir a AIU, e que uma AIU não precisa, necessariamente, ser um único arquivo de dados.

Quando um objeto de informação é admitido no OAIS, uma descrição de unidade (*unit description*, subtipo de descrição de pacote) é criada ao extrair informações a partir da informação de conteúdo (*content information*), da PDI (*preservation description information*) e por meio da adição de informações específicas. Além disso, a informação de empacotamento (*information packaging*) irá relacionar logicamente os componentes da AIU (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). A seguir a “Figura 4 – Unidade de informação de arquivamento” esquematiza as relações da AIU.

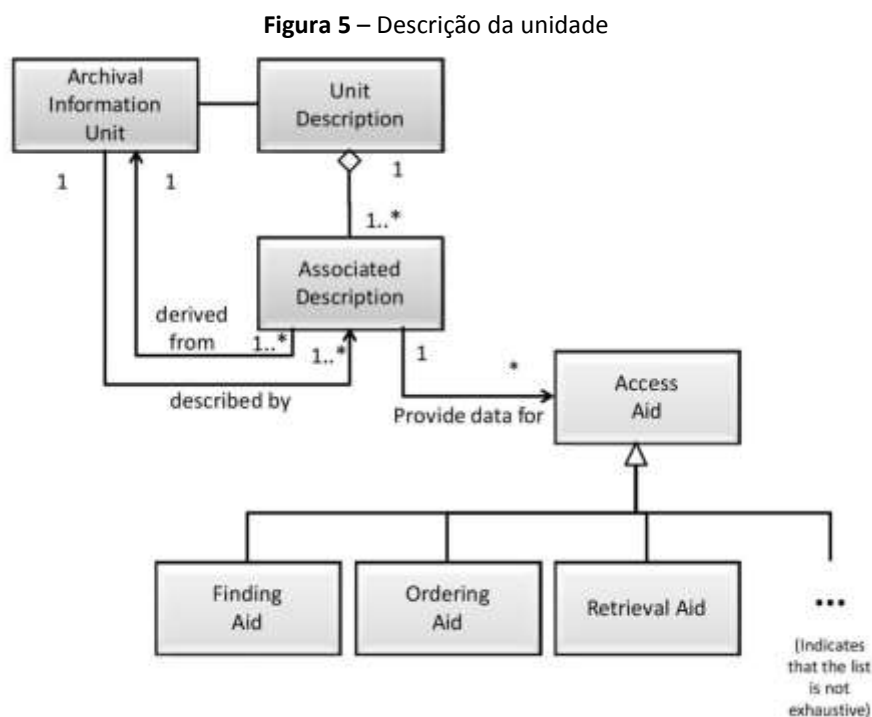
Figura 4 – Unidade de informação de arquivamento



Fonte: (CCSDS, 2012, p. 4-43)

Conforme a figura, a AIU comporta: o documento digital, os metadados de representação, os metadados de preservação, a informação de empacotamento e a descrição da unidade. Observa-se que a descrição da unidade é a principal diferença apresentada, é que sua estrutura será destinada especificamente para descrever uma AIP.

A descrição de unidade é uma especialização da descrição do pacote que conterà um conjunto de descrições associadas, assim, cada uma irá descrever as informações de conteúdo da AIU do ponto de vista de um único instrumento de acesso (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). A seguir, o conteúdo de descrição da unidade é apresentado na “Figura 5 – Descrição da unidade”.



Fonte: (CCSDS, 2012, p. 4-44)

Todas as descrições de unidade (*unit description*) devem fornecer uma descrição associada (*associated description*) para um instrumento de recuperação (*retrieval aid*), o qual permitirá que os usuários autorizados recuperem AIU detalhada pela descrição de unidade (*unit description*) da entidade armazenamento arquivístico (*archival storage*). Essa descrição inclui o identificador único que a entidade armazenamento arquivístico atribui ao AIP durante o processo de admissão (*ingest*) (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012).

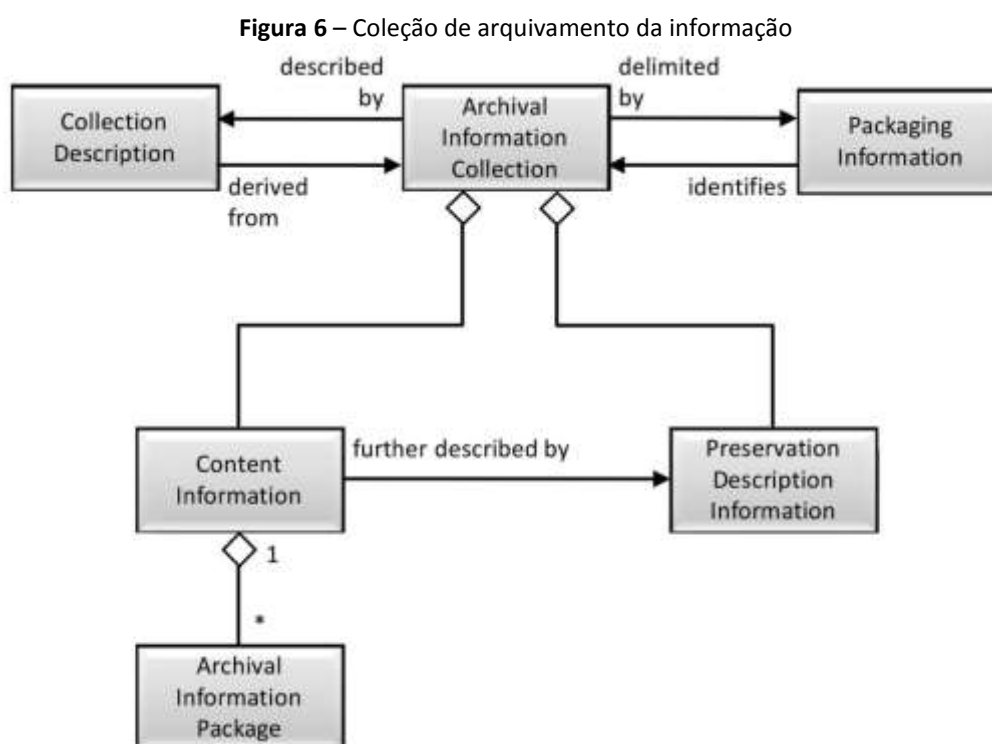
Uma única AIU pode ter diversas descrições associadas, que utilizam diferentes tecnologias, para descrever as informações de conteúdo. À medida que surgem novas tecnologias para extração de descrição e de apresentação, o OAIS pode requerer a atualização da descrição de unidade associada a cada uma de suas AIU's, e criar assim, novas descrições associadas que utilizam essas tecnologias (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). Destaca-se que um tipo importante de instrumento de acesso (*access aid*) para estes conteúdos é instrumento de busca (*finding aid*), um aplicativo que auxilia o consumidor na localização de informações de interesse; aumentando o processo de precisão da informação.

A descrição associada (*associated description*) também fornece dados aos instrumentos de solicitação (*ordering aid*), de modo que o consumidor possa conhecer os custos e solicitar as AIU's de interesse (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012).

Os instrumentos de solicitação (*ordering aid*) também permitem que os usuários especifiquem as transformações a serem aplicadas às AIU's antes da disseminação. Ressalta-se que essas transformações podem se aplicar ao objeto de dados, e até mesmo envolver a PDI da AIU.

## 5.2 Coleções de Informação de Arquivamento

A informação de conteúdo de uma AIC é composta por AIP completos, cada um contendo a sua própria informação de conteúdo, PDI, informação de empacotamento associada e descrição do pacote. Esses AIP's são agregados em AIC's considerando os critérios definidos pelo arquivista. Geralmente as AIC's são baseadas no AIU's com temas ou origens comuns e um conjunto comum de descrições associados (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). A seguir a "Figura 6 – Coleção de informação de arquivamento" apresenta a estrutura de uma AIC.



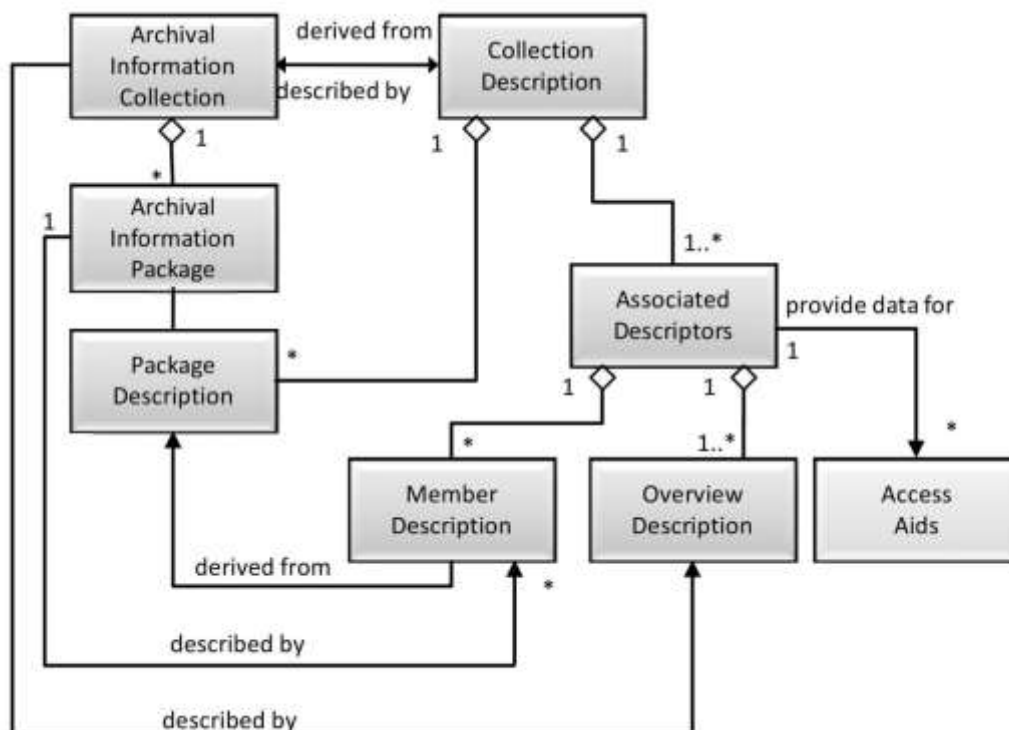
Fonte: (CCSDS, 2012, p. 4-46)

Vários AIP's podem ser reunidos, e caso possuam a mesma PDI, esta será enviada uma única vez, de modo que sejam registradas as suas relações com os AIP's correspondentes. Dessa forma, poderá ocorrer o agrupamento de AIP's em uma AIC conforme suas semelhanças sejam elas, temáticas ou em seus componentes de PDI. Por consequência, os AIP's de uma mesma AIC possuirão descrições semelhantes, fato que justifica o seu agrupamento.

A descrição da coleção (*collection description*) é um subtipo da descrição do pacote (*package description*), que contém estruturas para trabalhar melhor com a informação de conteúdo complexa de uma AIC. Existem dois tipos de descritores associados (*associated descriptors*) em uma descrição de coleção (*collection description*): uma descrição geral (*overview description*) que descreve a coleção como um todo; e zero ou mais descrições de membro (*member description*) que descreva separadamente cada membro da coleção

(ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). A seguir, as classes da descrição da coleção são apresentadas na “Figura 7 – Descrição da coleção”.

Figura 7 – Descrição da coleção



Fonte: (CCSDS, 2012, p. 4-47)

Os descritores associados (*associated descriptors*), necessários em uma descrição de coleção (*collection description*), fornecem informações para instrumentos de solicitação (*access aids*). Isso permite que o usuário acesse todo o conjunto de informações de conteúdo e PDI da AIC associada, mas não necessariamente os AIP's individuais contidos na AIC. A descrição da coleção (*collection description*) pode conter as descrições de pacotes AIP's contidos na AIC. Logo, a AIC pode incluir as descrições de pacotes (*package description*) dos pacotes de informação membro. Essa lista das descrições dos pacotes (*package description*) para AIP contidos em uma AIC pode fornecer instrumentos de acesso (*access aids*) para recuperar ou solicitar membros individuais da AIC (ABNT/NBR 15472:2007; CCSDS, 2012; ISO 14721:2012). Assim, é possível implementar instrumentos de acesso alternativos, onde o consumidor poderá solicitar AIP's de interesse que estejam contidos em uma AIC.

## 6 CONSIDERAÇÕES FINAIS

Este estudo teve a finalidade de analisar o modelo lógico da informação esquematizado pelo OAIS, para apontar as suas possíveis convergências com a Arquivística. Para tanto, verificou-se a estrutura do AIP, para então determinar sua capacidade de preservar documentos arquivísticos digitais em longo prazo, além de garantir a sua autenticidade.

Inicialmente, observou-se que o OAIS utiliza três pacotes de informação para os processos de submissão, arquivamento e disseminação de documentos, que são respectivamente, SIP, AIP e DIP. Desta forma, o documento e seus metadados podem ser

estruturados de formas diferentes, embora a informação de conteúdo que se deseja preservar seja a mesma. Tal fato se justifica, pois os pacotes SIP, AIP e DIP possuem finalidades e requisitos diferentes.

Para fins deste estudo, observaram-se as peculiaridades do pacote AIP, o qual apresenta uma estrutura satisfatoriamente confiável para preservar documentos arquivísticos digitais de forma confiável. Um componente fundamental do AIP que agrega confiabilidade ao processo de preservação é a PDI, subdividida em: informação de referência, informação de proveniência, informação contexto, informação de fixidez e informação de direitos de acesso.

Os componentes da PDI possibilitam manter aspectos fundamentais à preservação de documentos arquivísticos digitais, dentre estes, podem ser citados: a precisão no processo de busca, a linha de custódia documental, a organicidade, a integridade, e as possíveis restrições para preservação ou acesso ao conteúdo. Observa-se que a PDI é um elemento obrigatório apenas no AIP, fato que se justifica pela sua finalidade de preservação e manutenção da autenticidade. No entanto, isso não impede que a PDI seja submetida pelo produtor, no SIP, de forma facultativa; ou mesmo, seja solicitada pelo consumidor e entregue como parte do DIP.

Nota-se ainda, que o AIP possui uma especialização, subdividindo-se em AIU e AIC. Desta forma, os AIP's podem ser organizados individualmente ou em coleções que seguem um determinado aspecto em comum. Uma AIC ajuda a reduzir as informações duplicadas em um RDC-Arq, por exemplo, dois ou mais AIP's que possuam a mesma PDI podem ser reunidos para manter apenas uma versão da PDI.

Por fim, após descrever e avaliar a estrutura do pacote AIP observa-se a sua conformidade com a Arquivística, de modo que possibilita a preservação de documentos digitais em longo prazo com garantia de autenticidade, manutenção da proveniência, da organicidade, e do histórico de custódia. Desta forma, o presente estudo reafirma a necessidade de implementar RDC-Arq's em conformidade com o OAIS, tendo em vista a sua adaptabilidade aos documentos arquivísticos. A principal contribuição consiste em fornecer uma descrição suficientemente clara do pacote AIP, bem como os apontamentos relacionados à Arquivística, a fim de auxiliar pesquisadores recém-iniciados na sua compreensão.

## REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). **NBR 15472:2007**. Sistemas espaciais de dados e informações – Modelo de referência para um sistema aberto de arquivamento de informação (SAAI). Rio de Janeiro, 2007.

BRASIL. CONSELHO NACIONAL DE ARQUIVOS. Câmara Técnica de Documentos Eletrônicos. **Cenários de uso de RDC-Arq em conjunto com o SIGAD**. Rio de Janeiro: Arquivo Nacional, 2015a. Disponível em: [http://conarq.arquivonacional.gov.br/images/ctde/Orientacoes/Orientacao\\_tecnica\\_rdcarg\\_2\\_015\\_v8\\_pub.pdf](http://conarq.arquivonacional.gov.br/images/ctde/Orientacoes/Orientacao_tecnica_rdcarg_2_015_v8_pub.pdf). Acesso: 04 abr. 2020.

BRASIL. CONSELHO NACIONAL DE ARQUIVOS. Câmara Técnica de Documentos Eletrônicos. **Diretrizes para a implementação de repositórios arquivísticos digitais confiáveis – RDC-Arq**. Rio de Janeiro: Arquivo Nacional, 2015b. Disponível em: [http://www.conarq.gov.br/images/publicacoes\\_textos/diretrizes\\_rdc\\_arg.pdf](http://www.conarq.gov.br/images/publicacoes_textos/diretrizes_rdc_arg.pdf). Acesso: 04 abr. 2020.

CLOONAN, M. V. Preservando documentos de valor permanente. In: EASTWOOD, T.; MACNEIL, H. (Org.). **Correntes atuais do pensamento arquivístico**. Belo Horizonte: Editora UFMG, 2016, p.107-134.

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEM. **Reference model for an Open Archival Information System (OAIS)**. Washington: Management Council of the Consultative Committee for Space Data Systems, 2012. Disponível em: <https://public.ccsds.org/pubs/650x0m2.pdf>. Acesso: 04 abr. 2020.

GIL, A. C. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.

HEDSTROM, M. **Digital Preservation: Problems and Prospects**. University of Michigan: USA, 2001. Disponível em: [http://www.dl.slis.tsukuba.ac.jp/DLjournal/No\\_20/1-hedstrom/1-hedstrom.html](http://www.dl.slis.tsukuba.ac.jp/DLjournal/No_20/1-hedstrom/1-hedstrom.html). Acesso: 04 abr. 2020.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). **ISO 14721:2012**. Space data and information transfer systems: open archival information system – Reference model. ISO: Genebra, 2012.

LUNA, S. V. D. **Planejamento de pesquisa**: uma introdução. São Paulo: Educ, 1997.

MÁRDERO ARELLANO, M. Á. 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>. Acesso: 04 abr. 2020.

ROCHA, C. L. Repositórios para a preservação de documentos arquivísticos digitais. **Acervo**, Rio de Janeiro, v.28, n.2, p.180-191, 2015. Disponível em: <http://revista.arquivonacional.gov.br/index.php/revistaacervo/article/view/608/669>. Acesso: 04 abr. 2020.

SANTOS, H. M.; FLORES, D. Novos rumos da preservação digital: das estratégias aos sistemas informatizados. **Biblios: Journal of Librarianship and Information Science**, Brasília, n.70, p.31-43, 2018. Disponível em: <http://dx.doi.org/10.5195/biblios.2017.326>. Acesso: 04 abr. 2020.

SANTOS, V. B. Preservação de documentos arquivísticos digitais. **Ciência da Informação**, Brasília, v.41 n.1, p.114-126, jan./abr., 2012. Disponível em: <http://revista.ibict.br/ciinf/article/view/1357>. Acesso: 04 abr. 2020.

SARAMAGO, M. L. Metadados para preservação digital e aplicação do modelo OAIS. In: CONGRESSO NACIONAL DE BIBLIOTECÁRIOS, ARQUIVISTAS E DOCUMENTALISTAS, 8., Estoril, 2004. **Anais...** Lisboa: Cadernos BAD, 2004. Disponível em: <http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/640>. Acesso: 04 abr. 2020.

SAYÃO, L. F. 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://www.portalseer.ufba.br/index.php/revistaici/article/view/4709>. Acesso: 04 abr. 2020.

SILVA, M. **O arquivo e o lugar**: custódia arquivística e a responsabilidade pela proteção aos arquivos. Niterói: Eduff, 2017.

SILVA, E. L.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. 4. ed. rev. atual. Florianópolis: UFSC, 2005. Disponível em: [https://projetos.inf.ufsc.br/arquivos/Metodologia de pesquisa e elaboracao de teses e dissertacoes\\_4ed.pdf](https://projetos.inf.ufsc.br/arquivos/Metodologia_de_pesquisa_e_elaboracao_de_teses_e_dissertacoes_4ed.pdf). Acesso: 04 abr. 2020.

THOMAZ, K. P. Repositórios digitais confiáveis e certificação. **Arquivística.net**, Rio de Janeiro, v.3, n.1, p.80-89, jan./jun., 2007. Disponível em: [http://www.brapci.inf.br/repositorio/2010/05/pdf\\_fed0720dbb\\_0010726.pdf](http://www.brapci.inf.br/repositorio/2010/05/pdf_fed0720dbb_0010726.pdf). Acesso: 04 abr. 2020.

THOMAZ, K. P.; SOARES, A. J. A preservação digital e o modelo de referência Open Archival Information System (OAIS). **DataGramZero**, v. 5, n. 1, 2004, Rio de Janeiro. Disponível em: <http://www.brapci.inf.br/index.php/res/download/45229>. Acesso: 04 abr. 2020.

VOLPATO, G. L.; BARRETO, R. E.; UENO, H. M.; VOLPATO, E. D. S. N.; GIAQUINTO, P. C.; FREITAS, E. G. D. **Dicionário crítico para redação científica**. Botucatu: Best Writing, 2013.

---

Artigo recebido em 21/06/2019 e aceito para publicação em 03/04/2020

---