

USO DO PROTOCOLO DE INTEROPERABILIDADE OAI-ORE PARA PRESERVAÇÃO DIGITAL

USE OF THE OAI-ORE INTEROPERABILITY PROTOCOL FOR DIGITAL PRESERVATION



Resumo

Introdução: A preservação digital em Bibliotecas Digitais e Repositórios, no Brasil, tem sido um desafio, na medida em que grande parte desses sistemas estão implementados com o software livre DSpace, que não está preparado para atuar nesta atividade. **Objetivo:** Assim, o presente trabalho tem por objetivo apresentar uma nova opção para a preservação digital de documentos depositados em sistemas de informação implementados com o DSpace, com o uso do modelo Hipátia. **Metodologia:** Para tanto, o estudo faz uso de técnicas de pesquisas aplicadas em computação, voltados para inovação e resolução de problemas. **Resultados:** Os resultados apresentam os ajustes necessários, no DSpace, para utilizar o protocolo nativo *Open Archives Initiative - Object Reuse and Exchange* (OAI-ORE) para integração com o Modelo Hipátia. **Conclusão:** Com isso, contribui com a integração de ferramentas ao modelo Hipátia, voltado à preservação digital, de forma simples, utilizando um protocolo nativo do DSpace.

Palavras-chave: DSpace. Modelo Hipátia. OAI-ORE. Interoperabilidade. Preservação digital.

Abstract

Introduction: Digital preservation in Digital Libraries and Repositories in Brazil has been a challenge, as most of these systems are implemented with the free software DSpace, which is not prepared to act in this activity. **Objective:** Thus, the present work aims to present a new option for the digital preservation of documents deposited in information systems implemented with DSpace, using the Hypatia model. **Methodology:** To this end, the study makes use of applied research techniques in computing, returns for innovation and problem solving. **Results:** The results showed the necessary configurations, in DSpace, to use the native Open Archives Initiative - Object Reuse and Exchange (OAI-ORE) protocol for integration with the Hypatia Model. **Conclusion:** With this, it contributes to the integration of tools to the Hypatia model, aimed at digital preservation, in a simple way, using a native DSpace protocol.

Keywords: DSpace. Model Hypatia. OAI-ORE. Interoperability. Digital preservation.

 **Milton Shintaku**

Instituto Brasileiro de Informação em
Ciência e Tecnologia
E-mail: shintaku@ibict.br
Brasília – DF – Brasil

 **Lucas Ângelo Silveira**

Instituto Brasileiro de Informação em
Ciência e Tecnologia
E-mail: lucasangelo@ibict.br
Brasília – DF – Brasil

 **Alexandre Faria de Oliveira**

Instituto Brasileiro de Informação em
Ciência e Tecnologia
E-mail: alexandreoliveira@ibict.br
Brasília – DF – Brasil

 **Diego José Macedo**

Instituto Brasileiro de Informação em
Ciência e Tecnologia
E-mail: diegomacedo@ibict.br
Brasília – DF – Brasil

 **Ingrid Torres Schiessl**

Instituto Brasileiro de Informação em
Ciência e Tecnologia
E-mail: ingridschiessl@ibict.br
Brasília – DF – Brasil



LICENÇA DE USO

Os autores cedem à [Revista Brasileira de Preservação Digital](#) os direitos exclusivos de primeira publicação, com o trabalho simultaneamente licenciado sob a Licença Creative Commons Attribution (CC BY) 4.0 International. Esta licença permite que terceiros remixem, adaptem e criem a partir do trabalho publicado, atribuindo o devido crédito de autoria e publicação inicial neste periódico. Os autores têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada neste periódico (ex.: publicar em repositório institucional, em site pessoal, publicar uma tradução, ou como capítulo de livro), com reconhecimento de autoria e publicação inicial neste periódico.

PUBLISHERS

Universidade Estadual de Campinas – Sistema de Bibliotecas / Instituto Brasileiro de Informação em Ciência e Tecnologia – Rede Brasileira de Serviços de Preservação Digital – Cariniana. As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

EDITORES

Gildenir Carolino Santos, Miguel Angel Márdero Arellano.

CREDIT

RECONHECIMENTOS: Não aplicável.

FINANCIAMENTO: Não aplicável.

CONFLITOS DE INTERESSE: Os autores certificam que não têm interesse comercial ou associativo que represente um conflito de interesses em relação ao manuscrito.

APROVAÇÃO ÉTICA: Não aplicável.

DISPONIBILIDADE DE DADOS E MATERIAL: Não aplicável.

CONTRIBUIÇÕES DOS AUTORES: Conceituação, Metodologia: SHINTAKU, M.; SCHIESSL, I. T.; Supervisão: OLIVEIRA, A. F.; Software, Validação, Visualização: SILVEIRA, L. A.; MACÊDO, D. J.; Redação – rascunho original: SHINTAKU, M.; OLIVEIRA, A. F.; SILVEIRA, L. A.; MACÊDO, D. J.; Redação – revisão & edição: SCHIESSL, I. T.

1 INTRODUÇÃO

No Brasil, grande parte das Instituições de Ensino e Pesquisa implementaram Repositórios Digitais para gerenciar a sua produção acadêmica. Com isso, pode oferecer acesso à documentação de primeira fonte, como as teses e dissertações, além de facilitar o acesso à documentação de segunda fonte como os artigos científicos e de eventos, que foram publicados em outros canais. Assim, atua como dois canais de oferta de acesso: como canal de publicização e canal de facilitação de acesso, conforme modelo de comunicação científica proposto por Björk (2007).

A presença significativa de documentação de primeira mão em repositórios pode ser observada nas universidades, onde os Programas de Pós-Graduação (PPG) depositam teses e dissertações defendidas em formato digital nos seus repositórios. Isso confirma o que Shintaku e Vidotti (2016) defendiam sobre o uso de repositórios digitais de instituições de ensino no processo de tornar públicos documentos que não passam pelo processo tradicional de publicação.

O uso de repositórios para disseminar em texto completo as teses e dissertações em formato digital apresenta como solução a orientação da Capes, mas também apresenta certo desafio no que tange a preservação. Para a implantação dos repositórios, grande parte das instituições de ensino utilizam o DSpace, software livre de código aberto, que segundo Lampert (2016) tem objetivo de dar acesso, mas não atende a preservação a longo prazo. Formenton e Gracioso (2020) descrevem o DSpace como sistema de armazenamento, mas que não atua na preservação a longo prazo. Tanto que, Gonçalves *et al* (2022) apresentam solução para preservação de teses e dissertações eletrônicas dispostas em repositório implementado com o DSpace, por meio do uso da tecnologia *Lots of Copies Keep Stuff Safe* (LOCKSS), na Rede Cariniana de preservação digital.

Outra opção de preservação a longo prazo é o uso de Repositórios Arquivísticos Digitais Confiáveis (RDC-Arq), implementados com o software livre Archivematica, como proposto por Gava e Flores (2021). Em parte, o RDC-Arq têm papel importante na preservação da cadeia de custódia, voltada para documentos arquivísticos. Mesmo sem o total consenso, Teses e Dissertações podem ser consideradas como documentos arquivísticos por representarem a conclusão de um processo de capacitação, assim, preservada digitalmente como tal.

Para o uso de RDC-Arq, o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) desenvolveu o Modelo Hipátia para atender as premissas arquivísticas. Braga, Holanda e Pignataro (2022) apresentam o modelo, no qual integram sistemas produtores ou armazenadores de documentos ao RDC-Arq, apresentando-se como flexível para atender a diversas tecnologias dos sistemas fontes de documento. Para tanto, atende a questões arquivísticas e computacionais, de forma a garantir a cadeia de custódia.

Neste contexto, o presente trabalho tem por objetivo apresentar os ajustes no DSpace, de forma a integrá-lo ao modelo Hipátia, como opção de preservação para

repositórios e bibliotecas digitais. Nesse cenário, é importante destacar a possibilidade do uso de RDC-Arq para preservação de documentos oriundos de repositórios ou bibliotecas digitais implementadas com DSpace, com estratégia diferenciada do LOCKSS, em que a gestão da preservação se torna local.

2 REVISÃO DA LITERATURA

O conceito de interoperabilidade nasce com o movimento de Arquivos Abertos em muito voltado para integração de sistemas de informação em forma de ecossistemas, lançado na Convenção de Santa Fé, Novo México, Estados Unidos, em outubro de 1999. Algumas premissas criadas por esse movimento se mantiveram, sendo absorvidas por outros movimentos como o do Acesso Aberto (*Open Access*) e Ciência Aberta (*Open Science*), como o conjunto mínimo de metadados, muito baseados no padrão *Dublin Core* (DC) e a interoperabilidade por meio de protocolos de coleta automática de metadados (*Harvesting*).

No que tange aos protocolos, foram desenvolvidas duas iniciativas o *Open Archives Initiative - Protocol Metadata Harvesting* (OAI-PMH) e o *Open Archives Initiative - Object Reuse and Exchange* (OAI-ORE). Esses dois protocolos foram criados para possibilitar que sistemas de informação, inicialmente com bibliotecas digitais e revistas posteriormente repositórias, possam compartilhar metadados, facilitando, por exemplo, a recuperação de informação.

Van de Sompel e Lagoze (2002) relatam o rápido sucesso da interoperabilidade, por meio do OAI-PMH, com a criação de grandes concentradores de metadados, que possibilitam recuperar documentação digital disposta em bibliotecas digitais e revistas. No Brasil, um dos resultados desse movimento é a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), mantida pelo Ibict. Tanto que, Triska e Café (2001), descrevem as ações do instituto que culminaram na BDTD, ainda no âmbito do projeto Biblioteca Digital Brasileira (BDB), embasado no recente movimento dos Arquivos Abertos.

A estrutura de funcionamento dos protocolos conta com três grandes entidades, como mostra a **Figura 1**. Provedores de dados são as bibliotecas digitais, repositórios, revistas e outros sistemas informatizados que ofertam o metadados (OAI-PMH) ou metadados e objetos digitais (OAI-ORE), para a coleta automática de metadados (*harvesting*) e objetos digitais. Agregadores, por sua vez, coletam dos provedores de dados e possibilitam que outros sistemas possam consumir a partir deles, tendo um papel duplo. Provedores de serviço, por outro lado, só coletam tanto de agregadores, quanto de provedores de dados.

Figura 1. Estrutura de interoperabilidade



Fonte: Os autores embasados na documentação sobre protocolos de interoperabilidade.

Nota-se a possibilidade de criação de camadas aumentando no intuito de otimizar a disseminação da informação. Björk (2007) relata que os repositórios (provedores de dados) são facilitadores de acesso, enquanto, os agregadores são responsáveis por consolidar os metadados de repositórios. Assim, a interoperabilidade possibilita tanto o compartilhamento de metadados quanto a possibilidade de recuperação dos documentos depositados em provedores de dados.

2.1 Open Archives Initiative - Protocol Metadata Harvesting (OAI-PMH)

O protocolo Open Archives Initiative - Protocol Metadata Harvesting (OAI-PMH) foi o primeiro a ser desenvolvido para interoperabilidade, conforme as premissas do movimento dos arquivos abertos (open archives), logo no início da criação, principalmente, das bibliotecas digitais. Tanto que, é o principal protocolo para interoperabilidade, adotado em por vários softwares voltados para a comunicação científica como o DSpace (repositório), Open Journal Systems (revistas), entre tantas outras.

Este protocolo foi desenvolvido e é mantido pela instituição sem fins lucrativos Open Archives Initiative, com suporte ofertado pela biblioteca da Universidade de Cornell. Como em todos os protocolos informáticos, constitui-se de orientações que podem se materializar de diversas formas, em linguagens de programação diversas. Assim, cada instância cria a sua própria versão do protocolo OAI-PMH.

Neste protocolo, a interoperabilidade ocorre por meio da disponibilização e coleta de metadados, sendo um conjunto mínimo baseado no padrão DC. Isso permite a implementação de protocolos OAI-PMH com outros padrões de metadados, contanto que tenham oferecido o DC como base inicial. Isso possibilita que o protocolo atue apenas na camada de dados, que geralmente são armazenados em banco de dados. O protocolo utiliza seis verbos e as respostas são fornecidas em registros no formato Extensible Markup Language (XML), sendo o JavaScript Object Notation

(JSON) o idioma padrão atual para troca de informações na internet. Toda a comunicação na interoperabilidade é realizada em formato XML, com a estrutura de requisição e resposta. O provedor de dados faz a solicitação por meio de um dos verbos e recebe a resposta com os metadados em formato XML. Os verbos do OAI-PMH são:

- **GetRecord:** recupera um determinado registro, passando o identificador deste documento em um padrão metadados;
- **Identify:** recupera a identificação do provedor de dados;
- **ListIdentifiers:** lista todos os identificadores dos registros do provedor de dados;
- **ListMetadataFormats:** recupera todos os padrões de metadados a serem disponibilizados pelo provedor de dados;
- **ListRecords:** apresenta todos os registros do provedor de dados em um padrão de metadados;
- **ListSets:** lista todos os conjuntos de registros (sets) do provedor de dados.

Com o OAI-PMH, expondo unicamente os metadados, caso o usuário queira acessar o arquivo físico será remetido ao provedor de dados original, por meio de um link contido nos metadados. A possibilidade de recuperação do documento é incrementada, na medida que o provedor de dados, agregadores e provedores de dados geralmente são indexados pelos principais motores de busca, como nos casos dos sistemas que utilizam o DSpace, *Open Journal System* (OJS), *Open Monograph Press* (OMP), Omeka, CKAN etc.

2.2 *Open Archives Initiative - Object Reuse and Exchange (OAI-ORE)*

O *Open Archives Initiative - Object Reuse and Exchange* (OAI-ORE) foi a resposta para as necessidades da comunidade científica por um protocolo que oferecesse maiores recursos, principalmente às novas modalidades de representação da informação, que estendesse ao XML padrão, como o *Resource Definition Framework* (RDF). O OAI-ORE responde às requisições em XML-RDF, RDFa e JSON. Com isso, atende às novas necessidades surgidas com a evolução das tecnologias e utilização de novos padrões.

Possivelmente a grande diferenciação entre os protocolos está na questão das relações existentes entre os itens, que no caso do OAI-ORE é contemplado, numa visão mais moderna da comunicação científica. Entretanto, em muitos casos, o OAI-ORE é lembrado como o protocolo em que se compartilha não apenas os metadados, mas também os arquivos, ou seja, todo o item, podendo ter não apenas os metadados nos agregadores e provedores de dados, mas também os documentos digitais, reproduzindo os itens.

Evidentemente, a possibilidade de compartilhar recursos com as suas relações é expor os recursos com toda a sua riqueza das suas relações, facilitando o seu reuso. Assim, o OAI-ORE é baseado no chamado mapa do recurso (*resource map*) no qual revela como um recurso se relaciona com outros formando uma agregação que transcende ao que era tratado como um item (metadados mais objetos digitais).

Para atender ao OAI-ORE, quatro entidades foram estabelecidas:

- Agregação: é uma abstração, identificada por um link;
- Recurso agregado: qualquer recurso ligado a uma agregação, por meio do link;
- Mapa do recurso: um esquema (recurso) que apresenta as agregações por meio de afirmações;
- Proxy: semelhante ao conceito de proxy das redes, no qual apresenta ligações entre os recursos.

Esta organização proposta em rede possibilita a representação organizada dos objetos informacionais, com todas as suas relações. Outro ponto de interesse neste protocolo é a interoperabilidade que inclui os objetos digitais. Ao fazer coleta em um servidor de dados, coleta-se não apenas os metadados, mas também os objetos digitais, ou seja, os itens inteiros.

2.3 Modelo Hipátia

Mardero Arellano (2022) relata os desafios na preservação digital, na qual ainda pouco se sabe sobre o comportamento futuro e as tecnologias a virem a ser utilizadas ante ao rápido desuso de ferramentas, apresentando um alto risco à informação. Com isso, gestores de bibliotecas digitais, repositórios e outros sistemas ainda não apresentam o conhecimento necessário para atender a toda a complexidade da preservação a longo tempo.

Para atender a esse cenário desafiador, o Ibict propõe o modelo de integração de sistemas para preservação nomeado como Hipátia. Braga (2022), idealizador do projeto, relata que o modelo ainda não está totalmente desenvolvido, mas atende às necessidades atuais para promover a integração de sistemas de tramitação ou gestão documental com os de preservação digital. Para tanto, Braga (2022) relata que para atender a todo o processo, o modelo é estruturado da seguinte forma (**Figura 2**):



Fonte: Braga (2022).

O modelo atende a cadeia de custódia na medida em que atende a todas as orientações arquivísticas, por meio do fluxo iniciando na preparação até a disseminação. Todo o fluxo é informatizado de forma a manter os dados no formato original fazendo os processamentos necessários para a preservação, efetuada por meio de ferramentas próprias. A disseminação é uma novidade no modelo e atende a Lei de Acesso à Informação (LAI) de forma ativa, deixando uma cópia dos objetos digitais acessíveis.

Shintaku, Braga e Oliveira (2021) relatam que a informatização está alterando processos seculares de preservação de maneira abrupta. Modelos de preservação física tem evoluído desde a idade média, nas bibliotecas e arquivos, em que técnicas tem sido criada para conservação, restauro e custódia. Entretanto, a adoção cada vez maior do digital, com o seu uso exclusivo na criação e disseminação da informação tem apresentados desafios na preservação.

Para atendimento a necessidade de preservação de documentação de cunho arquivístico, o Ibict apresenta o Modelo Hipátia, uma opção viável para as instituições ou organizações a implementarem sistemas de preservação digital. Ao atender as premissas da arquivologia, na manutenção da cadeia de custódia, em que todo o trâmite do documento entre o sistema de informação e o sistema de preservação ocorre sem a manipulação humana, o modelo Hipátia garante a inviolabilidade do documento.

Entretanto, para que um sistema seja integrado ao modelo Hipátia, requer ajustes, na medida em que tudo ocorre na camada informatizada. Para integrar-se ao Hipátia o sistema deve enviar os metadados e arquivo digital, para que seja criado o pacote a ser depositado no repositório digital confiável. Todo o encapsulamento é feito no modelo, mas o sistema fonte deve enviar toda a informação para que seja efetuada o empacotamento, seguindo as características da preservação.

Por fim, o modelo Hipátia, por ainda ser uma novidade, está em constante desenvolvimento, passando por adaptações evolutivas, de forma a atender as novas necessidades. No entanto, todas as alterações no modelo seguem as premissas da arquivologia, no que se refere a preservação. Entretanto, essas premissas ainda não estão totalmente estabelecidas ante a evolução tecnológica, com surgimento de novos formatos e sistemas de informação.

3 MATERIAIS E MÉTODOS

O Estudo tem características aplicadas com nuances as efetuadas pela ciência da computação, em parte as alinhadas ao que Wazlawick (2009) descreve como estudos para apresentação de algo diferente, estudos voltados para inovações em resoluções de problemas. Em grande parte, o autor defende que essas pesquisas são efetuadas em temas emergentes. Da mesma forma, alinha-se a estudos de caso, na medida em que, conforme os objetivos, a pesquisa se propõe a solucionar um

problema exclusivo de integração do DSpace ao modelo Hipátia, por meio do uso do protocolo OAI-ORE.

Assim, este estudo teve início com a discussão de proposições de integração do DSpace ao modelo Hipátia, voltado a preservação digital, uma vez que, o DSpace não oferta essa funcionalidade. Em seguida, voltada a segunda etapa da pesquisa, é apresentada uma automatização, na qual responde à pergunta de “Como ajustar o DSpace para atendimento ao modelo Hipátia para preservação digital?”, de forma a responder a hipótese de: “Por meio do protocolo OAI-ORE é possível integrar o DSpace ao modelo Hipátia para a preservação digital de metadados e objetos digitais”.

O estudo foi dividido em três etapas:

- 1) Iniciando com a familiarização com o protocolo OAI-ORE, como funciona no âmbito do DSpace. Por ser um protocolo, pode ser implementado de várias formas, conforme as tecnologias implementadas na ferramenta. A forma de interação por meio do protocolo deve ser o resultado dessa etapa.
- 2) Posteriormente, o estudo de como o DSpace deve se comportar para atender ao modelo Hipátia e os requisitos necessários para efetivar a preservação, que deve ser própria de sistemas de informação implementados com o DSpace, como os repositórios e bibliotecas digitais.
- 3) Os resultados das etapas 1 e 2 são aplicados para orientar os ajustes a serem feitos no DSpace para atender ao modelo Hipátia. Esta estrutura de pesquisa encontra alinhamento com tantas outras das ciências tecnológicas, que tem na aplicação os maiores resultados. Assim como nos estudos tecnológicos voltados ao atendimento à Ciência da Informação, o amparo teórico da disciplina orienta o desenvolvimento do estudo.

4 RESULTADOS

Como em todas as pesquisas aplicadas, o presente estudo tem a meta de resolver um problema físico, que neste caso é a possibilidade de ajuste no DSpace para atender ao modelo Hipátia, por meio do uso do protocolo OAI-ORE para preservação dos itens armazenados no repositório. Assim, apresentou um modelo de ajuste do DSpace para atender a necessidade de preservação dos seus itens. Para tanto, os resultados vão ser apresentados conforme a estrutura proposta na metodologia: Protocolo OAI-ORE, Levantamento de Requisitos e Ajustes no DSpace.

4.1 Protocolo OAI-ORE no DSpace

O protocolo OAI-ORE pode ser implementado no DSpace tanto na versão para responder a coleta, quanto para fazer a coleta. Essa versão permite que o DSpace se torne um agregador, com configuração efetuada diretamente na sua interface, por

meio do usuário administrador. Assim, basta configurar para que o DSpace passe a responder a coletas por meio do OAI-ORE.

Tanto o OAI-ORE quanto o OAI-PMH respondem aos mesmos verbos com apenas algumas diferenças. Enquanto o OAI-PMH responde com apenas os metadados em formato *Extensible Markup Language* (XML), o OAI-ORE pode atender em outros formatos e entrega os objetos digitais, ou seja, todo o item. Para atender à coleta, pode-se fazer uso do verbo *getRecord*, que precisa do identificador do item (*identifier*) e em qual padrão de metadados deve responder (*metadataPrefix*).

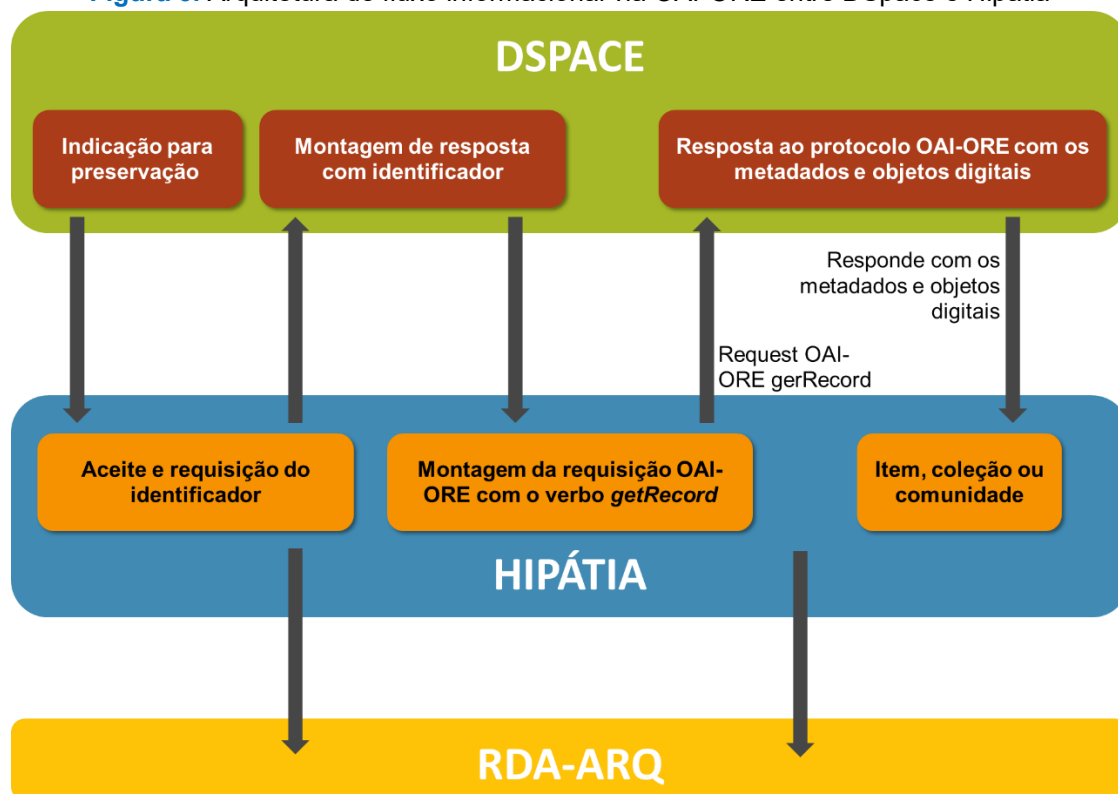
Por meio do protocolo OAI-ORE, o DSpace ao receber a requisição com o verbo *getRecord*, com um identificador válido e a indicação do padrão de metadados, responde com os metadados em vários formatos como o XML e o os objetos digitais. Assim, basta que se indique em que padrão os metadados serão apresentados e, o mais importante, o identificador, do item, coleção ou comunidade.

4.2 Requisitos de integração

A estrutura do DSpace consiste hierarquicamente em Comunidades (grandes temas), Coleções (conjuntos de documentos) e Itens (metadados mais objetos digitais). Para garantir a preservação desses elementos, é necessário ter a opção de solicitar a preservação de uma comunidade e suas subcomunidades, coleções e itens, de uma coleção com todos os seus itens, ou de um único item.

Nesse sentido, como mostra a Figura 3, o processo inicia-se no DSpace, com a indicação do desejo de preservar um determinado elemento, passando para que o processo seja iniciado também no Hipátia. Essa interação é necessária para sincronizar as duas ferramentas para início da transferência dos itens entre eles.

Figura 3. Arquitetura de fluxo informacional via OAI-ORE entre DSpace e Hipátia



Fonte: elaboração dos autores (2023).

Com a indicação que iniciará o processo de preservação, o Hipátia envia a requisição para a DSpace com o pedido do identificador do elemento a ser preservado. Como resposta ao Hipatia, o DSpace envia o identificador, de forma que Hipátia possa enviar a requisição do protocolo OAI-ORE com o verbo *getRecord*. Por fim, o DSpace responde com os itens a serem preservados. Ao final, os itens são preservados no RDC-Arq.

4.3 Ajustes no DSpace

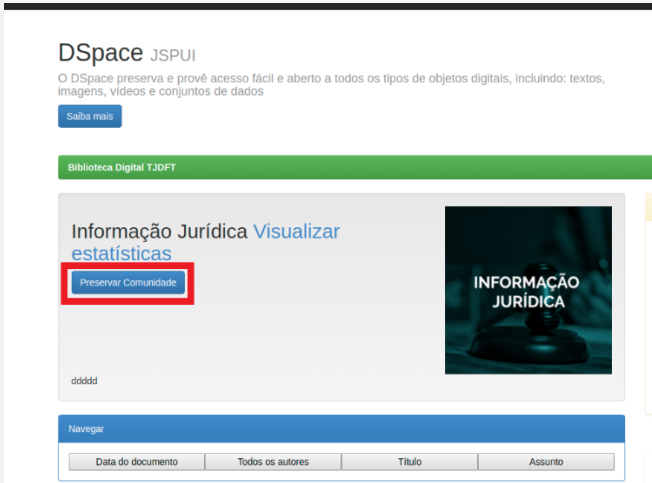
Para implementar a possibilidade de preservação digital, com a integração com o modelo Hipátia, como visto, precisa indicar em que elemento da hierarquia deseja-se preservar. Evidentemente que comunidades e coleções são abstrações sem objetos digitais, composto apenas de metadados. Assim, o que será realmente preservado serão os itens (metadados mais objetos digitais), podendo ser único, ou em conjunto, como nos casos das comunidades e coleções.

4.3.1 Preservação da Comunidade/subcomunidade

Uma comunidade no DSpace representa uma abstração que reúne coleções, ou seja, apenas os metadados descritivos. Entretanto, preservar uma comunidade, significa preservar todos os itens vinculados e não apenas os seus metadados, na medida em que representa uma entidade hierarquicamente em maior nível. Assim, é

preciso incluir um botão que realiza a chamada ao *endpoint* de API para proceder com a preservação, disponível apenas para os administradores (Quadro 1).

Quadro 1. Apresentação do botão de serviço de preservação para comunidade com o código XML gerado a partir dos dados da comunidade

| Página de uma comunidade com o serviço de preservação | XML gerado com os dados da comunidade |
|--|--|
|  | <pre> <community> <link> https://url/jspui/handle/123456789/1 </link> <handle> 123456789/1 </handle> <UUID> 7e245a06-766c-4ece-a7f2-63b29973c5d9 </UUID> <name> Nome da comunidade </name> <introductoryText> Texto introdutório sobre a comunidade </introductoryText> <shortDescription> Breve Descrição da comunidade </shortDescription> <type> community </type> <countItems> Quantidade de itens na comunidade </countItems> </community> </pre> |

Fonte: Elaboração dos autores, com captura de tela da interface do sistema DSpace

Para implementar a funcionalidade deste botão, é preciso alterar o arquivo *community-home.jsp*, de forma que consiga enviar os dados da comunidade, no formato XML, para o Hipátia, dando início ao processo de preservação. Todos os códigos alterados estarão disponíveis no servidor Git do Ibict (GIT..., 2023). Os dados no formato XML ficam disponíveis ao Hipátia pelo DSpace com o acionamento do botão de preservação da comunidade é apresentado no Quadro 1.


Para que o OAI-ORE faça a coleta do item, é necessário unicamente o identificador da comunidade. Entretanto, ao enviar todos os metadados em formato XML, o Hipatia pode validar a coleta, comparando o nome da comunidade e a quantidade de itens recebidos pelo protocolo OAI-ORE com os dados recebidos pelo DSpace alcançando maior confiabilidade do processo. Importa ressaltar, que o OAI-ORE identifica as atualizações realizadas no item, preservando o que é novo, isto é, o que foi modificado ou adicionado ao item, não sendo necessário sobrepor todo arquivo, tal operação é chamada de atualização incremental.

4.3.2 Preservação da Coleção

Uma coleção agrega itens diretamente, ao preservar uma coleção, todos os itens relacionados a essa entidade serão englobados. Da mesma forma que na

comunidade é preciso configurar o serviço no arquivo *collection-home.jsp* com o *endpoint* para iniciar o processo de preservação. O Quadro 2 apresenta a página de uma coleção implementado o serviço de preservação e mostra o XML gerado com as informações da coleção que será exposta ao Hipátia. Assim, ao pressionar o botão de preservar a coleção, na interface de usuário, aciona-se a API que envia ao Hipátia o XML indicando, entre outras informações, a identificação da coleção para que se possa fazer a coleta.

Quadro 2. Apresentação do botão de serviço de preservação para coleção na interface UI e o código XML gerado a partir dos dados da coleção

| Página de uma coleção com o serviço de preservação. | XML gerado com os dados da coleção. |
|---|--|
|  | <pre> <collection> <link> http://url/jspui/handle/123456789/1872 </link> <handle> 123456789/1872 </handle> <UUID> 7f6cd6d5-3834-4a32-974f-a0be265cb021 </UUID> <name> Nome da coleção </name> <introductoryText> Texto introdutório sobre a coleção </introductoryText> <shortDescription> Breve Descrição da coleção </shortDescription> <type> collection </type> <countItems> Quantidade de itens na coleção </countItems> <handleFatherCommunity> handle da comunidade pai </handleFatherCommunity> <nameFatherCommunity> Nome da comunidade pai </nameFatherCommunity> </collection> </pre> |

Fonte: Elaboração dos autores, com captura de tela da interface do sistema DSpace

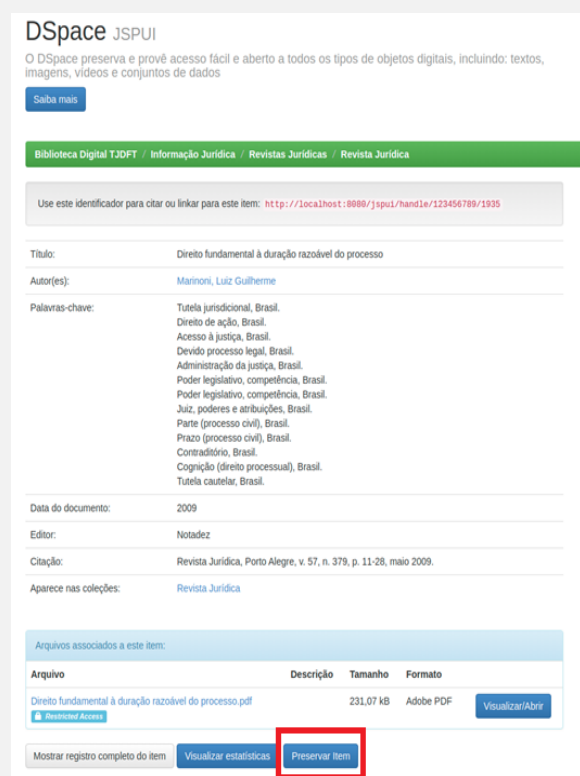
O protocolo OAI-ORE ao receber o identificador da coleção, coleta todos os itens da coleção. Assim, para facilitar a verificação, o XML contém a informação de quantos itens a coleção contém. Assim, ao terminar a coleta, o Hipátia pode verificar o sucesso pela comparação da quantidade de itens coletados, com as declaradas no XML. Em seguida, esses itens são enviados ao RDC-Arq. É importante ressaltar que, em caso de atualização da coleção, o OAI-ORE identifica quais são os novos itens ou modificações, enviando apenas essas alterações para a preservação.

4.3.3 Preservação do Item

Um Item é a unidade informacional do DSpace e consiste em vários campos de metadados descritivos, administrativos, técnicos, de uso aliados aos objetos digitais, que unidos formam uma unidade. Os Itens são depositados nas coleções, que por sua vez, estão contidas nas comunidades/subscomunidades. Dessa forma, o XML gerado quando o serviço de preservação é acionado, terá informações capazes de recuperar a coleção e a comunidade ao qual pertence. O serviço de preservação é implementado via chamada endpoint e deve ser customizado no *arquivo display-item.jsp*.

No Quadro 3 nota-se o botão criado para iniciar a preservação deste item, que ao ser pressionado aciona a API que gera e envia para o Hipátia a informação em formato XML (Quadro 3). Como em todos os casos, o mais importante é a identificação, que possibilita que o verbo GetRecord possa fazer a coleta do item identificado.

Quadro 3. Apresentação do botão de serviço de preservação para item na interface UI e o código XML gerado a partir dos dados do item

| Página de um item com o serviço de preservação | XML gerado com os dados de um item |
|---|---|
|  | <pre> <item> <link> http://url/jspui/handle/123456789/1935 </link> <handle> 123456789/1935 </handle> <UUID> ef979709-8596-4f2e-b506-46a4e57ba46c </UUID> <title> Título do item </title> <type> item </type> <fatherNameCollection> Nome da coleção no qual o item foi depositado </fatherNameCollection> <fatherHandleCollection> Handle da coleção no qual o item foi depositado </fatherHandleCollection> </item> </pre> |

Fonte: Elaboração dos autores, com captura de tela da interface do sistema DSpace

Como o contexto da estrutura hierárquica do item é importante, nos metadados há a informação sobre qual a coleção e comunidade o item pertence. Assim, pode-se ter de verificar em que parte do repositório ou biblioteca digital o item situa, de forma

a possibilitar entender qual a sua classificação temática no âmbito da biblioteca digital ou repositório.

6 CONSIDERAÇÕES FINAIS

Mesmo que seja um estudo de caso centrado no DSpace, os resultados apresentados neste estudo podem ser replicados, com ajustes, para outras ferramentas que fazem uso do protocolo OAI-ORE. Em todos os casos, pela mesma ideia de preservação ofertada pela metodologia do *Lots Of Copies Keep Stuff Safe* (Lockss), no qual a distribuição dos itens (metadados mais os arquivos) possibilita a preservação, o protocolo OAI-ORE pode ser útil para inúmeras casos, entre os quais a coleta de itens para serem enviados para a preservação digital em RDC-Arq.

Por fim, nota-se que o assunto ainda merece estudos com o uso na prática desse modelo, uma vez que foi ajustado apenas para ao DSpace, sendo ainda necessário os estudos no Hipatia. Todas as projeções foram feitas conforme as características apresentadas por Braga e Márdero Arellano (2022) e os diversos autores contidos nessa obra. Assim, o estudo alinha-se ao reuso de tecnologias existentes para resolução de novos problemas.

REFERÊNCIAS

BJÖRK, Bo-Christer. A model of scientific communication as a global distributed information system. **Information Research**, Borås, v. 12, n. 2, 2007. Disponível em: <http://informationr.net/ir/12-2/paper307>. Acesso em: 9 set. 2020.

BRAGA, Tiago Emmanuel Nunes. O modelo Hipátia: a proposta do Ibict para a preservação digital arquivística. In: BRAGA, Tiago Emmanuel Nunes; MÁRDERO ARELLANO, Miguel Ángel (orgs.). **Hipátia: modelo de preservação para repositórios arquivísticos digitais confiáveis**. Informação, Tecnologia e Inovação. Brasília: Ibict, 2022. p. 52–65. DOI [10.22477/9786589167501.cap4](https://doi.org/10.22477/9786589167501.cap4). Acesso em: 11 maio 2023.

BRAGA, Tiago Emmanuel Nunes; HOLANDA, Alex Pereira; PIGNATARO, Tatiana Canelhas. Resolução RDC-Arq Conarq: uma análise dos novos requisitos informacionais propostos. **Revista Brasileira de Preservação Digital**, Campinas, v. 3, p. e022004, 2022. DOI [10.20396/rebpred.v3i00.16583](https://doi.org/10.20396/rebpred.v3i00.16583). Acesso em: 10 maio 2023.

BRAGA, Tiago Emmanuel Nunes; MÁRDERO ARELLANO, Miguel Ángel (org.). **Hipátia: modelo de preservação para repositórios arquivísticos digitais confiáveis**. Brasília: Ibict, 2022

FORMENTON, Danilo; GRACIOSO, Luciana de Souza. Preservação digital. **RDBCi: Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 18, e020012, 8 jun. 2020. DOI [10.20396/rdbci.v18i0.8659259](https://doi.org/10.20396/rdbci.v18i0.8659259). Acesso em: 10 maio 2023.

GAVA, T. B. S.; FLORES, D. O papel do Archivematica no RDC-Arq e possíveis cenários de uso. **ÁGORA: Arquivologia em debate**, Florianópolis, v. 31, n. 63, p. 1–

21, 8 set. 2021. Disponível em: <https://agora.emnuvens.com.br/ra/article/view/1018>. Acesso em: 10 maio 2023.

GIT IBICT. Hipatia-Barramento-DSPaced. Project ID: 327. 2023. Disponível em: <https://git.ibict.br/publico/hipatia-barramento-dspaced.git>. Acesso em: 22 maio 2023.

GONÇALVES, Pâmella Benevides *et al.* Preservação digital distribuída para teses e dissertações: uma proposta para as bibliotecas universitárias da Unesp e UFRN. **Revista Brasileira de Preservação Digital**, Campinas, v. 3, e022006, 2022. DOI [10.20396/rebpred.v3i00.16580](https://doi.org/10.20396/rebpred.v3i00.16580). Acesso em: 10 maio 2023.

LAMPERT, Sérgio Renato. Os repositórios DSpace e Archivematica para documentos arquivísticos digitais. **Acervo**, Rio de Janeiro, v. 29, n. 2, seq. Artigos livres, p. 143–154, 17 nov. 2016. Disponível em: <https://revista.an.gov.br/index.php/revistaacervo/article/view/718>. Acesso em: 10 maio 2023.

MÁRDERO ARELLANO, Miguel Ángel. Preservação de documentos em formato digital. *In*: BRAGA, Tiago Emmanuel N.; MÁRDERO ARELLANO, Miguel Ángel (org.). **Hipátia**: modelo de preservação para repositórios arquivísticos digitais confiáveis. Informação, Tecnologia e Inovação. Brasília: Ibict, 2022. p. 10–25. DOI [10.22477/9786589167501.cap1](https://doi.org/10.22477/9786589167501.cap1). Acesso em: 11 maio 2023.

SHINTAKU, Milton; BRAGA, Tiago Emmanuel Nunes; OLIVEIRA, Alexandre de Faria. Hipátia: uma ferramenta livre no apoio à preservação digital. **Revista Brasileira de Preservação Digital**, Campinas, v. 2, e021006, 30 dez. 2021. <https://doi.org/10.20396/rebpred.v2i00.15987>. Acesso em: 10 maio 2023.

SHINTAKU, Milton; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Bibliotecas e repositórios no processo de publicação digital. **Biblos**, Rio Grande, v. 30, n. 1, 2016. Disponível em: <https://periodicos.furg.br/biblos/article/view/5762>. Acesso em: 13 abr. 2020.

TRISKA, Ricardo; CAFÉ, Lígia. Arquivos abertos: subprojeto da Biblioteca Digital Brasileira. **Ciência da Informação**, Brasília, v. 30, n. 3, seq. Artigos, 13 dez. 2001. DOI [10.18225/ci.inf.v30i3.917](https://doi.org/10.18225/ci.inf.v30i3.917). Acesso em: 11 maio 2023.

VAN DE SOMPEL, Herbert; LAGOZE, Carl. Notes from the interoperability front: a progress report on the Open Archives initiative. *In*: 6TH EUROPEAN CONFERENCE OF DIGITAL LIBRARIES, 6., 2002., collection-title: Lecture Notes in Computer Science DOI: 10.1007/3-540-45747-X_11. **Proceedings** [...]. Roma: Springer Berlin Heidelberg, 2002. v. 6, p. 144–157. DOI [10.1007/3-540-45747-X_11](https://doi.org/10.1007/3-540-45747-X_11). Acesso em: 11 maio 2023.

WAZLAWICK, Raul Sidnei. **Metodologia de pesquisa para ciência da computação**. 6ª reimpr. Rio de Janeiro: Elsevier Brasil, 2009.