



EIXO TEMÁTICO:

Organização e Representação da Informação e do Conhecimento

O PROTOCOLO OAI-PMH E OS PROVEDORES DE DADOS E PROVEDORES E SERVIÇOS

THE OAI-PMH PROTOCOL AND THE DATA PROVIDERS AND SERVICE PROVIDERS

Bruna Maria Campos da Cunha¹
Rachel Cristina Vesu Alves²

Resumo: Este trabalho busca entender como se caracterizam os provedores de dados e provedores de serviços do protocolo OAI-PMH nos repositórios digitais, tendo como enfoque apresentar as características que diferenciam um provedor de dados de um provedor de serviços, de forma a identificar o funcionamento básico do protocolo OAI-PMH e apresentar as características e pré-requisitos de cada tipo de provedor. A metodologia configura-se em uma abordagem qualitativa, exploratória e descritiva por meio de pesquisa bibliográfica realizada nas bases de dados *SciELO*, *Scopus* e *Brapci*, bem como nos documentos disponibilizados pela *Open Archives Initiative* sobre o desenvolvimento e implantação do protocolo OAI-PMH. Existem diferenças entre as formas de implementação para que um sistema possa ser um provedor de dados ou um provedor de serviços, porém ambos necessitam fazer uso comum de algumas ferramentas, como a linguagem de marcação XML e o padrão *Dublin Core*, tanto no ambiente *web* como no ambiente bibliográfico. O protocolo é amplamente utilizado pelos repositórios digitais, que expõem e coletam metadados para expandir seus resultados. Conclui-se que a exposição dos metadados por parte dos provedores de dados garante maior visibilidade e participação em comunidade; já aos provedores de serviços, a capacidade de combinar e agregar os metadados torna o sistema mais completo em termos de resultados de busca.

Palavras-chave: OAI-PMH. Provedores de dados. Provedores de serviços. Interoperabilidade. *Harvesting*.

Abstract: This research seeks to understand how are data providers and service providers of the OAI-PMH protocol in digital repositories, with the focus to provide the features that differentiate a data provider to a service provider, to identify the basic operation of the OAI-PMH protocol and present the features and prerequisites for each provider type. The methodology appears in a qualitative, exploratory and descriptive approach by means of bibliographical research held in *SciELO* databases, *Scopus* and *Brapci*, as well as in the

¹ Mestranda do Programa de Pós-Graduação da Universidade Estadual Paulista Júlio de Mesquita Filho (PPGCI/UNESP). E-mail: brunacamposcunha@gmail.com

² Professora do Departamento de Ciência da Informação da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP). E-mail: rachelvesu@gmail.com

documents made available by the Open Archives Initiative about the development and implementation of the OAI-PMH protocol. There are differences between the forms of implementation for a system can be a data provider or a service provider, but they both need to make common use of some tools, like the XML markup language and the Dublin Core standard, both in the environment Web as in the library environment.

Keywords: OAI-PMH. Data providers. Service providers. Interoperability. Harvesting.

1 INTRODUÇÃO

Para desempenhar da melhor forma possível o seu papel de dar acesso às informações, as unidades de informação sempre tiveram a necessidade de manter uma comunicação (SAYÃO, 2007). Essa comunicação entre as bibliotecas e o comum acordo entre elas consegue ampliar os resultados de uma busca ao possibilitar uma consulta em mais de uma unidade e consegue evitar o retrabalho da catalogação de um mesmo material.

A Internet e, em especial, a Web alavancaram a produção de novos tipos de recursos informacionais e o acesso à informação. A ampla utilização das Tecnologias de Informação e Comunicação (TIC) tem contribuído para a criação de um novo ambiente de acesso, promoção e cooperação de conhecimento em escala global (MARCONDES; SAYÃO, 2001).

Assim, a necessidade de comunicação entre diferentes unidades de informação também no meio digital permanece, e para tanto, é necessário que as máquinas também sejam capazes de interpretar e trocar as informações.

Nesse meio, surgem os repositórios digitais, que são plataformas que visam organizar, armazenar e disponibilizar as produções científicas ou temáticas (INSTITUTO BRASILEIRO DE INFORMAÇÃO E TECNOLOGIA, 2018).

Nesse contexto, discutem-se os aspectos sobre os provedores de dados e provedores de serviços na vertente dos repositórios digitais e o uso do protocolo OAI-PMH pra a realização dessa comunicação entre os sistemas informacionais, onde os repositórios provedores de dados expõe seus metadados para coleta e os repositórios provedores de serviços coletam esses metadados através do *harvesting* promovido pelo protocolo OAI-PMH.

O objetivo geral do trabalho é analisar o protocolo OAI-PMH e sua relação com provedores de dados e provedores de serviços.

A metodologia utilizada é de abordagem qualitativa, exploratória e descritiva por meio de pesquisa bibliográfica realizada com base na literatura disponível sobre

o tema, encontrada principalmente nos documentos disponibilizados pela *Open Archives Initiative* (OAI) sobre o desenvolvimento e implantação do protocolo OAI-PMH e nas bases de dados *Scientific Electronic Library Online* (SciELO), *Scopus* e Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (Brapci).

2 INTEROPERABILIDADE E A INICIATIVA DE DADOS ABERTOS (OAI – OPEN ARCHIVES INITIATIVE)

A comunicação e a troca de informações entre as diferentes unidades de informação, principalmente os repositórios digitais, sempre foram vantajosas na medida em que auxiliam na reunião de vários dados de uma pesquisa, expandindo as opções de seus usuários (GARCIA; SUNYE, 2003).

A interoperabilidade surge para manter a comunicação entre ambientes informacionais heterogêneos e as unidades de informação, em especial os repositórios digitais. Assim, a interoperabilidade pode ser definida como um “[...] processo contínuo de assegurar que sistemas, procedimentos e cultura de uma organização sejam gerenciados de modo a maximizar as oportunidades para intercâmbio e o reuso de informações.” (SAYÃO, 2007, p. 39). Ou, ainda, como a capacidade de sistemas de tecnologia de informação e comunicação (TIC) e todos os seus processos suportados trocarem dados e compartilhar informações (ARAÚJO, 2012).

A constante necessidade de comunicação e a falta de um padrão para pesquisas científicas no ambiente *web* levaram à criação da *Open Archives Initiative* (OAI), com o intuito de desenvolver um protocolo para simplificar e unificar as consultas aos ambientes informacionais digitais (GARCIA; SUNYE, 2003). O termo *open* (aberto), no caso do OAI, não significa acesso ilimitado e gratuito a todas as informações contidas nos repositórios, mas para se referir ao protocolo, que é aberto, ou seja, possui seu código aberto e não é proprietário (GARCIA; SUNYE, 2003).

Os primeiros periódicos eletrônicos surgiram na década de 1990, dando o pontapé inicial para o movimento de acesso livre, tido como um dos acontecimentos mais significativos da época, referente à comunicação científica (ALVES, 2008).

O movimento da OAI surgiu a partir da Convenção de Santa Fé, no Novo

México, no ano de 1999, para a discussão de um modelo de interoperabilidade entre bibliotecas e repositórios digitais (ALVES, 2008; GARCIA; SUNYE, 2003).

O resultado dessa reunião foi publicado em fevereiro de 2000, como uma solução mínima para intercambiar os metadados de *e-prints*, posteriormente abrangendo outros tipos de materiais. Os aspectos técnicos discutidos incluíam três pontos principais: um formato para metadados, um protocolo baseado em diretrizes mais antigas e um sistema de identificação (BARRUECO; COLL, 2003).

Ainda no ano de 2000, a *Digital Library Federation* (DLF) e a *Coalition for Networked Information* (CNI), ambas dos Estados Unidos, anunciaram seu suporte à iniciativa, dando início aos comitês e encargos pela coordenação.

Em janeiro de 2001 foi publicada a primeira versão (1.0) do *Open Archives Initiative - Protocol for Metadata Harvesting* (OAI-PMH) com as devidas revisões (BARRUECO; COLL, 2003).

Em 2002, o resultado da reunião de Budapeste foi a revisão do protocolo para interoperabilidade, acarretando na sua segunda versão (2.0), com a revisão de seus princípios para o intercâmbio das informações, melhorando questões de ambiguidade nos intercâmbios (BARRUECO; COLL, 2003). Atualmente o protocolo ainda encontra-se na versão 2.0, porém, muitos estudos por parte da OAI se desenvolvem para o melhoramento deste (OPEN ARCHIVES INITIATIVE, [201-]).

2.1 INTEROPERABILIDADE E NÍVEIS

A interoperabilidade não é a simples troca de informações (dados ou metadados) entre os sistemas, mas também diz respeito a como e quais dados serão compartilhados entre os sistemas (MILLER, 2011; MUCHERONI; SILVA, 2011; SAYÃO, 2007).

De acordo com Sayão e Marcondes (2008), a interoperabilidade pode ser dividida em níveis, os quais tratam do grau de compromisso e esforço por parte das instituições, bibliotecas e repositórios de tornar e manter sistemas interoperáveis. Os autores apresentam 3 níveis de interoperabilidade:

- **Federação:** pode ser considerado o nível mais alto de interoperabilidade. Nesse nível, os sistemas das organizações participantes devem estar em comum acordo e todos devem manter-se atualizados e padronizados de acordo com as normas estabelecidas. Porém, esse nível de interoperabilidade exige grandes esforços dos participantes para manter os sistemas de acordo e, por esta razão, torna-se

um nível mais difícil de ser estabelecido e pouco encontrado;

- **Harvesting:** pela difícil manutenção de grandes federações, a alternativa de manter um bom grau de interoperabilidade parte do conceito de *harvesting* (coleta) trazida pelo protocolo OAI-PMH. Esse conceito baseia-se em uma busca assíncrona dos dados, ou seja, os dados são coletados e armazenados em outro banco de dados onde será realizada a busca para o usuário final. Por tal motivo, as organizações não necessitam estabelecer um comum acordo e podem atualizar seus dados apenas quando necessário;
- **Agregação:** ainda que os sistemas de algumas organizações não estabeleçam nenhum grau de cooperação, um nível básico de comunicação entre os sistemas ainda é possível através da agregação automática de informações. Nesse caso, o sistema apenas coletará informações pública e/ou básicas do outro (SAYÃO; MARCONDES, 2008).

Além dos níveis, os autores também apresentam alguns aspectos nos quais a interoperabilidade precisa se basear para manter um sistema integrado e operante, ou seja, para que os níveis de interoperabilidade sejam atingidos.

- **Interoperabilidade técnica:** diz respeito ao tipo de interoperabilidade que assegura todos os aspectos técnicos dos sistemas (protocolos, padrões para transporte e representações) a serem utilizados, onde todos os sistemas estão em comum acordo, é a interoperabilidade técnica que garantirá que todos estejam utilizando os mesmos padrões e protocolos.
- **Interoperabilidade semântica:** está relacionada à adoção de ferramentas para a representação das informações que serão utilizadas, como por exemplo, guias de boas práticas, esquemas de codificação de valores e de conteúdo de dados.
- **Interoperabilidade política/humana:** Nesse caso, não há dependência sobre como a informação é descrita e disseminada, torna as informações amplamente disponíveis;
- **Interoperabilidade intercomunitária:** Tem enfoque na interdisciplinaridade cada vez mais presente entre as áreas, para um acesso a informações provenientes de organizações e áreas de conhecimentos distintas.
- **Interoperabilidade legal:** considera as implicações legais sobre disponibilizar livremente as informações.
- **Interoperabilidade internacional:** à nível internacional, deve-se levar em

consideração as diferentes normas e padrões, problemas de comunicação e barreira linguística (SAYÃO; MARCONDES, 2008).

2.2 PROTOCOLO OAI-PMH E OS PROVEDORES DE DADOS E PROVEDORES DE SERVIÇOS

Desenvolvido para facilitar a transmissão de informações e difundir conteúdos no ambiente *web*, principalmente conteúdo científico, o protocolo OAI-PMH é amplamente utilizado por sistemas como bibliotecas e repositórios digitais, para ampliar as buscas (GUAJARDO SALINAS, 2010).

Utilizando ferramentas gerais e abertas, como o padrão *Dublin Core* e a linguagem *eXtended Markup Language* (XML), o protocolo funciona, basicamente, através de perguntas e respostas por parte de seus provedores de dados e provedores de serviços, explicados a seguir (BARRUECO; COLL, 2003; GARCIA, SUNYE, 2003; GUAJARDO SALINAS, 2010).

Constituído pelo conceito de *metadata harvesting* (coleta de metadados) e cliente-servidor, a OAI divide os sistemas que utilizam o protocolo OAI-PMH em provedores de dados e provedores de serviços.

De modo geral, um repositório provedor de dados provê facilidades de busca para o acesso aos documentos armazenados. Ele também permitirá a coleta de seus metadados através do *harvesting*, caso utilize o protocolo OAI-PMH. Um provedor de serviços, por sua vez, agrega e dá acesso aos recursos informacionais disponibilizados pelos provedores de dados (MARCONDES; SAYÃO, 2001).

A criação do protocolo OAI-PMH teve como um dos pontos principais a ideia de que a interoperabilidade entre os repositórios “[...] é a chave para o aumento do seu impacto e no seu estabelecimento como uma alternativa viável ao modelo existente de comunicação” (MARTINS; FERREIRA, 2012, p. 434). Vale destacar que não só metadados descritivos podem ser disponíveis para coleta, mas também os próprios recursos, apesar de não ser seu uso mais comum ou uma exigência (CARDOSO JUNIOR, 2007).

Os provedores de dados são os repositórios que expõem metadados de seus recursos para coleta, mas não seus recursos em si. São os provedores de serviços que irão coletar esses metadados expostos e armazená-los numa base de dados centralizada, onde será realizada a busca no contexto do usuário final. O conteúdo do recurso não é, necessariamente, exposto, apenas seus metadados (CARDOSO JUNIOR, 2007).

Segundo Cardoso Junior (2007, p. 33), para que o sistema consiga ser um provedor de serviços, existem três pré-requisitos de infraestrutura técnica, apresentados a seguir.

- Um servidor conectado à Internet: por se tratar de um sistema Web, é pré-requisito básico que o Provedor de Serviço esteja conectado alocado em um servidor Web.
- Um Sistema de Banco de Dados: esse sistema serve justamente para agregarmos os diversos metadados coletados do repositório em uma única base de dados. Vale ressaltar que estas bases de dados, geralmente são Banco de Dados relacionais, mas também podem ser arquivos XML ou arquivos texto.
- Um Ambiente de Programação: este ambiente deve permitir que o Provedor de Serviços faça requisições HTTP para os repositórios, se comunique com a base de dados do repositório e inclua um parser XML, transformando o seu conteúdo em implementações [*Structured Query Language*] SQL para inserção em sua base de dados.

Existem maneiras para testar se um repositório é de fato um provedor de serviços através de uma ferramenta denominada *Repository Explorer*, que verifica se o repositório está em conformidade com as especificações do OAI-PMH e se possui a função *harvesting* instalada. Entretanto, os testes podem ser realizados apenas em sistemas registrados no *site* da OAI (CARDOSO JUNIOR, 2007).

O número real de provedores de dados e provedores de serviços que realizam essas tarefas por meio do protocolo OAI-PMH é desconhecido, uma vez que o cadastro no *site* da iniciativa é opcional (OPEN ARCHIVES INITIATIVE, [201-]).

Constam, no total, apenas trinta sistemas de bibliotecas e repositórios registrados em todo o mundo como provedores de serviços, um número pequeno quando comparado aos provedores de dados.

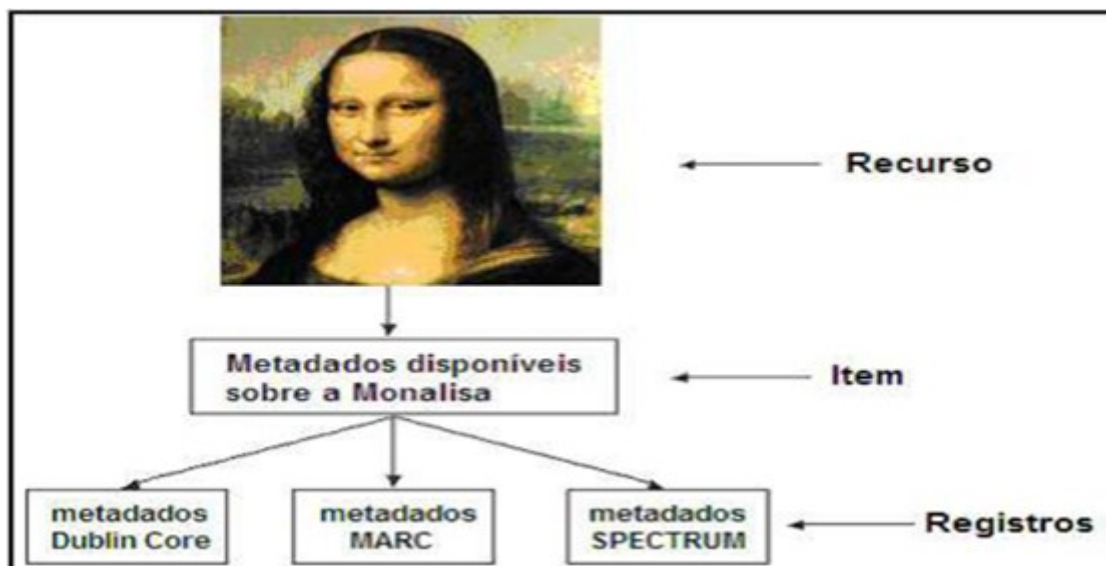
O registro de um sistema como provedor de dados também é opcional e seu número real também é desconhecido, devido ao cadastro opcional no site da iniciativa OAI. Entretanto, o aumento de provedores de dados (repositórios digitais) é perceptível na Web.

Alguns pré-requisitos, como um servidor Web, uma linguagem de marcação (XML), o recurso, item e os registros, também devem ser seguidos para que um repositório possa ser um provedor de dados e desempenhar a função corretamente (CARDOSO JUNIOR, 2007).

Assim, Os pré-requisitos para um provedor de dados, são:

- Servidor web: que é a base para tornar todos os outros possíveis e executáveis;
- Linguagem de marcação: que implemente a exportação dos metadados em formato XML e consiga estabelecer a conexão entre as bases de dados;
- Recurso: o recurso que contém a informação em si, pode ser um texto, imagem, som;
- Registros: São as formas de saída, expressos por diferentes padrões de metadados para a descrição do recurso (CARDOSO JUNIOR, 2007). Devem suportar, no mínimo, o padrão *Dublin Core*;
- Item: A OAI entende como “item” os objetos pertencentes a um repositório e seu registro metadados que poderão ser disseminados através de padrões diferentes, conforme apresentado na Figura 1 a seguir:

Figura 1 - Exemplo de recurso, item e registros



Fonte: Cardoso Junior (2007).

No exemplo apresentado pela Figura 1, o quadro da Mona Lisa é o recurso informacional, o conjunto de metadados que descreve o recurso é considerado o item e esses metadados constituem-se como o registro, que pode ser expresso por diferentes tipos de padrões de metadados: *Dublin Core*, MARC 21, SPECTRUM, entre outros (CARDOSO JUNIOR, 2007).

Além disso, em um repositório, cada recurso deve conter um identificador único, usado para retornar o item quando pesquisado e, no caso da utilização do protocolo OAI-PMH, o identificador único deve se apresentar no seguinte formato padrão:

oai-identifier = oai ":"namespace-identifier ":"local-identifier

Nesse caso, o *namespace-identifier* identifica em que domínio o repositório está localizado, enquanto o *local-identifier* diz respeito ao identificador único do próprio recurso e onde ele está armazenado fisicamente (CARDOSO JUNIOR, 2007).

Segundo Cardoso Junior (2007), os registros de um recurso dentro de um repositório são formados por três partes básicas:

- Cabeçalho (*Header*): o cabeçalho de um item deve conter os seguintes elementos:
 - *Identifier*: o identificador único;
 - *Datestamp*: informa a data de criação ou modificação do registro;
 - *SetSpec*: conjuntos que compõe o repositório. Permite selecionar conjuntos específicos dos repositórios;
- Padrões de metadados. deve suportar no mínimo o padrão *Dublin Core*, utilizado pelo OAI-PMH, identificado como (*oai_dc*);
- *About*: é opcional em um registro e carrega informações sobre os direitos de acesso que podem existir (*rights statement*) ou de quando e de qual repositório os metadados foram coletados (*provenance statement*).

A mesma ferramenta *Repository Explorer* pode ser utilizada para a verificação das funções de um provedor de dados, a qual fará a mesma verificação sobre a conformidade com o protocolo OAI-PMH; porém, nesse caso, o OAI-PMH requer que o provedor de dados forneça uma *Uniform Resource Locator* (URL) básica do provedor instalado, o que também estimula o cadastro dos sistemas na lista de provedores de dados da OAI.

A exposição dos metadados por parte dos provedores de dados garante uma maior visibilidade e a participação em comunidade; aos provedores de serviços, a capacidade de combinar e agregar os metadados torna o sistema mais completo em termos de resultados (OPEN ARCHIVES INITIATIVE, [201-]).

3 PROTOCOLO OAI-PMH E O HARVESTING

O protocolo OAI-PMH surge como alternativa para a integração entre os sistemas, partindo da premissa de que, com foco nos pesquisadores, seria interessante uma única interface de busca que apresentasse resultados de diferentes sistemas (MARCONDES; SAYÃO, 2001).

Assim, o protocolo parte da utilização da coleta automática de metadados, o *harvesting*, utilizando-se dos provedores de dados e provedores de serviços (SAYÃO, 2007).

3.1 HARVESTING

O protocolo OAI-PMH torna-se amplamente utilizado por utilizar-se de ferramentas já conhecidas e de fácil implantação nos sistemas (como a linguagem XML), porém a principal motivação pela sua utilização se dá pelo seu conceito de utilização de *harvesting* de metadados (GARCIA; SUNYE, 2003).

O conceito de *harvesting* não necessariamente exige padrões e normas por parte das instituições. Ele funciona num processo unilateral, onde os provedores de serviços realizam, periodicamente, uma busca aos provedores de dados e coletam os metadados expostos (GARCIA; SUNYE, 2003).

Através do protocolo OAI-PMH, os provedores de serviços “perguntam” aos provedores de dados sobre seus registros, através do *Hypertext Transfer Protocol* (HTTP), e esses “respondem” a partir de uma estrutura em XML (MARTINS; FERREIRA, 2012).

Nessa transmissão, existem seis verbos principais (os quais os provedores de serviços “perguntam” e os provedores de dados “respondem”), apresentados por Barrueco e Coll (2003):

- *GetRecord*: recupera um registro concreto e necessita dos identificadores do registro;
- *Identify*: recupera informações sobre o servidor;
- *ListIdentifiers*: recupera informações dos cabeçalhos dos registros, substituindo os registros completos;
- *ListRecords*: recupera informações dos cabeçalhos dos registros, porém, dos registros completos;
- *ListSets*: recupera um conjunto de registros. Utilizado para facilitar a recuperação

seletiva de registros;

- *ListMetadataFormats*: recupera a lista de formatos bibliográficos utilizados pelo servidor.

Na prática, os sistemas provedores de dados podem expor seus metadados em diferentes formatos, como por exemplo, XML, HTML ou TXT, e um provedor de serviços coleta esses metadados através do *harvesting*, armazena e dá acesso a eles quando solicitado por um usuário (MARCONDES; SAYÃO, 2001).

3.2 CARACTERÍSTICAS DO PROTOCOLO E FERRAMENTAS DE PERGUNTAS E RESPOSTAS

A utilização do conceito de *harvesting* é a principal característica do protocolo OAI-PMH. Sua divisão entre os sistemas em provedores de dados e provedores de serviços permite que sua ferramenta atenda tanto os sistemas que disponibilizam, quanto os que coletam os metadados (OPEN ARCHIVES INITIATIVE [201-]).

O funcionamento geral do *harvesting* pelo protocolo se dá por ferramentas relativamente simples, e o processo de solicitação e resposta por parte dos participantes é descrito a seguir.

A solicitação por parte dos provedores de dados é sempre realizada através do protocolo HTTP. Nesse caso, as solicitações devem ser enviadas em métodos *Get* ou *Post* HTTP, além de uma URL única para todos os pedidos, que especifica o local e a porta de internet (OPEN ARCHIVES INITIATIVE [201-]).

Na solicitação utilizando uma URL *Get* HTTP devem constar palavras-chave anexadas na URL base, separadas por um sinal de interrogação (?). Assim, uma URL de solicitação de *GetRecord*, por exemplo, é apresentada da seguinte maneira:

```
http://an.oa.org/OAI-script?verb=GetRecord&identifier=oai:arXiv.org:hep-th/9901001&metadataPrefix=oai_dc
```

Já na solicitação em *Post* HTTP, as palavras-chave são enviadas no corpo da mensagem HTTP. Utilizando o exemplo acima (*GetRecord*), tem-se, em *Post* HTTP:

```
POST http://an.oa.org/oai-script HTTP/1.0
Content-length: 82
Content-Type: application/x-www-form-urlencoded
verb=GetRecord&identifier=oai%3AarXiv.org%3Ahep-th%2F9901001&metadataPrefix=oai_dc
```

(OPEN ARCHIVES INITIATIVE [201-]).

As respostas às solicitações devem ser respondidas com um HTTP dizendo como está codificado o conteúdo, por exemplo, text/xml, além de identificar se é possível atender à solicitação, ou se o serviço está indisponível (OPEN ARCHIVES INITIATIVE [201-]).

O formato mais comum de resposta é feito em formato XML, como apresentado a seguir, respondendo, de forma bem sucedida, à solicitação *GetRecord*.

```
<? xml version = "1.0" encoding = "UTF-8"?>
<OAI-PMH xmlns = "http://www.openarchives.org/OAI/2.0/"
  xmlns:xsi = "http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation = "http://www.openarchives.org/OAI/2.0/
  http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd ">
<responseDate> 2002-05-01T19: 20: 30Z </ responseDate>
<request verbo = "GetRecord" identificador = "oai: arXiv.org: hep-th / 9901001"
  metadataPrefix = "oai_dc"> http://an.oa.org/OAI-script </ request>
<GetRecord>
<record>
...
</ record>
</ GetRecord>
</ OAI-PMH>
```

(OPEN ARCHIVES INITIATIVE [201-]).

Assim, ao perceber essas diferenças entre os provedores de dados e provedores de serviços, torna-se de grande importância essa divisão trazida por parte da OAI, uma vez que as formas de “perguntas” e “respostas” devem ser realizadas de formas previamente estabelecidas (OPEN ARCHIVES INITIATIVE, [201-]).

Um provedor de serviços, diferentemente de um provedor de dados (que necessita de um ambiente que exponha seus metadados), deve ter uma base de dados e um ambiente que propicie que este realize as *requests* (perguntas) aos provedores de dados, o que torna sua implementação um pouco mais trabalhosa e, conseqüentemente, um menor número de registros na OAI.

4 RELAÇÃO DO OAI-PMH COM OS PROVEDORES DE DADOS E PROVEDORES DE SERVIÇOS

A interoperabilidade é buscada constantemente pelos ambientes informacionais, pois garante uma visibilidade maior dos recursos a partir do momento em que permite que os registros com as representações por metadados sejam encontrados em outros ambientes. Além disso, trazem como vantagens a

diminuição do retrabalho nas representações dos recursos e a ampliação das possibilidades de busca para o usuário final.

O protocolo OAI-PMH tem o intuito de facilitar a troca de informações entre os sistemas, em especial, a coleta e troca de metadados entre os ambientes informacionais do tipo provedores de dados e provedores de serviços. Por ser um protocolo aberto, faz uso de ferramentas como o protocolo HTTP para transferências dos dados, da linguagem de marcação XML e, em sua maior parte, do padrão de metadados *Dublin Core* para representação dos recursos informacionais, tornando-se um protocolo simples e de alcance eficiente (GUAJARDO SALINAS, 2010).

Um importante aspecto do protocolo OAI-PMH é sua estrutura baseada em cliente-servidor, que divide seus participantes em provedores de dados e provedores de serviços e permite a coleta dos metadados (*harvesting*), expandindo ainda mais as possibilidades de resultados.

5 CONSIDERAÇÕES

A troca de informações entre ambientes informacionais traz vantagens tanto para as instituições quanto para seus usuários finais, permitindo que coletem informações em diversos sistemas heterogêneos, evitando o retrabalho, aumentando a visibilidade dos recursos informacionais e possibilitando o reuso dos dados.

O uso do protocolo OAI-PMH em provedores de dados e provedores de serviços possibilita um nível de interoperabilidade de agregação. Embora não rapidamente visíveis, as vantagens de ambos os provedores (dados e serviços) estão ligadas aos resultados para os usuários.

O estudo voltou-se aos provedores de dados e provedores de serviços, apresentando algumas de suas características e os pré-requisitos de cada um, além de abordar a relação entre esses provedores com o conceito de *harvesting*, utilizada pelo protocolo OAI-PMH. Estes provedores necessitam fazer uso comum de algumas ferramentas, como a utilização da linguagem de marcação XML e a utilização do padrão Dublin Core ao trocar os dados.

É evidente que, para um resultado plenamente satisfatório, padrões mais específicos, como por exemplo o MARC, poderiam ser utilizados, porém, ao se tratar de especificidade, as áreas diferem entre si e torna-se difícil a correspondência entre seus padrões. Por tal motivo, a decisão da utilização de um padrão mais abrangente

como o Dublin Core foi necessária, pois, embora exista a perda de especificidade e de valores de metadados, os sistemas, ao expor e coletar os metadados, conseguem realizar uma troca melhor entre si.

Para trabalhos futuros ligados à pesquisa, destaca-se a discussão sobre a interação do protocolo OAI-PMH com outro projeto da *Open Archives Initiative*, o *Open Archives Initiative Object Reuse and Exchange* (OAI-ORE), que possui uma outra abordagem, focando na descrição e na relação entre os recursos e como o OAI-PMH e o OAI-ORE podem melhorar os sistemas informacionais.

REFERÊNCIAS

- ALVES, V. B. A. Open Archives: via verde ou via dourada? **PontodeAcesso**, Salvador, v. 2, n. 2, p. 127-137, 2008. Disponível em: <https://portalseer.ufba.br/index.php/revistaici/article/view/1780>. Acesso em: 10 jul. 2019.
- ARAÚJO, C. J. A. **Um modelo para interoperabilidade entre instituições heterogêneas**. 2012. Tese (Doutorado em Ciência da Computação) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2012. Disponível em: <http://www.teses.usp.br/teses/disponiveis/45/45134/tde-08022013-111002/pt-br.php>. Acesso em: 12 jul. 2019.
- BARRUECO, J. M.; COLL, I. S. Open archives initiative. Protocol for metadata harvesting (OAI-PMH): descripción, funciones y aplicaciones de un protocolo. **EI Profesional de la Información**, v. 12, n. 2, p. 99-106, 2003. Disponível em: <http://eprints.rclis.org/4093/>. Acesso em: 14 jul. 2019.
- CARDOSO JUNIOR, M. J. M. **Clio-i: interoperabilidade entre repositórios digitais utilizando o protocolo OAI-PMH**. 2007. Dissertação (Mestrado em Ciência da Computação) - Centro de Informática, Universidade Federal de Pernambuco, Recife, 2007. Disponível em: <https://repositorio.ufpe.br/handle/123456789/2673>. Acesso em: 10 jul. 2019.
- GARCIA, P. A. B.; SUNYE, M. S. O protocolo OAI-PMH para interoperabilidade em bibliotecas digitais. *In*: CONGRESSO DE TECNOLOGIA PARA GESTÃO DE DADOS E METADADOS DO CONE SUL, 1., 2003, Ponta Grossa. **Anais** [...]. Ponta Grossa: UEPG, 2003.
- GUAJARDO SALINAS, A. Z39.50 y OAI-PMH: protocolos de transferencia y recuperación de información. *In*: CONFERENCIA INTERNACIONAL DE BIBLIOTECOLOGIA, 15., 2010, Santiago. **Anales** [...]. Santiago: Colegio de Bibliotecarios de Chile, 2010.
- INSTITUTO BRASILEIRO DE INFORMAÇÃO E TECNOLOGIA. **Repositórios Digitais**. 2018. Disponível em: <http://www.ibict.br/informacao-para-a-pesquisa/repositorios-digitais>. Acesso em: 20 maio 2019.

MARCONDES, C. H.; SAYÃO, L. F. Integração e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: a proposta da biblioteca digital brasileira. **Ciência da Informação**, Brasília, DF, v. 30, n. 3, p. 24-33, set/dez. 2001. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652001000300004&lang=pt. Acesso em: 25 maio. 2019.

MARTINS, D. L.; FERREIRA, S. M. S. P. Protocolo OAI-PMH e sistemas federados de informação: fundamentos de arquitetura da informação para análise de dados do portal de produção científica da área de ciência da comunicação. **Liinc em revista**, Rio de Janeiro, v. 8, n. 2, p. 431-447, 2012. Disponível em: <http://revista.ibict.br/liinc/%20article/view/479>. Acesso em: 3 maio 2019.

MILLER, S. J. **Metadata for digital collections: a how-to-do-it manual**. New York: Neal-Schuman, 2011.

MUCHERONI, M. L.; SILVA, J. F. M. A interoperabilidade dos sistemas de informação sob o enfoque da análise sintática e semântica de dados na web. **PontodeAcesso**, Salvador, v. 5, n. 1, p. 3-18, 2011. Disponível em: <https://portalseer.ufba.br/index.php/revistaici/article/view/3622>. Acesso em: 12 jul. 2019.

OPEN ARCHIVES INITIATIVE. **Open Archives Initiative Protocol for Metadata Harvesting**. [S. l.: s. n.], [201-]. Disponível em: <http://www.openarchives.org/pmh/>. Acesso em: 29 abr. 2019.

SAYÃO, L. F. Padrões para bibliotecas digitais abertas e interoperáveis. **Encontros Bibli**, Florianópolis, n. esp., p. 18-47, 2007. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2007v12nesp18/436>. Acesso em: 15 jul. 2019.

SAYÃO, L. F.; MARCONDES, C. H. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. **TransInformação**, Campinas, v. 20 n. 2, p. 133-148, maio/ago. 2008. Disponível em: http://www.scielo.br/scielo.php?pid=S0103-37862008000200002&script=sci_abstract&tlng=pt. Acesso em: 13 jul. 2019.