

# O PROTOCOLO OAI-PMH PARA INTEROPERABILIDADE EM BIBLIOTECAS DIGITAIS

## ***Patrícia de Andrade Bueno Garcia***

Departamento de Informática, Universidade Federal do Paraná  
Centro Politécnico - Jardim das Américas - Caixa Postal: 19081  
CEP 81531-990 - Curitiba-PR - Fone: 41-336-1472

[pgarcia@inf.ufpr.br](mailto:pgarcia@inf.ufpr.br) / [pabgarcia@uol.com.br](mailto:pabgarcia@uol.com.br)

Departamento de Sistemas de Informação, Faculdade Cenecista Presidente Kennedy

R. Rui Barbosa, 541 - Centro

CEP 83601-140 - Campo Largo - PR - Fone: 41-392-2211

## ***Marcos Sfair Sunye***

Departamento de Informática, Universidade Federal do Paraná  
Centro Politécnico - Jardim das Américas - Caixa Postal: 19081  
CEP 81531-990 - Curitiba-PR - Fone: 41-361-3410

[sunye@inf.ufpr.br](mailto:sunye@inf.ufpr.br)

## **Resumo**

O presente trabalho apresenta o protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*) da Iniciativa *Open Archives* (OAI), que proporciona interoperabilidade entre bibliotecas/repositórios digitais. Partindo de uma introdução à Iniciativa *Open Archives* e ao processo de *Harvesting*, são explicados os seis verbos que compõem o protocolo e sua implementação em bibliotecas/repositórios digitais novos ou já existentes.

Palavras-chaves: interoperabilidade, arquivos abertos, bibliotecas digitais, metadados, protocolo OAI-PMH.

## **1. Introdução**

Os professores e pesquisadores de países em desenvolvimento, que contam com recursos escassos para a manutenção de suas Universidades e Instituições de Pesquisa, têm dificuldade de acesso à literatura científica tradicional de sua área, sob a forma de periódicos, muitas vezes altamente especializados e com assinaturas caras. Recentemente, com o surgimento dos periódicos eletrônicos, esse problema ainda persiste, pois, muitos só oferecem acesso aos artigos mediante pagamento.

A cada dia surgem bibliotecas digitais e bases de dados públicas que se constituem de importante fonte de informação para pesquisadores. O problema é o tempo que se gasta para reunir informações relevantes a um dado assunto de pesquisa, seja visitando diversos portais de bibliotecas virtuais, seja

utilizando a busca convencional oferecida por *sites* como *Google*, *AltaVista*, *Cadê*, etc. A pesquisa através de portais de busca tradicionais é imprecisa e atinge apenas as páginas HTML, ignorando as bases de dados que se encontram por trás de algumas destas páginas.

Por outro lado, ao tentar divulgar seus trabalhos através dos periódicos, em busca do impacto de suas pesquisas, os pesquisadores esbarram em um processo burocrático e demorado, onde, desde a submissão até a publicação, devido à lenta arbitragem que por vezes ocorre, pode haver um espaço de tempo tão longo que os efeitos daquela pesquisa já não tenham valor quando de sua publicação.

Com a evolução da Internet, várias bibliotecas digitais começaram a surgir, algumas com a finalidade de expor a produção de teses e dissertações das grandes universidades dos Estados Unidos e Europa.

Tudo isso resultou em um grande avanço, onde informações científicas e acadêmicas já poderiam ser obtidas livremente pela Internet, e disponibilizadas através da publicação nas páginas de seus autores.

A falta de padrões para disponibilização e pesquisa de informações científicas na Internet levou à criação da Iniciativa *Open Archives* (Arquivos Abertos) e ao desenvolvimento de um protocolo com o intuito de oferecer simplicidade e eficiência na tarefa de unificar as consultas a bases de dados científicas/acadêmicas. Com os recursos oferecidos pela Iniciativa *Open Archives*, é possível melhorar significativamente a precisão das consultas eletrônicas e reduzir o tempo de procura, graças ao compartilhamento de informações (metadados) entre os participantes da Iniciativa. A interoperabilidade<sup>1</sup> entre bases de dados, dentre elas as bibliotecas digitais, tem o objetivo de promover o acesso simultâneo aos dados contidos nestes repositórios digitais, de forma a maximizar a pesquisa dos mesmos.

O protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*), vem se consolidando como a base para a interoperabilidade entre bibliotecas/repositórios digitais acadêmicos e científicos no mundo todo. Através do OAI-PMH é possível proporcionar visibilidade e integração de informações (metadados), com custos acessíveis à realidade de países em desenvolvimento, como o Brasil.

---

<sup>1</sup> Capacidade de troca de informações

Este trabalho contribui para a compreensão do protocolo e de suas aplicações, proporcionando fundamentos para a identificação de oportunidades de sua utilização em entidades públicas e privadas.

## **2. A Iniciativa Open Archives**

A Iniciativa *Open Archives* (Arquivos Abertos), ou OAI, surgiu a partir da Convenção de Santa Fé (Novo México) no final de 1999, com o intuito de "desenvolver e promover soluções de interoperabilidade que facilitem uma disseminação eficiente do conteúdo" [1]. Os padrões desenvolvidos independem do conteúdo que é disponibilizado, bem como de questões econômicas que possam restringir o acesso aos documentos armazenados.

O termo "*Open*" significa "aberto" do ponto de vista da arquitetura da proposta, não significando acesso gratuito ou ilimitado às informações constantes dos repositórios que fazem parte da Iniciativa. Em resumo, o protocolo da *Open Archives* é que é aberto. O livre acesso a documentos e questões envolvendo cobrança de valores, direitos autorais e outras questões similares, têm sido amplamente discutidas em outras Iniciativas [2, 3].

A base da Iniciativa é o protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*), que faz com que os participantes da Iniciativa possam compartilhar seus metadados. Metadados significam "dados sobre os dados", ou seja, são informações que descrevem o conteúdo dos registros dos repositórios, que por sua vez são documentos eletrônicos (artigos, teses, e outros). Os metadados do protocolo OAI-PMH seguem o padrão Dublin Core [4, 5], um conjunto de 15 elementos. Alguns exemplos são: título, autor, resumo, palavras-chaves.

Os participantes da Iniciativa são divididos em Provedores de Dados (*Data Providers* ou *DP*) e Provedores de Serviços (*Service Providers* ou *SP*). Os provedores de dados mantêm repositórios de documentos digitais que implementam o protocolo OAI-PMH como forma de expor os metadados de seus documentos. Já os provedores de serviços oferecem buscas a estes metadados ou outros serviços que visam agregar valor à Iniciativa.

## **3. O Protocolo OAI-PMH**

A versão inicial do protocolo OAI-PMH foi lançada no primeiro semestre de 2001. A versão atual, 2.0, foi disponibilizada em julho de 2002 e os

provedores de dados da versão 1.0 ou 1.1 têm sido incentivados a migrar para esta versão, sob pena de não serem listados na página da Iniciativa [6] até que implementem a versão atual do protocolo.

O Protocolo OAI-PMH consiste de 6 verbos ou *requests* (*Identify*, *ListIdentifier*, *ListMetadataFormats*, *ListSets*, *ListRecord*, *GetRecord*), e as respostas a todas as requisições estão codificadas em XML [7].

### 3.1. Harvesting

O protocolo OAI-PMH introduz o conceito de *Metadata Harvesting* (colheita de metadados), um processo unilateral (Figura 1), onde, os provedores de serviços, a partir da lista de repositórios (provedores de dados) registrados na OAI, realizam periodicamente uma busca a estes provedores de dados, "colhendo" os metadados para exibição sob a forma de consultas efetuadas pelos usuários.

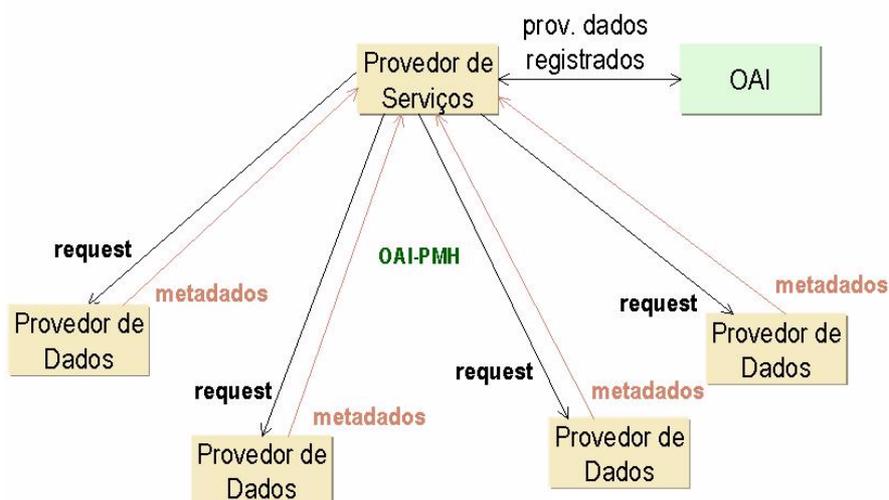


Figura 1 - Metadata Harvesting [8]

A coleta de metadados pode ser total ou baseada em critérios. Os critérios de coleta do *Harvesting* são:

- *Date-based* (baseado em data) - Serão coletados apenas os metadados incluídos/alterados após a data especificada.
- *Set-based* (baseado em conjuntos) – o protocolo define *set* como “uma estrutura opcional para agrupar itens num repositório para o propósito de uma coleta seletiva de registros” [7]. Esta estrutura é uma árvore de assuntos que representa a hierarquia do repositório.

### 3.2. Verbos do OAI-PMH

Os verbos do OAI-PMH são explicados a seguir, com exemplos de solicitações (*requests*) e suas respectivas respostas em XML. Estes exemplos são simulações do *Harvesting*, ou seja, como se a solicitação fosse enviada ao provedor de dados por parte do provedor de serviços. Nestes exemplos, está sendo simulada a atuação do provedor de serviços através de um comando *http* num *browser* (utilizou-se o Internet Explorer). Este processo é transparente para o usuário final, que receberá a informação já coletada e tratada pelo provedor de serviços, de maneira amigável.

O provedor de dados ao qual são submetidas as solicitações é o **UFPR Eprints** [8], repositório digital do Departamento de Informática da Universidade Federal do Paraná (<http://eprints.c3sl.ufpr.br>). O **UFPR Eprints** é um provedor de dados registrado na OAI e em conformidade com o protocolo OAI-PMH 2.0.

### 3.2.1. Identify

Este verbo é utilizado para recuperar as informações sobre o provedor de dados. A resposta mínima esperada consiste em: nome do provedor de dados, endereço do repositório (URL), versão do protocolo implementada e endereço eletrônico (e-mail) do administrador do repositório (Figura 2).

Solicitação: <http://eprints.c3sl.ufpr.br/perl/oai2?verb=Identify>

```
<?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>2003-07-08T07:35:45Z</responseDate>
  <request verb="Identify" resumptionToken="">http://eprints.c3sl.ufpr.br/perl/oai2</request>
  <Identify>
    <repositoryName>UFPR Eprints</repositoryName>
    <baseURL>http://eprints.c3sl.ufpr.br/perl/oai2</baseURL>
    <protocolVersion>2.0</protocolVersion>
    <adminEmail>mailto:admin@eprints.c3sl.ufpr.br</adminEmail>
    <earliestDatestamp>0001-01-01</earliestDatestamp>
    <deletedRecord>persistent</deletedRecord>
    <granularity>YYYY-MM-DD</granularity>
  </description>
  <description>
    <oai-identifier xmlns="http://www.openarchives.org/OAI/2.0/oai-identifier"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai-identifier
      http://www.openarchives.org/OAI/2.0/oai-identifier.xsd">
      <scheme>oai</scheme>
      <repositoryIdentifier>GenericEprints.OAI 2</repositoryIdentifier>
      <delimiter>:</delimiter>
      <sampleIdentifier>oai:GenericEprints.OAI 2:23</sampleIdentifier>
    </oai-identifier>
  </description>
  <description>
    <eprints xmlns="http://www.openarchives.org/OAI/1.1/eprints"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="http://www.openarchives.org/OAI/1.1/eprints
      http://www.openarchives.org/OAI/1.1/eprints.xsd">
      <content>
        <text>OAI Site description has not been configured.</text>
      </content>
      <metadataPolicy>
        <text>No metadata policy defined. This server has not yet been
          fully configured. Please contact the admin for more
        </text>
      </metadataPolicy>
    </eprints>
  </description>
</OAI-PMH>
```

```

information, but if in doubt assume that NO rights at all are
granted to this data.</text>
</metadataPolicy>
=< dataPolicy>
  <text>No data policy defined. This server has not yet been fully
  configured. Please contact the admin for more information, but
  if in doubt assume that NO rights at all are granted to this
  data.</text>
</dataPolicy>
=< submissionPolicy>
  <text>No submission-data policy defined. This server has not yet
  been fully configured.</text>
</submissionPolicy>
<comment>This system is running eprints server software (EPrints 2.2.1
(pepper) [Born on 2002-11-14]) developed at the University of
Southampton. For more information see
http://www.eprints.org/</comment>
</eprints>
</description>
</Identify>
</OAI-PMH>

```

**Figura 2 - Resposta do provedor de dados ao verbo Identify do OAI-PMH**

### 3.2.2. ListMetadataFormats

Este verbo recupera os formatos de metadados disponíveis no repositório (Figura 3). Opcionalmente, pode-se definir o identificador de um registro como argumento, que listará os metadados deste registro em particular. O formato padrão é *Dublin Core* (representado no protocolo por *oai\_dc*).

Solicitação: <http://eprints.c3sl.ufpr.br/perl/oai2?verb=ListMetadataFormats>

```

<?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>2003-07-31T22:52:41Z</responseDate>
  <request verb="ListMetadataFormats"
  resumptionToken="">http://eprints.c3sl.ufpr.br/perl/oai2</request>
  =< ListMetadataFormats>
    =< metadataFormat>
      <metadataPrefix>oai_dc</metadataPrefix>
      <schema>http://www.openarchives.org/OAI/2.0/oai_dc.xsd</schema>
      <metadataNamespace>http://www.openarchives.org/OAI/2.0/oai_dc/</metadata
      Namespace>
    </metadataFormat>
  </ListMetadataFormats>
</OAI-PMH>

```

**Figura 3 - Resposta do provedor de dados ao verbo ListMetadataFormats do OAI-PMH**

### 3.2.3. GetRecord

Recupera um único registro do repositório. Deve-se especificar o formato dos metadados (*metadataPrefix*), cujo padrão é *oai\_dc* (*Dublin Core*) e o identificador do registro (*Identifier*), que é único na comunidade OAI (Figura 4).

Solicitação:

[http://eprints.c3sl.ufpr.br/perl/oai2?verb=GetRecord&metadataPrefix=oai\\_dc&identifier=oai:GenericEPrints.OAI2:49](http://eprints.c3sl.ufpr.br/perl/oai2?verb=GetRecord&metadataPrefix=oai_dc&identifier=oai:GenericEPrints.OAI2:49)

```

<?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>2003-08-27T12:55:43Z</responseDate>

```

```

<request verb="GetRecord" identifier="oai:GenericEPrints.OAI 2:49" metadataPrefix="oai_dc"
  resumptionToken="">http://eprints.c3sl.ufpr.br/per1/oai2</request>
=<GetRecord>
  =<record>
    =<header>
      =<identifier>oai:GenericEPrints.OAI 2:49</identifier>
      <datestamp>2003-08-01</datestamp>
      <setSpec>7374617475733D707562</setSpec>
      <setSpec>7375626A656374733D646174616261736573</setSpec>
      <setSpec>7375626A656374733D696E666F726D6174696F6E5F72657472696576616C
        </setSpec>
    </header>
    =<metadata>
      =<oai_dc:dc xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
        xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
          http://www.openarchives.org/OAI/2.0/oai_dc.xsd"
        xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
        xmlns:dc="http://purl.org/dc/elements/1.1/">
        <dc:title>Provedores de Dados de Baixo Custo: Publicacao Digital ao Alcance
          de Todos</dc:title>
        <dc:creator>Garcia, Patricia A. B.</dc:creator>
        <dc:subject>Databases</dc:subject>
        <dc:subject>Information Retrieval</dc:subject>
        <dc:description>The availability of information by the electronic media has
          been improving the visibility of scientific researches, which was
          restricted to traditional periodicals. Research centers and academic
          databases, and more recently, digital libraries and e-prints archives have
          been contributed to a new scenery that offers fast and free access to
          scientific information. Now, several initiatives encourages the creation of
          open archives and digital libraries. The Open Archives Initiative (OAI)
          promotes the communication standardization between archives through
          OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting),
          improving the interoperability among themselves. There are software
          available at OAI which purpose is to facilitate the implementation of data
          providers (electronic content archives). However, kind of works that,
          through a Computer Science vision, analyses cost, complexity and the
          range of these solution are still rare. This work analyses the presented
          solutions, proposing a criteria for its evaluation and confirming the
          viability of installing a low-cost interoperable digital
          library.</dc:description>
        <dc:date>2003-07-01</dc:date>
        <dc:type>Thesis</dc:type>
        <dc:identifier>http://eprints.c3sl.ufpr.br/archive/00000049/</dc:identifier>
        <dc:format>pdf
          http://eprints.c3sl.ufpr.br/archive/00000049/01/dissertacao_patricia.p
          df</dc:format>
      </oai_dc:dc>
    </metadata>
  </record>
</GetRecord>
</OAI-PMH>

```

**Figura 4 - Resposta do provedor de dados ao verbo GetRecord do OAI-PMH**

### 3.2.4. ListRecords

É o verbo que "colhe" os metadados do repositório. Através de argumentos opcionais pode-se estabelecer a coleta seletiva baseada em data (*date-based*) ou em conjuntos (*set-based*). É obrigatório especificar o *metadataPrefix*, cujo padrão é *oai\_dc* (Figura 5).

Solicitação:

[http://eprints.c3sl.ufpr.br/per1/oai2?verb=ListRecords&metadataPrefix=oai\\_dc](http://eprints.c3sl.ufpr.br/per1/oai2?verb=ListRecords&metadataPrefix=oai_dc)

```

<?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>2003-07-31T23:07:07Z</responseDate>
  <request verb="ListRecords" metadataPrefix="oai_dc"
    resumptionToken="">http://eprints.c3sl.ufpr.br/per1/oai2</request>
  =<ListRecords>
    =<record>
      =<header>
        =<identifier>oai:GenericEPrints.OAI 2:36</identifier>

```

```

<datestamp>2003-06-27</datestamp>
<setSpec>7374617475733D707562</setSpec>
<setSpec>7375626A656374733D646174616261736573</setSpec>
</header>
<metadata>
  <= oai_dc:dc xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
      http://www.openarchives.org/OAI/2.0/oai_dc.xsd"
    xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
    xmlns:dc="http://purl.org/dc/elements/1.1/">
    <dc:title>Mapeamento de atributos complexos e multivalorados na
      extração de esquemas utilizando XML.</dc:title>
    <dc:creator>PI NTO, José Simão de Paula</dc:creator>
    <dc:creator>SUNYE, Marcos Sfair</dc:creator>
    <dc:subject>Databases</dc:subject>
    <dc:description>This work describes a way to map complex and / or
      multivalued attributes from a visual query environment to relational
      databases, using XML. These attributes are usually produced after
      the application of integration methodologies, like ERC+, whenever
      exists in original databases. The mapping was produced in
      databases already integrated, using XML DTD files to retain
      knowledge about pre-integration databases, and a tool, called
      VIQUEN, that generates a schema visualization by use the ERC+
      model. Queries and their results are both expressed in a visual
      environment. Express queries in a visual environment enables
      common users to create your own queries, since the tool hide
      complex database details and don't demand user knowledge about
      SQL statements.</dc:description>
    <dc:date>2001-08-01</dc:date>
    <dc:type>Thesis</dc:type>

    <dc:identifier>http://eprints.c3sl.ufpr.br/archive/00000036/</dc:ide
      ntifier>
    <dc:format>pdf
      http://eprints.c3sl.ufpr.br/archive/00000036/01/DissertacaoFinal.
      pdf</dc:format>
  </oai_dc:dc>
</metadata>
</record>
.
.
      (devido à extensão da resposta completa, alguns registros foram omitidos)
.
<= <record>
  <= <header>
    <identifier>oai:GenericEPrints.OAI 2:49</identifier>
    <datestamp>2003-08-01</datestamp>
    <setSpec>7374617475733D707562</setSpec>
    <setSpec>7375626A656374733D646174616261736573</setSpec>
    <setSpec>7375626A656374733D696E666F726D6174696F6E5F72657472696576
      616C</setSpec>
  </header>
  <= <metadata>
    <= <oai_dc:dc xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
        http://www.openarchives.org/OAI/2.0/oai_dc.xsd"
      xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
      xmlns:dc="http://purl.org/dc/elements/1.1/">
      <dc:title>Provedores de Dados de Baixo Custo: Publicacao Digital ao
        Alcance de Todos</dc:title>
      <dc:creator>Garcia, Patricia A. B.</dc:creator>
      <dc:subject>Databases</dc:subject>
      <dc:subject>Information Retrieval</dc:subject>
      <dc:description>The availability of information by the electronic media
        has been improving the visibility of scientific researches, which was
        restricted to traditional periodicals. Research centers and academic
        databases, and more recently, digital libraries and e-prints archives
        have been contributed to a new scenery that offers fast and free
        access to scientific information. Now, several initiatives encourages
        the creation of open archives and digital libraries. The Open Archives
        Initiative (OAI) promotes the communication standardization
        between archives through OAI-PMH (Open Archives Initiative
        Protocol for Metadata Harvesting), improving the interoperability
        among themselves. There are software available at OAI which
        purpose is to facilitate the implementation of data providers
        (electronic content archives). However, kind of works that, through
        a Computer Science vision, analyses cost, complexity and the range
        of these solution are still rare. This work analyses the presented
        solutions, proposing a criteria for its evaluation and confirming the
        viability of installing a low-cost interoperable digital
        library.</dc:description>
      <dc:date>2003-07-01</dc:date>
      <dc:type>Thesis</dc:type>
      <dc:identifier>http://eprints.c3sl.ufpr.br/archive/00000049/</dc:ide

```

```

ntifier>
  <dc:format>pdf
    http://eprints.c3sl.ufpr.br/archive/00000049/01/dissertacao_patri
    cia.pdf</dc:format>
  </oai_dc:dc>
</metadata>
</record>
<resumptionToken>0/475038/oai_dc</resumptionToken>
</ListRecords>
</OAI-PMH>

```

**Figura 5 - Resposta do provedor de dados ao verbo ListRecords do OAI-PMH**

### 3.2.5. ListIdentifiers

O *ListIdentifiers* é uma versão abreviada do verbo *ListRecords*, que retorna apenas os cabeçalhos dos registros. Através de argumentos opcionais pode-se estabelecer a coleta seletiva baseada em data (*date-based*) ou em conjuntos (*set-based*). É obrigatório especificar o *metadataPrefix* (*oai\_dc*) (Figura 6).

Solicitação:

[http://eprints.c3sl.ufpr.br/perl/oai2?verb=ListIdentifiers&metadataPrefix=oai\\_dc](http://eprints.c3sl.ufpr.br/perl/oai2?verb=ListIdentifiers&metadataPrefix=oai_dc)

```

<?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>2003-07-31T23:18:21Z</responseDate>
  <request verb="ListIdentifiers" metadataPrefix="oai_dc"
    resumptionToken="">http://eprints.c3sl.ufpr.br/perl/oai2</request>
  <ListIdentifiers>
    <header>
      <identifier>oai:GenericEPrints.OAI2:36</identifier>
      <datestamp>2003-06-27</datestamp>
      <setSpec>7374617475733D707562</setSpec>
      <setSpec>7375626A656374733D646174616261736573</setSpec>
    </header>
    .
    (devido à extensão da resposta, alguns itens foram omitidos)
    .
    <header>
      <identifier>oai:GenericEPrints.OAI2:49</identifier>
      <datestamp>2003-08-01</datestamp>
      <setSpec>7374617475733D756E707562</setSpec>
      <setSpec>7375626A656374733D6172746966696369616C5F696E74656C6C6967656E
        6365</setSpec>
    </header>
    <resumptionToken />
  </ListIdentifiers>
</OAI-PMH>

```

**Figura 6 - Resposta do provedor de dados ao verbo ListIdentifiers do OAI-PMH**

### 3.2.6. ListSets

Este verbo lista a estrutura do conjunto de um repositório (também conhecida como árvore de assuntos ou hierarquia do repositório), útil para o *harvesting* seletivo (Figura 7).

Solicitação: <http://eprints.c3sl.ufpr.br/perl/oai2?verb=ListSets>

```

<?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>2003-07-31T23:29:30Z</responseDate>
  <request verb="ListSets" resumptionToken="">http://eprints.c3sl.ufpr.br/perl/oai2</request>

```

```

=<ListSets>
  =<set>
    <setSpec>7375626A656374733D646174616261736573</setSpec>
    <setName>Subject = Databases</setName>
  </set>
  =<set>
    <setSpec>7375626A656374733D6F7065726174696E675F73797374656D73</setSpec>
    <setName>Subject = Operating Systems</setName>
  </set>
  =<set>
    <setSpec>7375626A656374733D68756D616E5F636F6D70757465725F696E74657261
      6374696F6E</setSpec>
    <setName>Subject = Human Computer Interaction</setName>
  </set>
  .
  . (devido à extensão da resposta, alguns itens foram omitidos)
  .
  =<set>
    <setSpec>7375626A656374733D696E666F726D6174696F6E5F72657472696576616C
      </setSpec>
    <setName>Subject = Information Retrieval</setName>
  </set>
</ListSets>
</OAI-PMH>

```

**Figura 7 - Resposta do provedor de dados ao verbo ListSets do OAI-PMH**

## 4. Implementação do OAI-PMH

A aplicação do OAI-PMH pode se dar tanto em bibliotecas digitais/bases de dados novas ou em já existentes. Com a utilização do protocolo, ocorre a exposição dos metadados do conteúdo da biblioteca digital/base de dados, aumentando assim sua visibilidade e facilitando a pesquisa, que poderá ser unificada através dos provedores de serviços da OAI.

### 4.1. Implementação em novas bibliotecas/repositórios digitais

É possível construir repositórios digitais com baixo custo compatíveis com a Iniciativa Open Archives, através dos programas GNU Eprints [9], Kepler [10,11,12] e DSpace [13,14]. Um estudo sobre as funcionalidades e custos de implantação destes programas é apresentado em [8]. Exemplos brasileiros de repositórios digitais construídos com suporte ao OAI-PMH: Coruja (IME/USP) [15], Sociedade Brasileira de Genética [16], UFPR Eprints [8].

### 4.2. Implementação em bibliotecas digitais já existentes

A implementação do protocolo OAI-PMH em bibliotecas digitais ou repositórios já existentes pode ser feita de duas maneiras: através do desenvolvimento completo da interface com o protocolo OAI-PMH ou através da utilização da ferramenta *OAI-Cat* da OCLC (*Online Computer Library Center*). O desenvolvimento completo da interface para o protocolo OAI-PMH deve se basear na documentação disponível na própria Iniciativa *Open Archives* [6,7]. A

ferramenta *OAI/Cat* é um programa de código aberto que implementa suporte ao protocolo OAI-PMH em repositórios de dados já existentes [17]. É utilizada pelo programa DSpace. O SciELO [18] é exemplo de uma implementação de interface com o protocolo OAI-PMH em uma base de dados já existente.

## 5. Conclusão

Atualmente existem quase 200 provedores de dados que implementam o protocolo OAI-PMH [6,8], e este número cresce a cada dia, consolidando a tendência do OAI-PMH como padrão de interoperabilidade entre bibliotecas digitais e bases de dados no mundo inteiro. Além disso, os programas desenvolvidos por integrantes e parceiros da OAI, como Eprints, Kepler, DSpace, OAI/Cat, facilitam a implementação do protocolo e diminuem os custos de sua implantação. Isto permite que informações de qualidade tornem-se disponíveis através de consultas gratuitas através dos provedores de serviços, sem necessidade de pagamento de taxas ou assinaturas para acessá-las. O protocolo OAI-PMH é uma grande oportunidade de oferecer, com baixo custo, integração e visibilidade à produção científica/acadêmica nacional.

## REFERÊNCIAS

- [1] LAGOZE, C. / VAN DE SOMPEL, H. **The Open Archives Initiative: Building a low-barrier interoperability framework.** Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, Roanoke VA, June 24-28, 2001, pp. 54-62. Disponível em <http://www.cs.cornell.edu/lagoze/papers/oai-jcdl.pdf> . Acesso em 22 de fevereiro de 2003.
- [2] HARNAD, S.. **Nature Debates: The Self-Archiving Initiative.** Disponível em <http://www.nature.com/nature/debates/e-access/Articles/harnad.html>. Acesso em 15 de abril de 2003.
- [3] Budapest Open Access Initiative. Disponível em <http://www.soros.org/openaccess> . Acesso em 17 de março de 2003.
- [4] Dublin Core Metadata Initiative. Disponível em <http://www.dublincore.org> . Acesso em 12 de abril de 2003.
- [5] Dublin Core Metadata Element Set, V.1.1:Reference Description. Disponível em <http://www.dublincore.org/documents/dces> . Acesso em 12 de abril de 2003.
- [6] Open Archives Initiative. Disponível em <http://www.openarchives.org> . Acesso em 02 de fevereiro de 2003.
- [7] LAGOZE, C. et al. **The Open Archives Initiative Protocol for Metadata Harvesting.** <http://www.openarchives.org/OAI/openarchivesprotocol.html> . Acesso em 17 de fevereiro de 2003.

- [8] GARCIA, P.A.B. **Provedores de Dados de Baixo Custo:** Publicação Digital ao Alcance de Todos. Dissertação de Mestrado. Curitiba, 2003. Setor de Ciências Exatas. Universidade Federal do Paraná. Disponível em <http://eprints.c3sl.ufpr.br/archive/00000049/> . Acesso em 01 de agosto de 2003.
- [9] Eprints.org. Disponível em <http://www.eprints.org> . Acesso em 12 de fevereiro de 2003.
- [10] MALY, K.; ZUBAIR, M.; LIU, X. **Enhanced Kepler Framework for Self-Archiving.** ICPP-02, pp. 455-461, Vancouver, August 2002. Disponível em <http://kepler.cs.odu.edu/publications/kepler.pdf> Acesso em 27 de janeiro de 2003.
- [11] MALY, K.; ZUBAIR, M.; LIU, X. **Kepler Proposal and Design Document.** January 2003. Disponível em <http://kepler.cs.odu.edu/publications/finaldes.doc> Acesso em 27 de janeiro de 2003.
- [12] MALY, K.; ZUBAIR, M.; LIU, X. **Kepler – An OAI Data/Service Provider for the Individual.** Disponível em <http://www.dlib.org/dlib/april01/maly/04maly.html> Acesso em 04 de dezembro de 2002.
- [13] DSpace.org. Disponível em <http://www.dspace.org> . Acesso em 26 de junho de 2003.
- [14] SMITH, M. et al. **DSpace – An Open Source Dynamic Digital Repository.** D-Lib Magazine, Janeiro/2003. Disponível em <http://www.dlib.org/dlib/january03/smith/01smith.html> . Acesso em 04 de junho de 2003.
- [15] VILA, H. **Coruja:** Um Provedor de Dados para a Iniciativa Open Archives. Instituto de Matemática e Estatística da USP. Disponível em [http://www.linux.ime.usp.br/~hvila/mac499/mono\\_final.htm](http://www.linux.ime.usp.br/~hvila/mac499/mono_final.htm). Acesso em 12 de maio de 2003
- [16] IBICT/SBG. **Arquivos Abertos - Sociedade Brasileira de Genética.** Disponível em <http://sbg.ibict.br>. Acesso em 1º de agosto de 2003.
- [17] OAIcat. Disponível em <http://www.oclc.org/research/software/oai/cat.shtm> . Acesso em 28 de junho de 2003.
- [18] MARCONDES, C. H. / SAYÃO, L. F. **The SciELO Brazilian Scientific Journal Gateway and Open Archives - A Report on the Development of SciELO - Open Archives Data Provider Server.** D-Lib Magazine, Março/2003. Disponível em <http://www.dlib.org/dlib/march03/marcondes/03marcondes.html> . Acesso em 10 de maio de 2003.