



## CLIO-I : PRIMANDO PELA USABILIDADE E ACESSIBILIDADE EM UM SISTEMA PARA GERENCIAMENTO E INTEROPERABILIDADE DE REPOSITÓRIOS DIGITAIS

*Marcos Cardoso Jr., Sandra de Albuquerque Siebra, Marcos Galindo, Vildeane Borba,  
Majóry Miranda*

**Resumo:** O interesse na criação de Repositórios Digitais cresceu significativamente a partir do surgimento e disseminação da Web, que trouxe consigo a necessidade de ferramentas que facilitassem a publicação, gerenciamento e recuperação da informação digital. Neste contexto, este artigo apresenta o Clio-i, um sistema para gerenciamento e interoperabilidade de Repositórios Digitais que possui um visualizador de documentos que prima por características de usabilidade e acessibilidade. Também, são apresentados estudos de caso de uso do sistema mostrando alguns contextos onde ele pôde ser aplicado com sucesso. Adicionalmente, é apresentado o resultado de uma pesquisa de satisfação com usuários do sistema. O resultado da pesquisa mostra que mesmo pequenas modificações já podem implicar em benefícios diretos para o usuário, tendo como consequência um maior interesse pelo uso do sistema.

**Palavras-chave:** Repositórios Digitais. Interoperabilidade. Usabilidade.

### 1 INTRODUÇÃO

De acordo com o glossário do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) / Ministério da Ciência e Tecnologia, repositórios digitais são sistemas de informação que armazenam, preservam, divulgam e dão acesso à produção intelectual de comunidades científicas. Incentivam e gerenciam a publicação pelo pesquisador (auto-arquivamento), utilizam tecnologia aberta e podem ser acessados por diversos provedores de serviços nacionais e internacionais. Em geral, repositórios digitais são desenvolvidos com a intenção de atender necessidades urgentes de uma comunidade específica, e assim, poucos desses serviços são planejados como sistemas que possam ser reusados em diferentes contextos (CARDOSO JUNIOR, 2007). Assim, torna-se importante a criação de sistemas de propósito geral que diminuam os custos de instalação e manutenção, tenham uma boa usabilidade, primem pela acessibilidade às informações e facilitem a implantação e gerenciamento de repositórios digitais.

Neste contexto, este artigo objetiva apresentar o Clio-i, um software livre para gerenciamento e acesso a Repositórios Digitais que oferece ao usuário uma ferramenta de indexação e busca automática por documentos e um módulo de visualização que permite operações sobre os documentos eletrônicos para melhorar sua visualização. Essas funcionalidades aumentam significativamente a usabilidade dos



documentos digitais. Adicionalmente, o Clio-i proporciona interoperabilidade com outros repositórios digitais que façam uso do protocolo OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting), que é um padrão internacional para interoperabilidade de repositórios digitais.

O Clio-i começou a ser desenvolvido em 2005, no Líber (Laboratório de Tecnologia do Conhecimento da Universidade Federal de Pernambuco)<sup>1</sup>, com apoio do Ministério de Educação (MEC) e convênio firmado entre o Centro de Documentação e de Estudos da História Brasileira (CEHIBRA) da Fundação Joaquim Nabuco (FUNDAJ)<sup>2</sup> e a UFPE. Dentro do contexto do projeto houve uma dissertação de mestrado defendida (CARDOSO JUNIOR, 2007) e em 2006 foi lançada a primeira versão do protótipo funcional que foi aprimorada e atualizada em 2007, com a inclusão dos protocolos de interoperabilidade. Desde então, não havia sido feita mais nenhuma modificação no sistema e o projeto ficou suspenso até junho de 2011, quando começou a ser retomado por uma nova equipe do Liber em parceria com o Centro de Informática da UFPE (Cin/UFPE)<sup>3</sup>. Entre os aprimoramentos a serem feitos no sistema estão a implementação das regras de acessibilidade sugeridas pela W3C (<http://www.w3.org/standards/webdesign/accessibility>), o aprimoramento da usabilidade de acordo com a ISO 9241 (ISO 9241, 1998 apud SANTOS, 2008), a taxonomia de Ferreira e Leite (2003) e as recomendações de Presmann sobre requisitos não-funcionais de usabilidade da interface de um sistema (PRESSMAN, 1992 apud FERREIRA; LEITE, 2003).

Atualmente, o Clio-i é utilizado em algumas instituições como, por exemplo, a FUNDAJ (Fundação Joaquim Nabuco) e a instituição Save The Children (<http://www.savethechildren.se/>) da Suécia e em diversos projetos como, por exemplo, o acervo sobre o Museu Murilo Lagreca (<http://www.liber.ufpe.br/lagreca>), o acervo do memorial popular Liedo Maranhão (<http://www.liber.ufpe.br/liedo>) e o acervo de obras holandesas sobre o Brasil (<http://liber.ufpe.br/visaoholandesa>).

O restante deste artigo está dividido como segue. Na seção 2 é dada a fundamentação teórica necessária para a completa compreensão do artigo. Na seção 3 são descritos: a arquitetura escolhida para desenvolvimento do Clio-i e o diferencial do Clio-i com relação a outras ferramentas e sistemas similares para a criação de repositórios digitais. Na seção 4 é apresentado o protótipo desenvolvido para avaliar a arquitetura proposta. Na seção 5 são apresentados dois estudos de caso de uso do Clio-i em diferentes contextos de utilização. Na seção 6 é descrita a pesquisa de satisfação realizada com usuários do Clio-i. Por fim, na seção 7 são apresentadas as conclusões deste artigo e são apontados melhorias e ajustes que estão programados para serem realizados no sistema, de forma a aprimorar sua usabilidade, enquadrá-lo dentro das regras de acessibilidade da W3C e torná-lo mais robusto.

---

1 O Liber ([www.liber.ufpe.br](http://www.liber.ufpe.br)) é um espaço de pesquisa ligado ao Departamento Ciência da Informação da UFPE, criado em 1996, como grupo de pesquisa e formalizado em 2006 como Laboratório de Tecnologia do Conhecimento. O Liber dedica-se a investigações no campo do gerenciamento eletrônico do conhecimento. Para este fim vem desenvolvendo, em ambiente controlado, repositórios e ferramentas que permitem a disponibilização, gerenciamento e pesquisa de conteúdos em formato digital.

2 URL: [www.fundaj.gov.br/](http://www.fundaj.gov.br/)

3 URL: <http://www.cin.ufpe.br>



## 2 REPOSITÓRIOS DIGITAIS

Segundo Viana, Márdero Arellano e Shintaku (2005), um repositório digital é uma forma de armazenamento de objetos digitais que tem a capacidade de manter e gerenciar material por longos períodos de tempo e prover o acesso apropriado. Essa estratégia foi possibilitada pela queda nos preços no armazenamento, pelo uso de padrões como o protocolo de coleta de metadados da Iniciativa dos Arquivos Abertos (OAI-PMH)<sup>4</sup>. (SOMPEL et al, 2004) e pelos avanços no desenvolvimento dos padrões de metadados que dão suporte ao modelo de comunicação dos arquivos abertos. Viana, Márdero Arellano e Shintaku (2005) e Souza, Cruz e Braga (2008) evidenciam ainda que os repositórios digitais incentivam a publicação na Rede gerenciada pelo pesquisador (auto-arquivamento); utilizam tecnologias abertas (open source); seu conteúdo está em regime de acesso aberto; há garantia de preservação digital do conteúdo a longo prazo (memória da produção científica); têm sistema de gestão integrado com outros serviços; promovem interoperabilidade com sistemas; usam padrões universais, como o protocolo OAI e as informações ficam disponíveis para serem acessadas, permanentemente, por diversos provedores de serviços a nível nacional e internacional.

Segundo o site OpenDOAR (Directory of Open Access Repositories)<sup>5</sup> nos últimos anos, tem aumentado o número de repositórios de acesso aberto, principalmente, dentro da comunidade universitária. É de responsabilidade e interesse das universidades que seus membros disponibilizem os seus trabalhos nesses repositórios institucionais. Esse interesse é também partilhado pela comunidade científica, pois ao disponibilizarem os seus trabalhos, estão contribuindo para que a sua instituição possa ganhar prestígio e credibilidade no meio da comunidade científica, além de contribuir para o progresso científico. (CARVALHO et al, 2006). Outros tipos de repositórios, além dos científicos/institucionais, também são fontes importantes de informação, porém, com finalidades e público alvo diferentes. Na verdade, os repositórios digitais podem ser implantados e operacionalizados em organizações do setor público ou privado, podendo ser acadêmicos, administrativos, técnicos ou mistos, mantendo a mesma confiabilidade de informação. (SHINTAKU; MEIRELLES, 2010)

### 2.1 Sistemas para criação de repositórios digitais

Existem diversos sistemas que aplicam-se à ampla maioria dos contextos institucionais, permitindo uma variedade de funcionalidades e recursos para gestão de documentos digitais. Entre eles, podemos citar: EPrints (<http://software.eprints.org/>); Fedora (<http://www.fedora.info>); Archimède (<http://archimede.bibl.ulaval.ca/>); MyCoRe (<http://www.mycore.de>); Digital Commons (<http://digitalcommons.bepress.com/>); Greenstone Digital Library Software (<http://www.greenstone.org/>); dLibra (<http://dlibra.psnc.pl/>) e DSpace (<http://www.DSpace.org>).

Segundo o OpenDOAR, o sistema mais utilizado, atualmente, é o DSpace (Ver Figura 1). O DSpace pertence à categoria dos Repositórios Digitais (RI) e é o resultado de um esforço conjunto do Massachusetts Institute of Technology (MIT) e da Hewlett-Packard (HP) e apoiado no Brasil pelo IBICT,

4 Maiores informações em: <http://www.openarchives.org/OAI/openarchivesprotocol.html>

5 URL: <http://www.opendoar.org>



desenvolvido com a finalidade de coletar, preservar e disseminar documentos digitais produzidos por docentes e pesquisadores do MIT. Posteriormente, adaptado por um conjunto de oito universidades americanas, o DSpace é hoje disponibilizado sob a forma de um produto open source, que pode ser livremente adaptado e expandido funcionalmente, nos termos da BSD (Berkeley Standard Distribution License open source license). Utiliza metodologia de arquivos abertos (Open Access Initiative) e padrões internacionais, como o Dublin Core (CARVALHO et al, 2006). Por ter como unidade os objetos digitais, o DSpace permite vários tipos de objetos e tipos de arquivos, tais como: texto, imagem, áudio e vídeo para os tipos de objetos digitais e todos os tipos mais comuns de arquivos como: PDF, DOC, TXT (para texto), JPG, GIF (para imagens), etc.

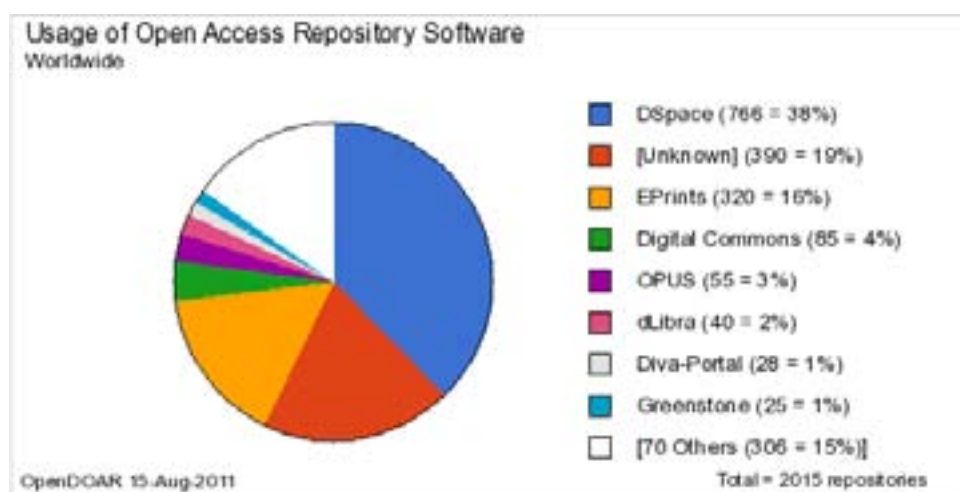


Figura 1: Estatísticas de uso dos softwares para criação de repositórios digitais  
Fonte: OpenDoar - 15/08/2011

O DSpace é o software adotado como referência no Brasil pelo IBICT e, a partir dessa adoção, várias iniciativas brasileiras passaram a utilizar o software DSpace como ferramenta de implementação de seus repositórios institucionais, tais como: Biblioteca Digital do Senado (<http://www2.senado.gov.br/bdsf/>); Repositório Institucional da Universidade de Brasília (<http://repositorio.bce.unb.br/>); Repositório Digital da UFRGS (<http://www.lume.ufrgs.br/>); Infoteca-e da EMBRAPA (<http://www.infoteca.cnptia.embrapa.br/>); entre outras.

Alguns problemas que puderam ser identificados no DSpace a partir da vivência dos autores com relação ao DSpace são:

- O DSpace foi desenvolvido com base na comunicação científica, seu projeto foi embasado na disseminação de literatura científica em formato, principalmente, de artigos que foram publicados anteriormente em periódicos. Por essa razão, sua organização está intimamente ligada a sua origem acadêmica. Assim, alguns conceitos orientadores dos repositórios remetem à comunicação científica. Logo, não há ferramentas diferenciadas para visualização e manipulação de, por exemplo, imagens digitalizadas (o que poderia ser útil, por exemplo, no contexto do acervo de museus e instituições).
- Como um software altamente configurável, o DSpace possui várias opções que lhe permitem ajustar-



se às diversas necessidades de uma instituição. Esse benefício é também uma dificuldade. Visto que a instalação e configuração do ambiente não é trivial e exige conhecimento técnico da equipe envolvida.

- As tarefas dos administradores requerem conhecimentos específicos do DSpace que exigem leitura de manuais ou treinamento apropriado.
- Não houve preocupação com acessibilidade e usabilidade na criação do DSpace
- É um software com muitas opções e muito amplo para atender a construção de qualquer tipo de repositório institucional. Por isso mesmo, se torna um pouco complexo para ser utilizado por iniciativas menores, como o acervo digital de museus, bibliotecas e instituições ou empresas públicas ou privadas.

### 3 O PROJETO DO SISTEMA CLIO-I

O Clio-i - sistema para gerenciamento e interoperabilidade entre Repositórios Digitais, oferece ao usuário uma ferramenta de busca automática por documentos, e um módulo de visualização que permite operações sobre os documentos eletrônicos para melhorar sua usabilidade. Outra característica importante do Clio-i é a adoção do protocolo OAI-PMH para prover a Interoperabilidade com outros repositórios. O uso do protocolo OAI-PMH implica que qualquer repositório que atenda as exigências do protocolo pode ter seus metadados coletados pelo Clio-i. Além disso, os metadados de um repositório local gerenciado pelo Clio-i poderão ser exportados para Provedores de Serviços OAI. A principal idéia por trás da criação do CLIO-I é que ele fosse um sistema fácil de instalar, configurar e manter, que não exigisse muita preparação da equipe envolvida na sua utilização e manutenção.

Na Figura 2, é possível visualizar o fluxo de informações do Clio-i. O sistema se comunica com uma base de dados, que armazena documentos digitais e seus metadados. Os metadados dos documentos podem ser inseridos localmente por um administrador ou podem ser coletados de outros repositórios (através de requisições a Provedores de Dados OAI externos). A definição dos repositórios acessados pelo Clio-i é realizada pelo Administrador do sistema. Os metadados e documentos inseridos do banco de dados são apresentados aos usuários através de serviços do Clio-i (e.g., módulo de Recuperação de Informação). Por fim, o sistema também pode disponibilizar os metadados depositados na sua base, para coleta por parte de Provedores de Serviço OAI.

#### 3.1 Arquitetura Geral

A arquitetura geral do sistema Clio-i conta com quatro módulos principais (Ver Figura 3) que serão descritos a seguir.

**Recuperação de Informação** - O módulo de Recuperação de Informação tem a finalidade de responder as consultas dos usuários por documentos referenciados na base de dados do Clio-i. Os registros são retornados em uma ordem de relevância, considerando tanto consultas aos metadados



que descrevem os documentos, como ao conteúdo dos documentos (no caso de documentos de texto).

**Visualização de Documentos** - Após o usuário realizar a pesquisa no Clio-i, os registros recuperados são listados com os metadados de cada documento. A partir daí, o usuário pode visualizar um documento específico de sua escolha, através do Visualizador de Documentos. O Visualizador de Documentos do Clio-i trabalha com arquivos do tipo áudio, imagem, vídeo e texto, e fornece ao usuário mecanismos para manipular o documento de texto e imagem de maneira rápida e simples.

**Clio-i Data Provider** - Um Provedor de Dados OAI é um sistema que exporta os metadados da sua base seguindo o protocolo OAI-PMH. Assim, esse módulo tem como função básica exportar os metadados da base de dados do Clio-i de acordo com o protocolo OAI-PMH. Os metadados poderão, assim, ser coletados por Provedor de Serviços OAI externos.

**Clio-i Service Provider** - Um Provedor de Serviços OAI realiza a coleta de diferentes repositórios que estejam de acordo com o protocolo padrão OAI-PMH (versão 2.0). Assim a função principal desse módulo é coletar metadados disponibilizados por outros Provedores de Dados OAI externos e armazená-los na base de dados do Clio-i.

## 4 O PROTÓTIPO IMPLEMENTADO

Esta seção descreve alguns detalhes de implementação do protótipo atual do Clio-i. Este protótipo foi utilizado nos estudos de caso apresentados na seção 4. O protótipo foi implementado em PHP (GILMORE, 2009), uma linguagem adequada para criação de páginas dinâmicas na Web. O sistema Clio-i pode ser utilizado facilmente em computadores que possuam servidores Web (e.g., Apache Server 2.0). Por se tratar de um sistema WEB, houve a preocupação de o sistema funcionar em diversos browsers, como o Internet Explorer e o Mozilla FireFox.

### 4.1 Base de Dados

O banco de dados escolhido para o armazenamento dos metadados foi o MySQL(<http://www.mysql.com>), um dos bancos de dados open-source mais populares do mercado. Esse software fornece recursos simplificados e apropriados para as suas aplicações, tendo um custo extremamente reduzido. Desta forma, destacamos as seguintes características para a escolha do MySQL: confiabilidade e bom desempenho, facilidade de utilização e distribuição, recursos e suporte para as mais diversas plataformas (WELLING; THOMSON 2004).



O sistema Clio-i usa como base de seu conjunto de metadados os campos descritores definidos no padrão Dublin Core (DC) (ROSETTO; NOGUEIRA, 2002) (ver quadro 1). O DC é um padrão internacional criado com o objetivo de fornecer um conjunto simples e intuitivo de metadados para a descrição de documentos digitais em contextos de um determinado domínio. Adicionalmente, o Dublin-Core é o padrão de metadados sugerido pelo protocolo OAI-PMH para permitir a unificação de metadados de diferentes repositórios. Além dos metadados do DC, cada documento é associado ainda a informações específicas dos objetos digitais como resolução, duração (no caso de áudio e vídeo) e seqüência de páginas (para texto e imagem).

<b>Elemento Dublin Core</b>	<b>Descrição</b>
<b><i>Title</i></b>	Título do documento
<b><i>Creator</i></b>	Autor do documento
<b><i>Description</i></b>	Resumo do documento
<b><i>Type</i></b>	A natureza ou gênero do conteúdo. Descreve categorias gerais de documentos.(e.g., iconografia, cordéis, etc.).
<b><i>Coverage</i></b>	Onde o documento foi produzido / impresso
<b><i>Publisher</i></b>	A entidade responsável por tornar o documento disponível na sua forma atual.
<b><i>Date</i></b>	A data da produção / impressão do documento
<b><i>Subject</i></b>	As palavras-chaves que identificam o assunto documento
<b><i>Contributor</i></b>	Entidade responsável por contribuições ao conteúdo do recurso (e.g., tradutor, revisor, ilustrador, etc.).
<b><i>Source</i></b>	Documento do qual o presente documento foi derivado.
<b><i>Language</i></b>	Principal idioma do conteúdo do documento.
<b><i>Format</i></b>	Manifestação física ou digital do documento (e.g., jpeg, tiff, mp3).
<b><i>Rights</i></b>	Copyright, direitos do autor, propriedade intelectual.
<b><i>Relation</i></b>	Referência para um documento relacionado.
<b><i>Identifier</i></b>	Uma referência não ambígua que identifica o documento (e.g., uma URL).

#### 4.2 Módulo de Recuperação de Informação

Nesse módulo, foi utilizado um recurso do banco de dados MySQL, chamado MySQL Full-Text Search (versão 3.23.23). Esse recurso permite o casamento de palavras-chave de consulta e



expressões booleanas com os campos do banco de dados. Os resultados da busca são ordenados segundo um critério de relevância. A Figura 4 apresenta um exemplo do uso das funções MATCH() e AGAINST() em uma cláusula WHERE de uma consulta SQL. Nesse exemplo, os campos de metadados pesquisados (i.e. título e resumo) são definidos na função MATCH() e as palavras-chave (i.e. Brasil) na função AGAINST().

Para cada resultado da consulta, a função MATCH() retorna um valor de relevância. Essa relevância é medida através da frequência com que os termos de consulta são encontrados nos campos do documento, ponderados pela frequência com que os termos estão presentes na base de documentos (BAEZA-YATES; RIBEIRO-NETO, 1999).

```
SELECT * FROM documentos WHERE MATCH (título,resumo) AGAINST ('Brasil')
```

id	título	resumo
5	O Brasil descoberto	...no brasil, o pau-brasil foi...
10	Tempos de crise	O Brasil de 1964...
7	A fonte de riqueza	...em meados de 80, o Brasil...

Figura 4: Uma busca utilizando o MySQL Full-Text Search

Fonte: Os autores, 2011

O MySQL Full-Text Search permite ainda que se trabalhe com buscas full-text booleanas, utilizando o modificador IN BOOLEAN MODE. Na Figura 5, temos um exemplo de consulta que recupera todos os documentos que contenham a palavra Brasil, mas que não possuam o termo ditadura.

### 4.3 Visualizador de Documentos

O visualizador de documentos foi implementado utilizando DHTML (Dynamic HTML). O DHTML é um método interativo de visualização de informação. A tecnologia foi empregada tanto nas operações sobre os documentos, quanto na construção das janelas móveis resultando num módulo satisfatório em termos de usabilidade e velocidade.

Na Figura 6, é apresentada a interface do módulo de Visualização de Documentos usada no Acervo Digital da Fundação Joaquim Nabuco. Os documentos são apresentados na tela central. À esquerda da tela, uma barra de ferramentas oferece opções para o usuário realizar operações sobre a imagem do documento sendo visualizado (e.g., zoom, giro vertical e horizontal, entre outros.). Em muitos casos, essas operações são essenciais para possibilitar uma boa leitura dos documentos apresentados (e.g., imagens de documentos manuscritos ou mal conservados), contribuindo assim para aumentar a usabilidade do repositório digital.



Figura 6: Apresentação de documento no *Visualizador*  
Fonte: Os autores, 2011

Os efeitos visuais disponíveis podem ser aplicados sobre documentos do tipo texto ou imagem, e incluem as seguintes opções: (1) operações de Zoom, para aumentar ou diminuir progressivamente o tamanho da imagem do documento; (2) negativar, para uma melhor visualização, em alguns casos é necessário negativar o documento; (3) girar vertical ou horizontalmente, para facilitar a leitura, por exemplo, de documentos que foram digitalizados em posição invertida; (4) tornar mais ou menos opaco, permite clarear ou escurecer progressivamente o documento; (5) restaurar, após aplicar diversas operações, o usuário tem a opção de restaurar a configuração original do documento. A barra de ferramentas apresenta ainda outras opções aplicáveis para qualquer tipo de documento. Essas opções dizem respeito às possibilidades de interação do usuário com o sistema, e incluem inserir notas sobre o documento, ler as notas inseridas, fazer download em formato PDF (no caso de textos e imagens) e um menu de ajuda. Este módulo oferece ainda uma barra de navegação (canto superior direito da figura 6), utilizada para “folhear” os documentos de texto ou imagem que contenham mais de uma página.

#### 4.4 Clio-i Data Provider

Para este módulo, um conjunto de funções foi implementado e as principais estão representadas na Figura 7, que apresenta um fluxograma do funcionamento deste módulo cujas principais funções são descritas a seguir

**Parser.** Conjunto de funções com a finalidade de validar as requisições em HTTP vindas de um Provedor de Serviços OAI.

**Gerador XML.** Caso a requisição HTTP tenha sido aprovada pelo parser, a função respostaXML retorna uma string no formato XML com todas as informações requisitadas. Como a quantidade de registros de metadados exportados pode ser grande, o módulo fornece a opção de retornar apenas 100 registros de cada vez na sua exportação. Ou seja, se a resposta da solicitação possuir mais do que esse limite estabelecido, o componente Controle de Fluxo do Clío-i Data Provider é acionado.

**Controle de Fluxo.** Como dito anteriormente, caso a requisição validada possua mais de 100 registros em sua resposta, esse componente é ativado através da função respostaResumption. Essa função retorna a mesma resposta em XML que é gerada pelo componente anterior. A diferença é que a função do Controle de Fluxo recebe um parâmetro a mais, que é justamente o resumptionToken, responsável por realizar a paginação dos resultados em XML das requisições.

**Gerador de erros.** Este componente é chamado caso o parser retorne falso para a validação da requisição HTTP. Através da função selecionaErro, é verificado que erro foi cometido e sua resposta em formato XML é apresentada.

**Clío-i Service Provider.** A implementação do Clío-i Service Provider foi feita com base nas funções apresentadas no fluxograma da figura 8 e detalhadas a seguir.

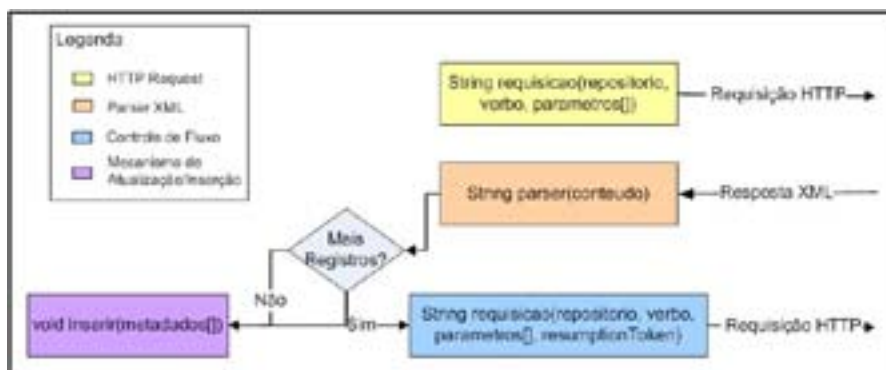


Figura 8: Principais funções e componentes do Clío-i Service Provider  
Fonte: Os autores, 2011

**HTTP Request.** Passo inicial para a coleta dos dados, este componente realiza a requisição para um Provedor de Dados previamente cadastrado no sistema. O HTTP Request possui a função requisição, onde passamos como parâmetro o repositório escolhido, o verbo e seus possíveis argumentos. Dessa maneira, a função seleciona a URL Básica para coleta do repositório e monta a requisição com base nos parâmetros passados pelo administrador do sistema.



**Parser XML.** Este componente valida o conteúdo XML enviado por um Provedor de Dados, e faz uma chamada ao componente Mecanismos de Inserção/Atualização para inserir os dados no banco MySQL. O Parser XML ainda verifica se há mais requisições a serem feitas (i.e. se o Provedor de Dados envia metadados usando Controle de Fluxo). Caso ainda haja porções de registros a serem coletadas do repositório selecionado, o componente Controle de Fluxo é chamado.

**Controle do Fluxo.** Após a confirmação que ainda existem registros a serem coletados, a função requisição é acionada, essa do componente Controle de Fluxo. Essa função é similar a encontrada no HTTP Request, tendo como diferença o parâmetro adicional `resumptionToken` que controla o fluxo de registros coletados.

**Mecanismos de Atualização/Inserção.** Recebe do parser XML os valores necessários para o componente realizar a atualização da base de dados do Clio-i Service Provider. Através da função `inserir`, os metadados coletados são inseridos ou atualizados, caso o registro coletado já tenha sido armazenado na base alguma outra vez. Ressaltamos que essa atualização é necessária para que a base do Clio-i Service Provider esteja de acordo com a base do Provedor de Dados em questão.

## 5 ESTUDOS DE CASO

Versões anteriores do Clio-i foram aplicadas em diferentes estudos de caso de Repositórios Digitais (CARDOSO JÚNIOR, 2007). A diferença do estado atual do Clio-i para suas versões anteriores é o uso do protocolo OAI-PMH para interoperabilidade de repositórios. Desta forma, nessa seção apresentamos dois os estudos de caso em que o Clio-i foi usado com as funcionalidades dos módulos Clio-i Data Provider e Clio-i Service Provider.

### 5.1 Estudo de Caso 1: Integrador de Repositórios Científicos

O primeiro estudo de caso corresponde ao Integrador de Repositórios Científicos (ver Figura 9). O objetivo principal desse estudo de caso foi testar o módulo de coleta de dados (Clio-i Service Provider). Para isto, foram coletados metadados sobre artigos científicos oriundos de diversos provedores de dados, formando uma base de metadados centralizada.

Para tanto, foi feita inicialmente uma consulta à página dos Provedores de Dados oficiais da OAI, a fim de selecionarmos os repositórios científicos que seriam coletados pelo Clio-i Service Provider. Ao todo foram selecionados 19 Provedores de Dados OAI de diferentes nacionalidades. A lista dos 19 repositórios pode ser vista em (CARDOSO JÚNIOR, 2007)

Após essa seleção manual de repositórios científicos, deu-se continuidade à construção da



base de metadados. Cada um dos 19 repositórios selecionados foi cadastrado no sistema, e em seguida foi realizada a coleta de todos os metadados disponíveis nos repositórios. Após a coleta de todos os metadados, um total de 138.510 registros de artigos científicos foi reunido na base de dados do Clio-i. Essa base pode ser consultada pelos usuários através do módulo de Recuperação de Informação na URL: <http://www.liber.ufpe.br/clioi>.



Figura 9: Clio-i Integrador de Repositórios Científicos  
Fonte: Os autores, 2011

Destacamos que o Provedor de Serviços desse estudo de caso foi registrado oficialmente no Open Archives Initiative, que conta atualmente com apenas 26 Provedores de Serviços cadastrados<sup>6</sup>. Vale salientar que o Clio-i foi o primeiro da América Latina a integrar a lista de sistemas que seguem o protocolo OAI-PMH.

## 5.2 Estudo de Caso 2: Integrador de Repositórios Multimídia

O segundo estudo de caso avaliou mais especificamente o módulo Clio-i Data Provider. Para isto, o Clio-i Data Provider foi utilizado para expor os metadados de três Bibliotecas Digitais que usam o Clio-i. Foram elas:

<sup>6</sup> Open Archives Initiative Service Providers - <http://www.openarchives.org/service/listproviders.html>



Acervo Digital FUNDAJ: O acervo digital da Fundação Joaquim Nabuco utiliza os serviços de Recuperação da Informação e Visualização de Documentos do Clio-i desde o início do ano de 2006. Adicionamos ao sistema o módulo Clio-i Data Provider para expor os seus metadados e os seus recursos.

Holandeses na Bahia: Base composta essencialmente por documentos do tipo texto sobre a conquista holandesa de Salvador no Brasil em 1624, e a reconquista da cidade por uma armada luso-espanhola em 1625.

Escrito nas Estrelas: Este repositório agrega vídeos e áudio de uma exposição sobre o Brasil na época do início do desenvolvimento dos transportes aéreos, e dos correios no início do século XX. Para o presente trabalho, foram selecionados alguns desses vídeos, utilizados em um Clio-i Data Provider.

Os metadados dessas três bases foram integrados por um Clio-i Service Provider com o nome de Integrador de Repositórios Multimídia. Dessas três bases, o Acervo Digital da FUNDAJ foi o primeiro a ser registrado, oficialmente, como Provedor de Dados no site oficial do Open Archives Initiative<sup>7</sup>, que conta com mais de 500 repositórios registrados. Os outros dois repositórios foram registrados posteriormente.

## 6 ESTUDO DE SATISFAÇÃO DO USUÁRIO

Em qualquer sistema que vá ser utilizado por diferentes perfis de usuário e que seja importante o acesso a informações através dele, torna-se ainda mais relevante aliar recursos de acessibilidade e usabilidade à interface, na medida em que essa ação pode aproximar mais os usuários potenciais e reais para uso efetivo do mesmo. Isso porque quando estão realizando pesquisas, os usuários buscam respostas concretas e esperam interagir com o sistema de maneira rápida, simples e efetiva, sem necessitar de treinamentos, consultas à ajuda ou da leitura de manuais. Assim, foi feito um estudo inicial, simplificado, da usabilidade do sistema CLIO-I e da satisfação dos usuários com o mesmo.

O ajuste da usabilidade e acessibilidade permite o uso eficiente, produtivo, satisfatório, útil e seguro para o usuário do computador e destinatário da informação do conhecimento disponibilizado na rede. Ou seja, as condições de usabilidade e de acessibilidade podem assegurar que a informação disponibilizada na internet possa ser compreendida e utilizada por todos, independente de limitações (IGNÁCIO; CARVALHO, 2008).

O teste da interface foi realizado com 20 estudantes de graduação de uma Instituição de Ensino Superior (IES) que foram convidados a participar do teste (nenhum deles tinha tido

<sup>7</sup> Open Archives Initiative Service Providers - <http://www.openarchives.org/Register/BrowseSites>



contato anterior algum com o sistema) e com 10 usuários dos acervos do Clio-i, atualmente, em funcionamento. Inicialmente, foi solicitado que cada usuário realizasse quatro atividades no Clio-i: 1) buscar documentos que tivesse a palavra Brasil; 2) acessar uma base de dados fora do sistema clio-i (disponibilizada através do protocolo OAI-PMH); 3) Manipular algum documento do tipo imagem com a ferramenta de Visualização e 4) fazer o download de algum documento recuperado.

O teste foi realizado individualmente por cada usuário e foi acompanhado por um observador que se valeu do método “pensando em voz alta”. Tal método consiste em uma observação direta de cada usuário utilizando o site. De acordo com Nielsen “ouvir os ‘pensamentos’ de cada usuário permite entender as razões de suas ações e essas informações são muito importantes”. (NIELSEN, 1993). Ao final do teste, foi feita uma entrevista com os usuários para levantar o grau de satisfação desses usuários. Foi pedido que eles pontuassem com notas de 0 a 10: 1) a qualidade dos resultados da busca usando o Clio-i; 2) a facilidade de baixar um arquivo pesquisado; 3) a forma de apresentação e manipulação das informações recuperadas; 4) a facilidade de utilização do sistema como um todo.

Todos os usuários envolvidos no teste conseguiram executar as ações solicitadas e comentaram que acharam a interface simples. Considerando a média aritmética das quatro questões pontuadas ao final do uso do sistema, o Clio-i ficou com média 8,69, que foi considerada muito boa, especialmente porque o sistema não sofre atualizações desde 2008 e possui diversas melhorias a serem feitas. Durante o teste os usuários comentaram que acharam muito diferente e interessante o módulo de visualização de documentos e as possibilidades que ele oferece. Alguns usuários durante a entrevista solicitaram novas funcionalidades e deram sugestões para o sistema que serão levadas em conta no novo ciclo de implementação que está apenas começando. O foco na simplicidade, usabilidade e acessibilidade, sem deixar de lado as funcionalidades básicas necessárias a uma ferramenta para criação de repositórios digitais são o grande diferencial do CLIO-I que pretendem ser mantidas na próxima versão a ser desenvolvida. Isso pode torná-lo interessante para ser utilizado por pessoas experientes e sem experiência na criação e utilização de repositórios digitais, especialmente no contexto de pequenas e médias iniciativas.

## 7 CONCLUSÕES E TRABALHOS FUTUROS

O Clio-i é uma ferramenta adequada, especialmente, para pequenas e médias iniciativas de repositórios digitais e para gerenciamento de acervos de museus. Isso devido a sua simplicidade de instalação e utilização para criação do repositório, a disponibilização de busca simples e avançada, a apresentação diferenciada da informação, ao seu módulo visualizador de documentos com opções de manipulação do mesmo e por seu desenvolvimento ter primado pela usabilidade. Adicionalmente, devido ao Clio-i ser um sistema implementado usando protocolos de interoperabilidade (OAI-PMH), é possível disponibilizar também, através dele, o acervo de outros repositórios que façam uso do protocolo OAI-PMH, o que amplia ainda mais as possibilidades de recuperação da informação.



O projeto CLIO-I está sendo retomado pela equipe do laboratório Líber em parceria como Centro de Informática da UFPE, a fim de aprimorar a ferramenta e torná-la uma opção simples e fácil de usar para pequenas e médias iniciativas que precisem montar o seu repositório digital. A idéia central será primar pela usabilidade, facilidade de uso e acessibilidade da ferramenta, tornando-a simples de usar inclusive por pessoas que não sejam da área de Ciência da Informação ou que não tenham grande conhecimento da tecnologia da informação e comunicação. Entre os trabalhos futuros a serem realizados no contexto da retomada do projeto CLIO-I estão:

- Possibilitar a leitura de metadados de imagens no formato JPEG. Isso se torna relevante a medida que várias imagens utilizadas hoje são originadas em câmeras ou filmadoras digitais que já imprimem na imagem os dados sobre o equipamento utilizado para obtenção da fotografia.
- Possibilitar a leitura de metadados nos formatos EXIF e IPTC
- Aprimorar o mecanismo de busca integrando conceitos de Web Social, Web Semântica e Findability. Neste contexto, instrumentos como ontologias ou taxonomias, possibilidade de criação de redes sociais dentro do repositório e variações das estratégias de pesquisa e uso da informação podem ser bastante úteis.
- Possibilitar a geração de tipos variados de relatórios gerenciais (ex: relatório de acesso e visualização de documentos) e infográficos (ex: pesquisas realizadas por perfil de usuário)
- Implementar a possibilidade de criptografar determinados arquivos que se classifiquem como confidenciais ou de acesso restrito, como medida de segurança da informação.
- Implementar as sugestões de acessibilidade da W3C.

**Abstract:** *The interest in creating digital repositories has grown significantly from the emergence and spread of the Web, which brought with it the need for tools to facilitate the publication, management and retrieval of digital information. In this context, this paper presents the Clio-i, a system for management and interoperability of digital repositories that has a document viewer that excels in usability and accessibility features. It also presents case studies of use of the system showing some contexts where it could be successfully applied. We also present the results of a survey of satisfaction with system users. The research result shows that even small changes may already result in direct benefits to the user, resulting in an increased interest in using the system.*

**Keywords:** Digital Repositories. Interoperability. Usability.



## REFERÊNCIAS

BAEZA-YATES, R. A.; RIBEIRO NETO, B. **Modern Information Retrieval**. Massachusetts: Addison-Wesley, 2010.

CARDOSO JUNIOR, Marcos José de Menezes. **Clio-i: Interoperabilidade entre repositórios digitais utilizando o protocolo OAI-PMH**. 2007. 137f. Dissertação (Mestrado em Ciência da Computação) - Programa de Pós-graduação em Ciência da Computação, Centro de Informática, Universidade Federal de Pernambuco, Recife, 2007. Disponível em: < <http://www.cin.ufpe.br/~rbcp/dissertacoes/dissertacaoCARDOSO.pdf>>. Acesso em: 15 ago. 2010.

CARVALHO, Maria Carmen Romcy de et al. O Repositório institucional como recurso de apoio à gestão do conhecimento da Universidade Católica de Brasília. In: CONFERÊNCIA IBEROAMERICANA DE PUBLICAÇÕES ELETRÔNICAS NO CONTEXTO DA COMUNICAÇÃO CIENTÍFICA, 1. 2006. **Anais...** Disponível em: < [http://DSpace.ibict.br/dmdocuments/Maria\\_Carmen\\_Romcy\\_de\\_Carvalho.pdf](http://DSpace.ibict.br/dmdocuments/Maria_Carmen_Romcy_de_Carvalho.pdf)>. Acesso em: 12 jun. 2011.

FERREIRA, Simone B. L.; LEITE, Julio C. S. do P. Avaliação da usabilidade em sistemas de informação: o caso do sistema submarino. **Revista de Administração Contemporânea – RAC**, São Paulo, v.7, n.2, p.115-137, abr./jun. 2003. Disponível em: <[http://www.anpad.org.br/rac/vol\\_07/dwn/rac-v7-n2-sbf.pdf](http://www.anpad.org.br/rac/vol_07/dwn/rac-v7-n2-sbf.pdf)>. Acesso em: 10 jun. 2011.

GILMORE, W. J. **Dominando Php e Mysql: do Iniciante ao Profissional**. Rio de Janeiro: Alta Books, 2009.

IGNÁCIO, Edilson A.; CARVALHO, José O. F. de. Avaliação da acessibilidade de sites oficiais de pesquisa no Brasil por pessoas com deficiência. **Encontros Bibli**, Florianópolis, v. 13, n.26, 2008. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/7193/6641>>. Acesso em: 10 jun. 2011.

NIELSEN, J. **Usability Engineering**. New York, NY: Academic Press, 1993.

OPEN DOAR. Disponível em: < <http://www.opendoar.org/>>. Acesso em: 10 jul. 2011.

ROSETTO, M.; NOGUEIRA, A. Aplicação de elementos metadados Dublin Core para descrição de dados bibliográficos on-line da biblioteca digital de teses da USP. In: SIMPÓSIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 12. São Paulo. **Anais...** 2002. Disponível em: < <http://www.sibi.ufrrj.br/snbu/snbu2002/oralpdf/82.a.pdf>>. Acesso em: 10 Mar. 2010.

SANTOS, Rodrigo C. dos. Revisão das métricas para avaliação de usabilidade de sistemas. In: GLOBAL BUSINESS AND TECHNOLOGY ASSOCIATION CONFERENCE, 2008, Madrid. **Anais...** Espanha: GBTA, 2008. Com o título original de: Systems Usability Evaluation Metrics Review. Disponível em: <[http://www.marcelomoraes.com.br/conteudo/marcelo/metricas\\_usabilidade.pdf](http://www.marcelomoraes.com.br/conteudo/marcelo/metricas_usabilidade.pdf)>. Acesso em: 10 jun. 2011.

SHINTAKU, Milton; MEIRELLES, Rodrigo. **Manual do DSpace** : administração de repositórios. Salvador : EDUFBA, 2010.



SOMPEL, H. V. de; NELSON, M.L.; LAGOZE, C.; WARNER, S. Resource Harvesting within the OAI-PMH Framework. **D-Lib Magazine**, v. 10, n. 12, dez. 2004. Disponível em: <<http://www.dlib.org/dlib/december04/vandesompel/12vandesompel.html>>. Acesso em: 04 jan. 2010.

SOUSA, M. C. P.; CRUZ, M. A. L.; BRAGA, M. F. A. Acesso livre e repositório institucional: uma ferramenta indispensável nas Instituições de Ensino Superior. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 15., 2008, São Paulo. **Anais**. Disponível em: <<http://www.repositorio.ufma.br:8080/jspui/bitstream/1/256/1/artigo%20R.I.pdf>>. Acesso em: 20 jan. 2009.

VIANA, C. L. M.; MÁRDERO ARELLANO, M. A.; SHINTAKU, M. **Repositórios institucionais em ciência e tecnologia**: uma experiência de customização do DSpace. 2005. Disponível em: <<http://eprints.radis.org/archive/00005563>>. Acesso em: 30 maio 2008.

WELLING, Luke; THOMSON, Laura. **Tutorial MYSQL**: Uma Introdução Objetiva aos Fundamentos do Banco de Dados MYSQL. Rio de Janeiro: Ciência Moderna, 2004.

W3C. **Recomendações de Acessibilidade para Conteúdo Web (WCAG) 2.0**. 2008. Tradução de Everaldo Bechara. Disponível em: <<http://www.ilearn.com.br/TR/WCAG20/>>. Acesso em: 10 jun. 2011.