SeeOER: Uma Arquitetura para Mecanismo de Busca na Web por Recursos Educacionais Abertos

Murilo G. Gazzola¹, Cristina D. A. Ciferri¹, Itana M. S. Gimenes ²

¹Instituto de Ciências Matemáticas e de Computação – Universidade de São Paulo (USP) Caixa Postal 668 – 13560-970 – São Carlos – SP

²Departamento de Informática – Universidade Estadual de Maringá (UEM) Caixa Postal 5790 – 87020-900 – Maringá – PR

{mgazzola, cdac}@icmc.usp.br, itana@din.uem.br

Abstract. SeeOER is an architecture for a specialized Open Educational Resources (OERs) Web search engine. The work was developed within the context of a master's degree project. The novel features of the proposed architecture will be introduced and compared to existing works. SeeOER uses intrinsic aspects of OER obtained from metadata. One of its distinticve features is the fact that the developed components take into account aspects of data provenance. This paper describes the scientific methodology, the developed architecture, standard metadata used, demonstrative seeds, related work and an analysis of the results.

Resumo. SeeOER é uma arquitetura para um mecanismo de busca na Web especializado para Recursos Educacionais Abertos (REAs). O trabalho foi desenvolvido no contexto de um projeto de mestrado. As características inovadoras da arquitetura proposta são introduzidas e comparadas com trabalhos existentes. SeeOER utiliza aspectos intrínsecos dos REAs obtidos a partir de seus metadados. Uma de suas características distintas dessa arquitetura é o fato de que os componentes desenvolvidos incorporam aspectos de procedência de dados. Este artigo descreve a metodologia científica utilizada, a arquitetura desenvolvida, os padrões de metadados utilizados, sementes demonstrativas, os trabalhos correlatos e uma análise dos resultados obtidos pelo SeeOER.

1. Introdução

O interesse por Recursos Educacionais Abertos (REAs) é cada vez mais crescente. Eles são um dos principais elementos da educação aberta que visa permitir o acesso gratuito ao conteúdo educacional em todo o mundo. Atualmente, diversas instituições de ensino e pesquisa têm investido no uso de REAs para a disponibilização de conteúdo relacionado à educação, como The Open University¹ [Little et al. 2011, Okada 2007], Stanford University², MIT³, Unicamp⁴, FGV⁵, UNESP⁶ e USP⁷. Outra vertente da educação aberta

Ihttp://openlearn.open.ac.uk/

²http://class2go.stanford.edu

³http://ocw.mit.edu

⁴http://www.ggte.unicamp.br/e-unicamp/public/

⁵http://www5.fgv.br/fgvonline/Cursos/Gratuitos

⁶http://www.unesp.br/unespaberta

⁷http://eaulas.usp.br/portal/home

que evidencia a propagação de REAs são os *Massive Open Online Courses* (MOOCs), os quais podem estar disponíveis como REAs e podem oferecer uma educação alternativa e de qualidade tanto na complementação da educação tradicional quanto na formação continuada [Matkin 2013]. Porém, os mecanismos atuais de busca na Web dificultam a identificação de REAs e, portanto, prejudicam a sua disseminação e incorporação em práticas educacionais. Alguns dos principais problemas específicos dos mecanismos atuais são descritos a seguir:

- Existem diferentes *padrões de metadados*, *repositórios e plataformas* disponíveis. Os REAs têm sido construídos sem a utilização adequada desses padrões, repositórios e plataformas, o que tem gerado diversos problemas de heterogeneidade, tanto em nível de esquema quanto em nível de instância.
- É difícil de se entender que *tipo de licença* está associada aos REAs.
- A granularidade dos recursos encontrados nem sempre atende às necessidades dos usuários.
- É difícil de se identificar a *qualidade dos recursos*. Pode-se associar a qualidade de um recurso à crítica que os usuários fizeram daquele recurso, e pode-se disponibilizar *rating* dos recursos. Mesmo assim, determinar a qualidade de um recurso de modo a incorporar esse parâmetro na recuperação de REAs ainda é uma questão em aberto.
- A *procedência de dados* de REAs não tem sido explorada na literatura para aprimorar o processo de busca. Como resultado, informações importantes tais como a origem do recurso, a sua qualidade e os processos de transformação aplicados aos recursos até a produção do recurso no estado atual não são considerados.

O SeeOER (*Search EnginE for Open Education Resources*) é uma arquitetura desenvolvida para um mecanismo de busca na Web por REAs que visa a resolução de conflitos em nível de esquema e em nível de instância, oriundos do uso de diferentes padrões de metadados, repositórios e plataformas de REAs. Além disso, a granularidade, a licença e a qualidade dos REAs são características importantes que foram consideradas quando da busca por REAs.

Foram usados no desenvolvimento desta arquitetura cinco padrões de metadados que facilitam a catalogação, pesquisa e reutilização de REAs, a saber: o Dublin Core Metadata Element Set (DCMES), IEEE Learning Object Metadata (IEEE/LOM), Protocolo *Open Graph* (OGP), MathML e Vídeo Sitemaps.

O presente artigo trará na seção 2 o processo seguido para desenvolvimento do trabalho e realização do experimento. Na seção 3 apresenta uma visão geral da arquitetura desenvolvida, enfatizando os principais componentes desenvolvidos. Na seção 4 descreverá os principais metadados no contexto de REAs que foram utilizados no trabalho. Também é apresentada uma breve lista de URLs que foram utilizadas como sementes para o SeeOER. Na seção 5 é feita uma síntese comparativa com trabalhos existentes. Na seção 6 é feita uma análise dos resultados obtidos pelo SeeOER. Por fim, na seção 7 é apresentada as conclusões do trabalho.

2. Metodologia

Inicialmente foi concebida a arquitetura SeeOER, conforme descrita na seção 3. A partir desta foi desenvolvido um protótipo. Sementes foram escolhidas para os testes realizados

a partir de pesquisas na Web, as quais foram verificadas manualmente. Foram considerados os repositórios brasileiros como também de outros países. A seção 4 mostra algumas das sementes utilizadas e descreve os metadados usados no SeeOER. Em seguida os REAs foram capturados e comparados aos trabalhos correlatos. Os dados obtidos dos trabalhos correlatos foram retirados, no caso do Jorum [Jorum 2014] do próprio mecanismo de busca, e do OERScout no artigo publicado [Abeywardena et al. 2013]. Os outros mecanismos de busca correlatos não foram mencionados na comparação, pois estes não publicaram resultados sobre a quantidade de REAs indexados obtidos. A próxima seção é apresenta a arquitetura SeeOER e seus componentes.

3. Arquitetura do SeeOER

O SeeOER é derivado da arquitetura de um mecanismo de busca na Web em grande escala. Portanto, sua arquitetura possui dois processos principais, o processo de *indexação* e o processo de *consulta*.

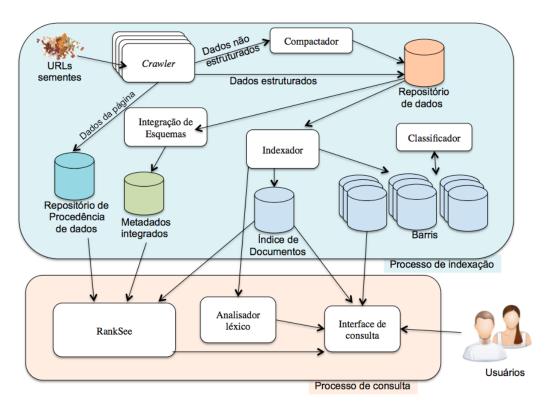


Figura 1. Arquitetura do SeeOER

As funcionalidades dos componentes da arquitetura são descritas a seguir. A busca inicia com uma lista de URLs sementes a serem rastreadas pelo *Crawler*, o qual faz o *download* das páginas referentes a esses URLs de forma distribuída. Os dados estruturados representam os metadados, enquanto que os dados não estruturados representam o próprio material, o qual pode estar em diversos formatos (por exemplo, HTML, PDF, DOC). Os dados não estruturados são enviadas ao *compactador*, o qual comprime e armazena-os no repositório, de forma que cada URL é associado a um número de identificação único. Os dados estruturados são armazenados no repositório, mas de forma que possam ser uniformizados futuramente. Os dados das páginas, como origem

do servidor, última atualização, entre outros dados, são armazenados no *Repositório de procedência de dados*. O componente de *integração de esquemas* tem como objetivo criar um esquema integrado representativo, que represente a maioria dos metadados. Esse esquema representativo é armazenado no repositório de *metadados integrados*. Por fim, o componente *RankSee* terá como funcionalidade recuperar e ranquear os REAs, levando em consideração características intrínsecas de REAs os quais podem auxiliar na busca e atendam às requisições dos usuários.

O **Indexador** tem várias funções. Ele lê o repositório de dados, descompacta os documentos e os analisa. Cada documento é convertido em um conjunto de ocorrências de texto chamado de *hits*. Cada *hit* possui a palavra, a posição no documento, uma aproximação do tamanho da fonte e a capitalização (maiúsculo ou minúsculo). Esses *hits* são distribuídos na forma de um índice para frente (*forward index*) em um conjunto de "*barris*". O **Classificador** realiza, dentre outras operações, a geração de um índice invertido que é usado pela **Interface de consulta** conjuntamente com o Analisador Léxico e o RankSee.

Um dos componentes chaves do mecanismo de busca na Web é o *Crawler*. Na arquitetura do SeeOER foi proposto um *Crawler* que obtivesse metadados das páginas dos repositórios REAs, como também os dados de procedências das páginas capturadas. O funcionamento do *Crawler* é descrito no algoritmo apresentado na Figura 2, e detalhado a seguir.

```
entrada: S = páginas sementes, P_n = profundidade, E_n = esquemas de
             metadados, S_m=estado da máquina
1 recupera-estado(S_m)
2 fila-de-URLs\leftarrow S
3 repeat
       página \leftarrow x \in fila-de-URLs
       armazena-Proc(pagina_{dados})
5
       metadados_{pagina} \leftarrow leitura-metadados(página)
6
       if metadados_{pagina} \in E_n then
7
            armazena-DEst (\{metadados_{pagina}\} \not\subset repositorio-DEst)
8
       end if
9
       conteúdo \leftarrow Download(página)
10
       (\text{textos}, links_p, \text{estruturas}, ...) \leftarrow \text{Analisar}(\text{conteúdo})
11
       links = \forall w \in links_p \mid w \leq P_n
12
       fila-de-URLs ← Adicionar-novos-links (fila-de-URLs, links)
13
       salvar-estado(S_m)
14
15 until (fila-de-URLs \neq \emptyset);
```

Figura 2. Algoritmo de um Crawler por REAs

O *Crawler* recebe como entradas algumas páginas sementes do site, a profundidade, os esquemas de metadados e o estado da máquina. A profundidade, neste contexto, significa a quantidade máxima de níveis permitidas ao *Crawler* percorrer. O diretório raiz (/) é o nível zero e os sub-diretórios seguintes são os níveis maiores. O estado da máquina

representa a possibilidade do *Crawler* ser pausado. Inicialmente, o *Crawler* retoma o estado da máquina, o qual pode ser vazio no começo (linha 1). As páginas sementes são incluídas na fila (linha 2). O laço de repetição, em seguida, desenfileira os URLs até que não exista nenhum URL para ser retirado da fila (linha 3 à 14). Nesse laço de repetição, são feitas as seguintes ações: uma página x é atribuída da fila-de-URLs para ser analisada (linha 4); são armazenados dados da página para o repositório de procedência (linha 5); é feita uma leitura dos metadados da página escolhida (linha 6); Se os metadados pertencerem em algum esquema de metadados (linha 7), então esses metadados que ainda não existem no repositório são armazenados, sendo considerados dados estruturados (*DEst*) (linha 8). Todas as informações não estruturadas também são armazenadas para que possam ajudar na recuperação dos REAs (linha 10 e 11). Em seguida, os links capturados da página passam por uma seleção. São selecionados todos os links exceto aqueles que tiverem profundidade maior que o permitido (linha 12). Esses links são adicionados na fila-de-URLs, nessa inclusão apenas os links que ainda não estão na fila são adicionados (linha 13). O estado da máquina é salvo (linha 14).

Na seção 4 é descrito os principais padrões de metadados no contexto de REAs e como um mecanismo de busca na Web poderá encontrá-lo, desde o formato do metadado até a sua instanciação. Também será listado algumas URLs que foram utilizadas como sementes para o SeeOER.

4. Padrões de Metadados e as Sementes Utilizadas no SeeOER

O reconhecimento da necessidade de reutilização de materiais educativos gerou o desenvolvimento de padrões de metadados para compartilhamento e armazenamento de documentos [McClelland 2003]. Outrossim, foram inspecionados alguns repositórios de REAs, considerando os repositórios que armazenam REAs de diversos formatos de arquivos, como: imagens, animações, arquivos de áudio, vídeos e outros. Na Tabela 1 está resumida uma pequena porção desses repositórios a partir dos quais foi possível evidenciar a diversidade entre os repositórios, desde os formatos de arquivos até os tipos de metadados utilizados. O tipo de metadados *Personalizado* refere-se aos tipos de metadados não catalogados.

Nome do repositório	Principal formato de arquivos	Padrões de metadados	
Banco de Imagens Geográficas	imagens	ISO 19115:2003	
Flickr	imagans	EXIF e	
	imagens	Personalizado	
Biblioteca Digital de Ciência	animações	Não encontrado	
Banco Internacional	imagens e	DCMES	
de Objetos Educacionais	arquivos de áudio	DCMES	
Connexions	textos	DCMES, IMS, CNX,	
		MathML e OGP	
OCW-MIT	vídeos e textos	OGP e Personalizado	
Teses USP	textos	DCMES	
Khan Academy	vídeos	OGP e MathML	
e-Aulas USP	vídeos	Vídeo Sitemaps	

Tabela 1. Repositórios REAs com diferentes padrões de metadados

Desse modo, foram usados no desenvolvimento desta arquitetura cinco padrões de metadados, a saber: o Dublin Core Metadata Element Set (DCMES), IEEE Learning Object Metadata (IEEE/LOM), Protocolo *Open Graph* (OGP), MathML e Vídeo sitemaps.

O padrão IEEE/LOM foi desenvolvido em um esforço conjunto do Comitê de Padrões de Tecnologia de Aprendizagem em colaboração com DCMI e outras organizações [McClelland 2003]. A *Instructional Management System* (IMS) continuou os trabalhos da IEEE/LOM [Koutsomitropoulos et al. 2010] e criou seu próprio padrão, o qual é chamado IMS Learning Resource Metadata. Esse padrão é equivalente ao IEEE/LOM, sendo que a principal diferença refere-se à taxonomia utilizada nos metadados [Gimenes et al. 2012].

No caso do DCMES (ou Dublin Core), os repositórios REAs podem representar seus metadados para o público com o formato variado, assim o formato como os metadados são instanciados não é sempre o mesmo [Powell et al. 2007]. Por exemplo, o Connexions instancia seus metadados por meio do formato XML, enquanto que o repositório Teses USP instancia seus metadados por meio do formato HTML/XHTML. Outros formatos incluem: HTML/XML, RDF/XML, XML e DC-DS-XML.

O Protocolo *Open Graph* (OGP) foi criado pelo Facebook. Ele foi inspirado no Dublin Core, Microformatos e RDFa (RDF em HTML com adição de atributos). O OGP requer apenas 4 informações essenciais: o título, o tipo, uma imagem miniatura (*thumbnail*) e um URL [Graham 2012]. Porém, ele possui muitos outros metadados opcionais como metadados para localização (latitude, longitude) e para vídeos (resolução e o tipo), dentre outros [Facebook 2013].

O MathML foi recomendado pela W3C em 2010. Ele é uma especificação para descrever as expressões matemáticas nas páginas da Web [W3C 2011]. Um dos potenciais do MathML é a capacidade de melhorar a busca de termos técnicos na literatura e nos materiais educativos. Ao integrar as expressões matemáticas com o material em uma forma altamente estruturada, MathML possibilita a busca por palavra-chave de pesquisa matemática. Por exemplo, pode-se digitar uma equação em um motor de busca e obter uma lista de materiais nos quais essa equação existe. Assim, MathML pode ter um papel no reforço existente nos sistemas de busca voltado para metadados [Miner 2005].

O Vídeo sitemaps representa uma forma de se incluir metadados em vídeos e áudios. Ele é um arquivo XML que lista os URLs de um site junto com metadados adicionais sobre cada URL [Sitemaps 2008]. No entanto, esses metadados não são suficientes para descrever vídeos e áudios. Como resultado, o Google estendeu os atributos tradicionais do Sitemap e adicionou novos metadados para descrição de vídeos e áudios.

Futuramente pretendemos estender o projeto para capturar o padrão de metadados do LRMI (Learning Resource Metadata Iniciative).

Na Tabela 2 sintetiza os principais padrões encontrados na Web dos repositórios REAs. Como também, a instanciação que refere-se à possibilidade que pode ser escrito o padrão de metadados. Além disso, são mostrados os elementos descritores o qual faz referência aos campos que possibilita descrever os REAs. Por fim, a possibilidade de extensão do padrão o que possibilita adicionar mais elementos descritores.

A Tabela 3 exibe algumas sementes que foram utilizadas no SeeOER. Com essas

Padrão de metadados	Instanciação	Elementos descritores	Possibilidade de estender?		
OGP	1 formato	4 elementos	Sim		
Vídeo Sitemaps	1 formato	6 elementos	Sim		
Dublin Core	5 formatos	15 elementos	Sim		
IEEE/LOM	Não descrevem	60 elementos divididos	Sim		
	formatos	em 9 categorias			

Tabela 2. Heterogeneidade dos padrões de metadados encontrado na Web

sementes é possível ilustrar a extensão que os REAs estão se propagando pelo mundo. A tabela mostra as sementes de diferentes países (16 países), com diferentes nacionalidades, línguas e culturas.

Título do URL	País	URL
Banco Internacional de Objetos Educacionais	Brasil	objetoseducacionais2.mec.gov.br
Matemática Multimídia	Brasil	m3.ime.unicamp.br
RRU Open Educational Resources	Canadá	oer.royalroads.ca/moodle/
Open Educational Resources for Typography	Argentina	www.oert.org
The Le@rning Federation	Austrália	www.ndlrn.edu.au
Educar Chile	Chile	www.educarchile.cl
Banco de Objetos de Aprendizaje	Colômbia	aplicaciones.virtual.unal.edu.co
Eduteka	Colômbia	www.eduteka.org
RVP Metodicky Portal	Checa	dum.rvp.cz
Materialeplatformen	Dinamarca	materialeplatform.emu.dk
Open Science Resources	União Europeia	www.osrportal.eu
Edu Fi	Finlândia	www.edu.fi
FREIburger Multimedia Object Repository	Alemanha	freimore.uni-freiburg.de
Open Educational Resources (OER) Africa	Quênia	www.oerafrica.org
VCILT	Maurícia	vcampus.uom.ac.mu
Aljazeera creative commons repository	Catar	cc.aljazeera.net
Maknaz	Arábia Saudita	maknaz.elc.edu.sa
University of Leicester OER Repository	Reino Unido	www2.le.ac.uk/projects/oer
Connexions	Estados Unidos	cnx.org

Tabela 3. Algumas sementes utilizadas pelo SeeOER

Na próxima seção são citados os mecanismos de busca na Web por REAs disponíveis na literatura usados como trabalhos correlatos, bem como uma breve comparação entre o SeeOER e esses trabalhos correlatos.

5. Trabalhos Relacionados

O SeeOER visa avançar no estado da arte por meio de uma nova arquitetura para um mecanismo de busca na Web especificamente voltado à recuperação de REAs que resolve limitações existentes na literatura. Ele apresenta diferenciais, na resolução de conflitos em nível de esquema e em nível de instância oriundos do uso de diferentes padrões de metadados, repositórios e plataformas. Os principais trabalhos relacionados encontrados na literatura são discutidos a seguir.

O trabalho de Warpechowski (2005) define técnicas para a recuperação de metadados de objetos de aprendizagem (OA), com a mínima intervenção do usuário, resultando

na indexação e recuperação desses objetos. Essas técnicas são definidas com base na estrutura e funcionamento do AdaptWeb, que é um ambiente de aprendizagem que disponibiliza material instrucional.

Bissell et al. (2009) introduzem um mecanismo de busca que recupera recursos a partir de um conjunto de repositórios REAs que utilizam *feeds*. Os *feeds* fornecem uma lista de URLs que indicam onde determinados recursos podem ser encontrados. O *Crawler* desse mecanismo utiliza os *feeds* para realizar um rastreamento direcionado dos recursos dentro de cada repositório incorporado ao mecanismo. Ele recupera cada recurso e adiciona o seu conteúdo a um índice, que pode então ser usado para recuperar resultados relevantes a partir dos termos buscados.

OERScout é uma proposta recente de mecanismo de busca por REAs na Web [Abeywardena et al. 2013]. Ele recupera os REAs a partir de uma lista estática de repositórios.

Rathod e Cassel (2013) descrevem um mecanismo de busca na Web usando classificadores. O objetivo do mecanismo consiste em recuperar especificamente planos de ensino para Ciência da Computação e que sejam REAs.

Jorum é um mecanismo de busca por REAs na Web proposto pelo Comitê de Sistemas de Informação Conjunta (JISC) do Reino Unido, baseado no DSpace, *Elastic-Search* e SOLR. Este mecanismo também disponibiliza um serviço de armazenamento de materiais, o qual é restrito à comunidade institucional do Reino Unido. Para função de repositório, o Jorum utiliza especificamente a plataforma DSpace [Jorum 2014], com suporte apenas à licença Creative Commons. O padrão de metadados utilizado pelo repositório é o Dublin Core estendido.

Na Tabela 4 faz um resumo comparativo do SeeOER com os outros trabalhos correlatos.

	AdaptWeb Warpecho- wnski (2005)	DiscoverEd Bissell et al. (2009)	OERScout Abeywarde- na et. al (2013)	Syllabus Rathod e Cassel (2013)	Jorum (2013)	SeeOER
Recupera REAs a patir de um conjunto de repositórios		X	X	X	X	X
Considera repositórios brasileiros						X
Utiliza os padrões de metadados	X	X			X	X
Integra metadados e/ou utiliza procedência de dados						X

Tabela 4. Comparativo com os trabalhos correlatos

6. Análise dos resultados

A Figura 3 exibe os resultados obtidos pelo protótipo do SeeOER. Pode ser observado que o SeeOER obteve 23.618 que representa a maior quantidade de REAs indexados em comparação com os outros trabalhos correlatos. O OERScout obteve a menor quantidade de REAs indexados, com 1.999 REAs. Em seguida, o Jorum obteve a segunda maior

quantidade de REAs indexados, com 15.955 REAs. Atualmente, o Jorum é um mecanismo de busca na Web disponível para o público em geral, além de publicar como foi desenvolvido seu mecanismo de busca na Web.

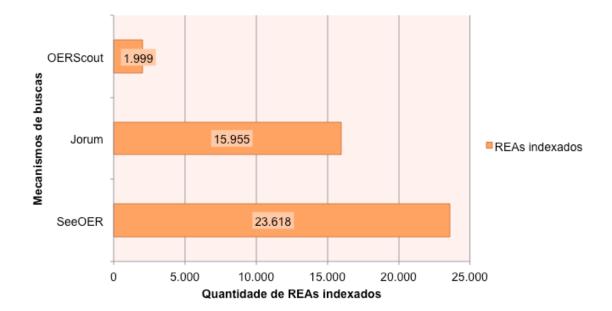


Figura 3. Resultado de indexação comparativo entre o SeeOER e os outros mecanismos de busca

7. Conclusões

De um ponto de vista mais abrangente, a pesquisa desenvolvida neste projeto de mestrado visa incentivar as práticas de utilização e produção de REAs na educação, pois propõe um mecanismo de busca especializado. Os mecanismos de busca na Web genéricos não consideram as especificidades de REAs e assim prejudicam a sua disseminação e incorporação em práticas educacionais. Vislumbra-se que os resultados deste projeto de mestrado terá um alto impacto social.

Os REAs possuem padrões de metadados, repositórios e plataformas com características particulares, assim introduzem heterogeneidades específicas. Os atuais mecanismos de busca na Web apresentam dois problemas importantes. Primeiro, eles são genéricos, assim buscam informação em qualquer lugar, desde páginas comerciais até definições escritas por pessoas anônimas (por exemplo, Wikipédia). Segundo, eles não levam em consideração as características intrínsecas de REAs. Os mecanismos de busca na Web específicos por REAs, como citado nos trabalhos correlatos, possuem diversas dificuldades para recuperar REAs na Web. Além disso, muitos repositórios REAs brasileiros não estão sendo identificados facilmente nos mecanismos de buscas genéricos e nem indexados nesses mecanismos de busca na Web específicos por REAs.

Portanto, o SeeOER foi desenvolvido para suprir as lacunas deixadas em aberto para recuperação de REAs na Web, como também incluir em cena os repositórios de REAs brasileiros. Como próximos passos, estão previstos testes de interação humanocomputador no mecanismo de busca e a publicação para o acesso do público em geral ao SeeOER.

Referências

- Abeywardena, I., Chan, C., and Tham, C. (2013). Oerscout technology framework: A novel approach to open educational resources search. *The International Review of Research in Open and Distance Learning*, 14(4).
- Bissell, A., Park, J., Yergler, N., and Linksvayer, M. (2009). Enhanced search for educational resources a perspective and a prototype from cclearn. Technical report, ccLearn.
- Facebook (2013). The open graph protocol. http://ogp.me/. Acessado: 2014-05-10.
- Gimenes, I. M., Barroca, L., and Feltrim, V. D. (2012). Tendências na educação a distância e educação aberta na computação. *CSBC*.
- Graham, W. (2012). Facebook developer tools. In *Beginning Facebook Game Apps Development*, pages 201–229. Apress.
- Jorum (2014). Find jorum. http://find.jorum.ac.uk. Acessado em: 2014-07-10.
- Koutsomitropoulos, D. A., Solomou, G. D., Papatheodorou, T. S., and Alexopoulos, A. D. (2010). The use of metadata for educational resources in digital repositories: Practices and perspectives. *D-Lib Magazine*, 16(1):3.
- Little, S., Mikroyannidis, A., Okada, A., and Scott, P. (2011). Formal metadata and shared experiences for discovering tools to adapt open educational resources. In *Signal-Image Technology and Internet-Based Systems (SITIS)*, 2011 Seventh International Conference on, pages 147–153.
- Matkin, G. W. (2013). Open educational resources in the post mooc era. *eLearn*, 2013(4).
- McClelland, M. (2003). Metadata standards for educational resources. *Computer*, 36(11):107–109.
- Miner, R. (2005). The importance of mathml to mathematics communication. *Notices of the AMS*, 52(5):532–538.
- Okada, A. (2007). Knowledge media technologies for open learning in online communities. *The International Journal of Technology, knowledge & Society*, 3.
- Powell, A., Nilsson, M., A., N., Johnston, P., and Baker, T. (2007). Dcmi abstract model. http://dublincore.org/documents/2007/06/04/abstract-model. Acessado em: 2013-09-11.
- Rathod, N. and Cassel, L. (2013). Building a search engine for computer science course syllabi. In *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries*, JCDL '13, pages 77–86, New York, NY, USA. ACM.
- Sitemaps (2008). What are sitemaps? http://www.sitemaps.org/. Acessado em: 2013-09-15.
- W3C (2011). W3c math home. http://www.w3.org/Math/. Acessado em: 2013-09-15
- Warpechowski, M. (2005). Recuperação de metadados de objetos de aprendizagem no AdaptWeb.