

Perspectivas em Ciência da Informação



Todo o conteúdo do artigo, exceto onde está identificado, está licenciado sob uma Licença Creative Commons. Fonte: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362015000300051&lng=en&nrm=iso. Acesso em: 5 abr. 2018.

REFERÊNCIA

SHINTAKU, Milton; DUQUE, Claudio Gottschalg; SUAIDEN, Emir José. Federações de repositórios: conceitos, políticas, características e tendências. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 20, n. 3, p. 51-66, jul./set. 2015. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362015000300051&lng=en&nrm=iso>. Acesso em: 5 abr. 2018. doi: <http://dx.doi.org/10.1590/1981-5344/2358>.

Federações de repositórios: conceitos, políticas, características e tendências

Milton Shintaku

Tecnólogo do Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT. Professor da Secretaria de Educação do Distrito Federal. Mestre em Ciência da Informação pela Universidade de Brasília. Doutor em Ciência da Informação pela Universidade de Brasília

Claudio Gottschalg Duque

Graduado em Licenciamento em Letras habilitação em Português e Alemão pela Faculdade de Letras da Universidade Federal de Minas Gerais. Mestre em Psicolinguística no Programa de Pós-Graduação em Estudos Linguísticos pela Faculdade de Letras da Universidade Federal de Minas Gerais. Doutor em Linguística Computacional no Angewandte Sprachwissenschaft und Computerlinguist - Justus-Liebig-Universität Giessen. Doutor em Produção e Gestão da Informação no Programa de Pós-Graduação em Ciência da Informação pela Escola de Ciência da Informação

Emir José Suaiden

Professor Titular da Faculdade de Ciência da Informação da Universidade de Brasília. Graduado em Biblioteconomia pela Universidade de Brasília. Mestre em Ciência da Informação pela Universidade Federal da Paraíba. Doutor em Ciência da Informação pela Universidad Complutense de Madri. Pós-doutor pela Universidad Carlos III de Madri

<http://dx.doi.org/10.1590/1981-5344/2358>

O surgimento das federações de repositórios, há mais de dez anos, tem apoiado a disseminação da literatura científica em acesso aberto. Entretanto, há pouca informação acadêmica sobre essas ferramentas. Nesse sentido, o presente estudo tem por objetivo apresentar informações sobre as federações, de forma a contribuir com esse tema, que ainda pouco explorado. Um estudo qualitativo, utilizando técnicas de pesquisa documental, para a verificação dos conceitos, políticas, características e tendências das federações disponíveis na internet. Os resultados revelaram consenso em vários aspectos, com um cenários colaborativo e pouco estudado, embora seja de interesse acadêmico, em que as federações fomentam

a criação de repositórios. Com isso, o estudo contribui com informações sobre as federações, apresentando um tema atual.

Palavras-chave: *Federação de repositórios; Conceitos; Políticas; Características; Tendências.*

Federation of repositories: concepts, policies, characteristics and trends

The emergence of federations of repositories, ten years ago, has supported the dissemination of scientific literature in open access. However, there is few scholarly information about these tools. Then, this study aims to present information about federations to contribute to this topic not yet explored. A qualitative study using techniques of documentary research, for the verification of concepts, policies, characteristics and trends of the federations available on the internet. The results revealed consensus in many aspects, with a collaborative scenarios and not too much studied, although it is of academic interest, in which the federation promoting the development of repositories. Thus, the study provides information on the federations, with a actual theme.

Keywords: *Federation repositories; Concepts; Policies; Features; Trends.*

Recebido em 15.03.2015 Aceito em 16.08.2015

1 Introdução

O movimento de acesso aberto tem alterado significativamente o processo de disseminação da informação científica, na medida em que possibilita o acesso ao conteúdo integral sem custos. A oferta aberta à informação científica depende da utilização de ferramentas apropriadas para isso, dentre as existentes, duas ferramentas tem se destacado, os periódicos de acesso aberto e os repositórios, no que Harnad *et al.* (2004) denominam de via dourada e via verde.

Os periódicos de acesso aberto publicam os novos conhecimentos, seguindo o processo tradicional de avaliação pelos pares, diferenciando dos outros tipos de periódicos por disponibilizar livremente a sua publicação via internet. A oferta de novos conhecimentos fica garantida de forma livre, alterando o negócio relacionado à publicação de artigos científicos.

Os repositórios por sua vez preservam e dão acesso a uma variedade de documentos científicos, inclusive os artigos publicados nos periódicos. Eles se tornaram importantes instrumentos para a disseminação da informação científica, visto que disponibilizam documentações científicas, que até recentemente apresentavam restrições de acesso, como as teses e dissertações.

As universidades e instituições de pesquisas brasileiras, devido as características citadas, têm criado repositórios, com o intuito de divulgarem e conservarem a produção acadêmica da instituição. O *Ranking Web of Repositories* apresenta 40 repositórios científicos brasileiros, sendo 37 desenvolvidos com o DSpace, que conforme Macedo *et al.* (2014), totalizam cerca de 48% dos repositórios brasileiros.

Os repositórios brasileiros de universidades e institutos de pesquisa disponibilizavam, em janeiro de 2014, mais de 420 mil documentos científicos em texto integral (Quadro 1), contabilizados por meio de *harvesting* nos 40 repositórios levantados por Macedo *et al.* (2014), utilizando o software Repox (desenvolvido e mantido pela Europeia - União Européia) e categorizados por meio do software VuFind (desenvolvido e mantido pela Universidade de *VillaNova* - Pensilvânia - Estados Unidos). O Brasil se destaca na América Latina em relação ao número de repositórios, visto que a Colômbia possui 22 repositórios, Argentina 19, México 14 e Chile 9, conforme levantamento dos autores.

Quadro 1 – Distribuição, por tipo, dos Documentos Científicos em Repositórios Brasileiros

Quantitativo de documentos em repositórios brasileiros por tipo								
Dissertação	Artigo	Tese	Artigo de evento	Outro	Livro	Relatório	Capítulo de livro	Total
176.922	115.423	63.648	30.751	23.189	1.324	839	692	412.788

Fonte: Dados da pesquisa.

O cenário atual é promissor para o desenvolvimento e consolidação de repositórios, promovendo a livre disseminação da produção científica brasileira. Entretanto, cada repositório disponibiliza a produção da sua instituição, de maneira independente, criando suas próprias políticas de armazenamento e indexação. Os repositórios mantêm certo isolamento, criando ilhas de conhecimentos. Uma forma de solucionar esse isolamento seria integrar os repositórios em uma federação por meio da interoperabilidade, preceito dos arquivos abertos.

2 Federações de repositórios

Federação de repositórios científicos, mesmo que seu conceito ainda não esteja bem estabelecido, possui algumas características próprias, ofertando serviços de busca consolidada, por meio de uma base de dados coletadas, oriundas dos repositórios. Cada repositório que fornece metadados à federação possui estrutura própria e oferta serviços do seu acervo, ou seja, é independente, por isso, a denominação de federação.

No âmbito dos sistemas comerciais de informações de turismo, Garita *et al.* (2013) definem federação de repositórios como uma base de dados que mantém toda a informação necessária aos serviços federados. Mesmo com certa diferença contextual, esse conceito possui certo alinhamento com o entendimento de Putmann (2001), que na tecnologia considera a federação como uma comunidade que contempla vários domínios, juntando para compartilhar recursos.

Van de Sompel *et al.* (2006) sugerem que a criação federações de repositório aumente a oferta serviços acadêmicos, serializando repositórios num fluxo de obtenção, coleta e disposição de objetos digitais, apoiado na interoperabilidade. Com isso, garante uma infraestrutura de compartilhamento de objetos digitais acadêmicos, de forma mais globalizada.

Assim, a federação é uma rede, em que seus partícipes possuem autonomia, mas que se agregam em uma entidade maior para ofertar serviços mais amplos. No que concerne à Ciência da Informação, Weitzel (2006) considera as federações como fontes terciárias, em que os periódicos são fontes primárias por publicarem os artigos, os repositórios são secundários por disponibilizarem uma cópia e as federações são terciárias por disseminarem o que estão nos repositórios.

Federação está intimamente relacionada aos sistemas federados, contemplando bases de dados heterogêneas, que, por meio de esquemas federados, integra informações oriundas de outras bases de dados (NICOLLE; SIMON; YÉTONGNON, 2005). Nota-se que segundo os autores, a federação transcende de serviço federado para sistema federado, escala que aumenta a complexidade e oferta de serviços.

A Ciência da Informação e a Ciência da computação, quando envolvidas na discussão de sistemas colaborativos, declaram que, a federação de repositórios é um *locus* que armazena informações, apoiando o desenvolvimento e distribuição de produtos informatizados (ELMARAGHY, 2006). Essa assertiva dá margem a dois entendimentos, um mais prático, tecnicista, informático e outro mais filosófico, teórico mais informacional.

Em um aspecto mais estrutural, Aschenbrenner *et al.* (2010) delineiam as federações em camadas, discernindo entre as ferramentas e o desenho da federação. Destacam a possibilidade de oferta de serviços que transcendem aos repositórios, tais como a análise de conteúdo inter-repositórios, rastreamento de objetos digitais, verificação de colaboração, entre outros.

O presente artigo tem por objetivo identificar e discutir os principais conceitos, políticas, características e tendências das federações com o intuito de contribuir com a discussão sobre essas ferramentas atuais de apoio à disseminação da informação científica. Assim, apoiar trabalhos futuros sobre as federações, pois como os repositórios científicos, até certo ponto, se apresentam estáveis do ponto de vista metodológico e conceitual, a integração por meio de federações torna-se a evolução natural em um ambiente dinâmico com tendências a integração como da internet.

3 Metodologia

O presente estudo possui abordagem qualitativa, objetivos com viés de análise conceitual, alinhando ao que Denzin e Lincoln (2006) enfatizam. Destaca-se que a pesquisa qualitativa observa a natureza social contextualizada dos fenômenos. Com isso, permite a verificação multifacetada do fenômeno social, em um processo de estudo mais profundo.

O estudo utiliza a técnica da pesquisa documental, considerando as páginas dos portais das federações como documentos, unindo marcos nas pesquisas sociais e humanas com o contexto tecnológico atual. A pesquisa documental é característica das ciências sociais e humanas por ter nas coleções de documentos fontes importantes de pesquisa, pois como ressaltam Sá-Silva, Almeida e Guindani (2009) são na maioria base de trabalho de investigação. Da mesma forma que atualmente a internet se apresenta como uma fonte importante de informações, na medida em que grande parte das instituições, organizações, órgãos e outros possuem portais disponíveis na internet.

Da mesma forma, o presente estudo reuniu documentação publicada oriunda do Portal da Capes, *Google scholar*, *Google*, *Library Information Science Abstract (LISA)* e *Web of Science*. Utilizou-se argumentos de busca exatos como “*repository federation*”, “*Repository integration*” e “*repository confederation*”. Descartaram-se os textos que não tratavam de federações acadêmicas, como as de objetos educacionais ou de softwares, por exemplo, por não se tratar do foco da pesquisa.

Outros argumentos buscados foram sobre as federações estudadas, criando um corpus documental utilizado para análise das variáveis qualitativas como: conceito, características, políticas, problemas e tendências. Em ambas as buscas foram recuperadas poucas publicações, de forma que os documentos mantidos nos portais das federações foram as maiores fontes de informação.

Nesse contexto, o presente estudo analisa os portais das federações como coleção de documentos, em que identifica a auto-conceituação, complementado com informações de textos acadêmicos. Dessa forma, possibilita verificar a auto-conceituação, políticas aplicadas, problemas e tendências, objetivos do estudo.

4 Resultados e discussão

Federações de repositórios científicos não se apresentam comuns na internet. O levantamento efetuado revelou dez federações de repositórios científicos, localizadas em vários países, sendo três no Brasil. Esta constatação reforça questões relacionadas à inovação do tema e os desafios técnicos e tecnológicos em agregar repositórios. Uma vez que indica que se tornam iniciativas quase únicas por país. Assim, o presente estudo analisou as seguintes federações, apoiados em algumas publicações e nos portais:

- a) *Networked Digital Library Theses and Dissertation (NDLTD)*¹: federação americana criada em 1996 agrega repositórios e outras federações, restrita a trabalhos de conclusão de cursos. Publicações como: Fox *et al.* (1997) e Hussain *et al.* (2001);
- b) *Biblioteca Digital Brasileira de Teses e Dissertações (BDTD)*²: federação brasileira criada em 2002 agrega repositórios de teses e dissertações. Publicações como: Kuramoto (2006), Southwick (2006) e Macedo *et al.* (2014);
- c) *Portal Brasileiro de Acesso Aberto à Informação Científica (oasisbr)*³: federação brasileira criada em 2006 agrega a BDTD e outras fontes de informações científicas brasileiras. Publicações como: Silva, Alencar e Souza (2014);
- d) *Repositório Científico de Acesso Aberto de Portugal (RCAAP)*⁴: federação pertencente a um projeto amplo de integração de repositórios portugueses criada em 2008 agrega principalmente repositórios acadêmicos portugueses. Publicações como: Ribeiro e Pinto (2009) e Carvalho *et al.* (2010);
- e) *LaReferencia*⁵: federação latinoamericana, formada por instituições nacionais de nove países criada em 2010 (Argentina, Brasil, Chile, Colômbia, Equador, El Salvador, México, Peru e Venezuela) agrega a produção científica dos países membros;
- f) *Sistema Nacional de Repositorios Digitales (SNRD)*⁶: federação argentina criada em 2011, , agrega repositórios científicos da Argentina. Publicações como: Bongiovani e Nakano (2011) e Ministério da Ciência, Tecnologia e Inovação Produtiva (ARGENTINA. Ministério da Ciência, Tecnologia e Inovação Produtiva, 2011).
- g) *Biblioteca Digital Colombiana (BDCOL)*⁷: Federação Colombiana criada em 2007 agrega repositórios científicos da Colômbia. Publicações como: Otero e Gutiérrez (2011);
- h) *Red Mexicana de Repositorios Institucionales (REMEDI)*⁸: Federação mexicana criada em 2012 agrega repositórios científicos mexicanos. Publicações como: Tapia (2013);
- i) *Acceso Libre a Información Científica para la Innovación (ALICIA)*⁹: federação peruana criada em 2013 agrega

¹ Disponível em: <<http://www.ndltd.org/>>. Acesso em: 11 set. 2015.

² Disponível em: <<http://bdtb.ibict.br/>>. Acesso em: 11 set. 2015.

³ Disponível em: <<http://oasis.ibict.br/>>. Acesso em: 11 set. 2015.

⁴ Disponível em: <<http://www.rcaap.pt/>>. Acesso em: 11 set. 2015.

⁵ Disponível em: <<http://lareferencia.redclara.net/>>. Acesso em: 11 set. 2015.

⁶ Disponível em: <<http://repositoriosdigitales.mincyt.gob.ar:8380/dnet-web-generic/>>. Acesso em: 11 set. 2015.

⁷ Disponível em: <<http://www.bdcol.org/>>. Acesso em: 11 set. 2015.

⁸ Disponível em: <<http://www.remedi.org.mx/>>. Acesso em: 11 set. 2015.

⁹ Disponível em: <<http://alicia.concytec.gob.pe/alicia/>>. Acesso em: 11 set. 2015.

repositórios científicos Peruanos. Publicações como: Peru (2015);

j) Repositório da Produção Científica do Conselho dos Reitores das Universidades Estaduais Paulistas (CRUESP)¹⁰: Federação restrita brasileira criada em 2013 agrega os documentos dos repositórios Institucionais da Universidade de São Paulo, Universidade de Campinas e Universidade Estadual Paulista Júlio de Mesquita. Publicações como: Ferreira *et al.* (2013).

4.1 Sobre os conceitos

O estudo revelou uma variada presença de nomenclatura, como biblioteca digital, rede, sistema, repositório e portal que reflete a própria condição da internet, em que a variação terminológica pode apresentar um problema, mas também apresenta uma riqueza a ser estudada. Com isso, apresenta conceitos diversos, mas similares e complementares. Em relação a conceitos, tem-se que a federação pode ser considerada como:

Esforço internacional (NDLTD), regional (LaReferencia) ou unificação de iniciativas isoladas (BDTD), formando colaborativamente uma rede de repositórios ou bibliotecas digitais (LaReferencia, SNRD e REMERI), utilizando um sistema para coletar os metadados (BDCOL), para ofertar em um ponto único de acesso (RCAAP) a documentação, com a utilização de uma ferramenta de busca (oasisbr, CRUESP).

Nota-se que as federações mais antigas como a NDLTD (criada em 1996), BDTD (criada em 2002), BDCOL (criada em 2007) denominam-se de bibliotecas digitais, termo muito em voga no período. Já as mais novas assumem nomenclatura mais genérica como: sistema para SNRD (criado em 2011), rede para REMERI (2012) e acesso livre para o ALICIA (criada em 2013).

Independente da nomenclatura ou ponto de destaque, todas as federações têm por objetivo potencializar a visibilidade dos documentos referenciados em seu acervo, dos autores e das instituições parceiras. Para tanto, coletam, indexam, preservam e dão acesso à informações coletadas de repositórios, por meio de um portal de serviços. As federações também ofertam esses metadados para outras federações.

4.2 Sobre as características

Quanto às características, as federações estudadas apresentaram uma variedade de pontos. Destaca-se que em grande parte dos casos as características são as mesmas, por serem federações com objetivos muito próximos, praticamente similares. Entretanto, cada federação apresenta uma contextualização que a difere, assim, podem-se apresentar algumas características distintas, como:

¹⁰ Disponível em: <<http://cruesp.sibi.usp.br/>>. Acesso em: 11 set. 2015.

- a) ser criada por regulamentação governamental (ALICIA e SNRD);
- b) ser mantida por instituição pública (ALICIA, SNRD, BDTD, oasisbr, RECAAP, CRUESP);
- c) ter uma organização própria para manutenção (NDLTD, REMERI, LaReferencia, BDCOL)
- d) utilização de padrões internacionais para interoperabilidade como: Dublin Core (LaReferencia, oasisbr, SNRD, ALICIA, REMERI, RECAAP, CRUESP), ETD-MS (NDLTD), ou baseados em Dublin Core (BDTD e BDCOL);
- e) fomentar a criação de repositórios (ALICIA, BDTD, BDCOL, CRUESP, NDLTD, RECAAP, REMERI, SNRD, LaReferencia, oasisbr)
- f) instituições parceiras são colaborativas e voluntárias, requerendo apenas atender aos critérios determinados (ALICIA, BDTD, BDCOL, NDLTD, RECAAP, REMERI, SNRD, LaReferencia, oasisbr);
- g) possuem portal de busca consolidada (ALICIA, BDTD, BDCOL, NDLTD, RECAAP, REMERI, SNRD, LaReferencia, oasisbr, CRUESP);
- h) ofertam dados para outras federações (ALICIA, BDTD, BDCOL, RECAAP, REMERI, SNRD, oasisbr);
- i) hospedam repositórios (ALICIA e RECAAP).

A colaboração é a tônica das federações, conforme identificado nesse estudo. Esse é um dos pontos que mais caracterizam as federações, visto que a adesão a uma federação é totalmente voluntária e que cada membro federado mantém a sua independência. Com isso, tem-se um modelo em que a federação é constituída de membros independentes e autônomos, formando uma rede colaborativa, ofertando serviços que aumentam a visibilidade mútua.

Assim, o estabelecimento da federação se dá por meio da interoperabilidade e um portal de busca consolidada, que também a caracteriza no que tange a questões mais específicas com relação a técnicas e tecnológicas adotadas. Nesse sentido, o uso de esquemas mais simplificados como o Dublin Core ainda se faz presente na interoperabilidade com o uso do protocolo OAI-PMH, revelando o mesmo modelo de operacionalização do início das federações com a NDLTD.

O estudo identificou que as federações se tornam promotoras da criação de repositórios e bibliotecas digitais, ofertando apoio técnico. Com isso, esses sistemas apoiam a disseminação da informação, não apenas por agregar a informação coletada, mas no apoio ao desenvolvimento de repositórios institucionais. Assim, tornam-se centros de divulgação do acesso aberto, apoiando os repositórios membros.

Algumas federações apresentam pontos comuns, porém, em poucos casos há certa unanimidade nas características. Observou-se que é nas diferenças que se apresentam as riquezas desses sistemas, o que pode revelar-se um problema, tendendo a crescer com o avanço da tecnologia e do volume de documentos armazenados.

4.3 Sobre as políticas

As políticas são orientações de cada uma das federações, que remete ao seu funcionamento. Nesse sentido, a política mais presente nas federações diz respeito à tipologia documental a ser coletada e indexada. Nesse ponto, mesmo sendo federações de cunho científico, há certa divergência, visto que atualmente não há consenso na definição e entendimento do que venha a ser 'documentação científica', ainda mais com o movimento da Ciência Aberta. Assim a tipologia documental segundo as políticas apresentadas pelas federações é:

- a) documentos científicos tradicionais avaliados pelos pares (ALCIA, BDTD, BDCOL, NDLTD, RECAAP, REMERI, SNRD,
- b) LaReferencia, oasisbr, CRUESP), como os artigos de periódicos;
- c) teses e dissertações (ALCIA, BDTD, BDCOL, NDLTD, RECAAP, REMERI, SNRD, LaReferencia, oasisbr, CRUESP);
- d) dados de pesquisa (ALICIA, REMERI);
- e) documentação técnica (ALICIA, LaReferencia, REMERI, BDCOL, SNRD);
- f) documentos culturais e históricos (BDCOL, REMERI);
- g) objetos educacionais (REMERI, BDCOL).

Em todas as federações ocorrem os tradicionais documentos avaliados pelos pares (artigos, teses e dissertações, etc), uma vez que são federações de repositórios institucionais acadêmicos, sendo um dos critérios para criação dos acervos. Entretanto, outras tipologias também são coletadas e disseminadas por parte das federações, apresentando-se mais flexíveis, o que reflete diretamente na forma de coleta e aos tipos de repositórios membros. Essas características novamente fomentam questões sobre o que é e o que não é documentação científica atualmente, principalmente no contexto do Movimento da Ciência Aberta.

Algumas federações já coletam registros sobre dados de pesquisa, alinhados ao movimento de Dados Abertos, da mesma forma que coletam documentação técnica e objetos educacionais, que possuem menor relevância científica no sentido tradicional. Entretanto, essas outras tipologias documentais revelam-se importantes como produção acadêmica e intelectual.

Coletar e disseminar tipologias não científicas tradicionais coloca em cheque algumas políticas mais restritivas das federações, tendo em vista

as mudanças nos movimentos científicos, como os Dados Abertos, Ciência Aberta e e-science, que expande significativamente os conceitos do Acesso Aberto, centrado nos artigos, teses e dissertações. Revelam a necessidade de estudos para adequação das federações aos novos movimentos.

Nota-se, no entanto, que a política da federação relacionada à tipologia documental pode ser revista e alterada a qualquer momento, incluindo novas tipologias, adequando-se às tendências ou necessidades e, assim, afetar os repositórios membros. Da mesma forma, alterações nas políticas dos repositórios podem, em alguns casos, afetar as federações, ao passo que passam a gerenciar e ofertar registros de outros tipos de documentos. Esse sistema de inter-relacionamento revela-se útil a ambos.

4.4 Sobre os problemas

O levantamento dos problemas apresentados nas federações possibilita planejar modelos ou ajustes que proponha soluções adequadas. Assim, contribuir com as federações existentes, apresentando soluções a problemas, que em alguns casos são comuns. Entretanto, na maioria dos casos, muitos dos problemas apresentados são oportunidades de estudos ou desafios a serem suplantados. Os problemas identificados foram:

- a) qualidade dos dados coletados (BDTD, NDLTD);
- b) questões de indexação linguística (NDLTD, LaReferencia, RECAAP);
- c) navegabilidade (REMEDI, CRUESP);
- d) problema de busca (SNRD);
- e) duplicidade de registros (oasisbr).

Como as federações possuem fluxo de informação simplificado, composto pelos processos de coleta, indexação e recuperação, pode-se categorizar os problemas conforme essas etapas. Com isso, facilitar a análise, ao ponto que os problemas podem ter implicações se estiver relacionados a um mesmo processo.

Nota-se que não foi apresentado nenhum problema relacionado à coleta de metadados, visto que é um processo simples. Por meio do uso de um protocolo de comunicação comum, a federação coleta os metadados dos repositórios membros, num processo de transferência. Assim, os problemas que afetam essa etapa geralmente impedem a operacionalização da federação.

A indexação, por outro lado, revela problemas, pois determina como será a recuperação da documentação, afetando diretamente a precisão e a revocação dos sistemas que indexam os repositórios. Três dos principais problemas apresentados tem relação com esse processo são:

1)quanto à qualidade dos dados: a padronização de campos e a normalização de conteúdos se apresentam como fatores de problemas na indexação. Registros que não apresentam campos obrigatórios, por exemplo, impedem a plena indexação, gerando descarte do registro em muitos casos. Variações de conteúdos em muitos campos influenciam negativamente na recuperação ou categorização. Campos que necessitam de formatos específicos (data, números etc) ou restrição de valores (sim/não, tese/dissertação etc) muitas vezes apresentam grandes variações, influenciando na indexação.

2)quanto às questões linguísticas: a diversidade de registros com idiomas distintos revela-se como problemas às federações, que intrinsecamente possuem registros de várias fontes. Da mesma forma que as homônimas são problemas complicados, que mesmo com ferramentas de web semântica ainda não tem solução condizente. Assim, indexadores mais semânticos resolveriam em parte o problema, mas atualmente grande parte dos indexadores é sintática.

3)quanto à duplicidade de registros: Esse problema ocorre geralmente quando dois registros iguais vêm de fontes distintas, por isso não sendo problema do processo de coleta. Nesse sentido o indexador pode resolver com ferramentas mais sofisticadas de indexação que detectam a duplicidade e apresenta como um único registro, com mais que uma fonte.

A recuperação na forma de busca e apresentação é o processo de maior interação com os usuários e apresentou nas seguintes funcionalidades:

a)quanto à busca: geralmente os portais de busca ofertam a busca simples e a avançada, cada qual com suas especificidades, com oferta de operadores booleanos e outros, não apresentando problemas;

b)quanto à navegabilidade: os portais devem ofertar a possibilidade dos usuários navegar pelo acervo, ofertando listas categorizadas e facetadas que apoiem o usuário a navegação pelo acervo. Entretanto, nem todas as federações ofertam essa funcionalidade.

Os problemas de recuperação, independente do tipo, são resolvidos com ferramentas de descoberta e entrega que, na maioria dos casos, ofertam todas as funcionalidades para o processo de recuperação. Por isso, grande parte das federações tem adotado essas ferramentas em seus portais.

Por fim, um dos problemas não mencionado explicitamente, mas que se apresenta latente em todas as federações, refere-se à habilidade

de se manter tecnologicamente atualizado, visto a dinamicidade da Ciência da Computação. Requer, na maioria dos casos, equipe de informática dedicada à federação, com processo de prospecção contínua. Assim, esse problema tem relação íntima com a sustentabilidade da federação.

4.5 Sobre as tendências

As tendências são altamente dinâmicas devido às mudanças técnicas e tecnológicas. Da mesma forma, as políticas adotadas podem apoiar certas tendências, visto que muitas vezes possuem alinhamento teórico. Nesse ponto, a identificação e análise das tendências possibilitam verificar um cenário mais provável e quais os planejamentos e políticas que algumas federações estão deliberando. Assim, as tendências identificadas são:

- a) uso de sistema de descoberta e entrega em software livre para o desenvolvimento do portal de busca (ALICIA, BDCOL, LaReferencia). A CRUESP utiliza uma versão paga;
- b) manter objetos de aprendizagem alinhando ao movimento de objetos educacionais livres (BDCOL, REMERI);
- c) uso de mais de um esquema de metadados, apresentando flexibilidade na interoperabilidade (BDTD);
- d) operação em ambiente heterogêneo, com uso de padrões internacionais (ALCIA, BDTD, BDCOL, NDLTD, RECAAP, REMERI, SNRD, LaReferencia, oasisbr);
- e) modelo de sustentabilidades amplo (REMERI, LaReferencia).

O uso de software livre tem se tornado comum e, por isso, uma grande quantidade de tipos de ferramentas tem sido disponibilizada, incluindo os softwares de descoberta e entrega que as federações utilizam. Com isso, os portais de busca consolidada das federações ofertam formas de recuperação mais adequada a sistemas de informação que tratam com grandes quantidades de registros.

Softwares de descoberta e entrega, como defendem Vaughan (2011) e Sutradhar (2014), são apropriados para manipular grandes quantidades de informação, próprios de sistemas atuais, evolução que as bibliotecas desejam e de forma intuitiva como desejam os usuários. Assim, tornam-se naturalmente as ferramentas mais apropriadas para as federações que agregam grandes quantidades de informações e requerem serviços ágeis de busca.

Outro ponto importante é a integração em cenário heterogêneo, em que cada repositório membro da federação pode ser desenvolvido por um tipo de ferramenta. A federalização é possível utilizando ou não padrões mais rígidos, requerendo apenas que sejam conhecidos. Essa questão

permite a flexibilização necessária à operacionalização mais ampla, com a possibilidade de adesão de uma variedade de tipos de ferramentas.

Dentre todas as tendências, provavelmente uma das mais controversas repousa na questão da sustentabilidade. Mesmo que as federações sejam mantidas por instituições públicas, há vários pontos relacionados à sustentabilidade que merecem discussão. No que se refere à questão financeira, a REMERI e LaReferencia revelam modelos de sustentabilidade em que membros e entidades públicas e privadas contribuem com cotas. A NDLTD e BDTD, por exemplo, que possuem mais tempo revelam sustentabilidade, no sentido de longevidade e ser interessante aos membros, por ofertar apoio, modelo que as outras federações têm seguido.

5 Conclusão

As federações se apresentam altamente colaborativas, visto que a BDTD teve influência e apoio da NDLTD e o LaReferencia fomentou a criação da BDCOL, SNRD e ALICIA, ofertando orientações e compartilhando ferramentas. O oasisbr e o RCAAP possuem relação de compartilhamento de acervo, numa relação bem mais interativa. Revela-se um ambiente, até certo ponto, amistoso, pois possuem a mesma finalidade de facilitar o acesso à documentação científica.

No que se refere à abrangência, tem-se federações globais (NDLTD), regionais (LaReferencia, oasisbr e RCAAP), nacionais (BDTD, BDCOL, SNRD e ALICIA) ou mesmo mais restritas como o da CRUESP. Da mesma forma que se tem federações mais restritas em relação aos tipos de documentos disseminados como a BDTD e a NDLTD e outras mais flexíveis como a BDCOL. Nesse ponto, o estudo verificou que as federações mais antigas têm características mais restritivas, no que diz respeito à tipologia da documentação disseminada.

Por fim, o tema "federação de repositórios científicos" se apresentou pouco estudado e sem muita literatura acadêmica ou técnica/comercial. no Brasil e no exterior. Revela-se um assunto recente, como objeto de estudo científico. Destaca a pouca presença nos modelos de comunicação científica, mesmo sendo uma realidade. Em modelos muito detalhados como o de Bjork (2005) ou de previsão de futuro como a de Hurd (2000), as federações ainda não são representadas.

Com isso apresenta-se como desafio e oportunidade de estudos futuros, ante a necessidade de formalizações de iniciativas na internet. Marcondes e Sayão (2008) consideram a pesquisa sobre integração de sistemas autônomos um grande desafio, por envolver quase todos os tópicos estudados pela Ciência da Informação. Torna-se, dessa forma, oportunidades de estudos futuros.

Referências

ARGENTINA. Ministério da Ciência, Tecnologia e Inovação Produtiva. Resolução nº 469, de 17 de maio de 2011. *Boletim Oficial Republica Argentina*, Buenos Aires, 4 out. 2014.

ASCHEBRENNER, A. *et al.* Towards an open repository environment. *JoDI. Journal of digital information*. University of Texas Libraries. v. 11, n.1, 2010. Disponível em: <<https://journals.tdl.org/jodi/index.php/jodi/article/view/758>>. Acesso em: 1 dez. 2014.

BONGIOVANI, P. C.; NAKANO, S. Acceso abierto en Argentina: la experiencia de articulación y coordinación institucional de los repositorios digitales en ciencia y tecnología. *Revista de ciencia, educación, innovación y cultura apoyadas por Redes de Tecnología Avanzada*, v. 1, n. 2, p. 163-179, Julio/Diciembre 2011.

BJÖRK, B. C. *Scientific communication life-cycle model*. 2005. Disponível em: <<http://oacs.shh.fi/publications/Model35explanation2.pdf>>. Acesso em: Dez. 2014.

CARVALHO, J. *et al.* Repositório científico de acesso aberto de Portugal. In: GOMES, M. J.; ROSA, F. (Orgs). *Repositórios institucionais democratizando o acesso ao conhecimento*. Salvador: EdUFBA, 2010. p. 127-152.

DENZIN, N. K.; LINCOLN, Y. S. *Planejamento da pesquisa qualitativa: teorias e abordagens*. Porto Alegre: Artmed, 2006, p. 367-388.

ELMARAGHY, H. A. A complexity code for manufacturing systems. In: INTERNATIONAL CONFERENCE ON MANUFACTURING SCIENCE AND ENGINEERING, MSEC 2006, 2006. Ypsilanti, Michigan. *Anais...* Ypsilanti, Michigan. American Society of Mechanical Engineers, 2006.

FERREIRA, S. M. P. *et al.* Repositório da produção científica CRUESP: mais do que um consórcio, um trabalho integrado USP, UNESP e UNICAMP. In: CONFOA, 4., São Paulo, 2013. *Anais...* São Paulo: SIBI-USP, 2013.

FOX, E. *et al.* NDLTD: encouraging international collaboration in the academy. *DESIDOC Bulletin of Information Technology*, v. 17, n. 6, p. 45-56, Nov. 1997.

GARITA, C. *et al.* A service interface definitions catalogue for virtual Interprises in tourism. In: MARIN, V.; CAMARINHA-MATOS, L. M.; AFSARMANESH, H. (Eds). *Knowledge and technology integration in production and services: balancing knowledge and technology in product and service life cycle*. Nova Iorque: Springer, 2013, p. 97-108.

HARNAD, S. *et al.* *The access/impact problem and the green and gold roads to open access*. 2004. Disponível em: <<http://eprints.soton.ac.uk/265852/2/serev-revised.pdf>>. Acesso em: mar. 2013.

HURD, J. M. The transformation of scientific communication: a model for 2020. *Journal of the American Society for Information Science and Technology*, v. 51, p.1279-1283, 2000.

HUSSAIN, S. *et al.* Networked Digital Library of Theses and Dissertations. *Dlib*, v. 7, n. 9, 2011. Disponível em: <<http://www.dlib.org/dlib/september01/suleman/09suleman-pt1.html>>. Acesso em: Dez. 2014.

KURAMOTO, H. Informação científica: proposta de um novo modelo para o Brasil. *Ciência da Informação*, v. 35, n. 2, p. 91-102, 2006.

MACEDO, D. J. *et al.* Theses and dissertations digital library: ten years of Open Access and Open Archives in Brazil. In: OPEN REPOSITORIES 2014, 2014. Helsink. *Anais...Helsink*, National Library of Finland, 2014. Disponível em: <<http://www.doria.fi/handle/10024/97679>> . Acesso em: 1 dez. 2014.

SAYÃO, L. F.; MARCONDES, C. H. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. *Transinformação*, Campinas, v. 20, n. 2, 2008.

NICOLLE, C. SIMON, J. C.; YÉTONGNON. K. Interoperability of B2B Applications : methods and tools, advanced topics. In: KENG, S. (Ed.). *Database research*. Hershey, Pensilvania: Idea Group Publishing, 2005. v. 4. p. 297-299.

OTERO, A. P. S.; GUTIÉRREZ, M. M. B. Biblioteca Digital Colombiana (bdcol): ¿unportal de acceso abierto a la información?. *Revista de Derecho, comunicaciones y nuevas tecnologías*, Bogota, v5, n.1, p. 1-29. 2011.

PERU. Ley nº 30035, de 24 de janeiro de 2015, que regula el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto. *Diario Oficial El Peruano*, Lima, 24/01/2015. Disponível em: <http://portal.concytec.gob.pe/images/stories/images2013/portal/areas-institucion/dsic/reglamento_repositorio_nacional_alicia.pdf>. Acesso em: 1 fev. 2015.

RIBEIRO, F.; PINTO, M. M. G. A. *O acesso aberto à investigação em ciência da informação em Portugal: alcance e impacto*. 2009. Disponível em: <<http://repositorio-aberto.up.pt/bitstream/10216/57297/2/63193.pdf> 2009>. Acesso em: 1 out. 2014.

SÁ-SILVA, J. R.; ALMEIDA, C, D.; GUINDANI, J. F. Pesquisa documental: pistas teóricas e metodológicas. *Revista Brasileira de História & Ciências Sociais*, ano 1, n. 1, página inicial e final?, jul. 2009. Disponível em: <www.rbhcs.com>. Acesso em: 01/10/2014.

MARCONDES, C. H.; SAYÃO, L. F. Softwares livres para repositórios institucionais: alguns subsídios para a seleção. In: SAYÃO, L. *et al.* (Org.).

Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 2009. p. 23-54.

SILVA, A. P. L. da; ALENCAR, L. L.; SOUZA, J. S. de. Iniciativas de acesso aberto à informação científica: uma abordagem sobre o portal brasileiro de acesso aberto à informação científica - Oasisbr. *In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS (SNBU), 18., Belo Horizonte, 2014. Anais...* Belo Horizonte: SNBU, 2014. Disponível em: <https://www.bu.ufmg.br/snbu2014/trabalhos/index.php/sn_20_bu_14/sn_20_bu_14/paper/view/305>. Acesso em: 23 nov. 2014.

SOUTHWICK, S. B. The Brazilian electronic theses and dissertations digital library: providing open access for scholarly information. *Ciência da Informação*, v. 35, n. 2, p. 103-110, 2006.

SUTRADHAR, B. One Day National Workshop on Web-Scale Discovery Services: Single Window Access to Library e-Resources (workshop), 2014. Disponível em: <http://www.library.iitkgp.ernet.in/sites/bookFair/wkp/OneDayNationalWorkshop.pdf>. Acesso em: 1 nov. 2014.

TAPIA, R. V. La Red Mexicana de Repositorios Institucionales – REMERI: Un proyecto nacional para compartir y dar visibilidad al conocimiento. PKP Scholarly Publishing Conference 2013.

VAN DE SOMPEL, H., Lagoze *et al.* An Interoperable Fabric for Scholarly Value Chains. *D-Lib Magazine*, 12(10). 2006. doi:10.1045/october2006-vandesompel ; <http://dx.doi.org/10.1045/october2006-vandesompel>

VAUGHAN, J. Web scale discovery services. *Library Technology Reports*, v. 47, n.1, p. 5-21, January 2011.

WEITZEL, Si. da R. O papel dos repositórios institucionais e temáticos na estrutura da produção científica. *Em Questão*, Porto Alegre, v. 12, n. 1, p. 51-71, jan./jun. 2006.