

VOCABULÁRIO CONTROLADO EM REPOSITÓRIOS DE INSTITUIÇÕES DE ENSINO SUPERIOR BRASILEIRO: estudo comparativo

Guilhermina de Melo Terra
Tainá Rebelo Vasconcelos

RESUMO: O propósito deste artigo é apresentar o resultado do estudo comparativo estabelecido entre dois repositórios institucionais já conhecidos pela comunidade acadêmica, a fim de subsidiar a construção do repositório do Departamento de Arquivologia e Biblioteconomia da Universidade Federal do Amazonas (DAB). Tal necessidade se deu devido à agregação do curso de Arquivologia ao Departamento de Biblioteconomia. Assim, o número de professores e alunos cresceu e, com isso, o número de produções acadêmicas. Devido à natureza dos cursos que passaram a compor o DAB, além da guarda informacional, cita-se a divulgação das informações produzidas como sendo uma das atividades ligadas ao cotidiano acadêmico dos mesmos. Portanto, torna-se cada vez mais importante, para o DAB, contar com a criação de uma ferramenta que tanto consiga representar as informações contidas nas produções técnico-científicas, quanto permita a sua recuperação de forma segura, eficiente e eficaz. Nesta perspectiva, visou-se como objetivo geral, estabelecer um estudo comparativo entre dois sistemas de representação da informação, subsidiando a construção de um vocabulário controlado próprio para a produção técnico-científica do DAB. Para isso, tomaram-se como base os seguintes objetivos específicos: apresentar a importância da linguagem documentária para o processo de recuperação bibliográfica; demonstrar a relevância do vocabulário controlado para a representação e recuperação da informação e; analisar os sistemas de recuperação da informação desenvolvidos por duas Instituições de Ensino Superior, tomando por base o vocabulário controlado utilizado nos repositórios *online*. Para isso, optou-se por um estudo de abordagem qualitativa, baseado na pesquisa bibliográfica para a construção da fundamentação teórica. Para a parte operacional, pelo método comparativo, a partir do estudo de dois repositórios já existentes, deu-se início à parte prática desta investigação, a fim de servir como base para a construção do repositório do DAB. Conclui-se o quanto é relevante para a UFAM manter-se ativa perante a comunidade acadêmica, em nível nacional e internacional.

Guilhermina de Melo Terra
guilherminaterra@gmail.com
<http://lattes.cnpq.br/3223150635170651>

Professora Assistente da Universidade Federal do Amazonas (UFAM) do Departamento de Biblioteconomia. Doutora em Museologia pela Faculdade de Letras da Universidade do Porto (Portugal). Mestre em Sociedade e Cultura na Amazônia pela UFAM. Graduada em Biblioteconomia pela UFAM.

Tainá Rebelo Vasconcelos
tainarebelo@hotmail.com
<http://lattes.cnpq.br/4947974731898284>
Graduação em Biblioteconomia

Submetido em: 11/03/2015
Publicado em: 14/06/2015

PALAVRAS-CHAVE: Sistemas de representação da informação. Vocabulário controlado. Repositório. Prateleira virtual.

1 INTRODUÇÃO

O curso de Biblioteconomia da Universidade Federal do Amazonas (UFAM) foi criado no ano de 1966, por meio da Resolução n. 29/66, de 14.11.66, do Conselho Universitário. Desde então, a sua estrutura curricular vem se modificando, com o propósito de não só acompanhar seu contexto histórico, mas também atender as novas demandas do mercado.

Atualmente, o curso de Biblioteconomia encontra-se inserido no Departamento de Arquivologia e Biblioteconomia (DAB). Isso implica frisar que a quantidade de alunos, professores e produções acadêmicas do departamento aumentou significativamente, pois além do aumento quantitativo de sujeitos produtores do conhecimento, as estruturas curriculares dos dois cursos incluíram o Trabalho de Conclusão de Curso (TCC) como uma exigência obrigatória.

Diante o exposto, caberá ao DAB à responsabilidade de guardar a produção técnico-científica dos alunos de graduação e pós-graduação, bem como dos seus professores. Devido à natureza dos cursos que compõe o departamento em tela, além da guarda informacional, cita-se a divulgação informacional como sendo uma das atividades ligadas ao cotidiano acadêmico. Portanto, torna-se cada vez mais importante, para o DAB, contar com uma ferramenta que consiga tanto representar as informações contidas nas produções técnico-científicas, quanto permita a sua recuperação de forma segura, eficiente e eficaz.

Seguindo esta linha de raciocínio, esta investigação, partiu da seguinte problemática: “De que forma o DAB poderá disseminar a produção técnico-científica de seus alunos e professores?”. A resposta a tal questionamento liga-se à seguinte hipótese: “Com a criação de um repositório online, acredita-se que o DAB conseguirá, qualitativamente, disseminar a produção técnico-científica de seus alunos e professores”. Neste sentido, buscou-se com esta investigação apresentar ao DAB a análise de repositórios *online* de renomes e confiáveis, tomando por base o vocabulário controlado apresentado por tais repositórios, a fim de se compreender como as organizações responsáveis pelos repositórios faziam uso do processo de indexação.

A indexação é aqui mencionada, devido seu papel junto à representação da informação a ser construída nos instrumentos de busca, já que o processo linguístico necessitará ser compatível com a linguagem dos usuários, ou seja, apresentar o mesmo processo de significação, possibilitando uma busca mais precisa da informação que, por sua vez, levará para outros universos os resultados dos estudos desenvolvidos pelos alunos e professores dos cursos de Arquivologia e Biblioteconomia. Nesta perspectiva, objetivou-se estabelecer um estudo comparativo entre dois sistemas de representação da

informação já conhecimentos pela comunidade acadêmica, subsidiando a construção de um vocabulário controlado próprio para a produção técnico-científica do DAB.

Para isso, apresentou-se um estudo sobre linguagem documentária, a fim de demonstrar sua importância junto ao processo de recuperação bibliográfica. Também se sentiu a necessidade de estabelecer um estudo sobre vocabulário controlado para a representação e recuperação da informação, no sentido de trabalhar sua importância para, depois, trabalhar os repositórios institucionais, a fim de ser iniciado o processo de análise dos sistemas de recuperação da informação desenvolvidos por duas Instituições de Ensino Superior, tomando por base o vocabulário controlado utilizado nos seus repositórios *online*, de modo a ser apresentado ao DAB não só o funcionamento, mas também os aspectos positivos dos instrumentos estudados.

Procurou-se apresentar tal estudo, pois desde o lançamento da proposta em tela, buscou-se contribuir com a construção de um instrumento próprio a ser utilizado tanto pela comunidade acadêmica da UFAM, quanto pelos demais interessados na temática voltada para os campos da Arquivologia e Biblioteconomia, pode-se afirmar o quanto este estudo foi relevante para área científica, uma vez que o resultado deste projeto irá permitir a continuação do processo operacional do departamento, no que diz respeito à construção de seu próprio repositório, o qual também será disponibilizado por meio da internet. Para que este estudo fosse realizado e atingisse o resultado esperado, o desenho metodológico adotado tomou como base os seguintes procedimentos:

- a) Levantamento bibliográfico sobre linguagem documentária e sistema de representação da informação;
- b) Estudo bibliográfico sobre o material bibliográfico selecionado;
- c) Seleção dos repositórios *online* que se tornariam objeto de estudo;
- d) Estudo, individualmente, os repositórios escolhidos como objeto de estudo;
- e) Identificação da metodologia de representação da informação que mais se adequava ao contexto do DAB;
- f) Esboço da metodologia de representação da informação a ser aplicada no futuro repositório do DAB.

Diante o desenho metodológico apresentado, sentiu-se a necessidade de construir a fundamentação teórica a partir de estudos sobre: 1. Linguagem documentária, por estas serem consideradas um tipo de ferramenta mediadora entre a linguagem dos sistemas de informação e os usuários; 2. Vocabulário controlado, devido a sua importância no tratamento dos documentos informacionais, juntamente com a representação de assuntos, possibilitando a recuperação

padronizada da informação; 3. Arquitetura da informação, devido à importância de se construir e aplicar sistemas de representação e recuperação do conhecimento e; 4. Repertórios institucionais, a fim de trabalhar os aspectos conceituais, a apresentação dos instrumentos escolhidos, juntamente com o processo de análise dos mesmos. Em seguida, foram abordados as questões metodológicas, os resultados e discussões e, por fim, os aspectos conclusivos deste estudo.

2 LINGUAGEM DOCUMENTÁRIA

Com base no pensamento de Cintra et al. (2002), observou-se que com o término da II Guerra Mundial, mais precisamente a partir da segunda metade do século XX, o processo referente à "explosão informacional" se tornou algo significativo. Mediante tal fato, as pesquisas científicas ganharam mais espaço e com elas, também, os estudos sobre as Linguagens Documentárias (LD), uma vez que tais linguagens são consideradas como uma ferramenta mediadora entre a linguagem dos sistemas de informação e os usuários (VEIGA, 2006).

Em nível conceitual, as LD são linguagens artificiais, controladas, criadas dentro dos objetivos de uma organização/setor, a partir de um conjunto de documentos e domínio, para serem utilizadas na indexação e recuperação da informação em um determinado sistema de recuperação da informação (SOUZA, 2007). Isso implica frisar que as LD apresentam uma ponte de comunicação entre o sistema utilizado e seu usuário, visando sua funcionalidade como instrumento fundamental no processo da representação da informação disponibilizando-a de modo conciso ao usuário. Onde as LDs desenvolvem uma função importante no que se diz respeito à representação das informações, pois as mesmas se tornam documentos para que possam ser recuperados e utilizados gerando a disseminação e formando novos conhecimentos.

Diante do exposto, podem-se considerar as LDs como sendo aquelas utilizadas na organização das informações, tendo por base a integração entre a linguagem natural e a linguagem artificial, de modo a representar áreas temáticas, possibilitando a representação, identificação e acesso às referidas informações. Melhor dizendo, as LDs passam a ser concebidas pela literatura como "instrumentos utilizados para representar o conhecimento de uma dada área do saber", permitindo sua busca por parte do pesquisador (CAMPOS, 2001, p.17).

No modo de facilitar a comunicação entre a linguagem natural dos usuários e a unidade de informação, visando representar conteúdo dos documentos, para estabelecer um padrão em um sistema de recuperação. Para que tal fato aconteça, é sabido que

o vocabulário inserido no sistema de representação da informação necessita ser consoante ao vocabulário dos usuários que farão uso do referido sistema, já que as LD possuem duas funções: representar o conteúdo dos documentos e mediar à recuperação desta informação (BOCCATO, 2005).

Desta forma, as linguagens documentárias permitem a comunicação entre sistema e usuário, sendo o principal instrumento no processo da representação da informação, disponibilizando a informação de modo adequado ao usuário. A LD traz consigo a definição dos vocabulários controlados em sistemas de classificação bibliográfica, as listas de cabeçalhos de assuntos e os tesouros que são as estruturas organizadas em termo, símbolos ou regras para intitular um documento, ou seja, formalizar o uso para a linguagem.

Sales (2007) complementa que as linguagens documentárias são linguagens artificiais, constituídas de sistemas simbólicos, que traduz de forma sintética, o assunto dos documentos para a indexação, armazenamento e recuperação da informação.

Segundo ele as linguagens visam da uniformidade as linguagens de especialistas, visando uma representação padronizada do conteúdo documental, e trazendo a recuperação mais pertinente. A definição mais completa de linguagem documentária proposta por Tálamo (1997, p. 10) é a que segue:

De maneira geral, define-se linguagem documentária (LD) como uma linguagem construída, oposta à natural, portanto, que tem como objetivo específico tratar a Informação para fins de recuperação. Atualmente as questões relativas à sua construção são tratadas pela Lingüística Documentária e aquelas relativas ao seu uso encontram-se integradas às questões mais amplas relativas ao tratamento e à recuperação da informação, discutidas no âmbito da Análise Documentária (AD).

Portanto, a LD é fundamental, uma vez que a sua qualidade de disseminação da informação depende da linguagem que é utilizada pelo sistema que indexa e recupera os documentos. Isso se da quando a linguagem captada traduz a linguagem do autor sem que se perca a ideia principal.

3 VOCABULÁRIO CONTROLADO

Os primeiros vocabulários controlados começaram a surgir para atender a demanda na organização do acervo, pois com o passar do tempo esse acervo crescia cada vez mais e entendeu-se que precisaria organizá-los para facilitar seu processo de recuperação (MOREIRA; MOURA, 2006). E diante da necessidade de verificação de diversos documentos em diversas áreas do conhecimento foi que surgiu essa necessidade de

diversificação do mesmo. O vocabulário controlado ajuda no tratamento dos documentos informacionais, que tem por função representar os assuntos de documentos e possibilitar recuperação mais precisa da informação, pois é uma lista de termos padronizados. Além de ser um instrumento que aparece da necessidade de instituições informacionais, onde os centros de documentação constroem linguagens próprias e específicas.

Kobashi (2008) justifica que o vocabulário controlado é uma linguagem artificial constituída de termos organizados em estrutura relacional, que é usado na elaboração que padroniza e facilita a entrada e a saída de dados em um sistema de informações. Ele por sua vez atua como uma ferramenta no processo de indexação dos documentos, que após a leitura e a interpretação dos textos utilizará descritores para descrever concisamente o conteúdo dos documentos.

Com o intuito de esclarecer a principal função do vocabulário controlado, podemos determinar que a sua estrutura de indexação destina-se, segundo Lancaster (2004, p. 14), a: “1) controlar sinônimos optando por uma única forma padronizada, com remissivas de todas as outras formas; 2) Diferenciar homógrafos; 3) reunir ou ligar termos cujos significados apresentem uma relação mais estreita entre si”.

Para que o mesmo seja um instrumento de locomoção de conceitos, visando sua eficácia tanto no tratamento, quanto na recuperação, é necessária a definição de seu universo notacional, pois Segundo Cintra et al. (2002, p. 49), “todo e qualquer conhecimento corresponde a um conjunto de noções que lhe é próprio”. Por isso o vocabulário controlado deve estar inserido em um universo conceitual muito bem delimitado, para que sua ação de tradução de conceitos seja determinada em sua unidade onde será feita de forma contextualizada.

Motta (1987) aborda que o vocabulário controlado é estruturado em 3 (três) níveis: o tesauro, o cabeçalho de assunto e os sistemas de classificação. Nesta perspectiva, o primeiro nível de linguagem documentária nomeado tesauro que, de acordo com Campos (2001), surgiu da palavra “thesaurus” que é de origem latina e significa “tesouro”, tendo sido utilizada no sentido de “tesouro de palavras”. Um Tesauro de Recuperação e um instrumento que reúne conceitos de uma determinada área do conhecimento relacionados entre si.

O uso de um tesauro em um sistema de informação contribui muito para a diminuição da inconsistência na recuperação da informação em uma base de dados, segundo Souto (2003), ao se fazer uma pesquisa pelo termo não aceito, o usuário, utilizando um tesauro, além de ter maiores chances de encontrar um documento, conhecerá outras formas de pesquisa. Pois, a estrutura hierárquica do tesouro facilita a compreensão do

assunto e de seu contexto, que esclarece os relacionamentos de conceitos e termo dentro de uma área do conhecimento. Tendo em vista o grande número de atualizações de um instrumento de indexação e que por mais eficiente que seja, há de ser sempre monitorado, conforme ressalta Kobashi (2008, p. 3):

É preciso reconhecer, no entanto, que as LDs embora úteis, são imperfeitas. Sua atualização permanente é sempre um desafio. É necessário, desse modo, encontrar formas de atualização e adaptação que sigam mais de perto a velocidade e a dinâmica da criação terminológicas para que, de fato, seja garantida a sua função comunicacional.

Contudo, os tesouros são elementos fundamentais em um sistema de recuperação da informação, contribuindo tanto para os usuários quanto para os indexadores no processo de classificação, portanto ambos conseguem se auxiliar por meio de uma estrutura de termos, para localizá-los de uma maneira que se adéquem na representação do assunto. Já o segundo nível citado temos os cabeçalhos de assuntos que representam os assuntos sob a forma de cabeçalhos estruturados. Eles também podem ser listas de termos, mas com um maior domínio sobre os termos adicionando relações diretas entre eles.

Almeida (2011) ressalta que o cabeçalho de assunto é uma LD pré-coordenada exigindo como regras para o estabelecimento de uma padronização adequada, ajudando na classificação e indexação dos assuntos sendo como sua principal função a organização determinada do acervo por assunto. Por sua vez, o fato de ser pré-coordenada permite atingir maior precisão, dispensando artifícios complexos, dispendiosos e nem sempre eficientes, necessário em linguagens pós-coordenadas, tais como ligações de coordenação e indicadores de função. O cabeçalho de assunto, sendo uma linguagem controlada, antecipa regras fixas abordadas em listas de controle que dizem a respeito, sobretudo na forma de entrada direta ou invertida e uso de singular e plural.

Por isso por ser uma linguagem pré-coordenada, quanto mais específica for uma lista, maior será o grau de pré-coordenada que irá usar. Portanto, deve-se apresentar um vocabulário, uma lista de termos aprovados para a representação de assuntos e regras formais para o uso deste termo. E no terceiro nível de vocabulário controlado, encontram-se os sistemas de classificação, que tem por objetivo organizar os acervos de bibliotecas e facilitar o acesso dos usuários à informação. Pode-se, então, mencionar os diversos tipos de esquemas utilizados, que são: Cutter, Ranganathan e o da Biblioteca do Congresso. Sendo que os mais utilizados nas bibliotecas são: a Classificação Decimal Dewey (CDD) e a Classificação Decimal Universal (CDU).

Para isso, os usuários necessitam encontrar um sistema de informação que os permita navegar pelo mesmo, de modo a manejar facilmente as informações, recuperando-as a fim de satisfazer suas necessidades informacionais. Neste aspecto, os descritores que serão utilizados no sistema necessitam manter a mesma linguagem do público, cujo instrumento se destina, pois à medida que a linguagem apresentada no sistema for apropriada à linguagem da clientela que dele fizer uso, o resultado obtido será muito mais qualitativo.

4 REPOSITÓRIOS INSTITUCIONAIS

Como é sabido, nos dias atuais, os sistemas de informação e de comunicação, mais do que nunca, passaram a fazer parte do cotidiano das pessoas, uma vez que o papel destes sistemas foram se modificando com o passar do tempo. Afirma-se isso, pois, inicialmente, tais sistemas eram utilizados, exclusivamente, a partir de seus aspectos técnicos, passando para o aspecto gerencial, no sentido de contribuir com o processo decisório, pois se voltam para o processo de coleta, recuperação, processamento, armazenamento e disseminação das informações.

Mediante tal fato, corrobora-se com Souza (2000, p. 162), quando se afirma que tais sistemas “hoje permeiam e mesmo viabilizam virtualmente todas as atividades humanas, e não mais se pode conceber a sociedade sem a acentuada imbricação das tecnologias de informação que nela surgem e que a modificam”. Em nível conceitual, os sistemas de informação e comunicação passam a ser concebidos como instrumentos mediadores entre a informação e as pessoas que dela necessitam, pois a tarefa primordial dos mesmos volta-se para a representação, armazenamento, organização e acesso da informação, como defendido por Baeza-Yates e Ribeiro Neto (1999).

Nesta perspectiva, García Jiménez (2004) frisa o quanto estudos sobre os sistemas de representação e recuperação da informação se tornam urgentes, no que se refere à construção e aplicação destes sistemas no processo de recuperação do conhecimento produzido pelo meio acadêmico. Devido aos objetivos esperados por este estudo, corrobora-se com as ideias do autor, pois no âmbito do DAB da UFAM, a riqueza bibliográfica produzida pelos alunos e professores dos cursos de Arquivologia e Biblioteconomia não pode ficar disponível apenas ao contexto local, valendo a pena divulgá-la em um repositório *online*, com o propósito de disseminar para outros contextos a produção técnico-científica dos corpos docente e discente do DAB e, sobretudo, da UFAM.

Por repositórios digitais (RDs), entendem-se como sendo as bases de dados *online* que reúnem de maneira organizada a produção científica de uma instituição ou área temática, a qual pode ser armazenada e disseminada em diversos formatos. Por esta razão, pode-se afirmar que os RDs proporcionam uma série de benefícios tanto para os pesquisadores, quanto para as instituições ou sociedades científicas, proporcionando maior visibilidade aos resultados de pesquisas e possibilitam a preservação da memória científica de sua instituição.

Cabe aqui salientar que os RDs podem ser de dois tipos: institucionais ou temáticos. Os repositórios institucionais lidam com a produção científica de uma determinada instituição, enquanto que os repositórios temáticos se voltam para a produção científica de uma determinada área, sem limites de institucionais. É crescente o número de instituições brasileiras e estrangeiras que têm criado seus repositórios digitais como intuito de garantir o armazenamento e acesso permanente e confiável de sua produção científica gerada em seu âmbito para aumentar a visibilidade e uso dos respectivos documentos. No Brasil, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) é o responsável por estabelecer uma política nacional de acesso livre à informação científica.

Segundo Waetei, Mouraai e Mangue (2012, p. 196), a principal meta do acesso livre volta-se para o processo o qual permita “a livre utilização dos resultados da pesquisa científica contida em revistas especializadas”, no sentido de permitir uma maior circulação do conhecimento produzido pela comunidade científica, contribuindo positivamente, com as melhorias do meio e não o enfraquecimento da qualidade mínima exigida das produções apresentadas, sobretudo, nos formatos tradicionais, pois mesmo recebendo a titulação “acesso livre” que os repositórios passarão a ser simples depósitos onde, qualquer pessoa possa inserir suas produções, sem o menor critério e exigência científica.

Afirma-se isso, pois, com base na literatura, os repositórios de acesso livre, na sua essência, manterão todas as exigências inerentes à publicação científica contidas nos periódicos científicos tradicionais. A única diferença é que este instrumento permitirá aos pesquisadores obter acesso às produções de forma aberta e sem restrições, em decorrência da Internet (SANZ-VALERO; VEIGA DE CABO; CASTIEL, 2007; CANESSA; ZENNARO, 2008). Nesta perspectiva, defende-se a ideia do quanto à política de acesso livre é importante para o meio científico, pois garantirá que a produção científica seja alcançada de todos, só trazendo, portanto, benefícios à sociedade como um todo, pois tem em definitivo, a eliminação das barreiras, sobretudo financeiras que, muitas vezes, distanciavam as produções científicas da comunidade acadêmica. Assim, optou-se em estudar os 02 (dois)

repositórios a seguir, a fim de servir como base para a construção do futuro repositório do DAB.

4.1 Lume Repositório Digital da UFRGS

O Lume Repositório Digital da Universidade Federal do Rio Grande do Sul (UFRGS) foi criado em Janeiro de 2008 e é o portal de acesso às coleções digitais produzidas no âmbito da Universidade, é de interesse da Instituição centralizar sua preservação e difusão. O Lume objetiva: reunir, preservar, divulgar e garantir o acesso confiável e permanente aos documentos acadêmicos, científicos e administrativos produzidos na Universidade, ampliando o uso desses recursos. Os documentos digitais que integram as coleções podem conter texto, imagem, vídeo e áudio e são, em sua maioria, de acesso aberto.

O repositório é composto por metadados e objetos digitais, sendo que os metadados são informações estruturadas que possuem semântica padronizada e são utilizados para preservar as informações digitais de maneira bibliográfica. Oferece serviço de busca por autor, título e assunto. A busca feita por assunto pode ser digitando o assunto desejado ou as primeiras letras do mesmo. E isso acontece com a busca por autor e título.

Oferece também uma busca mais detalhada que apresenta uma espécie de questionário para que seja feita a pesquisa específica como o acervo, que pertence às diversas unidades acadêmicas e demais órgãos da universidade, os trabalhos acadêmicos e técnicos que são trabalhos de conclusão de curso de especialização e graduação produzido no âmbito da UFRGS, a produção intelectual que inclui artigos de periódicos, livros, capítulos de livros e trabalhos de eventos realizados na UFRGS e/ou por autores da universidade, assim como por teses e dissertações e os trabalhos de conclusão de mestrado profissional produzidos na universidade e fora dela.

O Lume é um repositório que integra documentos digitais oriundos de diversas fontes. O principal sistema de informação que têm seus documentos disponibilizados para a comunidade por meio do Lume é o Sistema de Automação de Bibliotecas (SABi), onde realiza a gerência integrada das atividades e serviços das 33 bibliotecas da UFRGS, seu sistema é composto por módulos que são responsáveis pela aquisição, catalogação, periódicos, circulação de coleções, empréstimo, renovação, devolução e reservas, além de todo o processo de automação das rotinas das bibliotecas que por sua vez conta com a coordenação de uma comissão técnica específica voltada para tais atividades. Seu padrão adotado é voltado para as normas internacionais para registro de dados bibliográficos, o

MARC 21 que se volta para um conjunto de códigos e designações de conteúdos definidos para codificar registros que serão interpretados por máquinas. Além de analisar três bases de dados: bibliográfica, de autoridades e administrativas.

4.2 Biblioteca Digital de Teses e Dissertações da Universidade de São Paulo

A Biblioteca Digital de Teses e Dissertações (BDTD) da Universidade de São Paulo (USP) foi inaugurada em 2001, com o intuito de viabilizar via Internet o conhecimento produzido pelos trabalhos defendidos na Universidade, permitindo que outras comunidades brasileira e internacional pudesse ter acesso tanto na versão impressa, quanto a digital completa das teses e dissertações.

A Biblioteca Digital é uma oportunidade para exibição dos trabalhos, de forma rápida e fácil. Podendo ser acessada em qualquer lugar do mundo com acesso disponível à Internet. Cabe frisar que a mesma está associada a uma iniciativa global reconhecida pela UNESCO, a *Networked Digital Library of Theses and Dissertations* (NDLTD), garantindo maior confiabilidade e abrangência, ela também está associada ao IBICT do Ministério da Ciência e Tecnologia, através da Biblioteca Digital Brasileira de Teses e Dissertações.

A USP é a maior universidade brasileira, e por sua vez é reconhecida tanto nacional, quanto internacionalmente pela sua excelência em pesquisa e desenvolvimentos, méritos que são reconhecidos na comunidade acadêmica. Seu repositório digital possui a opção de pesquisa simples e avançada, onde a opção de busca simples aparece somente uma janela para a pesquisa de qualquer forma que você deseja colocar o assunto, já na opção de busca avançada surge para o usuário um questionário com os operadores booleanos e, ou, e não em seguida o assunto e depois onde você quer encontrar esse assunto como em autor, título, resumo e etc. Essa pesquisa pode ser feita em todos os documentos, ou em uma unidade específica como em teses, dissertações e outras.

Seu acervo está dividido em ano, área do conhecimento, autor, documentos, orientador e unidades onde você pode fazer sua pesquisa também por essa ferramenta, sendo que de forma limitada, pois ela irá mostrar apenas os documentos conforme o acervo selecionado.

5 METODOLOGIA

A investigação foi estabelecida a partir de uma abordagem quantitativa, por ter procurado estabelecer um estudo sobre os dois repositórios, no sentido de extrair o melhor dos dois. Segundo Pinto e Pinheiro (2003, p. 323), a abordagem qualitativa:

é pautada no pensar a realidade estudada visando a sua compreensão e a sua descrição de modo rigoroso, onde o pesquisador busca produzir, entendendo o indivíduo como a parte ativa no processo de investigação e, permitindo-lhe a interferência com maior profundidade nas respostas.

Quanto aos fins, o estudo construiu-se a partir das pesquisas exploratória e descritiva, devido estabelecerem um estudo, baseado na análise, no registro, na interpretação e na análise das características, fatores ou variáveis que se relacionam com o fenômeno, mas sem a interferência do grupo de pesquisadores (MEZZAROBA; MONTEIRO, 2005). A sua valorização está baseada na premissa que os problemas podem ser resolvidos e as práticas podem ser melhoradas através de descrição e análise de observações objetivas e diretas. Portanto sendo aqui uma pesquisa de cunho, pois a mesma tem o intuito apenas de observação e análise, sem o objetivo de intervir na realidade.

No que se refere aos procedimentos operacionais, à investigação fez uso da pesquisa bibliográfica. De acordo com Lakatos e Marconi (2001), a pesquisa bibliográfica obedece ao levantamento de toda a bibliografia já publicada, em forma de livros, revistas, publicações avulsas e imprensa escrita. Quando em relação a esta forma de pesquisa, é importante salientar que a mesma, tem a finalidade de fazer com que o pesquisador entre em contato direto com todo o material escrito sobre um determinado assunto, auxiliando o mesmo na análise de suas pesquisas ou na manipulação de suas informações. Ressaltando que a pesquisa bibliográfica pode ser considerada como o primeiro passo de toda a pesquisa científica.

6 ANÁLISE DOS RESULTADOS

Em face dos resultados das atividades desenvolvidas na busca de contribuir com a apresentação de informações voltadas para a construção do futuro repositório a ser desenvolvido pelo DAB, apresentar-se-á o quadro abaixo, com a devida análise:

Quadro 1: Estudo comparativo dos repositórios

REPOSITÓRIOS	USP	UFRGS
Software / Plataforma	Dspace	DSpace
Comitê Gestor	Conselho Supervisor do SIBi	Possui comitê gestor não especificado na política.
Obrigatoriedade de Depósito	Não. É uma recomendação	Não cita obrigatoriedade
Auto Arquivamento	Possui o mecanismo de auto arquivamento	Possui o mecanismo de auto arquivamento
Objetivos	Promover e assegurar coleta, tratamento e preservação da produção intelectual, bem como a disseminação e acessibilidade.	Promover livre acesso às informações no âmbito da Universidade, maximizar visibilidade, uso e impacto da produção, facilitar gestão da informação em meio digital, preservar memória institucional
Responsabilidade	Conselho Supervisor do Sistema Integrado de Bibliotecas (SIBi): Estabelecer e validar normas para coleta, tratamento, preservação da produção intelectual, definir tipos de documentos para depósito. Bibliotecas em relação à BDPI: registrar tecnicamente a produção intelectual, edição, revisão, validação e disponibilização online da mesma, capacitação, fornecer dados, estatísticas, informações. Departamento técnico do SIBi: gerencia, atualiza sistema de gestão, gera dados e indicadores sobre produção intelectual, estatísticas, cria diretrizes para garantir controle e preservação.	Equipe técnica do Lume - Comitê Gestor do Lume
Tipo de Documento	Resolução	Portaria
Povoamento	Pela equipe da Biblioteca, importação de dados ou autor (autoarquivamento)	Autoarquivamento, importação do Sistema de Automação de Bibliotecas da UFRGS (SABi)
Metadados	Quando a produção não estiver disponível em formato digital, os metadados deverão ser registrados na BDPI e um exemplar deve ser depositado na biblioteca.	Estão associados aos objetos digitais encontrados nas coleções. Seguem padrões internacionalmente aceitos, definidos pela equipe do Lume. No caso de autoarquivamento, os metadados são registrados pelo autor ou responsável pela comunidade, sob a orientação da equipe do Lume.

Fonte: Elaborado pelas autoras, 2015

Tomando por base o estudo comparativo entre os repositórios em tela, chegou-se a seguinte conclusão:

6.1 Software/ Plataforma

Os repositórios da UFRGS e USP utilizam a plataforma Dspace para reunir, organizar e armazenar materiais em formato digital, sendo que esta plataforma permite o depósito de documentos em qualquer formato (texto, vídeo, áudio e dados), disponibilizando os conteúdos na Internet e o mantendo preservado a longo prazo.

6.2 Comitê gestor

O responsável pelo repositório da Universidade de São Paulo é Conselho Supervisor do Sistema Integrado de Bibliotecas (SIBi). O conselho supervisor estabelece e valida normas para coleta, tratamento e preservação da produção intelectual geradas na universidade, sejam elas impressas ou digitais, definindo os tipos de documentos para depósito. A Biblioteca Digital da Produção Intelectual (BDPI), é responsável por reunir a produção intelectual da USP, aumentando a visibilidade, acessibilidade e difusão da produção intelectual.

O comitê gestor do Lume é responsável por incluir comunidades e coleções no repositório além de realizar procedimentos de avaliação. A política não especifica por quem é composto o comitê gestor, citando apenas as atribuições do mesmo.

6.3 Obrigatoriedade de depósito

Na USP existe uma política que recomenda a todos os membros da universidade para publicar seus resultados de pesquisa, preferencialmente em fontes que se encontrem em livre acesso ou que façam constar seus contratos de publicação a permissão para depósito na BDPI.

Já na Universidade Federal do Rio Grande do Sul não cita obrigatoriedade de depósito, onde apresenta apenas quesitos a serem obedecidos para que o depósito seja realizado, devendo o autor poder e estar disposto a conceder à universidade o direito não exclusivo de dar acesso ao público via internet e de preservar seu trabalho integral no LUME.

6.4 Auto arquivamento

No repositório da USP a formação e o desenvolvimento da memória da produção intelectual podem ser realizados através

do auto depósito do conteúdo integral da produção intelectual. E na UFRGS ao realizar o auto arquivamento, o autor também deve registrar os metadados.

6.5 Objetivos

Os objetivos são parecidos em ambos os repositórios analisados, onde seus objetivos principais é fomentar, assegurar coletar, tratar e preservar a produção científica, promovendo livre acesso às informações no âmbito da comunidade acadêmica e maximizar a visibilidade, uso e impacto da produção.

6.6 Base

A Universidade de São Paulo não cita nenhum documento específico como base para a criação da política do repositório. Já Universidade Federal do Rio Grande do Sul, baseia-se pela Creative Commons, organização não governamental sem fins lucrativos voltada para a expansão da quantidade de obras criativas disponíveis à sociedade (CREATIVE..., s.d., p. 1).

6.7 Responsabilidade

O Conselho Supervisor do Sistema Integrado de Bibliotecas, SIBi, é o responsável por estabelecer e validar normas para coleta, tratamento, preservação da produção intelectual, definir tipos de documentos para depósito. As bibliotecas são responsáveis por registrar tecnicamente a produção intelectual, edição, revisão, validação e disponibilização online da mesma, capacitação, fornecimento de dados, estatísticas e informações. O departamento técnico do SIBi gerencia, atualiza o sistema de gestão, gera dados e indicadores sobre produção intelectual, estatísticas, cria diretrizes para garantir controle e preservação da produção.

Já a responsabilidade pelo repositório da Universidade Federal do Rio Grande do Sul é da equipe técnica do Lume e do Comitê Gestor.

6.8 Tipo de documento

As políticas instituídas para a criação e regimento dos repositórios analisados foram identificadas como Resolução, no caso da Universidade de São Paulo e Portaria, na Universidade Federal do Rio Grande do Sul.

6.9 Povoamento

Na USP, esta atividade é realizada pela equipe da Biblioteca, por meio de importação de dados ou por auto arquivamento. No Lume, além do auto arquivamento, pode ser realizada a importação de dados do SABi.

6.10 Metadados

Em relação ao repositório da Universidade Federal do Rio Grande do Sul, a produção digital arquivada segue padrões internacionalmente definidos pela equipe do Lume, que orientam o autor ou o responsável pela comunidade no caso do auto arquivamento e estão associados aos objetos digitais encontrados nas coleções.

Na USP, quando uma produção não estiver disponível em formato digital, os metadados deverão ser registrados na BDPI e um exemplar deve ser depositado na biblioteca.

7 CONCLUSÃO

A respeito da relevância do vocabulário controlado dentro de repositórios digitais, o estudo constatou o quanto é importante ser trabalhado um vocabulário que permita a identificação dos documentos, visando à recuperação da informação por parte dos usuários.

Pode-se enfatizar que os vocabulários controlados apresentados pelos repositórios institucionais estudados estariam a cumprir sua finalidade, ou seja, a satisfação informacional de seus usuários, apesar das diferenças de estrutura. Afirma-se isso, pois, a estrutura apresentada pela Universidade de São Paulo visa disseminar as teses e dissertações da comunidade acadêmica, enquanto que o repositório da Universidade Federal do Rio Grande do Sul, além das teses e dissertações, visa difundir os demais documentos produzidos pelos membros da comunidade acadêmica, bem como pela universidade, enquanto editora, sendo possível, portanto, acessar as teses e dissertações dos alunos e professores, além dos trabalhos acadêmicos e técnicos; dos trabalhos completos e resumos apresentados em eventos institucionais promovidos pela universidade; dos artigos, livros ou capítulos de livros produzidos na universidade ou pelos membros da universidade, bem como dos acervos das diversas Unidades Acadêmicas e demais órgãos da Universidade.

Cabe salientar, também, a diferença no processo de busca. Para a Universidade de São Paulo, a localização das obras pode ocorrer de forma simples ou avançada. A primeira permite a

recuperação apenas pelo Título das obras, independente do que for digitado no campo de busca. Somente na pesquisa avançada é que o usuário obterá a recuperação das obras por meio das variáveis Autor, Título, Resumo, Palavras-chave, Orientador, Unidade e Ano de defesa.

O repositório da Universidade Federal do Rio Grande do Sul, apesar de apresentar separadamente os *links* para cada tipo de documento que compõe o acervo, é interessante frisar que, independente da escolha do usuário, este obterá apenas uma tela contendo várias janelas, dando a possibilidade de localizar a obra pelo Autor/Orientador/Outros, Título, Assunto, Data, Formato, Idioma ou Pesquisa Geral. Nesta perspectiva, considerou-se este sistema de informação mais fácil e rápido para o usuário.

Não se pode deixar de frisar é que os dois repositórios se preocupam com o processo de representação e de comunicação, estando estruturados conforme o processo de significação das comunidades que atendem, apresentando um perfeito controle do vocabulário, por meio do estabelecimento de relações conceituais consistentes, constituindo-se, os dois, em um bem para a sociedade em pesquisas e desenvolvimento científico.

VOCABULARY IN CONTROLLED REPOSITORIES BRAZILIAN HIGHER EDUCATION INSTITUTIONS: comparative study

ABSTRACT: The purpose of this paper is to present the results of the comparative study established between two institutional repositories known by the academic community in order to subsidize the construction of the repository Department of Archival and Library Science at the Federal University of Amazonas (DAB). This need occurred, due to aggregation of the course of Archival of the Department of Library. Thus, the number of teachers and students has grown, and with it, the number of academic productions. Due to the nature of the courses that were included in the DAB, beyond the informational guard, it is cited as an informational disclosure of activities related to daily academic thereof. Therefore, it becomes increasingly important for the DAB, to have a tool which can represent both the information contained in the technical and scientific productions, as enabling their recovery in a safe, efficient and effective manner. As a general goal, establish a comparative study between two representation of information systems, supporting the construction of a controlled vocabulary to the technical-scientific production of DAB. For this, they were taken based on the following specific objectives:

to present the importance of documentary language to the process of bibliographic recovery; demonstrate the relevance of the controlled vocabulary for the representation and retrieval of information; analyze the information retrieval system developed by two institutions of higher education, based on the controlled vocabulary used in the online repositories. To this end, we opted for the comparative method, from the study of two existing repositories, in order to serve as a basis for the construction of the DAB repository. The conclusion is relevant to how UFAM remain active before the academic community at national and international level.

KEYWORDS: Systems of representation of information. Controlled vocabulary. Repository. Virtual Shelf.

REFERÊNCIAS

ALMEIDA, Rogerio Carlos Petrini de. **Paralelo entre tesouro e listas de cabeçalhos de assuntos.** 2011. Disponível em: <<http://rogeriopetrinialmeida.blogspot.com.br/2011/09/paralelo-entre-tesouro-e-listas-de.html>>. Acesso em: 02 fev. 2015.

CANESSA, Enrique; ZENNARO, Marco. **Science dissemination using open access:** a compendium of selected literature on Open Access. [S.l.]: ICTP, 2008. Disponível em: <<http://sdu.ictp.it/openaccess/SciDissOpenAccess.pdf>>. Acesso em: 04 jul. 2014.

CREATIVE Commons: cite note-ccboard1-3. Disponível em: <http://pt.wikipedia.org/wiki/Creative_Commons_-_cite_note-CCBoard1-3>. Acesso em: 27 abr. 2015.

BAEZA-YATES, Ricardo; RIBEIRO NETO, Berthier. **Modern Information Retrieval.** Harlow: Addison Wesley, 1999.

BOCCATO, Vera Regina Casari. **Avaliação de linguagem documentária em fonoaudiologia na perspectiva do usuário:** estudo de observação da recuperação da informação com protocolo verbal. 2005. 239f. Dissertação (Mestrado em Ciência da Informação)– Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Marília, 2005.

CAMPOS, Maria Luiza de Almeida. **Linguagem documentária:** teorias que fundamentam sua elaboração. Niterói: EdUFF, 2001.

CINTRA, Ana Maria et al. **Para entender as linguagens documentárias**. 2. ed. rev. e ampl. São Paulo: Polis, 2002.

GARCÍA JIMÉNEZ, Antonio. Instrumentos de representación del conocimiento: tesauros versus ontologías. **Anales de documentación: Revista de Biblioteconomía y Documentación**, n. 7, 2004. Disponible em: <<http://revistas.um.es/index.php/analesdoc/article/view/1691>>. Acesso em: 12 abr. 2013.

KOBASHI, Nair Yumiko. Vocabulário controlado: estrutura e utilização. REUNIÃO DA REDE DE ESCOLAS DE GOVERNO, 1., São Paulo, 2008. **Anais...** Disponível em: <www2.enap.gov.br/rede_escolas/arquivos/vocabulario_controlado.pdf>. Acesso em: 04 jul. 2014.

LAKATOS, Eva Maria, MARCONI, Marina de Andrade. **Fundamentos de metodologia científica**. São Paulo: Atlas, 2001.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. 2. ed. Brasília: Briquet de Lemos, 2004.

MEZZAROBA, Orides; MONTEIRO, Cláudia Servilha. **Manual de metodologia da pesquisa no direito**. São Paulo: Saraiva, 2005.

MOREIRA, Manoel Palhares; MOURA, Maria Aparecida. Construindo tesauros a partir de tesauros existentes: a experiência do TCI: Tesauro em Ciência da Informação. **DataGramazero - Revista de Ciência da Informação**, v. 7, n. 4, 2006. Disponível em: <http://dgz.org.br/agoo6/Art_01.htm>. Acesso em: 06 jul. 2014 .

MOTTA, Dilza Fonseca da. **Método relacional como nova abordagem para a construção de tesauros**. Rio de Janeiro: SENAI/DN/DPEA, 1987.

PINTO, Virginia Bentes; PINHEIRO, Edna Gomes. Ensinar e aprender: reflexões acerca da pesquisa em Ciência da Informação. **Transinformação**, Campinas, v. 15, n. 3, p. 319-331, set./dez., 2003.

SALES, Rodrigo de. Suportes teóricos para pensar linguagens documentárias. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 5, n. 1, p. 95-114, jul./dez. 2007.

SANZ-VALERO, Javier; VEIGA DE CABO, Jorge; CASTIEL, Luis D. The open access initiative in access to technical and scientific information in health sciences. **Revista Eletrônica**

de Comunicação, Informação & Inovação em Saúde,
Rio de Janeiro, v. 1, n. 1, p. 19-26, 2007.

SOUTO, L. F. Recuperação de informações em bases de dados: usos de tesauro. **Transinformação**, Campinas, v. 15, n. 1, p. 73-81, jan./abr.2003.

SOUZA, Jóice Cleide Cardoso Ennes de. **Avaliação de linguagem de indexação aplicada à informação jornalística**: estudo de caso. 2007. 156f. Dissertação (Mestrado em Ciência da Informação)– Universidade Federal Fluminense, Instituto Brasileiro de Informação em Ciência e Tecnologia, Niterói, 2007.

SOUZA, Renato Rocha. Sistemas de recuperação de informações e mecanismos de busca na web: panorama atual e tendências. **Perspectiva em Ciência da Informação**, Belo Horizonte, v. 11, n. 2, p. 161-173, maio/ago. 2000.

TÁLAMO, Maria de Fátima Gonçalves Moreira. **Linguagem documentária**. São Paulo: APB, 1997.

VEIGA, Francisco José de Abrantes. **Evolução dos sistemas se informação**. 2006. Disponível em:
<http://student.dei.uc.pt/~fveiga/GSI/Evolucao_Sist_Inf.pdf>. Acesso em: 04 jul. 2014.

WAETEI, Ranito Zambo et al.; MOURAII, Maria Aparecida; MANGUE, Manuel Valente. Acesso livre à informação científica em países em desenvolvimento: o caso dos repositórios “saber” (Moçambique) e digital da UFMG (Brasil). **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 17, n. esp., 2. p. 195-210, 2012.