



SISTEMA DE GESTÃO PARA LINGUAGEM DOCUMENTÁRIA: METADADOS E REDE COLABORATIVA NO VOCABULÁRIO CONTROLADO DO SIBI/USP.

Cibele Araújo Camargo Marques dos Santos¹, Silvia Regina Saran Della Torre², Roseli Koizimi Matsuda³, Sonia Regina Yole Guerra⁴, Juliana de Souza Moraes⁵, Vânia Mara Alves Lima⁶

¹Doutora em Ciência da Informação, Bibliotecária da Divisão de Biblioteca e Documentação da Faculdade de Medicina da USP, São Paulo, SP;

²Graduada em Biblioteconomia, Bibliotecária da Divisão de Biblioteca da Escola Politécnica da USP, São Paulo, SP;

³Graduada em Biblioteconomia, Bibliotecária do Departamento Técnico do Sistema Integrado de Bibliotecas da USP, São Paulo, SP,

⁴Graduada em Biblioteconomia, Bibliotecária do Serviço de Biblioteca do Instituto de Geociências da USP, São Paulo, SP;

⁵Mestre em Ciência da Informação, Bibliotecária da Biblioteca Professor Achille Bassi, do Instituto de Ciências Matemáticas e da Computação, São Carlos, SP,

⁶Profa. Dra. do Departamento de Biblioteconomia da Escola de Comunicações e Artes da USP.

Resumo

Sistemas de gestão desenvolvidos para a web, a partir de metadados, permitem manutenção eficiente de grandes quantidades de informação. Um vocabulário controlado como o utilizado pelo Sistema Integrado de Bibliotecas da USP (SIBi/USP) necessita de atualização contínua realizada através de uma rede colaborativa com a participação de bibliotecários indexadores de todas as áreas do conhecimento. Este trabalho apresenta os resultados obtidos com o sistema de gestão desenvolvido pelo Grupo de Gerenciamento para a manutenção do Vocabulário Controlado do SIBi/USP. O fluxo deste sistema consiste em filtros de validação realizados pelos componentes do Grupo de Gerenciamento do Vocabulário. A metodologia de gestão do Vocabulário possui além deste sistema, uma política de governança. Os resultados obtidos nos seis anos desde a ativação do sistema de gestão pela Base de Sugestões consistiram em: 1192 inclusões de descritores, 240 alterações, 61 exclusões, totalizando 1493 operações. A gestão e o controle de qualidade do Vocabulário permitiram o aprimoramento do tratamento e da recuperação da informação no Banco de Dados Bibliográficos da USP – DEDALUS.

Palavras-Chave: Linguagens documentárias, gerenciamento; Vocabulários Controlados; Metadados; Rede de bibliotecas; Redes sociais.

Abstract

The management systems developed for the web, using metadata, enables efficient maintenance of large amounts of information. A controlled vocabulary as the one used by the



Integrated Library System of the University of Sao Paulo (SIBi/USP) requires continuous updates done through a collaborative network with the participation of indexing librarians from all areas of knowledge. This paper presents the results obtained from the management system developed by the Management Group for the maintenance of the Controlled Vocabulary developed by SIBi/USP. The workflow of this system consists in filters to validate changes, which are performed by the components of the Vocabulary Management Group. The methodology for managing the vocabulary beyond this system has a policy of governance. The results obtained in the six years since the activation of the system by the creation of a Data Base for Suggestions were: the inclusion of 1192 descriptors, 240 changes, 61 deletions, totaling 1493 operations. The management and quality control of vocabulary allowed the improvement of treatment and recovery information in the Bibliographic Database of USP called DEDALUS.

Keywords: Indexing languages, management; Controlled Vocabularies; Metadata; Network Libraries, Social networks.

1 Introdução

Os vocabulários controlados são usados para promover efetiva organização e busca da informação em sistemas de recuperação, sistemas de navegação da internet e outros ambientes de pesquisa para identificação e localização de conteúdos (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2005).

Devido a alterações nas terminologias das áreas do conhecimento, decorrentes da própria evolução das áreas e suas relações com seus contextos, linguagens documentárias como estes vocabulários necessitam de manutenção permanente.

Os sistemas de gestão de vocabulários controlados constituem uma área em desenvolvimento, sendo que os sistemas desenvolvidos para a web, a partir de metadados, permitem manutenção eficiente de grande quantidade de informação.

Um vocabulário controlado como o utilizado pelo Sistema Integrado de Bibliotecas da USP (SIBi/USP) necessita de atualização contínua realizada através de uma rede colaborativa com a participação de bibliotecários indexadores de todas as áreas do conhecimento.

Este trabalho apresenta os resultados obtidos com o sistema de gestão desenvolvido pelo Grupo de Gerenciamento do Vocabulário Controlado do

SIBi/USP. Este sistema constitui-se em ferramenta de controle para atualizações e revisões terminológicas no Vocabulário através de uma base de dados com formulário online e padrão de metadados para solicitação de inclusão, alteração e exclusão de descritores, desenvolvido em PHP, com base de dados em MySQL. O fluxo deste sistema consiste em filtros de validação realizados pelos componentes do Grupo de Gerenciamento do Vocabulário.

2 Vocabulários Controlados

A função principal dos vocabulários controlados é obter consistência na descrição de conteúdo de um objeto através da linguagem para facilitar a recuperação. Em 2005, a NISO apresentou a revisão mais recente de sua norma para Construção, Formatação e Gestão de Vocabulários Controlados Monolíngues (ANSI/NISO Z39.19-2005) com ênfase na organização terminológica, na interoperabilidade e na manutenção (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2005).

Um vocabulário controlado, segundo esta norma, pode ser uma lista de termos, uma taxonomia ou até um extenso tesauro com complexa estrutura hierárquica e diversos tipos de relacionamentos entre os termos, mas sempre será necessário um sistema de gestão, pois a linguagem apresenta conceitos emergentes, mudanças terminológicas e obsolescência de termos. Na medida em que um vocabulário controlado é utilizado torna-se vital conduzir testes periódicos e avaliações para garantir o melhor uso possível (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2005).

A norma ANSI Z39.19 -1974 sobre Estrutura, Construção e Uso de Tesauros foi revisada em 1980 e em 1993, sendo nesta última equiparada a norma internacional ISO 2788 e a britânica BS 8733, passando a constituir-se nas Diretrizes para Construção, Formatação e Gestão de Tesauros Monolíngues. Porém com as mudanças ocorridas na área da informação foi necessário expandir o escopo da norma para inclusão de vocabulários controlados, taxonomias e outras formas de linguagens documentárias (FAYEN, 2004).



Segundo a autora, com o surgimento de novos usuários da informação tornou-se importante explicar o funcionamento dos vocabulários controlados, incluindo novas formas de acesso como navegação, *browsing*, busca por palavra-chave e outras aplicações para web, além da pesquisa pós-coordenada com operadores booleanos.

Assim, o surgimento de recursos digitais, metadados, redes sociais e indexação colaborativa sinalizam novas possibilidades de gestão de vocabulários controlados (MACGREGOR e MCCULLOCH, 2006).

A Association of College and Research Libraries - ACRL *Research Planning and Review Committee* (2010) apresenta como uma das dez maiores tendências para as bibliotecas acadêmicas o crescimento da colaboração expandindo o papel das bibliotecas na instituição e nos sistemas próximos.

O trabalho colaborativo de bibliotecários e especialistas no desenvolvimento de um vocabulário controlado com abrangência em diversas áreas do conhecimento é fundamental para uma adequada estruturação hierárquica, bem como uma excelente oportunidade para os participantes e para as respectivas instituições.

Para os bibliotecários indexadores surge a necessidade do desenvolvimento de novas habilidades além das relacionadas à identificação de conceitos e seleção de termos, aspectos centrais na indexação, como a utilização de metadados e novas ferramentas para melhor captação de termos candidatos à inserção nos vocabulários controlados (LEISE, 2004).

O desenvolvimento e manutenção de um vocabulário controlado de forma colaborativa representam importante possibilidade de aprimoramento para um sistema de informação na era digital.

3 Metodologia

Conforme apresentado por Lima, et al (2006), o Vocabulário Controlado do Sistema de Bibliotecas da USP (SIBi/USP) foi construído em duas etapas: a primeira como resultado de um projeto do qual participaram os profissionais das Bibliotecas da Universidade, pesquisadores/docentes dos cursos de Biblioteconomia da USP e





Universidade Estadual Paulista (UNESP) para orientação metodológica e professores especialistas nos assuntos integrados ao Vocabulário, para garantia da consistência terminológica e dos sistemas conceituais de cada área do conhecimento.

O projeto contou com o suporte de informática do Departamento Técnico do SIBi/USP (DT/SIBi) para a primeira versão da base de dados em CD-ROM, na implementação do uso do Vocabulário em 2001 pelo Banco de Dados Bibliográficos da USP (DEDALUS), catálogo on-line das bibliotecas do SIBi/USP e na disponibilização para acesso on-line do Vocabulário ao público pela WEB, apresentado a seguir na Figura 1.

VOCABULÁRIO CONTROLADO DO SIBI/USP



MACROESTRUTURA

LISTA DE ASSUNTOS

- [Geral Alfabética](#)
- [Geral Hierárquica](#)
- [Locais Geográficos e Históricos](#)
- [Gênero e Forma](#)
- [Profissões e Ocupações](#)

TABELA DE QUALIFICADORES

[Pesquisar no DEDALUS](#) [Relatório de Atualização](#) [Contadores](#)

© 2006-2008 SIBi/USP - Departamento Técnico - SIBi/DT

Figura 1: Vocabulário Controlado do SIBi/USP

Fonte: <http://143.107.73.99/Vocab/Sibix652.dll>

A segunda etapa correspondeu a definição do modelo de gestão para manutenção e gerenciamento do Vocabulário com a criação de um grupo de gerenciamento incorporado aos processos sistêmicos do DT/SIBi. O Grupo de Gerenciamento do Vocabulário Controlado do SIBi/USP é constituído por 3 bibliotecários de cada área do conhecimento que atuam em bibliotecas do SIBi/USP, um bibliotecário coordenador de conteúdo, um bibliotecário coordenador do





processo, um analista de sistemas e bibliotecários do Departamento Técnico (DT/SIBi/USP) e por professor especialista em Linguagens Documentárias do Departamento de Biblioteconomia da Escola de Comunicações e Artes da USP (ECA/CBD).

A indicação de membros do Grupo ocorre a cada dois anos, visando à renovação do grupo para garantia da continuidade dos trabalhos e incorporação de novos membros.

Recentemente foi aprovado o regimento do Grupo. Entre as atividades previstas para o Grupo, constam a gestão e manutenção do Vocabulário, a definição da política de indexação, colaboração nas revisões hierárquicas de áreas do conhecimento e na consistência de descritores, treinamentos para os bibliotecários do SIBi/USP, e a elaboração de manuais e tutoriais.

O processo de manutenção e aprimoramento do Vocabulário necessita, do ponto de vista terminológico, da participação dos bibliotecários de todas as Bibliotecas do Sistema, que cooperam com sugestões de novos descritores e alterações dos descritores existentes, através da Base de Sugestões desenvolvida pelo DT/SIBi.

Esta base foi implementada em 2006 a partir do fluxo colaborativo de bibliotecários indexadores do SIBi/USP, pelo preenchimento de dados em um formulário on-line, desenhado para a web com metadados e desenvolvido na linguagem PHP em banco de dados MySQL utilizando navegadores da internet (LIMA, 2006).

Este sistema contém um módulo de usuários e senhas, além de uma base de dados que permite as sugestões de inserção, alteração e exclusão de descritores sempre devidamente justificadas e embasadas em fontes bibliográficas confiáveis. O fluxo de trabalho é controlado pelos níveis de validação propostos pelo modelo de gestão, com sugestão de bibliotecários do SIBi/USP na primeira fase, análise dos bibliotecários representantes de áreas do conhecimento em segunda fase, validação do bibliotecário coordenador de conteúdo na terceira fase e sendo o descritor aprovado, a alteração é executada pela coordenação de processo na quarta fase.



Na Figura 2 apresentamos a interface da Base de Sugestões, que conta no Menu Principal com as opções de Sugestão de Descritores, Verificação de Status dos Pedidos realizados. Em Sugestões de Descritores é possível escolher Inclusão, Alteração e Exclusão. Na opção de Verificação do Status dos Pedidos está disponibilizada a Lista de Registros Pendentes (que foram sugeridos e aguardam análise), as Pendências nas respectivas fases/níveis de análise, os Registros Indeferidos (que não foram aceitos e sua justificativa) e os Registros Executados com Sucesso.



Figura 2: Base de Sugestões do Vocabulário Controlado do SIBi/USP
Fonte: Área Técnica do site do SIBiNET.

O histórico de sugestões pode ser acompanhado pelo solicitante através dos relatórios disponibilizados pelo sistema de sugestões com indicação da fase em andamento, o que permite a elaboração das estatísticas da gestão do Vocabulário (Figura 3).

Sibix-653 v4.3 - Sistema de Sugestões do Vocabulário Controlado

Registros Executados com Sucesso

Fase

Fase Atual	Data	Tabela	# Reg.	Área	Descritor Proposto	Descritor Atual	Operação	Unidade	Responsável	Zoom
4	25.06.2010	ASS	1867	CH	INDICADORES SOCIAIS		Inclusão	SIBIDT	Roseli Koizumi Matsuda	
4	25.06.2010	ASS	1834	CH	AUTONOMIA DE GOVERNO		Inclusão	SIBIDT	Roseli Koizumi Matsuda	
4	25.06.2010	ASS	1835	CH	INTERROGAÇÃO		Inclusão	SIBIDT	Roseli Koizumi Matsuda	
4	25.06.2010	ASS	1865	CH	INSUCESSO EMPRESARIAL		Inclusão	SIBIDT	Roseli Koizumi Matsuda	
4	25.06.2010	ASS	1868	CH	FRACASSO EMPRESARIAL ver INSUCESSO EMPRESARIAL		Inclusão	SIBIDT	Roseli Koizumi Matsuda	
4	25.06.2010	ASS	1881	CH	STAKEHOLDER		Inclusão	SIBIDT	Roseli Koizumi Matsuda	
4	25.06.2010	ASS	1882	CH	BALANCED SCORECARD		Inclusão	SIBIDT	Roseli Koizumi Matsuda	
4	25.06.2010	ASS	1883	CH	MENSURAÇÃO DE DESEMPENHO ORGANIZACIONAL		Inclusão	SIBIDT	Roseli Koizumi Matsuda	

Figura 3: Relatório de Registros Executados da Base de Sugestões
 Fonte: Área Técnica do site do SIBiNET.

A metodologia de gestão do Vocabulário possui além deste sistema, uma política de governança que inclui o modelo descentralizado, a sensibilização dos profissionais para colaboração e treinamentos necessários, como ocorre em sistemas de gestão na web (VIBERTI, 2010).

4 Resultados

Os resultados obtidos nos seis anos desde a ativação do sistema de gestão e da Base de Sugestões consistem em: 1192 inclusões de descritores, 240 alterações, 61 exclusões, totalizando 1493 operações.



Foram realizadas 745 modificações em descritores da área de Ciências Humanas, 230 da área de Ciências Exatas e 518 da área de Ciências Biológicas. Em relação ao período, foram realizadas 184 modificações em 2005, 242 em 2006, 442 em 2007, 439 em 2008, 160 em 2009 e 38 até abril de 2010.

Estas informações estão disponíveis em relatórios gerados pela Base de Sugestões para os usuários credenciados, para os bibliotecários indexadores responsáveis nas respectivas bibliotecas pela sugestão de descritores para indexação e para o público em Relatório de Atualização no site do Vocabulário Controlado do SIBi/USP.

Além destas alterações, foram realizadas revisões de áreas do conhecimento, tanto na estrutura hierárquica como na adequação da terminologia.

Na consistência de descritores e de registros indexados no DEDALUS, realizada entre 2008 e 2009 foi analisada uma lista com 5.100 descritores que constavam no DEDALUS e eram anteriores à implantação do Vocabulário nas bases, portanto não constavam do Vocabulário. A partir desta análise, os descritores de 519.380 registros foram reindexados com os descritores atualizados, sendo 506.355 registros corrigidos via máquina e 13.025 registros corrigidos manualmente (SANTOS et al 2009).

Os resultados obtidos no processo de gestão do Vocabulário referentes aos fluxos de alterações e inserções e nos processos de revisões hierárquicas foram facilitados pela utilização das ferramentas desenvolvidas para a Base de Sugestões.

5 Considerações Finais

A gestão e o controle de qualidade do Vocabulário permitiram o aprimoramento do tratamento e da recuperação da informação no Banco de Dados Bibliográficos da USP - DEDALUS, bem como a formação de profissionais especializados em gestão de linguagens documentárias e o treinamento constante para as equipes bibliotecárias do SIBi/USP.

O sistema de gestão mostrou-se eficiente e rápido, possibilitando modificações e inserções de descritores no prazo de uma semana a quinze dias.





Outras instituições demonstraram interesse pelo uso do Vocabulário e em contribuir para sua atualização, permitindo o início de parcerias, que ampliam a rede colaborativa, o que sugere maior rapidez nas atualizações terminológicas e ampliação da abrangência das áreas de conhecimento do Vocabulário uma vez que levarão em consideração outros pontos de vista.

Existem lições para aprender com o trabalho colaborativo e inovações que podem ser implementadas para melhoria da gestão.

Assim, pode-se promover o incentivo à participação dos bibliotecários no aprimoramento terminológico do Vocabulário, o aperfeiçoamento da divulgação e comunicação das atividades realizadas, tais como as fontes de dados para a supervisão do processo e realização de melhorias, se necessárias, e também o desenvolvimento de novos treinamentos para as equipes.

6 Referências

ACRL RESEARCH PLANNING AND REVIEW COMMITTEE. 2010 top ten trends in academic libraries: a review of the current literature. **C&RL News**, June 2010. p. 286-92.

FAYEN, E. A new standard for controlled vocabularies. **The Indexer**, v.24, n.2, 2004.

LEISE, F. Metadata and content management systems: an introduction for indexers. **The Indexer**, v.24, n.2, 2004.

LIMA, V. M. L.; KOBASHI, N. Y.; COUTTO, M. L. M.; SANTOS, C. A. C. M.; AMARAL, M. C.; TOKAREVICZ, S.; DELLA TORRE, S. R. S.; GUERRA, S. R. Y.; BOCCATO, V. R. C.; BARCELLOS, J. C. H. Estudos para implantação de ferramenta de apoio à gestão de linguagens documentárias: Vocabulário Controlado USP. **Transinformação**, v.18, n.1, 2006.

MACGREGOR, G.; McCULLOCH, E. Collaborative tagging as a knowledge organization and resource discovery tool. **Library Review**, v. 55, n.5, 2006.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **Guidelines for the construction, format, and management of monolingual controlled vocabularies**, 2005. (ANSI/NISO Z39.19-2005.)

SANTOS, C. A. C. M.; LIMA, V. M. L.; DELLA TORRE, S. R. S.; GUERRA, S. R. Y. Consistência de descritores na base de dados DEDADUS da USP: uma abordagem terminológica. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO, no., 2009, Bonito, MS. **Anais...** Bonito, MS: editora, 2009.

VIBERTI, F. Conteúdo: centralizar ou descentralizar, eis a questão. **Intranet Portal**. Disponível em: http://www.intranetportal.com.br/comunicacao/cc_1. Acesso em: 8 abr. 2010.

