

1765778

**Área Temática: Redes de conhecimento – Informação, conhecimento e sociedade em rede**

CONSISTÊNCIA DE DESCRITORES NA BASE DE DADOS DEDALUS DA USP: UMA ABORDAGEM TERMINOLÓGICA

**Santos, Cibele Araújo Camargo Marques dos** – doutoranda da Escola de Comunicações e Artes da USP. Representante de área no Grupo de Gestão do Vocabulário Controlado do SIBi/USP. Bibliotecária da FSP/USP. [cibeleac@usp.br](mailto:cibeleac@usp.br)  
Faculdade de Saúde Pública da USP. Biblioteca/CIR – Av. Dr. Arnaldo, 715 01246-904 – São Paulo, SP Brasil

**Lima, Vânia Mara Alves** – Profa. Dra. da Escola de Comunicações e Artes da USP. Consultora técnica do Grupo de Gestão do Vocabulário Controlado do SIBi/USP. [vamal@usp.br](mailto:vamal@usp.br)  
Escola de Comunicações e Artes da USP. Departamento de Biblioteconomia e Documentação. Av. Prof. Lúcio Martins Rodrigues, 443 05508-900 - São Paulo, SP Brasil

**Della Torre, Sílvia Regina Saran** – Coordenadora de conteúdo do Grupo de Gestão do Vocabulário Controlado do SIBi/USP. Bibliotecária da Poli/USP [silvia.torre@poli.usp.br](mailto:silvia.torre@poli.usp.br)  
Escola Politécnica da USP. Av. Prof. Luciano Gualberto, tv 3, n.380 - 05508-970 – São Paulo, Brasil

**Guerra, Sonia Regina Yole** - Representante de área no Grupo de Gestão do Vocabulário Controlado do SIBi/USP. Bibliotecária do IGC/USP [syog@usp.br](mailto:syog@usp.br)  
Instituto de Geociências da USP. Biblioteca/Seção de Processamento Técnico – Rua do Lago, 562 05508-080 – São Paulo, SP Brasil

**Resumo:** O objetivo deste trabalho, norteado por parâmetros documentários, linguísticos e terminológicos, é apresentar a metodologia e os resultados obtidos quando do processo de consistência realizado nos descritores de assuntos dos registros da Base de Dados Bibliográficos da USP – DEDALUS. Essa consistência, visando melhorar a recuperação da informação, compatibilizou os cabeçalhos de assuntos, pertencentes à antiga Lista de Assuntos USP, utilizada até 2001, com o Vocabulário Controlado do SIBi/USP, linguagem documentária construída a partir do referencial terminológico de diversas áreas do conhecimento cujos documentos constituem o acervo bibliográfico da USP. Nessa compatibilização foram analisados 5.100 descritores e reindexados 519.380 registros no DEDALUS, identificadas inconsistências, assim como a necessidade de aprimoramento de algumas áreas do Vocabulário, e a importância da utilização de uma metodologia de indexação com diretrizes e normas de uso, além da necessidade de treinamento contínuo dos indexadores. Os resultados reforçaram que a abordagem terminológica é adequada para desenvolvimento e manutenção de linguagem documentária, para os procedimentos de indexação e processos de análise e revisão de um sistema de informação.

**Descritores:** BASES DE DADOS; TERMINOLOGIA; LINGUAGENS DOCUMENTÁRIAS; INDEXAÇÃO (BIBLIOTECONOMIA).

In: Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação – CBBD, 23.  
2009. Brasília, 2009.

**Abstract:** The objective of this work, guided by documentary, linguistic and terminological parameters, is to present the methodology and results obtained with the process of consistency achieved in the subjects' descriptors of records of the Bibliographic Database of University of Sao Paulo - DEDALUS. This consistency, to improve the retrieval of information, compared the subjects headings, belonging to the former Subject List of USP and used until 2001, with the Controlled Vocabulary of SIBi/USP, index language built from the reference terminology in various fields of knowledge which documents are the bibliographic collection of USP. In this alignment were analyzed 5100 descriptors and reindexed 519,380 records of DEDALUS, identified inconsistencies, the need for improvement in some areas of vocabulary and the importance of using a method of indexing with guidelines and standards in addition to the need for continued training of the indexers. The results reinforced that the terminology approach is suitable for development and maintenance of index language to the indexing procedures and processes of analysis information system and review.

**Descriptors:** DATABASES; TERMINOLOGY, INDEX LANGUAGES; SUBJECT INDEXING.

## **Introdução**

Através do desenvolvimento de redes informacionais com modelo de trabalho cooperativo, as bibliotecas conseguiram organizar, armazenar e promover o acesso à informação. A formação de redes é uma questão fundamental para o tratamento da informação e a convergência da tecnologia com as comunicações, afeta a criação, gestão e uso da informação de modo inédito (McGARRY, 1999).

O Sistema Integrado de Bibliotecas da Universidade de São Paulo - SIBi/USP, está organizado no modelo de rede informacional cooperativa e foi instituído pela Resolução da Reitoria n. 2.226 de 08.07.1981, iniciando suas atividades a partir de março de 1982. É composto, atualmente, por 43 bibliotecas, distribuídas nos seis "campi" universitários com acervos bibliográficos que atingem aproximadamente 6 milhões de volumes (SIBi/USP, 2009). Estes acervos estão reunidos na Base de Dados Bibliográficos DEDALUS.

O Vocabulário Controlado do SIBi/USP é a linguagem documentária utilizada pelo sistema para a representação de conteúdo dos registros da Base de Dados Bibliográficos DEDALUS, disponível na SIBiNET (<http://www.usp.br/sibi/>) e contém a informação dos acervos de livros e outros materiais, periódicos, produção científica e teses da Universidade. Implantado no sistema em 2001, passou a ser utilizado na

indexação dos novos registros, sendo que registros anteriores a esta data seriam revistos posteriormente.

As linguagens documentárias utilizadas pelos sistemas de informação para a representação do conteúdo dos documentos atuam, de forma decisiva, no resultado da recuperação da informação no sistema, oferecendo ferramentas para resolver as dificuldades que ocorrem quando utilizamos a linguagem natural para essa representação, principalmente no que se refere à polissemia, sinonímia, modos de expressão e as relações entre palavras. Toda linguagem documentária tem um conjunto terminológico escolhido dentro do sistema nocional da área que representa.

Para Lancaster (1996), a avaliação de um sistema de informação implica na necessidade de análise do acervo registrado; análise do catálogo ou outra fonte para verificar se os usuários encontram o que procuram; estudo de custo-eficiência e outros para um diagnóstico do sistema, como prevê a teoria da administração. O autor enfatiza que esta avaliação será útil para melhorar o serviço prestado quando for uma avaliação analítica, criteriosa, objetiva e com resultados quantificáveis.

O modelo cooperativo mostrou ser uma forma viável de tratamento e organização de grandes quantidades de informação pelas bibliotecas. No entanto, o tratamento interinstitucional desenvolvido pelas redes cooperativas necessita da adoção de políticas e metodologias de indexação visando a padronização dos registros gerados no sistema.

Svenonius (2001) afirma que um sistema de organização da informação, para ser efetivo precisa estar baseado em proposições com objetivos a serem atingidos pelo sistema e princípios ou diretrizes que guiem o projeto, enfatizando a importância da organização inteligente.

Portanto, um sistema de informação desenvolvido por uma rede cooperativa necessita de avaliação criteriosa, quantificável, baseada na política de indexação definida para seus registros.

Além disso, os processos de organização da informação precisam estar formalizados no que concerne às conceituações lingüísticas e aos modelos de relações entre atributos e entidades. A revisão das formas de organização da

informação passa, invariavelmente, pelo diagnóstico e avaliação dos sistemas de informação existentes.

Após análise, visando a melhoria da recuperação da informação na Base de Dados Bibliográficos DEDALUS, o Grupo de Gestão do Vocabulário do SIBi/USP decidiu, em 2008, realizar a consistência de seu índice de assunto, através da compatibilização dos termos que permaneciam nos registros da base, mas não pertenciam ao Vocabulário, norteadas por parâmetros terminológicos para a revisão, considerando-se que o vocabulário é uma linguagem construída a partir do referencial terminológico das diversas áreas do conhecimento que o constituem.

O objetivo deste trabalho é apresentar processos e resultados desta consistência a partir da abordagem terminológica.

### **Histórico**

A criação, em 1985, de um catálogo global online que reunia o acervo das bibliotecas da USP, deu início à informatização do acesso aos recursos bibliográficos da Universidade. Batizado de DEDALUS – Banco de Dados Bibliográficos da USP utilizava a Lista de Assuntos USP que, desenvolvida a partir das terminologias consagradas de cada especialidade em seus catálogos locais, era comum para todas as bibliotecas da USP com aproximadamente 8.000 entradas para a representação de conteúdo do mencionado acervo. Entre 1988 e 1989, com a inclusão de termos novos, a Lista de Assuntos USP passou a ter 8.300 assuntos autorizados.

Em 1992, com o início da modernização do Sistema Integrado de Bibliotecas da USP (SIBi/USP), o aprimoramento da atual Base de Dados Bibliográficos DEDALUS foi definido como atividade prioritária para o Sistema; e em 1993, foi montado um grupo de bibliotecários cuja atribuição específica era expandir e atualizar a Lista de Assuntos então usada.

A orientação metodológica da estruturação do Vocabulário esteve sob a responsabilidade de docentes do Departamento de Biblioteconomia e Documentação da Escola de Comunicações e Artes da USP (CBD/ECA/USP) que ministraram um curso sobre Princípios de Compatibilização de Linguagens Documentárias, de

maneira a estabelecer um patamar comum de conhecimentos e procedimentos para o grupo envolvido na tarefa. Com a participação de cerca de 40 (quarenta) bibliotecários do Sistema e de inúmeros docentes especialistas das várias Unidades da USP, foi desenvolvido um projeto para a estruturação dos sistemas conceituais e adequação das terminologias específicas das áreas do conhecimento – Projeto para Aprimoramento da Lista de Assuntos USP, com suporte metodológico dos docentes participantes da linha de pesquisa "Análise Documentária e Terminologia" do CBD/ECA/USP. Também participaram do Grupo, uma bibliotecária do Sistema, responsável pela Coordenação Técnica e uma docente da ECA, responsável pela Coordenação Metodológica, além da equipe de bibliotecários, analistas e técnicos do Departamento Técnico do SIBi/USP que criaram a base de dados para viabilizar o Vocabulário Controlado do SIBi/USP.

Com metas claras estabelecidas, procedimentos sistemáticos, controlados e submetidos a ajustes periódicos e uma equipe sintonizada com os objetivos globais do trabalho ficou assegurada a realização do projeto.

Em 2001, o Vocabulário Controlado do SIBi/USP foi implantado na Base de dados bibliográficos DEDALUS através de um software denominado Sibix, desenvolvido por analista de sistemas do DT/SIBi, distribuído em CD-ROM, e depois acessado apenas pela intranet da USP. A partir de 2005, o Vocabulário foi disponibilizado via Web, podendo ser usado por qualquer instituição interessada.

O modelo de gestão do Vocabulário foi apresentado por Lima e col (2006) quando se verificou a necessidade de finalizá-lo como projeto e inseri-lo no Sistema de Bibliotecas da USP como um processo sistêmico de responsabilidade do Departamento Técnico do SIBi/USP.

Este modelo definiu os objetivos do Grupo de Gestão do Vocabulário, sua estrutura e participantes. Atualmente, é constituído por 3 bibliotecários do SIBi/USP para cada grande área do conhecimento (Ciências Humanas, Exatas e Biológicas), por um bibliotecário coordenador de conteúdo, por um bibliotecário coordenador de processo, por bibliotecários do Departamento Técnico e por um professor especialista em Linguagens Documentárias do Departamento de Biblioteconomia da Escola de Comunicações e Artes da USP. A indicação de membros do Grupo é feita

a cada dois anos, podendo os membros ser reconduzidos e procura-se também, renovar o grupo com a indicação de novos membros para garantir a continuidade das atividades de gerenciamento e controle do Vocabulário.

O processo de manutenção e aprimoramento do Vocabulário Controlado do SIBi/USP demanda também a colaboração de todas as Bibliotecas do Sistema, que apresentam sugestões de descritores baseadas em estudos da terminologia das áreas, através de um software online desenvolvido para esta atividade.

### **Importância da consistência de descritores**

Analisando alguns artigos sobre revisão de linguagens documentárias e/ou de registros, foi possível perceber a utilização de parâmetros terminológicos, reforçando o modelo metodológico utilizado para a consistência de descritores realizada na Base de Dados Bibliográficos DEDALUS da USP em 2008.

Harris (1979) estudou as mudanças na terminologia em Ciência da Informação durante 11 anos para determinar os efeitos destas mudanças no vocabulário da área. Chegou a uma taxa de mudança de 4% do total de termos excluídos e incluídos. Indicou neste estudo, a importância de que sejam previstas mudanças e atualizações nas linguagens documentárias com a retirada dos descritores em desuso e a adição de novos.

Para a disponibilização online da base de dados CINAHL – Nursing and Allied Health pelo DIALOG, foi realizada uma revisão da lista de cabeçalhos de assuntos utilizada na indexação da base, visando melhorar a qualidade da recuperação pelo novo sistema de acesso. Como resultado, manteve-se a política de atender às necessidades dos usuários e as mudanças ocorridas na literatura em Enfermagem, portanto na terminologia da área. (FISHEL e col. 1985).

Em 1987, Janik e Brunet analisaram a metodologia desenvolvida pelo Quebec Office for Handicapped People para a atualização e expansão do Quebec Thesaurus: Handicapped people, quando se aproveitou 50% da versão original, com a expansão dos termos médicos e educacionais, desenhado com base nos novos recursos políticos e sociais. A terminologia e a estrutura foram analisadas detalhadamente.

O Vocabulário Controlado do SIBi/USP prevê, pelo seu sistema de gestão, que as alterações, inclusões e exclusões de descritores baseadas na terminologia de cada área, sejam feitas com periodicidade constante, pretendendo-se que as sugestões enviadas pelas bibliotecas do Sistema sejam resolvidas no menor prazo possível o que, para descritores sugeridos com adequada base terminológica, deve ocorrer em até uma semana.

Analisando a obra de Ranganathan, Gomes, Motta e Campos (2006) lembram que a última lei fundamental proposta pelo bibliotecário indiano é a Lei da Osmose que determina que, quando ocorre mudança no sistema de classificação, os itens alterados serão afetados nas estantes, no catálogo e na base de dados, o que implica na necessidade de mudá-los. Porém, se o custo da mudança for alto, deverá prevalecer a Lei da Parcimônia, ou seja, da economia, e soluções deverão ser encontradas para indicação ao usuário das mudanças ocorridas.

Assim, estando encaminhada a questão do gerenciamento do Vocabulário em nível de controle terminológico, percebeu-se a necessidade de revisar os índices de descritores do DEDALUS que continham, além dos descritores do Vocabulário, palavras pertencentes à lista antiga, o que dificultava um resultado adequado nas pesquisas realizadas por assunto na base.

### **Consistência de descritores: abordagem terminológica e metodologia**

A Terminologia fornece, para a indexação e para as linguagens documentárias, importante reforço conceitual para a compreensão do conteúdo dos documentos e para a definição dos termos utilizados como descritores.

Segundo Cabré (1999), a Terminologia tem por objetivo estudar os termos de uma especialidade e são unidades de conhecimento do mundo ou de representação da idéia de organização do mundo.

Para as disciplinas técnico-científicas, a terminologia é um conjunto de unidades de expressão, comunicação e denominação que permitem transferir o pensamento especializado. Os termos são importantes para o intercâmbio de informação.

Para as áreas de especialidade, os termos relacionam-se com outros no âmbito da comunicação, configurando uma série de conexões ou campos terminológicos para representar a realidade do mundo especializado (CABRÉ, 1999).

O *corpus* para o estudo terminológico parte da linguagem real e documentada para dar conta de uma área especializada. A Terminologia é um instrumento poderoso para os processos de indexação, elaboração de linguagens documentárias e, conseqüentemente, para os processos de busca e recuperação da informação.

A Terminologia ao ter como ponto de partida o conceito, pratica a ordenação sistemática de suas entradas como na proposta de organização das linguagens documentárias. Apresenta-se como uma disciplina importante para a análise documentária e, portanto, para a indexação, devido ao interesse pela padronização dos termos, pela análise conceitual, e pela tentativa de estabelecimento de univocidade dos termos para resolver a polissemia e a sinonímia, sendo uma ferramenta imprescindível para a elaboração, gestão e atualização das linguagens documentárias.

A Base de Dados Bibliográficos da USP, o DEDALUS, antes da implementação do Vocabulário Controlado do SIBi/USP, já possuía uma lista de assuntos, porém, apesar do cotejamento realizado entre as duas listagens, muitos desses descritores não foram incluídos na nova linguagem, por serem muito genéricos ou partes de cabeçalhos de assuntos, ou por questões terminológicas que necessitariam de uma análise mais aprofundada.

Depois de alguns anos de gerenciamento do Vocabulário a necessidade de compatibilizar os registros cujos assuntos não foram contemplados pelo Vocabulário tornou-se premente.

Foi realizado então um novo cotejamento entre os descritores da antiga lista de assuntos e o Vocabulário constatando-se incidência de cabeçalhos de assunto (como Títulos Nominativos – Direito Comercial, Sistemas informacionais e computacionais de Transporte, Organizações – Área Cultural), terminologias antigas (como Tireóide – Câncer), uso inadequado de qualificadores (como Parasitologia – técnicas) e junção de conceitos diferentes (Probabilidade e Estatística, Odontologia Social e Preventiva), portanto, muitos registros tiveram que ser reindexados, através

da substituição, via máquina, pelos descritores atualizados e pela reindexação manual quando havia dúvidas sobre a terminologia adequada.

### **Resultados e considerações finais**

Foram analisados 5.100 descritores da lista antiga, sendo reindexados 519.380 (quinhentos e dezenove mil trezentos e oitenta) num universo de aproximadamente 2.000.000 (dois milhões) de registros existentes no DEDALUS, no período compreendido entre março de 2008 e março de 2009 sendo destes, 506.355 por máquina e 13.025 registros acertados manualmente por bibliotecários e técnicos do Grupo de Gestão do Vocabulário e do DT/SIBi.

Com certeza a base de dados ganhou em qualidade de indexação e recuperação da informação, passando a disponibilizar um índice compatibilizado e com mais consistência.

Este trabalho permitiu identificar inconsistências na indexação, tanto de registros antigos quanto mais recentes, estes, já indexados com o Vocabulário.

Notou-se em alguns casos pouco conhecimento da terminologia da área, pois registros com assuntos que possuíam descritores consolidados na especialidade foram indexados com descritores gerais acrescidos de qualificadores, por ex. Diagnóstico (Odontologia) no lugar de Diagnóstico Bucal. Verificou-se também a tendência a uma indexação genérica, o que não é a política de indexação do sistema, que recomenda a especificidade, e registros indexados erroneamente, portanto perdidos em uma pesquisa por assunto.

Cabe salientar que, para melhorar a qualidade da indexação realizada de forma colaborativa, onde nem todos os indexadores conhecem adequadamente as terminologias de diversas áreas do conhecimento ao mesmo tempo e onde a interdisciplinaridade é uma constante, deve-se aprimorar a linguagem documentária, os manuais e o treinamento dos indexadores.

Observou-se com isso a necessidade de treinamento tanto no que se refere à parte terminológica quanto às normas de indexação adotadas pelo sistema.

Foram identificados descritores importantes nas áreas que não haviam sido contemplados no Vocabulário, portanto, foi necessária a inclusão de vários deles através da base de sugestões online. Percebeu-se também que algumas áreas precisam ser aprimoradas no Vocabulário.

Torna-se importante reforçar a utilização de uma metodologia de indexação com diretrizes e normas de uso adequadas, o desenvolvimento de uma linguagem documentária com estrutura lógico-semântica clara para o indexador na representação da informação e, para o usuário, na recuperação.

O treinamento do indexador na metodologia do sistema de informação deve acrescentar, além dos aspectos teóricos, práticas de indexação e uso da linguagem documentária, e a aquisição de conhecimentos específicos da área interdisciplinar através de cursos sobre a temática.

Os resultados alcançados reforçaram que a abordagem terminológica é adequada para o desenvolvimento e manutenção de uma linguagem documentária, bem como para os procedimentos de indexação e para os processos de análise e revisão previstos em um sistema de informação.

### **Referências Bibliográficas**

- CABRÉ, M. T. **La terminologia**: representación y comunicación, elementos para una teoría de base comunicativa y otros artículos. Barcelona: Institut Universitari de Lingüística Aplicada; Universitat Pompeu Fabra, 1999. 369 p.
- FISHEL, C. C et al. List of subject headings: a nursing thesaurus revised. **Bulletin of the Medical Library Association**, v.73, n. 2, p. 153-159, Apr. 1985.
- GOMES, H. E.; MOTTA, D. F.; CAMPOS, M. L. A. **Revisitando Ranganathan**: a classificação na rede. Ago. 2006. Disponível em: <http://www.conexaorio.com/biti.revisitando/revisitando.htm>.
- HARRIS, J. Terminology change: effect on index vocabularies. **Information Processing and Management**, v.15 n. 2, p.77-88, 1979.
- JANIK, S; BRUNET, L. La mise a jour d'un thesaurus. **Documentaliste**, v. 24, n. 6, p. 215-229, Nov.-Dec. 1987.
- LANCASTER, F. W. **Avaliação de serviços de bibliotecas**. Brasília: Briquet de Lemos/Livros, 1996.

LIMA, V. M. A et al. Estudos para implantação de ferramenta de apoio à gestão de linguagens documentárias: Vocabulário Controlado da USP.

**Transinformação**, Campinas, v. 18 n.1 p.17-25, jan./abr., 2006.

McGARRY, K. **O contexto dinâmico da informação**: uma análise introdutória. Brasília: Brique de Lemos/Livros, 1999.

SISTEMA INTEGRADO DE BIBLIOTECAS DA USP. DEPARTAMENTO TÉCNICO. **SIBiNet**. Abr. 2009. Disponível em: <http://www.usp.br/sibi>.

SVENONIUS, E. **Intellectual foundation of information organization**. Cambridge: MIT Press, 2001.