



SENADO FEDERAL
DIRETORIA GERAL
SECRETARIA DE BIBLIOTECA

PADRONIZAÇÃO DO TRATAMENTO DOCUMENTAL NO SENADO FEDERAL

PROJETO

Brasília

Novembro de 2012

Elaboração do projeto Padronização do Tratamento Documental no Senado Federal (nov. 2012):

Simone Bastos Vieira

(Diretora da Secretaria de Biblioteca)

Elaine Ricevich Bastos de Oliveira

(Diretora da Subsecretaria de Processamento Técnico de Informações Bibliográficas)

Maria de Fátima P. Jaegger

(Chefe do Serviço de Gerência da Rede Virtual de Bibliotecas - RVBI)

Osmar Arouck

(Analista Legislativo - Bibliotecário)

SUMÁRIO

1. Setor proponente	3
2. Objeto	3
2.1 Relação do objeto com o planejamento estratégico do Senado	3
3. Objetivo	3
4. Justificativa.....	3
4.1 Processo de análise, descrição e representação de conteúdos	4
5. Motivador do projeto	6
6. Benefícios esperados.....	6
6.1 Qualidade da informação	7
6.2 Eficiência na padronização dos procedimentos	7
7. Público-alvo.....	8
7.1 Usuários dos sistemas de informação do Senado Federal	8
7.2 Unidades do Senado Federal que gerenciam os sistemas de informação	8
7.3 Usuários dos sistemas de informação com os quais o Senado se integra	8
8. Histórico	9
8.1 Vocabulário Controlado Básico (VCB)	9
8.2 Thesaurus (THES)	9
8.3 Resumo comparativo	10
9. Estratégia de implementação.....	10
10. Unidade responsável	10
11. Valor do programa	11
12. Indicadores.....	11
13. Referências bibliográficas	13

PLANO DE AÇÃO

1. SETOR PROPONENTE

Diretoria-Geral do Senado Federal

2. OBJETO

Integração dos sistemas de informação do Senado Federal para garantir a gestão da qualidade da informação em acervos e bases de dados, por meio de padronização do tratamento documental e de definição de terminologia comum às diversas áreas de atuação.

2.1 Relação do objeto com o planejamento estratégico do Senado

De acordo com a missão, objetivos, estratégias e metas da Diretoria-Geral do Senado Federal.

3. OBJETIVO

Padronizar os procedimentos de tratamento documental relativos à descrição física dos documentos e à análise do conteúdo e representação de assuntos adotados nos diversos sistemas de informação do Senado Federal.

4. JUSTIFICATIVA

Informação de qualidade é matéria-prima para o pleno desenvolvimento da democracia, exercida pelos representantes do povo e seus Estados. Assim, a informação de qualidade é subsídio estratégico à ação parlamentar, contribuindo com a credibilidade e a eficácia do Legislativo junto ao cidadão brasileiro. A gestão da informação busca garantir a precisão, objetividade e imparcialidade no tratamento da informação para sua recuperação eficiente e isenta de ambiguidades.

Essa gestão compreende diversos processos que visam garantir o acesso pleno, simples e eficiente aos conteúdos informacionais de interesse dos usuários, independente de seu nível de conhecimento sobre um determinado assunto ou padrões de linguagem. Dentre esses processos, destacam-se os procedimentos técnicos e metodológicos de representação descritiva e temática da informação, de acordo com padrões internacionais, com o propósito de integrar e compartilhar acervos em seus diversos meios, suportes e conteúdos.

No âmbito da representação temática, os fenômenos próprios da linguagem podem interferir na qualidade da recuperação da informação. Situações como ambiguidades (dois sentidos para uma mesma palavra) e sinônimas (duas palavras para um mesmo sentido) devem ser gerenciadas cuidadosamente para que não ocorra a recuperação de dados não pertinentes, ou perda de dados relevantes. As generalizações (palavras que podem assumir um sentido mais abrangente) e as particularizações (palavras que podem assumir um sentido mais específico) precisam estar explicitadas na estrutura do vocabulário para que o conteúdo analisado seja indexado no nível mais adequado para sua adequada recuperação.

A representação dos documentos e seus assuntos são um desafio constante para os bibliotecários, na medida em que as necessidades dos usuários, expressadas em linguagem natural, precisam ser traduzidas para uma linguagem controlada e atualizada e, ao mesmo tempo, acessível a todos.

Esses e outros fenômenos mostram a pluralidade e complexidade existentes no tratamento documental. O processo de recuperação da informação em bases e sistemas requer a unidade entre a descrição física e a descrição do conteúdo. Na linguagem natural o signo linguístico é plural; para a recuperação da informação ele deve ser unívoco. Neste cenário, surgem as linguagens documentárias para dirimir as ambiguidades, como vocabulários controlados e tesouros, que traduzem a linguagem natural utilizada nos documentos.

No Senado Federal, existem diversas ferramentas para a descrição, tratamento de conteúdo dos documentos e seu controle vocabular, que comprometem a acessibilidade aos conteúdos de informação existentes nos diversos sistemas gerenciados pelas unidades afetas à gestão da informação e do conhecimento.

4.1 Processo de análise, descrição e representação de conteúdos

O processo de análise e representação de assuntos inclui procedimentos que consistem em extrair os conceitos que caracterizam o conteúdo de um documento, por meio da leitura documentária, e a representação desses conceitos com termos expressivos do seu conteúdo e/ou descritores padronizados, com o fim de facilitar a recuperação das informações.

A análise de assuntos é a fase inicial desse processo, requerendo a compreensão do conteúdo do texto para a identificação dos conceitos, que serão selecionados segundo as diretrizes adotadas. Estreitamente relacionada a essa fase, encontra-se a leitura documentária, ou leitura técnica, que é a leitura profissional para a consecução dos objetivos e elaboração de produtos documentários.

A leitura documentária, diferentemente da leitura comum, exige técnicas para sua execução, bem como capacitação profissional, conhecimento e experiência (FUJITA, 2004)¹. Lancaster (1993, p. 79)² identifica alguns erros que ocorrem no processo de indexação e que devem ser evitados:

- a) O indexador infringe a política, especialmente a política relativa à exaustividade da indexação.
- b) O indexador deixa de empregar os elementos do vocabulário de forma como devem ser utilizados (por exemplo, uma combinação incorreta de cabeçalho principal/subcabeçalho).
- c) O indexador deixa de utilizar um termo no nível correto de especificidade. Na maioria dos casos isto significará que o termo selecionado não é o mais específico existente.
- d) O indexador emprega um termo evidentemente incorreto, talvez devido à falta de conhecimento especializado (por exemplo, combustíveis líquidos para foguetes quando o documento trata mesmo é de combustíveis gasosos).
- e) O indexador omite um termo importante.

Dentre as diversas linguagens documentárias existentes na Casa, comparamos as duas mais antigas e utilizadas nas bases de dados com maior volume de registros: o Vocabulário Controlado Básico (VCB), desenvolvido pela Biblioteca do Senado, e o Thesaurus (THES), elaborado pela Subsecretaria de Informações.

O VCB tem, hoje, 9.692 termos e 523 especificadores, e o THES tem 10.430 termos. Ocorre que diversos descritores coincidem em ambos os instrumentos, e alguns termos são adotados em um e não o são em outro instrumento. Isto para exemplificar casos que prejudicam a coerência da indexação feita com diferentes instrumentos de controle vocabular.

A *coerência na indexação*, conforme Lancaster (1993, p. 61), refere-se à extensão com que existe concordância quanto aos termos a serem usados para indexar um documento, ressaltando que “um vocabulário controlado deve melhorar a coerência da indexação em relação a um grupo de documentos”.

¹ FUJITA, Mariângela Spotti Lopes. A leitura documentária na perspectiva de suas variáveis: leitor-texto-contexto. DataGramZero: revista de ciência da informação, v. 5, n. 4, ago. 2004. Disponível em: <http://www.dgzero.org/ago04/Art_01.htm>. Acesso em: 1 mar. 2007.

² LANCASTER, F. W. Indexação e resumos: teoria e prática. 2. ed. Brasília: Brique de Lemos, 1993. 347 p.

Atente-se também ao chamado *princípio da especificidade* preconizado por Cutter (1876), conforme registra Lancaster (1993, p. 27). Esse princípio, considerado o mais importante da indexação de assuntos, é aquele segundo o qual um tópico deve ser indexado sob o termo mais específico que o abranja completamente.

Atualmente, com a adoção de duas ferramentas similares, mas não integradas, ocorrem discrepâncias terminológicas entre as ferramentas em tela, como exemplificado abaixo, a partir de uma pesquisa feita no SICON e no catálogo coletivo da RVBI:

ASSUNTO	VCB		THES	
	Termo	Documentos recuperados	Termo	Documentos recuperados
Servidão ambiental ³	1	24	0	0
Pregão eletrônico ⁴	1	98	0	0
Termo de ajustamento de conduta	1	11	0	0

No Anexo 1, resultado de pesquisa sobre “pregão eletrônico” nas bases de dados da Biblioteca e do Sicon.

5. MOTIVADOR DO PROJETO

Reforma administrativa do Senado Federal e a necessidade de integração das bases de dados gerenciadas pelas unidades administrativas que têm como responsabilidade a gestão da informação.

6. BENEFÍCIOS ESPERADOS

A uniformização da representação dos documentos quanto aos aspectos físicos e de conteúdo, registrados nos diversos sistemas de informação do Senado Federal, pode contribuir para a melhoria da qualidade da informação e para a eficiência na padronização dos procedimentos.

³ Lei nº 11.284, de 2 de março de 2006.

⁴ Lei nº 10.520, de 17 de julho de 2002.

6.1 Qualidade da informação

Identificam-se três dimensões da qualidade da informação: acessibilidade, utilidade e credibilidade (WANG *et al.*, 1995)⁵. Esses autores postulam que, para ser acessível, a informação deve estar disponível; para ser útil, deve ser relevante e oportuna; para ser crível, o usuário considerará se a informação é completa, coerente e precisa.

6.2 Eficiência na padronização dos procedimentos

A avaliação de um sistema de recuperação de informações deve considerar os resultados alcançados ao final de todos os procedimentos técnicos efetuados, ou seja, “uma base de dados bibliográficos não pode ser avaliada de forma isolada, mas somente em função de sua utilidade ao responder a várias necessidades de informações” (LANCASTER, 2004, p. 135).

Destacam-se como vantagens da unificação de procedimentos documentais:

- a) Eficiência no uso de recursos humanos especializados;
- b) Recuperação eficaz dos recursos informacionais;
- c) Integração e harmonização das políticas de indexação de conteúdos dos diversos acervos e bases de informação do Senado Federal;
- d) Integração e harmonização com padrões de indexação adotados em outros sistemas de informação com os quais o Senado colabora.

Os profissionais que lidam com a informação devem considerar que a tradução ou representação dos conceitos é semelhante tanto na entrada da informação (fase de processamento), como na saída (pesquisa); para que as pesquisas nas bases de dados sejam positivas, deve haver uma total correspondência entre as linguagens utilizadas nessas fases, pois a recuperação automática funciona por coincidência de caracteres das palavras e não por busca de conceitos.

O bibliotecário não deve desempenhar apenas sua função técnica, mas imaginar-se no papel do usuário. É necessário que haja uma correspondência entre as linguagens utilizadas nas etapas de processamento e de recuperação das informações, onde o produto da indexação represente a expectativa dos usuários. O intercâmbio permanente entre os bibliotecários das áreas de processamento e de

⁵ WANG, R. Y. *et al.* Toward quality data: an attribute-based approach. *Decision Support Systems*, v. 13, n. 3-4, March 1995, p. 349 – 372.

recuperação, para aprimorar o nível de qualidade das informações analisadas e recuperadas é, também, fundamental nesse processo.

É importante o intercâmbio entre os profissionais que trabalham nas áreas de descrição e aqueles que atuam nas áreas de recuperação e atendimento aos usuários, para o aumento do nível de qualidade das informações analisadas e recuperadas. Neste sentido, vale recordar a orientação de LANCASTER (1993, p. 22) quando afirma que “uma função importante do indexador é contribuir para o aperfeiçoamento do vocabulário controlado, comunicando suas deficiências aos responsáveis pela sua manutenção”⁶.

7. PÚBLICO-ALVO

7.1 Usuários dos sistemas de informação do Senado Federal

Senadores
Servidores
Unidades do Senado Federal

7.2 Unidades do Senado Federal que gerenciam os sistemas de informação

Secretaria Geral da Mesa (SGM)

Secretaria de Ata (SATA)
Serviço de Sinopse (SESINO)

Diretoria-Geral (DGER)

Secretaria de Arquivo (SARQ)
Secretaria de Biblioteca (SBIB)
Secretaria de Informação e Documentação (SIDOC)
Subsecretaria de Informações (SSINF)
Secretaria Especial de Editoração e Publicações (SEEP)
Subsecretaria de Anais (SSANS)

7.3 Usuários dos sistemas de informação com os quais o Senado se integra

Usuários e pesquisadores do portal do Senado
Consórcio BDJur
LeXML
Rede Virtual de Bibliotecas – Congresso Nacional – RVBI

⁶ LANCASTER, F. W. Indexação e resumos: teoria e prática. 2. ed. Brasília: Briquet de Lemos, 1993. 347 p.

8. HISTÓRICO

A compatibilização dos procedimentos documentais, e do VCB e THES, é um projeto que remonta à década de 1980, conforme registra Lemos (1986)⁷. O VCB, adotado pela Rede Virtual de Bibliotecas – Congresso Nacional – RVBI, e o THES, mantido pela Subsecretaria de Informações, da Secretaria de Documentação e Informação, têm funções e objetivos convergentes que permitem a proposição de um instrumento único que integre e harmonize o conteúdo e estrutura de ambos.

8.1 Vocabulário Controlado Básico (VCB)

O VCB foi desenvolvido na década de 1980 com estrutura de tesouro. A utilização do VCB visa manter a uniformidade da indexação e recuperação das informações nas bases de dados bibliográficos da RVBI, compostas por livros, publicações seriadas, obras raras, artigos de periódicos e de jornais, documentos digitais e eletrônicos.

O VCB é mantido pela Biblioteca do Senado Federal (SBIB), com a colaboração das bibliotecas da RVBI por meio de sugestões e indicações de necessidades de termos novos. Desde 2002 a Biblioteca do Senado coordena o Grupo de Estudo do Tesouro da RVBI, formado por representantes de algumas bibliotecas da Rede, com o objetivo de reformular, atualizar e sistematizar o Vocabulário Controlado Básico - VCB, aplicando normas para a elaboração de um tesouro multidisciplinar monolíngue para a RVBI.

Um dos principais produtos do Grupo é o manual “Diretrizes para construção do tesouro da Rede Virtual de Bibliotecas – Congresso Nacional – RVBI”, publicado em 2007, utilizado como instrumento de trabalho para a manutenção do VCB.

A atualização terminológica do VCB é um trabalho de rotina da Biblioteca do Senado e, atualmente, o Grupo está aguardando a implantação de um software gerenciador de tesouros para dar continuidade ao trabalho de revisão. No período de 2001 a 2012, o VCB foi atualizado com 676 novas entradas, sendo: 451 termos/descriptores, 178 nomes geográficos e 47 especificadores (ANEXO 2).

8.2 Thesaurus (THES)

O **THES** foi desenvolvido na década de 1970, para tratamento de informações jurídicas, normas jurídicas federais, matérias em tramitação, jurisprudência e discursos

⁷ LEMOS, M. L. V. de. Desenvolvimento de um vocabulário controlado na Biblioteca do Senado Federal. *Ciência da Informação*, v. 15, n. 2, jul./dez. 1986, p. 155-158.

parlamentares, pela Subsecretaria de Informações (SSINF). Em 1976 o THES tinha 8.484 termos. Nessa época, havia uma “previsão de hierarquização e refinamento do léxico [...]”⁸, para a estruturação do tesouro, projeto que não foi implementado.

O THES é atualmente usado na indexação das seguintes bases legislativas:

- NJUR – Normas jurídicas
- MATE – Matérias legislativas
- DISC – Discursos
- NADM – Normas administrativas

8.3 Resumo comparativo

Base de dados	Vocabulário	Manutenção do vocabulário	Indexação dos documentos	Quantidade de documentos
RVBI	VCB	SBIB	SBIB / RVBI	960.000
BDSF	VCB	SBIB	SBIB	222.790
NJUR	THES	SSINF	SSINF	227.051
MATE	THES	SSINF	SESINO	(⁹)
NADM	THES	SSINF	SIDOC	(¹⁰)
DISC	THES	SSINF		124.519

9. ESTRATÉGIA DE IMPLEMENTAÇÃO

- a) Identificação de procedimentos comuns às áreas de gestão da informação.
- b) Identificação e integração dos termos equivalentes nas linguagens documentárias.
- c) Análise e proposição de macroestrutura para a padronização de procedimento de tratamento da informação.
- d) Análise e proposição de áreas temáticas para a linguagem documentária.

10. UNIDADE RESPONSÁVEL

Secretaria de Informação e Documentação

Secretaria de Biblioteca

Subsecretaria de Informações

⁸ DUTRA, Yamil e Sousa. *Indexação de textos legislativos* : a experiência da Subsecretaria de análise do Senado Federal. Revista de Informação Legislativa, Brasília, v. 14, n. 55, p. 227-244, jul./set. 1977.

⁹ Sem informação

¹⁰ Sem informação

11. VALOR DO PROGRAMA

- a) Custos administrativos relativos aos servidores envolvidos no projeto.
- b) Cursos de capacitação.
- c) Software gerenciador de tesouros.

12. INDICADORES

Relação percentual entre a quantidade de termos analisados e integrados em relação à quantidade total de termos existentes no início do projeto. Considerando-se a quantidade de termos do VCB e do THES, estima-se o total de 20.122 termos.

13. REFERÊNCIAS BIBLIOGRÁFICAS

DUTRA, Yamil e Sousa. Indexação de textos legislativos: a experiência da Subsecretaria de Análise do Senado Federal. *Revista de informação legislativa*, Brasília, v. 14, n. 55, p. 227-244, jul./set. 1977.

FUJITA, Mariângela Spotti Lopes. A leitura documentária na perspectiva de suas variáveis: leitor-texto-contexto. *DataGramaZero: revista de ciência da informação*, v. 5, n. 4, ago. 2004. Disponível em: <http://www.dgzero.org/ago04/Art_01.htm>. Acesso em: 1 mar. 2007.

LANCASTER, F. W. *Indexação e resumos: teoria e prática*. 2. ed. Brasília: Brique de Lemos, 1993. 347 p.

LEMOS, M. L. V. de. Desenvolvimento de um vocabulário controlado na Biblioteca do Senado Federal. *Ciência da informação*, v. 15, n. 2, jul./dez. 1986, p. 155-158.

WANG, R. Y. et al. Toward quality data: an attribute-based approach. *Decision support systems*, v. 13, n. 3-4, March 1995, p. 349 – 372.

ANEXO 1

The screenshot shows the SICON website interface in Internet Explorer. The browser title is "SICON : Senado Federal - Portal de Legislação - Windows Internet Explorer". The address bar shows "http://legis.senado.gov.br/sicon/". The page content is titled "Resultado da Pesquisa Básica" and displays search options and results for "MPV 2026-001".

PESQUISA

Básica

Avançada:

- Padronizada
- Tradicional

Legislação:

- Rápida
- Textual

Biblioteca

Cesta de Ítems

Preferências

Ajuda

Fale Conosco

Resultado da Pesquisa Básica

Opções de visualização [Melhore sua Pesquisa](#) | [Ver argumentos](#) | [Imprimir](#) | [Relatório](#)

Página 1 de 3 26 documentos [Próxima](#)

Legislação Federal - MPV

Autor PODER EXECUTIVO

Título MPV 2026-001 de 01/06/2000 - MEDIDA PROVISÓRIA

Data 01/06/2000

Resumo INSTITUI, NO AMBITO DA UNIÃO, NOS TERMOS DO ARTIGO 37, INCISO XXI, DA CONSTITUIÇÃO FEDERAL, MODALIDADE DE LICITAÇÃO DENOMINADA PREGÃO, PARA AQUISIÇÃO DE BENS E SERVIÇOS COMUNS.

Observação EDIÇÃO ORIGINAL: MPV-002026. EDIÇÃO ANTERIOR: MPV-002026. TOTAL DE REEDIÇÕES ANTERIORES: 1. NOTA: PERDEU A EFICACIA (PEF).

Editor Imprensa Nacional

Fonte Diário Oficial da União

Publicação DOFC PUB 02/06/2000 000028 1 Diário Oficial da União

Catálogo LICITAÇÃO.

Indexação FIXAÇÃO, MODELO, LICITAÇÃO, OBJETIVO, AQUISIÇÃO, BENS, SERVIÇO. CRITERIOS, COMPETENCIA, CONVOCACÃO, EDITAL, LICITAÇÃO.

Idioma por

Formato text/xml

Código 229.861

Relações MPV-002026 002 2000 DOFC 30/06/2000 000031 2 REEDIÇÃO COM ALTERAÇÃO # REEDIÇÃO COM ALTERAÇÃO

Fim 01/07/2000

Início 02/06/2000

[Detalha](#) [Texto Original](#) [Texto Atualizado](#) [Inclui](#) [Nautius](#)

Intranet local 100%

12:58

The screenshot shows a Windows Internet Explorer browser window displaying a catalog record. The title bar reads 'Catálogo Coletivo - Registro Completo - Windows Internet Explorer'. The address bar shows the URL: <http://biblioteca2.senado.gov.br:8991/F/GU1KDEXB3HIAAEIGAH3A7K4A9C1HU7MVNSXD4Y2PC>. The browser interface includes a menu bar with 'Arquivo', 'Editar', 'Exibir', 'Favoritos', 'Ferramentas', and 'Ajuda'. Below the menu bar are several tabs, with the active one being 'Catálogo Coletivo - Registro Completo'. The main content area is titled 'Registro Completo' and features a navigation bar with 'Registro Ant.' and 'Próx. Registro'. The record details are as follows:

Formato:	OPAC Ficha Reduzido Nomes MARC Campos MARC
Registro 1 de 1	
Base	Artigo de revista
Autor	Andrade, Fernanda Alves.
Título	O critério de seleção de licitantes à fase de lances no pregão : extensão e limites à interpretação do artigo 4º, IX, da lei nº 10.520/02 / Fernanda Alves Andrade. --
In	Fórum de contratação e gestão pública, v. 9, n. 100, p. 45-49, abr. 2010. Revista JML de licitações e contratos, v. 5, n. 20, p. 16-21, set. 2011.
Nota bibliografia	Inclui notas explicativas, bibliográficas e bibliografia.
N. outro form disp	Disponível, também, na Internet.
Assunto	Pregão, legislação, Brasil.
Assunto	Pregão eletrônico, Brasil.
Assunto	Pregão presencial, Brasil.
Assunto	Licitação, empresa, classificação, Brasil.
Legis. relacionada	LEI-010520 (2002) ART 00004 PAR 09. LEI-008666 (1993).
Versão do recurso	Texto Integral (PDF) na BDJur/Superior Tribunal de Justiça. Disponível para servidores do STJ. Acesso em: 4 ago. 2010.
No. sistema	000880586
Acervo Geral	Todos os itens
Itens na Biblioteca	AGU
Itens na Biblioteca	CAM
Itens na Biblioteca	CI

The bottom of the browser window shows the Windows taskbar with the 'Iniciar' button and several open applications. The system tray on the right displays the date and time as 13:07.

