

# INDEXAÇÃO DE TEXTOS LEGISLATIVOS

A experiência da Subsecretaria de Análise  
do Senado Federal

YAMIL E SOUSA DUTRA

## SUMARIO

1. A AUTOMATIZAÇÃO DO PROCESSAMENTO DE DADOS NO SENADO
2. O ACERVO INICIAL
3. PRIMEIRAS ATIVIDADES DA SUBSECRETARIA DE ANÁLISE
4. A ESCOLHA DE UMA METODOLOGIA
5. O MÉTODO DE ANÁLISE
6. A NORMA EXPERIMENTAL PARA A EXTRAÇÃO DE DESCRITORES
7. O **THESAURUS** DA CONSTITUIÇÃO
8. A ATUAL SITUAÇÃO DO TRABALHO
  - 8.1. A METODOLOGIA
  - 8.2. O ÍNDICE CONTROLADO
  - 8.3. A SUBSECRETARIA E O PRODASEN
    - 8.3.1. RECURSOS COMPUTACIONAIS DISPONÍVEIS
    - 8.3.2. INTER-RELAÇÃO DE RECURSOS HUMANOS
  - 8.4. O DESENVOLVIMENTO DE UM SISTEMA DE INFORMAÇÃO LEGISLATIVA
  - 8.5. O TREINAMENTO DOS ÓRGÃOS EM CONVÊNIO

## 1. A AUTOMATIZAÇÃO DO PROCESSAMENTO DE DADOS NO SENADO

Como resultado da Reforma Administrativa por que passou o Senado Federal a partir de novembro de 1972, e também devido à instalação de um Centro de Processamento de Dados, foi criada, na estrutura dessa Casa Legislativa, a Secretaria de Informação, órgão competente para operar um sistema de informações que seria desenvolvido tendo por base o tratamento automatizado de dados de natureza bibliográfica e jurídica.

Esta Secretaria, responsável pela manutenção, atualização e desenvolvimento de importantes bancos de dados no Serviço de Processamento de Dados do Senado (PRODASEN), é composta dos seguintes órgãos:

- Subsecretaria da Biblioteca
- Subsecretaria de Análise
- Serviço de Controle de Informações

No presente trabalho vamos enfocar especialmente as atividades e experiências levadas a efeito pela Subsecretaria de Análise, a que compete, sumariamente, manter e desenvolver os arquivos referentes a dados jurídico-legislativos, estruturando a metodologia de análise e recuperação destes dados.

## 2. O ACERVO INICIAL

Tão logo foi criado o Centro de Processamento de Dados do Senado Federal (PRODASEN), e utilizando-se de um convênio com o Serviço Federal de Processamento de Dados (SERPRO), que deu suporte técnico para a iniciativa, e através de um contrato com a subsidiária brasileira da IBM, foram implantados, nos bancos de dados, parte dos documentos existentes nos arquivos manuais do setor de Referência Legislativa da então Biblioteca do Senado Federal.

Os dados introduzidos no computador foram transcritos do Arquivo Manual da Seção de Referência Legislativa da Biblioteca do Senado Federal, o qual, elaborado ao longo de vários anos, não possuía um padrão de uniformização definitivo nem para sua aparência formal nem para sua terminologia de indexação.

De qualquer modo, cada um dos aproximadamente 68.000 documentos implantados, através de cartões perfurados e arquivados em fita magnética, continha os seguintes dados básicos:

- Tipo e nº de documento
- Indicação do órgão e data de publicação
- Referências:
  - sobre republicação ou retificação
  - sobre remissões, tais como:
    - regulamentação
    - insubsistência

alteração  
derrogação  
revogação

— Cabeçalho de assuntos

Estes documentos, que passaremos a chamar genericamente de **normas**, abrangiam o período de 1946 a outubro de 1972 e incluíam especificamente:

Atos Complementares  
Atos do Comando Supremo da Revolução  
Atos Institucionais  
Leis  
Leis Constitucionais  
Leis Delegadas  
Decretos Executivos  
Decretos Legislativos  
Decretos do Conselho de Ministros  
Decretos-Leis  
Decretos Reservados  
Emendas Constitucionais  
Resoluções do Senado Federal  
Resoluções da Câmara dos Deputados  
Resoluções do Congresso Nacional

### 3. PRIMEIRAS ATIVIDADES DA SUBSECRETARIA DE ANÁLISE

Instalada definitivamente a partir de janeiro de 1973, a Subsecretaria de Análise vem tendo por encargo específico operacionalizar o sistema de análise e recuperação dos dados jurídico-legislativos através da criação de instrumentos metodológicos apropriados e do treinamento adequado do pessoal.

Inicialmente, coube à Subsecretaria a tarefa de viabilizar a utilização do acervo implantado na fase inicial, o que tinha sido feito sem sua responsabilidade direta e, em consequência de prazos contratuais convencionados com a IBM, sujeito a inevitáveis deficiências decorrentes da falta de uma necessária e conveniente conferência, correção e padronização de dados.

Assim, a Subsecretaria teve que enfrentar duas importantes frentes de trabalho ao mesmo tempo. Por um lado, com a cooperação de programadores e analistas do PRÓDASEN, iniciou o desenvolvimento de **lay-outs** dos formulários de indexação e dos arquivos magnéticos, prevendo, desta forma, encontrar uma padronização formal para os documentos. Por outro lado, resolvida a questão formal, restava o aspecto mais importante do

ponto de vista da recuperação, ou seja, o controle da linguagem a ser utilizada na indexação dos documentos, já que havia sido decidido que, do ângulo da economia, a implantação do **full-text** de todos os documentos seria dispendiosa e desnecessária para mais de 90% dos mesmos, além de ser ineficiente para os casos de conceitos implícitos na norma implantada.

#### 4. A ESCOLHA DE UMA METODOLOGIA

Além de pessoal treinado e experimentado no campo da recuperação legislativa, a Subsecretaria, em seu início, contava com a cooperação e assistência de um grupo de assessores do Senado Federal, especializados em diferentes setores do Direito, assim como em outras ciências, tais como Economia, Sociologia, História, Lingüística, Biblioteconomia etc.

No primeiro momento, este pessoal partiu para a análise dos problemas suscitados no campo da indexação de legislação, tomando conhecimento dos esforços realizados em outras áreas e levando em conta, sobretudo, a experiência internacional. Assim, foram desenvolvidos estudos sobre organismos e instituições que estivessem levando a efeito trabalhos no setor, bem como sobre projetos semelhantes em vias de execução. Num segundo momento, teve-se a cooperação direta de alguns **experts** no assunto, dentre os quais destacamos o Barão Edouard de Houtart, Diretor do CREDOC, Serviço Belga de Documentação Jurídica, do Prof. Mario Losano, da Universidade de Milão, do bibliotecário francês Jean Vief e do Prof. Antônio Briquet, do Departamento de Biblioteconomia da Universidade de Brasília.

Mesmo com esta importante contribuição, mantinha-se um obstáculo, ou seja, o estabelecimento de uma metodologia própria para a análise do texto legal.

Entretanto, um ponto ficara bastante claro, ou seja, o de que a manutenção de um arquivo de dados jurídicos, e mais especificamente, de um arquivo de referência legislativa, é, fundamentalmente, ao contrário do que se possa supor, um problema de documentalista, mais do que uma questão a ser resolvida exclusivamente com o uso de juristas como analistas dos documentos.

Basicamente, o instrumental para indexar um documento bibliográfico comum serve para indexar uma norma jurídica. Porém, a diferença colocase no momento da recuperação dos dados indexados, pois enquanto um **recall** significativo de 90% para dados bibliográficos comuns pode ser considerado excelente, o mesmo índice não teria qualquer validade no campo da informação legislativa caso, naqueles 10% não atingidos, estivesse contida a norma que revogasse ou alterasse substancialmente os 90% recuperados. Daí a necessidade de, na indexação de documentos legislativos, criar instrumentos que permitam um **recall** significativo tão próximo dos 100% quanto possível.

Tendo em conta esta necessidade, admitiu-se que, para alcançar aquele padrão mínimo de recuperação, seria necessário o desenvolvimento de um instrumento que viabilizasse este tipo de demanda, e a estruturação de um **Thesaurus** pareceu um objetivo a ser atingido.

Terminada a fase inicial de “conhecimento de terreno”, os assessores e os técnicos optaram pela realização de versões experimentais de análise de textos.

Foram, então, organizados quatro grupos de trabalho, os quais analisaram simultaneamente a legislação trabalhista, a legislação tributária, a Constituição Federal (Emenda Constitucional nº 1, de 1969), e a legislação esparsa publicada diariamente no **Diário Oficial** da União.

Este trabalho experimental, sustentado no objetivo de criar um instrumental de indexação e recuperação capaz de sobrepor-se aos problemas típicos de documentação jurídico-legislativa, visava primordialmente dois aspectos:

a) criar a linguagem documental que servisse de base conceitual para o desenvolvimento normal e permanente da metodologia de indexação do acervo de todos os documentos já implantados nos arquivos magnéticos e daqueles que viessem a fazer parte dos mesmos;

b) dar treinamento prático, para um grupo de leigos, no método de elaboração de um **Thesaurus**.

## 5. O MÉTODO DE ANÁLISE

A tarefa básica dos grupos de trabalho foi, através de análise, levantar do universo de documentos sob sua responsabilidade aqueles conceitos explícitos ou implícitos nos mesmos.

No primeiro momento houve uma tendência, principalmente por parte dos analistas treinados na profissão jurídica, de indexar os documentos seguindo a ortodoxia da classificação do Direito. Entretanto, esta alternativa mostrou-se inadequada tanto em face dos documentos como do que seria exigido do sistema em termos de informação.

Que campo do Direito poderia abrigar uma norma que tratasse da obrigatoriedade de utilização de “registros de segurança” em caldeiras de média e alta pressão? Seria Direito do Trabalho, seria Norma da Higiene do Trabalho, seria Garantias e Direitos do Cidadão, seria Direito Comercial? A situação agrava-se ainda mais quando é levado em conta o fato de que o cliente da informação não questionaria o banco de dados para saber quantas normas do mesmo pertenciam a este ou àquele campo do Direito, mas suas questões incidiriam sobre situações práticas e reais, tais como: “Há alguma norma a respeito de instalação de caldeiras?”, “Vocês têm alguma informação sobre aspectos específicos da segurança de equipamentos de pressão?”, e assim por diante.

Este tipo de problema apresentou-se para aqueles analistas que, por formação, tendiam a ver a solução na aplicação de classificações gerais, tais como a Classificação Decimal de Dewey, a Classificação da Biblioteca do Congresso (EUA), ou a Classificação Decimal Universal.

Ficou claro para todos os grupos de trabalho que a solução estava muito mais na linguagem documental de estrutura combinatória, próxima

da linguagem natural, do que na aplicação pura e simples de alguma linguagem de estrutura hierárquica como a utilizada nas classificações acima citadas.

Verificado que, nas linguagens de estrutura combinatória, a indexação efetua-se a nível de cada conceito elementar, sendo cada um destes expresso por um termo que chamamos de descritor, partiu-se para a análise sem a preocupação de localizar cada conceito neste ou naquele campo do Direito, nesta ou naquela hierarquia de classificação, mas objetivando aproximar a linguagem criada, tanto quanto possível, da linguagem natural. Assim o dispositivo constitucional que prescrevia “o prazo de convocação do Congresso para apreciação do relatório sobre o Estado de Sítio” foi indexado com os seguintes descritores:

**PRAZO, CONVOCAÇÃO, CONGRESSO NACIONAL, APRECIAÇÃO, RELATÓRIO, ESTADO DE SÍTIO.**

Um outro dispositivo constitucional que tratava da “igualdade de todos perante a lei, sem distinção de sexo, raça, trabalho, credo religioso e convicções políticas” teve como indexadores, além dos conceitos explícitos em seu texto, um descritor que lhe era doutrinariamente implícito, ou seja, **ISONOMIA CONSTITUCIONAL**.

Desta tarefa experimental desenvolvida pelos quatro grupos de trabalho resultou o seguinte:

- a) uma lista de aproximadamente 2.600 descritores que, apesar de um trabalho de refinamento posterior, tornou-se a base do atual Índice Controlado dos Descritores da Legislação Brasileira (IC);
- b) desenvolvimento de uma “Norma Experimental para a Extração de Descritores”;
- c) execução de um projeto-piloto de elaboração de um **Thesaurus** Experimental da Constituição com o qual pretendeu-se preparar os analistas para a tarefa da organização de um **macro-Thesaurus** da Legislação Brasileira.

## **6. A NORMA EXPERIMENTAL PARA A EXTRAÇÃO DE DESCRITORES**

Os problemas e respectivas soluções encontradas durante o trabalho experimental foram condensando-se e forneceram a estrutura para a formação de uma “Norma Experimental” para a extração de descritores. Esta norma foi qualificada de experimental tendo-se em vista que as regras nela contidas, resultado de um universo limitado de documentos, poderiam ser alteradas em razão do crescimento tanto quantitativo como qualitativo das normas analisadas. Assim, o elenco de regras, que visava criar um comportamento homogêneo e padronizado no trabalho analítico, estava sustentado na compreensão de que estas regras deveriam ser modificadas em função da realidade documental, evitando-se uma estratificação que impusesse uma análise de tal modo artificial e tão afastada da realidade por ela descrita, que perdesse o sentido.

A atual “Norma Experimental para a Extração de Descritores”, resultado da flexibilidade acima descrita, possui setores bem caracterizados. Primeiramente ela traz definições instrumentais, ou seja, traz os conceitos básicos a serem utilizados no método de análise. Designa como **documento** “o texto que contenha uma informação completa, constituindo por isso mesmo a menor unidade de indexação”, desta forma, para o banco de dados das normas jurídicas (NJUR), o documento pode ser representado por um sumário do texto — a ementa, ou por cada artigo, com seus parágrafos, itens, alíneas e incisos, porventura existentes, quando o texto é implantado integralmente. A “Norma Experimental” define descritor como sendo “a palavra ou conjunto de palavras de valor substantivo e sem homógrafos”, classificando-o em descritor simples — composto por uma só palavra, e descritor composto — composto por mais de uma palavra. Entretanto, uma característica importante do descritor é aquela de representar, por si só, um assunto ou conceito preciso e definido, o que lhe permite ser usado sozinho para indexar um documento ou a informação nele contida.

Um exemplo de descritor seria a expressão SOCIEDADE ANÔNIMA.

Outra figura de indexação, criada através da “Norma”, é a do especificador, “o termo... que, por si só, não representa um assunto preciso e definido, necessitando, portanto, de um descritor ao qual esteja associado e complete a idéia”, por isto mesmo o especificador não pode, sozinho, indexar um documento ou a informação nele contida.

Um exemplo de especificador seria o termo CRIAÇÃO.

Um exemplo de combinação analítica seria: CRIAÇÃO, SOCIEDADE ANÔNIMA.

A “Norma Experimental”, entre outras coisas, define o método a ser executado na extração e uso dos descritores e especificadores, menciona o caso dos conceitos implícitos, determina a obrigatoriedade do uso de descritores compostos em função da especificidade e do uso consagrado ou comum da expressão, limita os caracteres a serem utilizados pelo descritor, resolve o problema resultante da indexação de unidades federativas existentes e extintas, explica a forma de indexação de órgãos públicos e privados, determina o caso do uso de siglas, indica a forma do uso dos numerais e determina as características necessárias ao documento para que seja incluído em texto integral (**full-text**).

## 7. O THESAURUS DA CONSTITUIÇÃO

Para desenvolver o projeto-piloto de estruturação de um **Thesaurus** experimental, escolheu-se a Constituição da República Federativa do Brasil (Emenda Constitucional nº 1, de 1969), por um lado, por ser um texto com amplo limite de abrangência relativamente ao Direito e, por outro, pelo fato de tratar-se de um texto jurídico organizado internamente, o que facilitaria o trabalho dos analistas, então aprendizes da técnica de estruturar um **Thesaurus**.

A elaboração do **Thesaurus** da Constituição foi do tipo que podemos chamar a **posteriori**, ou seja, levantamento dos descritores à medida em que aparecem no texto indexado e na estruturação hierárquica após análise de todo o universo documental.

Este procedimento determinou que, ao mesmo tempo em que se criava um descritor, fosse criado outro tipo de relacionamento na estrutura tesáurica, ou seja, o do **não-use**, que nada mais é do que um sinônimo não autorizado de termo já levantado como descritor. Desta forma, na fase de indexação criaram-se dois tipos de termos relacionados, o descritor (subdividido em descritor e especificador) e o **não-use**, este nada mais sendo do que um ou mais sinônimos daquele descritor.

Outro tipo de relacionamento montado ainda na fase de indexação foi o de tipo **combine**, que, nos casos de termos compostos não autorizados, indicava quais os descritores ou especificadores a serem utilizados a fim de indicar o conceito.

Vejamos a seguir um exemplo de como se montava a listagem de termos.

Um artigo que mencionou ser "atribuição do Presidente da República a legislação financeira" poderia dar origem aos seguintes termos:

ATRIBUIÇÃO

PRESIDENTE DA REPÚBLICA

LEGISLAÇÃO FINANCEIRA

Entretanto, verificar-se-ia que ATRIBUIÇÃO era um sinônimo perfeito de COMPETÊNCIA, já escolhido como descritor em artigo anterior, que PRESIDENTE DA REPÚBLICA já havia sido levantado como descritor e que LEGISLAÇÃO FINANCEIRA poderia ser formada combinando os descritores já autorizados LEGISLAÇÃO e FINANÇAS PÚBLICAS.

Desta forma, ATRIBUIÇÃO passaria a compor a listagem como termo do tipo **não-use**, o qual remeteria o analista ao descritor COMPETÊNCIA, o qual teria uma informação indicativa de que ATRIBUIÇÃO era um de seus termos não autorizados. No que diz respeito a LEGISLAÇÃO FINANCEIRA, passaria a compor a listagem como termo do tipo **não-use** e remeteria o analista para a combinação dos descritores LEGISLAÇÃO, FINANÇAS PÚBLICAS, os quais, por sua vez, teriam uma informação sobre a possibilidade de combinação entre os mesmos para formar o termo não autorizado.

Na listagem alfabética dos termos, esta situação se apresentaria da seguinte forma:

N/USE ATRIBUIÇÃO

USE COMPETÊNCIA

USE COMPETÊNCIA



N/USE ATRIBUIÇÃO  
 USE FINANÇAS PÚBLICAS  
 N/USE LEGISLAÇÃO FINANCEIRA  
 COMBINE LEGISLAÇÃO  
 USE LEGISLAÇÃO  
 N/USE LEGISLAÇÃO FINANCEIRA  
 COMBINE FINANÇAS PÚBLICAS  
 N/USE LEGISLAÇÃO FINANCEIRA  
 USE LEGISLAÇÃO  
 FINANÇAS PÚBLICAS  
 USE PRESIDENTE DA REPÚBLICA

Na fase inicial o trabalho foi realizado através de fichas catalográficas comuns e, numa segunda fase, através de listagens alfabéticas e do tipo "KWIC" emitidas pelo computador.

Terminada a fase de levantamento de descritores, passou-se para a tarefa de hierarquização, ocasião em que procurou-se dar a cada descritor levantado três níveis de relacionamento hierárquico dentro do universo documental de onde o mesmo se originava. Havia um nível de relacionamento superior, um inferior e um de correlação. Este relacionamento tridimensional, somado à nota explicativa e aos outros tipos já mencionados, dava a seguinte apresentação aos documentos do **Thesaurus** nas listagens produzidas pelo computador:

PARAGRAFO	DENOMINAÇÃO	CONTEUDO
001	USE	DESCRITOR AUTORIZADO
100	NOTA (Scope)	SENTIDO RESTRITIVO DO DESCRITOR
200	NÃO USE	SINÓNIMOS NÃO AUTORIZADOS
300	SUPERIOR	DESCRITOR DE MAIOR ABRANGÊNCIA SEMÂNTICA
500	INFERIOR	DESCRITOR DE MENOR ABRANGÊNCIA SEMÂNTICA
600	VER TAMBÉM	DESCRITOR AUTORIZADO DE SENTIDO CORRELATO
700	COMBINE	DESCRITOR AUTORIZADO QUE COMBINADO AO "USE" FORMA UM CONCEITO NÃO AUTORIZADO

As informações do “**Thesaurus da Constituição**” podiam ser recuperadas utilizando-se o programa STAIRS, estando o usuário capacitado a atingi-las sem restrições ou, caso quisesse, limitando a pesquisa a um tipo de parágrafo determinado. Isto significava que o usuário, mesmo questionando sobre a existência de um termo não autorizado, encontraria o termo autorizado, seu relacionamento vertical e horizontal e suas possíveis combinações. Poderia, então, subir ou descer na escala hierárquica, buscando uma maior generalidade ou especificidade, além de poder alargar sua pesquisa através do nível VIDE TAMBÉM.

Apesar de alguns aspectos falhos, verificados posteriormente através da experiência prática, o **Thesaurus** mostrou-se altamente positivo no sentido de que demonstrou plenamente a possibilidade de que sistemas de linguagem estruturada podem substituir, no campo do Direito, as classificações enciclopédicas, e de que propiciou treinamento e preparação especializada para os analistas que trabalharam no projeto.

## 8. A ATUAL SITUAÇÃO DO TRABALHO

Após quase quatro anos de existência, os resultados das atividades da Subsecretaria de Análise se têm apresentado extremamente positivos, sob vários aspectos, entre os quais destacamos os seguintes:

- a) criação de uma metodologia documental especificamente apropriada ao acervo legislativo brasileiro;
- b) inter-relacionamento frutífero e construtivo entre indivíduos treinados em processamento de dados, ciências da informação e ciências jurídicas;
- c) iniciação de um vasto sistema de implantação e recuperação de dados normativos, no qual se incluem, além das normas de nível federal, aquelas relativas aos órgãos que, mantendo convênio com o PRODASEN, utilizam-se da metodologia desenvolvida pela Subsecretaria de Análise;
- d) projeção de planos que abarcariam uma cobertura informativa total à clientela do sistema.

A seguir faremos uma análise mais detalhada de cada um dos aspectos acima mencionados.

### 8.1. A METODOLOGIA

Desenvolvida e refinada a metodologia da Subsecretaria de Análise no campo do tratamento dos dados jurídicos legislativos, a situação mostra-se da seguinte forma:

Estão implantados no sistema 202.448 documentos, sendo que a Subsecretaria é responsável diretamente por 166.221 dentre os mesmos, os quais fazem parte dos bancos de dados abaixo discriminados:

NJUR (Normas Jurídicas)	119.965
DISC (Discursos)	4.007

JURI (Jurisprudência) 33.765

THES (Thesaurus) 8.484

O crescimento do número de documentos incluídos nestes bancos de dados no período de outubro de 1975 a outubro de 1976 foi o seguinte:

NJUR de 109.408 para 119.965

DISC de 2.680 para 4.007

JURI de 24.279 para 33.765

THES (Thesaurus) 8.484

TOTAL 143.739 166.221

Constata-se desta forma que, percentualmente, o montante de documentos analisados e implantados nos bancos de dados aumentou em aproximadamente 15% durante um ano, prevendo-se aumentos ainda mais acentuados em próximos períodos, principalmente tendo-se em conta que muitos órgãos em convênio, somente agora, enfatizaram seus contatos com a Subsecretaria visando incrementar sua participação na implantação de dados.

O sistema de análise utilizado atualmente não se diferencia grandemente daquele desenvolvido na fase inicial, ou seja, levantamento de descritores com sentido substantivo e estruturados da forma mais próxima possível da linguagem natural. Estes descritores, na indexação, são separados por vírgulas e formam sentenças, as quais, ao mesmo tempo em que indexam o documento, dão ao pesquisador uma idéia imediata do conteúdo do mesmo, podendo aquele verificar a relevância, ou não, do documento para a pesquisa que está sendo efetuada.

É importante que se tenha em mente a existência de uma tradicional deficiência na sumarização do conteúdo das normas através das ementas. Estas, em sua maioria, não definem nem indicam precisamente o de que trata o texto do documento. Assim, a indexação, através de sentenças compostas por descritores, permite a recuperação dos documentos por meio de um descritor componente e a imediata verificação de sua relevância para o aspecto específico pretendido pelo pesquisador.

Esta situação pode ser demonstrada pelo exemplo seguinte.

Um pesquisador interessado na situação do pessoal que trabalha para o Serviço Nacional de Informações sob o regime das Leis Trabalhistas questionaria o sistema sobre documentos em que os descritores PESSOAL e SNI constassem.

Em resposta, o sistema apresentaria uma série de documentos nos quais ocorreriam os termos PESSOAL e SNI, como nos dois casos seguintes:

Documento 1.

**Ementa**

**DISPÕE SOBRE A TABELA PROVISÓRIA DO PESSOAL TEMPORÁRIO DO SERVIÇO NACIONAL DE INFORMAÇÕES.**

**Indexação**

**REGULAMENTAÇÃO, PESSOAL, LEGISLAÇÃO TRABALHISTA, (SNI).**

Documento 2.

**Ementa**

**ALTERA A REDAÇÃO DO PARÁGRAFO PRIMEIRO DO ARTIGO SEXTO DA LEI 4.341, DE 13 DE JUNHO DE 1964, E DÁ OUTRAS PROVIDÊNCIAS.**

**Indexação**

**ALTERAÇÃO, LEGISLAÇÃO, CRIAÇÃO, (SNI).**

**REQUISIÇÃO, PESSOAL, MINISTÉRIO, EXECUTIVO, (SNI).**

**ADMISSÃO, PESSOAL, (SNI).**

Em vista das indexações providas para os documentos acima transcritos, o pesquisador pode decidir-se imediatamente pelo primeiro como sendo aquele que responde à sua necessidade de informação.

Com este exemplo pretendemos indicar que o sistema de análise e indexação utilizado pela Subsecretaria de Análise, ao mesmo tempo em que levanta os descritores significativos de um documento, transcreve este mesmo documento utilizando a linguagem controlada, mas o faz de tal forma que permite ao pesquisador, ao enfrentar um documento, decidir sobre a relevância, ou não, do mesmo para a pesquisa que esteja realizando.

Assim, nosso sistema está numa posição equidistante entre uma classificação convencional por palavras-chaves, e uma transcrição do texto integral **full-text**. Com isto conseguimos manter os aspectos mais positivos de ambos, pois ao utilizarmos descritores conseguimos uma padronização léxica e, ao usarmos uma ordenação semelhante à da linguagem natural, obtemos, pelo menos em parte, os benefícios da construção linear e as características da lógica complexa desta linguagem.

A experiência tem demonstrado a aplicabilidade da utilização de uma linguagem documental controlada — presentemente em vias de ser estruturada hierarquicamente — combinada com a ampla possibilidade combinatória da linguagem natural.

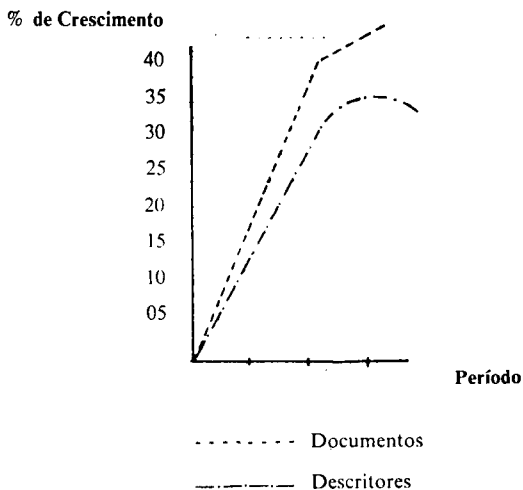
O controle da terminologia atualmente levado a efeito reduz substancialmente os fatores negativos apresentados pela linguagem natural, quais sejam, as características semânticas, as polissemias, as sinonímias, os modos e a indefinição de certas relações complexas e idiomáticas.

Por outro lado, a estruturação e a hierarquização deste léxico controlado, por meio de um **Thesaurus** ou outro instrumento de referência, organizado de modo conceitual e que apresente os termos agrupados por afinidades semânticas e por indicações de relações, ampliará enormemente a capacidade de recuperação.

## 8.2. O ÍNDICE CONTROLADO

Na lista dos descritores utilizados (ÍNDICE CONTROLADO) os termos estão controlados basicamente a nível de NOTA (**scope-note**), NÃO-USE (sinônimo) e VER TAMBÉM (correlação). Estão projetados a hierarquização e refinamento do léxico, mas entendemos, como o Prof. F. W. Lancaster, que aquela somente deverá tomar início a partir do momento em que o crescimento do universo documental e o crescimento do número de descritores coloquem-se em acentuada desproporção, ou seja, que para um aumento no número de documentos analisados, não haja um aumento proporcional de descritores.

Graficamente poderemos representar esta situação na forma seguinte:



Nota-se que há um quase paralelismo no crescimento percentual dos documentos e dos descritores por um determinado período de tempo. Entretanto, chega-se a um momento em que se inicia uma sensível diferença entre as linhas, tendendo aquela dos descritores a refluir a níveis mais baixos, até chegar a um ponto em que os descritores existentes abrangem praticamente todos os conceitos contidos nos documentos que compõem ou venham a compor o universo documental.

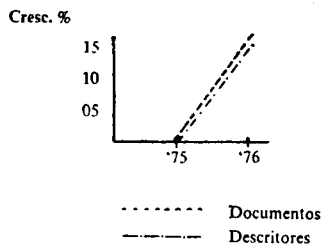
Verificada esta tendência, deverá então ser iniciada a hierarquização do léxico, já que se evidencia estar o mesmo completo. Claro está que esta é uma situação hipotética, pois, devido à própria criatividade humana, sempre haverá alguma nova idéia ou comportamento passível de ser conceituado ou conceito a ser designado por nova terminologia.

O crescimento dos documentos e dos descritores na área de responsabilidade da Subsecretaria no período de 1975 a 1976 poderá ser representado no quadro e gráfico seguintes:

QUADRO DE CRESCIMENTO

Ano	1975	1976	Cresc. %
DOCUMENTOS	136.739	157.737	15
DESCRITORES	7.372	8.484	15

GRÁFICO DE CRESCIMENTO



Verifica-se a existência de um paralelismo no crescimento dos documentos e nos descritores e pode prever-se que esta situação deverá perdurar por mais algum tempo, visto que novos documentos serão analisados e implantados no sistema. Estes documentos, provenientes de fontes normativas diversas e especializadas, determinarão, sem qualquer dúvida, a manutenção do crescimento proporcional entre documentos e descritores. Conclui-se, desta forma, que a hierarquização da terminologia, apesar de ser um projeto importante, não se mostra ainda como de execução imediata.

### 8.3. A SUBSECRETARIA E O PRODASEN

A criação e a estruturação paralelas da Subsecretaria de Análise e do PRODASEN demonstram claramente o caráter complementar dos dois órgãos. Por um lado, a Subsecretaria fornecendo os dados e desenvolvendo um sistema de análise e recuperação e, de outro, o PRODASEN, oferecendo a tecnologia, em termos de processamento daqueles dados.

#### 8.3.1. RECURSOS COMPUTACIONAIS DISPONÍVEIS

O principal instrumento oferecido pelo PRODASEN como recurso a nível de **soft-ware** é o Programa-Produto STAIRS, que consiste basicamente em três grupos de programas:

- os utilitários que criam os bancos de dados;
- os utilitários que executam as tarefas de manutenção dos bancos de dados; e
- o sistema "AQUARIUS".

Este último sistema consiste de uma série de funções operacionais, através das quais o pesquisador e o analista podem trabalhar com os documentos existentes nos bancos de dados.

Em ordem de importância para a Subsecretaria de Análise, destacamos algumas destas funções:

- SEARCH:** permite a pesquisa dos documentos por intermédio de termos e possibilita pesquisa a nível de parágrafo ou campo formatado de uso livre.
- BROWSE:** permite a exibição do documento na tela do terminal de vídeo.
- SELECT:** permite que sejam efetuadas pesquisas a nível de campo formatado não livre, tais como o tipo da norma jurídica, a data de publicação da mesma, o número da norma, o órgão legislativo ou normativo de onde se originou a norma, etc.
- SORT:** permite a classificação e ordenação dos documentos cronológica e alfabeticamente, tanto em ordem ascendente como descendente.
- PRINT:** permite a impressão do documento no terminal de impressão ou no Centro de Processamento de Dados.
- SAVE:** permite o arquivamento de uma pesquisa ou de uma estratégia (conjunto de "queries") para utilização posterior.

Sob o ponto de vista de **hard-ware** o PRODASEN dispõe e fornece, direta ou indiretamente, à Subsecretaria o seguinte material:

Sistema IBM/370 — 158 com as seguintes características:

Memória Principal 1024 K

6 unidades de discos magnéticos	— Mod	33-30
5 unidades de fita magnética	— Mod	34-20
6 leitoras e perfuradoras de cartão	— Mod	129
1 leitora e perfuradora de cartão	— Mod	25-40
8 perfuradoras Diskett	— Mod	37-42
Leitora Diskett	— Mod	35-4
Terminais de vídeo	— Mods	3277, 3275, TCV 275
Terminais de impressão	— Mods	3284, 3286

A Subsecretaria dispõe para sua utilização direta de 5 terminais de vídeo Mod. 3277 e de dois terminais de impressão Mod. 3284, que estão sendo utilizados para a implantação e correção de dados diretamente através dos terminais de vídeo, criando-se assim a possibilidade de abandonar gradativamente o atual sistema de implantação e manutenção dos dados através de formulários.

Com este instrumental a Subsecretaria capacita-se, não só a exercer suas atividades de manutenção dos bancos de dados, mas também a fornecer informações à sua clientela.

### 8.3.2. INTER-RELAÇÃO DE RECURSOS HUMANOS

O relacionamento entre o PRODASEN e a Subsecretaria não está resumido na utilização de recursos computacionais, mas numa integração de esforços que objetiva solucionar problemas, desenvolver projetos e aumentar a confiabilidade e eficácia do sistema.

O trabalho integrado dos dois órgãos tem sido indispensável em todos os passos do desenvolvimento da experiência. Desde a fase inicial, quando a discussão sobre as características da documentação legislativa determinou a criação de formulários especiais, até a situação atual, em que estão sendo introduzidos os programas especiais de controle crítico sobre características gráficas da terminologia utilizada nos documentos.

Esta inter-relação é demonstrada pelas constantes reuniões entre os dois órgãos, e o treinamento do pessoal de ambas as entidades, respectivamente nas áreas de Processamento de Dados, da Ciência do Direito e da Informação, além da contratação, pelo PRODASEN, de pessoas especializadas para trabalhar diretamente na Subsecretaria de Análise, após testes e cursos aplicados por esta às mesmas.

### 8.4. O DESENVOLVIMENTO DE UM SISTEMA DE INFORMAÇÃO LEGISLATIVA

Vários órgãos da administração pública, tanto federal como estadual, vêm assinando, junto ao PRODASEN, documentos de convênio pelos quais habilitam-se a utilizar os serviços de informação constantes dos vários bancos de dados disponíveis. Uma boa parte dessas instituições está contratualmente habilitada a implantar dados no campo da legislação a elas diretamente referentes, sendo em relação a elas que a Subsecretaria de Análise tem desenvolvido um trabalho especial.

Dentre as competências específicas da Subsecretaria está aquela de analisar e implantar a documentação legislativa federal de hierarquia superior, originada tanto do Poder Executivo como do Poder Legislativo, as normas do Governo do Distrito Federal e aquelas normas de nível hierárquico inferior que abrangem interesses de mais de uma unidade da Administração Pública.

O quadro a seguir indica quais as normas cuja análise e implantação são de competência exclusiva da Subsecretaria.

<b>Origem</b>	<b>Denominação</b>	<b>Sigla</b>
GDF	Decreto Executivo do Governo do DF	DFE
	Decreto Normativo do Governo do DF	DFN
EXECUTIVO	Ato Complementar	ACP
	Ato do Comando Supremo da Revolução	ACR
	Ato Institucional	AIT



	Decreto do Conselho de Ministros	DCM
	Decreto Executivo (numerado)	DEC
	Decreto-Lei	DEL
	Decreto Reservado	DER
	Portaria Interministerial	PRI
LEGISLATIVO	Ato Adicional	AAD
	Decreto Legislativo	DLG
	Emenda Constitucional	EMC
	Lei Complementar de Ato Adicional	LCA
	Lei Complementar	LCP
	Lei Constitucional	LCT
	Lei Delegada	LDL
	Lei Ordinária	LEI
	Resolução da Câmara dos Deputados	RCD
	Resolução do Congresso Nacional	RCN
	Resolução do Senado Federal	RSF

Os órgãos em convênio alimentadores são atualmente em número de 13:

Ministério da Aeronáutica  
 Ministério da Educação  
 Ministério do Exército  
 Ministério da Fazenda  
 Ministério do Interior  
 Ministério da Justiça  
 Ministério da Marinha  
 Ministério das Minas e Energia  
 Ministério do Trabalho  
 Secretaria de Planejamento da Presidência da República  
 Superior Tribunal Militar (\*)  
 Tribunal de Contas do Distrito Federal  
 Tribunal Superior do Trabalho (\*)

Estes órgãos, com exceção feita aos dois Tribunais Superiores, estão interessados na implantação de uma variada tipologia de normas, que vão desde importantes portarias ou circulares ministeriais, até específicas instruções de serviço.

Entretanto, a Subsecretaria, em seus contatos prévios e no treinamento em indexação aplicado a estes órgãos, vem esclarecendo que, para evitar um desvirtuamento do sistema, a norma particular ou de hierarquia inferior deverá integrá-lo somente quando possuir uma das seguintes características:

a) complementar, regulamentar ou alterar norma de nível hierárquico superior;

(\*) Implantando jurisprudência sem análise.

- b) responder a uma tendência do interesse parlamentar;
- c) atingir uma área razoável de interesses da comunidade dos usuários do sistema.

Com isto pretende-se ampliar o universo de dados de forma controlada, tendo em vista que uma posição crítica face aos documentos a serem implantados pelos órgãos em convênio é necessária, pois o que se está visando é transformar o sistema num instrumento a serviço do processo legislativo, proporcionando sua maior racionalização e eficácia, e, em nível de menor prioridade, mas assim mesmo de grande importância, atender às necessidades documentais dos órgãos em convênio.

### 8.5. O TREINAMENTO DOS ÓRGÃOS EM CONVÊNIO

Aos órgãos em convênio são fornecidos dois tipos de treinamento. Um oferecido pelo PRODASEN e no qual as entidades são treinadas a utilizar o instrumental e os programas disponíveis *on line*. O outro treinamento é oferecido pela Subsecretaria de Análise e tem por objetivo a padronização da forma documental e da metodologia de indexação, o que evita a criação de índices paralelos e permite uma estratégia de pesquisa comum, independentemente do banco de dados visado.

Neste treinamento oferecido pela Subsecretaria, através de sua Seção de Pesquisa, exige-se dos treinandos a formação em nível superior, principalmente nas áreas de Direito, Biblioteconomia e Ciências Sociais.

O curso desenvolve-se em duas fases. Na primeira são dadas aulas teóricas sobre o sistema SICON (Sistema de Informação do Congresso), o banco de dados NJUR (Normas Jurídicas), as normas de indexação e extração de descritores, o preenchimento dos formulários e o relacionamento das normas. Na segunda fase, os treinandos são acompanhados em seus trabalhos práticos por um período que se estende até o instante em que, no entendimento da Subsecretaria, o órgão possa desenvolver seu trabalho autonomamente.

Durante esta segunda fase de treinamento, a Subsecretaria define junto ao órgão o programa de trabalho e a documentação a ser processada pelo mesmo. Assim, são feitos levantamentos relativamente aos tipos de documentos de interesse do órgão, à existência de alguma forma de classificação ou indexação e à possibilidade de compatibilização entre o léxico utilizado pelo órgão em convênio e aquele desenvolvido pela Subsecretaria. Este é um aspecto importante, pois através do mesmo assegura-se um enriquecimento contínuo da terminologia em vista dos novos documentos acrescentados pelos órgãos ao sistema, e, por outro lado, graças ao treinamento em indexação, é exercido um controle sobre a linguagem utilizada.

Em vista do objetivo da Subsecretaria de desenvolver um instrumento de tipo *Thesaurus*, será grande a responsabilidade dos órgãos em convênio, já que cada um deles deverá, sob o controle e a coordenação da Subsecretaria, estruturar os descritores que se referem à sua área específica de atuação, ou seja, desenvolver setores de um todo que poderá ser chamado de *Thesaurus* da Legislação Brasileira.