

LexML – visão unificada da informação legislativa e jurídica do Brasil

João Alberto de Oliveira Lima¹

Edilenice Passos²

O Projeto LexML Brasil se propõe a organizar a informação legislativa e jurídica disponibilizada na internet pelos vários órgãos do governo, considerando os três Poderes da República (Legislativo, Executivo e Judiciário), a Advocacia Geral da União, o Ministério Público e as diversas esferas administrativas (federal, estadual, distrital e municipal). Para atingir esse fim, implementou-se um portal de acesso unificado às informações, criou-se o identificador unívoco e persistente de documentos legislativos e jurídicos, e ainda instituiu-se o Comitê Gestor de Informação com o objetivo de aprimorar o tratamento e a recuperação da informação legislativa e jurídica contidas no Portal. Todos esses aspectos serão apresentados neste estudo de caso.

Palavras-chave: LexML (Brasil). Informação legislativa (Brasil) . Informação jurídica (Brasil).

The LexML Brasil Project proposes to organize the legislative and legal information made available on the internet by the various government agencies, considering the three Branches of the Republic (Legislative, Executive and Judiciary), the Federal Attorney General's Office, the Public Prosecutor's Office and the various administrative spheres (federal, state, district and municipal). In order to achieve this goal, a unified access portal was implemented, the univocal and persistent identifier of legislative and legal documents was created, and the Information Management Committee was created with the objective of improving treatment and recovery of the legislative and legal information contained in the Portal. All these aspects will be presented in this case study.

Keywords: LexML (Brazil). Legislative information (Brazil). Legal information (Brazil).

¹ Servidor do Senado Federal. joaolima@senado.gov.br.

² Servidora do Senado Federal. edilenic@senado.gov.br.

El Proyecto LexML Brasil propone organizar la información legislativa y legal disponible en Internet por las distintas agencias gubernamentales, considerando las tres Ramas de la República (Legislativo, Ejecutivo y Judicial), la Oficina del Fiscal General Federal, la Oficina del Fiscal General y el Diversos ámbitos administrativos (federales, estatales, distritales y municipales). Para lograr este objetivo, se implementó un portal de acceso unificado, se creó el identificador unívoco y persistente de los documentos legislativos y legales, y se creó el Comité de Gestión de la Información con el objetivo de mejorar el tratamiento y la recuperación de la información legislativa y legal contenida en el portal. Todos estos aspectos serán presentados en este estudio de caso.

Palabra-claves: LexML (Brazil). Información legislativa (Brazil). Información jurídica (Brazil).

1 INTRODUÇÃO

A Constituição Federal de 1988, no art. 5^o, inciso XIV, define o acesso à informação como um direito fundamental do cidadão.

A preocupação em franquear o acesso à informação jurídica é anterior à Constituição Cidadã. Em meados da década de 70, do século passado, o Senado Federal, pioneiro na iniciativa, criou bases de dados que proporcionaram o controle informatizado das informações jurídica e legislativa brasileiras. Desde então, foram desenvolvidas, de forma independente, inúmeras bases de dados pelos vários órgãos governamentais. Com a mudança para o ambiente da internet, tais bases de dados proliferaram e facilitaram o acesso à informação. Entretanto, como foram desenvolvidas de forma autônoma, sentiu-se a necessidade de criar uma infraestrutura que permitisse organizar e integrar a grande quantidade de informações legislativas e jurídicas existente facilitando o seu acesso e que contribuísse para a realização do direito fundamental supracitado. Nesse sentido, foi criado o Projeto LexML Brasil, iniciativa conjunta de diversos órgãos participantes da Comunidade TI Controle em busca do estabelecimento de padrões abertos, integração de processos de trabalho e compartilhamento de dados de interesse comum, que permitam a identificação e a estruturação de informações legislativas e jurídicas.

Um primeiro resultado do Projeto LexML é o Portal da Rede de Informação Legislativa e Jurídica (<http://www.lexml.gov.br>), inaugurado em 30 de junho de 2009, com 1,28 milhão de documentos. Trata-se de um ponto de acesso unificado às diversas fontes de informação legislativa e jurídica do governo, criando assim a infraestrutura necessária para a difusão em larga escala do identificador unívoco e persistente. Além do portal, estão sendo desenvolvidas outras ferramentas, tais como, editor, compilador e consolidador com o objetivo de auxiliar a gestão da informação documental.

As recomendações seguidas pelo Projeto LexML foram reunidas das melhores práticas de iniciativas para identificação e estruturação de documentos legislativos e jurídicos, adaptando as experiências bem sucedidas do exterior às necessidades do nosso ordenamento jurídico e da técnica legislativa. Devido à influência do Direito Romano na organização do nosso ordenamento jurídico, as iniciativas de países que adotam o sistema *civil law* são mais aplicáveis à nossa realidade do que as de países que adotam outros sistemas legais. Duas grandes influências vieram da Itália, um dos países com maior evolução na aplicação de novas tecnologias para a estruturação e organização de normas jurídicas no formato XML. O primeiro é o Projeto *Norme in Rete* (ARCHI *et al.*, 2000), coordenado pelo CNIPA (Centro Nazionale per Informatica nella Pubblica Amministrazione), e realizado com o suporte do ITTIG (Istituto di Teoria e Tecniche dell'Informazione Giuridica) e da Universidade de Bolonha, foi implantado desde 2002 e tem tido resultados expressivos. O segundo é o Akoma Ntoso (VITALI, 2010) promovido pelas Nações Unidas e realizado com o suporte de pesquisadores da Universidade de Bolonha com o objetivo de capacitar os parlamentos da África com novas tecnologias de informação e comunicação.

2 PORTAL LexML – REDE DE INFORMAÇÃO LEGISLATIVA E JURÍDICA

A página inicial do Portal LexML é simples e apresenta um campo para digitação do argumento de pesquisa. É possível, nessa tela, restringir a pesquisa à legislação, à jurisprudência ou às proposições legislativas. A Fig. 1 apresenta a página inicial onde foi digitado o termo “Lei Maria da Penha” para pesquisa. A Fig. 2 apresenta o resultado da pesquisa formulada que encontrou seis documentos.



Figura 1: Página Inicial do Portal LexML.



Figura 2: Resultado de pesquisa do termo “Lei Maria da Penha”.

O resultado da pesquisa vem acompanhado de filtros que permitem o seu refinamento de acordo com as suas necessidades de informação. Cinco filtros estão disponíveis: localidade, autoridade emitente, tipo de documento, data e siglas.

LEXML [Página Anterior](#) | [Página Inicial](#) | [Pesquisa Avançada](#)

Localidade Brasil
Autoridade Federal
Título Lei nº 11.340, de 7 de Agosto de 2006
Data 07/08/2006
Apelido LEI MARIA DA PENHA
Ementa Cria mecanismos para coibir a violência doméstica e familiar contra a mulher, nos termos do § 8º do art. 226 da Constituição Federal, da Convenção sobre a Eliminação de Todas as Formas de Discriminação contra as Mulheres e da Convenção Interamericana para Prevenir, Punir e Erradicar a Violência contra a Mulher; dispõe sobre a criação dos Juizados de Violência Doméstica e Familiar contra a Mulher; altera o Código de Processo Penal, o Código Penal e a Lei de Execução Penal; e dá outras providências.
Nome Uniforme urn:lex:br:federal:lei:2006-08-07:11340
Mais detalhes [Câmara dos Deputados](#) (text/html)
Mais detalhes [Senado Federal](#) (text/html)

Publicação Oficial

Publicação Original 2006-08-08 Diário Oficial da União. Seção 1. 08/08/2006. p. 1
[Imprensa Nacional](#) (application/pdf)

Outras Publicações

Texto Atualizado Multivigente [Presidência da República](#) (text/html)
Publicação Original 2006-08-08 [Câmara dos Deputados](#) (text/html) [Linker](#)
Publicação Original 2006-08-08 [Senado Federal](#) (text/html) [Linker](#)

Figura 3: Detalhe do primeiro documento recuperado na pesquisa.

3 IDENTIFICADOR UNIFORME – URN

Na organização de um acervo é necessário definir, entre outras coisas, um critério de identidade para as classes de interesse, ou seja, para cada tipo de item devemos escolher as características que identificam univocamente cada instância. Por exemplo, no caso de normas jurídicas, é possível considerar os seguintes tipos de itens: a) a norma de uma forma geral, considerando toda a sua evolução no tempo; b) uma versão do texto da norma para uma determinada data; c) um dispositivo de uma versão específica da norma.

Para cada um destes três níveis de granularidade listados, é possível definir um grupo de elementos (datas, tipos, números seqüenciais etc.) para a composição de um identificador unívoco. Raciocínio análogo aplica-se às proposições legislativas e aos julgados. Uma das principais vantagens de se definir claramente identificadores unívocos é a possibilidade de se estabelecer remissões textuais precisas.

A remissão textual é muito utilizada nos documentos legislativos e jurídicos. Por exemplo, uma norma jurídica pode fazer referências a outras normas, seja para realizar alterações, revogar um dispositivo ou regulamentar uma matéria; um acórdão pode referenciar normas jurídicas bem como outros acórdãos e súmulas.

As remissões textuais podem ser classificadas de diversas formas:

- quanto à localização do alvo de uma remissão, podemos identificar dois tipos: as remissões internas e as externas. As remissões internas têm como alvo o próprio documento. As externas referenciam outros documentos.
- quanto à granularidade de uma remissão externa, podemos identificar dois tipos: a remissão ao documento no todo e a remissão a uma parte do documento.
- quanto à dimensão temporal de uma remissão externa, podemos identificar dois tipos: a remissão a uma versão específica de uma norma e a remissão à norma de uma forma genérica, sem especificar uma versão específica.

Para representar corretamente as remissões é necessário criar um identificador que qualifique unicamente cada possível alvo de uma remissão textual. Na impossibilidade de criar um sistema único de numeração para todo o acervo de normas, julgados e proposições legislativas do Brasil, fez-se necessária a criação de um identificador que seja compatível com os diversos sistemas de numeração existentes.

O Projeto LexML define um identificador unívoco e persistente para os documentos legislativos e jurídicos. A persistência possibilita que os alvos sejam encontrados, mesmo que ocorra mudança de endereços da Internet (URLs – Universal Resource Locator), evitando assim o conhecido “Erro 404” (Recurso não encontrado) do protocolo HTTP. Para que a persistência se realize é necessário, apenas, que os provedores de dados atualizem a nova localização do recurso quando da mudança de localização.

O identificador LexML permite a integração entre as fontes de informações legislativas e jurídicas pois define um referencial único para esse tipo de documento.

A lista abaixo relaciona endereços da internet (URLs) onde é possível consultar o texto da Lei nº 8.666, de 1993, nos sítios do Senado Federal e da Presidência da República, respectivamente. É possível notar que cada identificador possui regras de

formação diferentes e está diretamente relacionado à solução tecnológica adotada pela instituição em um determinado momento.

<http://www6.senado.gov.br/legislacao/ListaPublicacoes.action?id=102446>

<http://www.planalto.gov.br/ccivil/Leis/L8666cons.htm>

Caso ocorra mudança na tecnologia utilizada ou alteração na convenção de nomes por parte dos órgãos gestores desses recursos, todos os links dos documentos que referenciavam esses identificadores irão apresentar Erro 404 (recurso não encontrado) ao serem ativados.

O identificador LexML utiliza a notação URN que permite a criação de identificadores persistentes, pois estes não estão atrelados a detalhes de tecnologia ou à uma convenção de nomes definida de forma unilateral. As URNs seguem um esquema previamente acordado entre os participantes da rede. Abaixo, é apresentada a URN da Lei nº 8.666, de 1993.

<urn:lex:br:federal:lei:1993-06-21;8666>

3.1 Estrutura Básica de URNs e Exemplos

O nome uniforme do Projeto LexML Brasil é prefixado por “urn:lex:” e é seguido, basicamente, por mais quatro elementos, delimitados por “:” (símbolo de dois pontos), conforme lista abaixo:

- localidade – contém a jurisdição da autoridade emitente;
- autoridade – indica a autoridade emitente do documento;
- tipo do documento – indica a espécie da tipologia documental;
- descritor – formado pela data representativa e um descritor alfanumérico.

Os três primeiros elementos são definidos em vocabulários controlados que estão sendo construídos à medida em que as instituições aderem ao Projeto LexML.

A lista abaixo relaciona alguns exemplos de nomes uniformes de documentos jurídicos e legislativos disponíveis no Portal LexML.

- Lei nº 11.705, de 19 de Junho de 2008.
 - urn:lex:br:federal:lei:2008-06-19;11705
- Súmula Vinculante nº 11
 - urn:lex:br:supremo.tribunal.federal:sumula.vinculante:2008-08-13;11
- Lei nº 13.745 de 10 de fevereiro de 2006 (Lei Municipal – São Carlos – SP)
 - urn:lex:br;sao.paulo;sao.carlos:municipal:lei:2006-02-10;13745
- Acórdão TCU 395/1997
 - urn:lex:br:tribunal.contas.uniao;camara.1:acordao:1997-09-02;395

3.2 Tipos de Nomes Uniformes

Considerando os cenários apresentados de uso da URN, foram definidos dois tipos:

- URN Canônica – normalizada segundo o vocabulário e regras mais restritas, existindo apenas uma para cada entidade identificável (Documento, Documento Complexo e Fragmento). Ela é obrigatoriamente informada pelo provedor de informações e relacionada a um documento específico.
- URN de Referência – utilizada pelos usuários ou sistemas de informação nas referências a documentos, podendo utilizar a forma compacta, omitir elementos que assumirão valores default ou ainda utilizar nomes alternativos.

3.3 Princípios do Nome Uniforme

O nome uniforme deve ser unívoco, deve identificar uma e apenas uma entidade, e é construído, tanto quanto possível, alinhado com os princípios apresentados a seguir:

Princípios Gerais da URN:

- auto-explicativo para os usuários;
- dedutível por meio de regras simples e claras;
- alinhado constantemente aos demais padrões do projeto.

Princípios da URN Canônica:

- unívoco para o espaço das URNs canônicas;
- registrável com o mínimo de informação (o suficiente para ser unívoca)
- representativo dos aspectos formais do documento;
- em conformidade, à data de sua emissão, com a estrutura/organização da autoridade emitente e com a tipologia do documento.

Princípios da URN de Referência

- representativo dos aspectos formais e/ou substanciais do documento;
- mapeável em URNs canônicas cabíveis;
- compatível com a prática em uso para criar referências;
- reduzido ao essencial, para simplificar os links com outros documentos;
- capaz de ser gerado de forma automática por analisadores de remissões textuais.

4 COMITÊ GESTOR DE INFORMAÇÃO DO PORTAL LexML

Na primeira fase, os profissionais da área de informática criaram a infra-estrutura para o portal LexML. Na segunda fase, os bibliotecários e analistas de informação precisam resolver, por exemplo, questões ligadas ao conteúdo e à compatibilização dos vários tesouros e vocabulários controlados dos quais fazem uso.

Foi, então, criado o Comitê Gestor de Informação do Portal LexML, com o objetivo de aprimorar o tratamento e a recuperação da informação legislativa e jurídica contidas no Portal, aplicando as teorias e técnicas da Ciência da Informação. É formado basicamente por bibliotecários, mas aberto a todos os profissionais que queiram contribuir com sua experiência e conhecimento.

Para a realização de suas atribuições foram estabelecidas sete metas e, para a realização das metas foram criados grupos de trabalho, a saber:

GT para definir as Categorias e Critérios para Formação do Acervo do Portal LexML – tendo como primeiro objetivo de estudar as categorias existentes no Portal (atualmente são cinco: Localidade; Autoridade; Tipo de Documento; Data e Sigla) de

modo que favoreçam a usabilidade e que tornem o Portal LexML mais amigável. Seu segundo objetivo é definir critérios para a formação do acervo LexML, estabelecendo de forma transparente e inequívoca os tipos de documentos que serão incluídos ou não no acervo do Portal.

GT para o Tratamento de Linguagem – deverá atingir quatro objetivos: 1) complementar/revisar o dicionário da língua portuguesa comum (existem algumas falhas, como por exemplo, falta a entrada idoso/idosos/idosas/idosas); 2) criar o dicionário jurídico-legislativo considerando as variações das flexões (gênero/número/grau) dos termos próprios da linguagem jurídica (ex.: herdeiro/herdeira); 3) verificar variações da ortografia da língua portuguesa (ex.: objecto, objeto); 4) verificar palavras em línguas estrangeiras (ex.: latim, inglês, espanhol).

GT para o Tratamento de Assunto – deverá promover estudos e esforços para compatibilizar os diversos vocabulários utilizados por cada entidade com a finalidade de otimizar a recuperação da informação contida no Portal.

GT para elaborar Manual de Pesquisa Resumido e Detalhado – com o objetivo de publicar manual de pesquisa com orientações sobre os operadores, conteúdo, navegação e tipos de pesquisa.

GT para definir Modelo de Requisitos de Qualidade da Informação Legislativa e Jurídica – com o propósito de definir os requisitos mínimos de qualidade das informações contidas nas bases de dados a serem incluídas no acervo do Portal LexML.

4.1 Resultados obtidos

Os grupos de trabalho realizam reuniões mensais e os frutos dos seus trabalhos começam a aparecer. Alguns já estão disponíveis e outros ainda em fase de homologação.

O Comitê Gestor produziu texto de apresentação do Portal e o mandou traduzir para inglês, francês e espanhol.

O GT para definir as categorias e critérios para formação do acervo do Portal LexML fez estudos e propôs mudanças na apresentação das categorias, principalmente a apresentação das datas por décadas, a FRBRização dos documentos referenciais de

doutrina e a criação de uma check list com critérios e regras a serem cumpridas por outros tipos de documentos que queiram integrar, futuramente, o Portal.

O GT para o tratamento de linguagem alcançou seus objetivos ao fazer estudos para complementar os dicionários existentes, que facilitam a recuperação da informação. Assim ao pesquisar uma palavra com a grafia equivocada – como, por exemplo, “usocapião” – o usuário recebe os documentos onde consta a grafia correta – no caso, “usucapião”.

O GT utilizou a Constituição de 1988, todos os códigos brasileiros em vigor para extrair termos jurídicos e palavras comuns que não constavam do dicionário. Além disso, utilizou a Constituição de 1824 para localizar as palavras com grafia arcaica.

O GT para elaborar Manual de Pesquisa Resumido e Detalhado, recentemente implantado, está estudando as formas mais amigáveis graficamente de apresentar o manual. Enquanto os estudos prosseguem, o GT preparou uma lista de perguntas mais frequentes (FAQs) para atender as necessidades mais básicas de informação sobre o conteúdo do Portal e as diversas formas de pesquisar.

5 CONCLUSÃO

Para a gestão da informação legislativa e jurídica foram criados, por diversos órgãos governamentais, bancos de dados de forma independente. Para facilitar a busca e contribuir para a realização do direito fundamental de acesso à informação foi criado um ponto unificado de acesso: o Portal LexML.

A informação legislativa e jurídica é altamente caracterizada por remissões textuais, sejam remissões para o próprio texto da norma jurídica, sejam para outros diplomas legais. Podem ainda ser remissões em diferentes categorias de documentos: a doutrina remetendo para o texto de uma lei, por exemplo. A regra geral é que qualquer documento deste domínio está sempre inserido em um contexto, parte do nosso ordenamento jurídico.

O Projeto LexML adapta para o Brasil experiências bem sucedidas do exterior que inclui a definição de um identificador persistente implementado como uma URN. Tal identificador será utilizado como instrumento para representar as referências (e, mais genericamente, quaisquer tipos de relações) entre os documentos eletrônicos, com a

finalidade de tornar disponível um ambiente hipertextual global sobre os recursos de informação, operacional em um ambiente distribuído. Mais detalhes sobre o Projeto LexML e a definição das URN, incluindo a gramática em formato EBNF (*Extended Backus Naur Form*), podem ser consultadas no Portal de Documentação do Projeto (<http://projeto.lexml.gov.br>).

A unificação do acesso, tendo como origem diferentes bases de dados criadas de forma autônoma, gerou alguns aspectos que precisam ser estudados, como por exemplo, a compatibilização dos vocabulários. Para estudar esses aspectos e propor soluções foi instituído o Comitê Gestor de Informação do Portal LexML, que se propôs cumprir sete metas, no biênio 2009–2010, por meio da criação de grupos de trabalho (GT).

Os GTs são compostos, basicamente, por bibliotecários jurídicos que, com seu trabalho, experiência e conhecimento, têm contribuído para melhorar a recuperação da informação jurídica e legislativa contida no Portal.

REFERÊNCIAS

Archi, A., Biagioli, C., Cherubini, M., Ciampi, C., Fameli, E., Guidotti, P., Marinai, E., Mercatali, P., Nannuci, R., Ragona, M., Serroti, L., Socci, F. and Spinosa, P. Studio di fattibilità per la realizzazione del progetto “Accesso alle norme in rete”. *Informatica e Diritto*, v. 24, n. 1, p. 1–200, 2000.

VITALI, F. 2008. **Akoma Ntoso Release Notes**. Disponível em <http://www.akomantoso.org>. Acesso em: 5 mar. 2010.

Como citar este artigo:

LIMA, João Alberto de Oliveira; PASSOS, Edilenice. LexML: visão unificada da informação legislativa e jurídica do Brasil. In: SEMINÁRIO NACIONAL DE DOCUMENTAÇÃO E INFORMAÇÃO JURÍDICA, 2., 2010, Brasília. *Anais...*Brasília: 2010. 1 CD. Republicado em: *Cadernos de Informação Jurídica*, Brasília, v. 6, n. 1, p. 248–259, jan./jun. 2019. Disponível em: <http://www.cajur.com.br/>.