

REVISTA de INFORMAÇÃO LEGISLATIVA

Brasília • ano 48 • nº 192
outubro/dezembro – 2011

SUBSECRETARIA DE EDIÇÕES TÉCNICAS DO SENADO FEDERAL

Apuração do texto original da Lei Geral de Orçamento (Lei nº 4.320/1964) a partir das Bases de Legislação Federal

Estudo de Caso

João Alberto de Oliveira Lima

Sumário

1. Considerações iniciais. 2. A precisão como atributo de qualidade da informação. 3. Apuração do texto da Lei Geral de Orçamento. 4. Metodologia do estudo de caso. 5. Considerações finais e recomendações

1. Considerações iniciais

Este estudo de caso foi concebido a partir de pedido de informação realizado pelo Ministério Público Federal (MPF) que questionou o Senado Federal “sobre quais mecanismos e providências podem ser adotadas para minimizar a incidência de erros na divulgação de textos legislativos, por meios eletrônicos”. Essa consulta foi motivada pela questão de ordem da decisão judicial proferida pelo Exmo. Juiz Federal Leandro Paulsen na qual constatou que as partes (Empresa Privada e União) trabalhavam com uma redação equivocada do § 4º do art. 1º do Decreto nº 64.833/1969 que estava disponível na internet, inclusive em sítios de órgãos oficiais. Segundo a decisão, exarada na Ação Ordinária nº 87.00.01353-6/RS em 18 de dezembro de 2009, o entendimento equivocado poderia ter “implicado prejuízos vultosos à União”. No intuito de detalhar as nuances do problema, realizamos um estudo de caso com o objetivo de apurar o texto original da Lei Geral de Orçamento (Lei nº 4.320/1964) a partir dos textos disponíveis nas bases de legislação federal.

João Alberto de Oliveira Lima é Analista de Informática Legislativa do Senado Federal, mestre em Ciência da Computação e doutor em Ciência da Informação.

O “Diário Oficial da União” (DOU), periódico editado pela Imprensa Nacional, é a publicação oficial de atos resultantes do processo legislativo previsto na Constituição Federal. Imprecisões na publicação de atos no DOU são corrigidas por retificações ou republicações. Os textos publicados nas bases de dados de legislação na internet no Brasil, mesmo que armazenadas em sítios oficiais das instituições do Estado, não possuem nenhum caráter oficial. Essa ressalva é apresentada pelas bases de dados como alertas. Por exemplo, na base de legislação da Presidência da República, consta o seguinte alerta, em letras vermelhas: “Este texto não substitui o publicado no DOU de dd.mm.aaaa”.

Mesmo não tendo valor oficial, é importante ressaltar que as bases de textos legislativos desempenham um importante papel de auxílio na compreensão do nosso ordenamento jurídico. Edilenice Passos (2009), no trabalho intitulado “Requisitos recomendáveis para as bases de dados de legislação estadual e distrital”, descreve o cenário de acesso à legislação como sendo suportado pelo tripé: (a) obrigação de publicação das normas jurídicas; (b) obrigação de se conhecer a legislação brasileira, que, embora não possa significar um dever de conhecimento efetivo do conjunto da legislação brasileira, certamente resulta no dever estatal em disponibilizá-la para acesso amplo e facilitado; (c) volume de normas jurídicas editadas diariamente.

“A obrigação de publicação das normas jurídicas está determinada no art. 1º da Lei de Introdução ao Código Civil que postula: ‘Salvo disposição contrária, a lei começa a vigorar em todo o país quarenta e cinco dias depois de oficialmente publicada’. (...)

A obrigação de conhecer a legislação brasileira está descrita em duas normas: no art. 3º da Lei de Introdução ao Código Civil, que claramente dispõe que ninguém se escusa de cumprir a lei, alegando não a conhecer; e no art.

21 do Código Penal, que contém disposição similar: ‘O desconhecimento da lei é inescusável’.

Entretanto, o volume de informações jurídicas normativas pode dificultar a sustentação do segundo lado do tripé, pois, tanto o excesso quanto a falta de informações, pode causar a desinformação e, no caso de acesso à informação jurídica, impedir o exercício da cidadania. (...)

Se de um lado, o Estado sempre cumpriu bem seu papel publicando religiosamente toda a legislação, por outro, nem sempre o Estado consegue completar o processo de comunicação. Isto ocorre porque a mensagem geralmente não é reconhecida pelo receptor, o cidadão brasileiro, que, em teoria, deveria ler diariamente o Diário Oficial da União, Diário Oficial do seu Estado (DOE) e o Diário Oficial do município onde reside. Além disso, precisa decifrar a mensagem, entendendo seu conteúdo e significado.

Até pouco tempo, a única forma de acesso ao texto das normas jurídicas se restringia ao texto impresso de diários oficiais. Mais recentemente, com a popularização dos computadores e o advento da Internet, o Estado passou a ter uma atitude pró-ativa, proporcionando alternativas ao acesso à informação jurídica com a criação de bases de dados que reúnem e tornam disponível a legislação brasileira. Trata-se de um serviço prestado pelo Estado para o exercício da cidadania” (PASSOS, 2009, grifo nosso).

Com essas considerações, podemos afirmar que a Subsecretaria de Informações do Senado Federal, o Centro de Documentação e Informação da Câmara dos Deputados e a Subchefia de Assuntos Jurídicos da Presidência da República, gestores de bases de informações de legislação federal disponíveis na internet, contribuem de forma

decisiva com o cumprimento do preceito constitucional que define o cidadão como titular do direito de acesso à informação (CF, art. 5º, XIV).

O resto do artigo está organizado em mais quatro seções: inicialmente, apresentaremos o processo histórico de formação das principais bases de legislação federal e identificaremos a principal causa dos problemas de precisão; na sequência, como forma de detalhar as nuances do problema, mostraremos os resultados do estudo de caso da apuração do texto original da Lei de Orçamentos, bem como a metodologia utilizada; por fim, serão apresentadas as considerações finais e sugeridas recomendações para minimizar a incidência de erros em bases de dados de legislação.

2. A precisão como atributo de qualidade da informação

Segundo Knight e Burn (2005), a “precisão” é o principal atributo de qualidade da informação. No estudo *“Developing a Framework for Assessing Information Quality on the World Wide Web”*, as autoras definem “precisão” como uma medida na qual se verifica se os dados estão corretos, confiáveis e asseguradamente livres de erros. O indicador de “precisão” qualifica a informação como “precisa” ou “imprecisa”, não existindo um meio termo.

A situação da precisão das bases de legislação informatizadas é derivada de um processo histórico que se iniciou na década de 1970 com a instalação do Centro de Informática e Processamento de Dados do Senado Federal – PRODASEN. Já nessa época, o sistema de informação NJUR (Normas Jurídicas) realizava o registro das normas jurídicas, dos assuntos e relacionamentos (vides), no intuito de representar a intrincada rede de conexões derivada da constante mutação do ordenamento jurídico.

No início da década de 1990, com o advento da microinformática, o PRODASEN realizou um trabalho de transcrição do

texto de normas jurídicas do Diário Oficial da União para o formato digital e lançou quatro edições do CD-ROM com a íntegra da Legislação Brasileira desde o império.

Em meados da década de 1990, diversas instituições criaram suas próprias bases de textos de legislação e as disponibilizaram na internet. Nesse processo, textos que haviam sido digitalizados pelo PRODASEN foram replicados em outras bases. Daí a explicação da ocorrência de um mesmo erro em diversas bases de legislação.

Apesar dos esforços envidados pelas instituições para uma revisão contínua nas bases de informação, a ocorrência de erros em textos de bases de legislação é mais comum do que se imagina, principalmente quando consideramos os textos anteriores ao ano de 2000 que foram transcritos por processo de digitação ou reconhecimento ótico de caracteres (OCR – *Optical Character Recognition*).

Faz-se necessário destacar que os erros se multiplicam pelo processo da cópia, mas, infelizmente, a sua identificação e correção, não. As eventuais correções, normalmente, são realizadas apenas na base de legislação da instituição que identificou o erro. Inexiste uma cooperação sistemática para a melhoria conjunta dos textos das bases de legislação federal.

A transcrição por digitação está no núcleo do problema da precisão dos textos legislativos. Todos estamos sujeitos aos erros de transcrição. Por exemplo, o texto da decisão do Exmo. Juiz Federal Leandro Paulsen apresenta o nome do sítio de legislação da Presidência da República transcrito – *data venia* – com quatro erros. Digitou-se:

www.presidencia.gob.br/legislacao
quando o correto seria:

www.presidencia.gov.br/legislacao

Nos dias atuais, a incidência de erros em textos de novas normas é menor pois não ocorre o processo de transcrição por digitação ou por reconhecimento ótico de caracteres. Na transcrição por digitação, o

operador realiza a leitura em um suporte, interpreta os símbolos e os recodifica por meio da digitação. No caso do uso da técnica de OCR, a imagem digitalizada de um texto é interpretada por um *software* que gera uma sequência de caracteres. Ambos os métodos podem gerar erros ocasionados por diversos motivos.

3. Apuração do texto da Lei Geral de Orçamento

Com o auxílio de técnicas da Filologia e de ferramentas desenvolvidas no âmbito do Projeto LexML (www.lexml.gov.br), comparamos os textos da Lei nº 4.320/1964 disponíveis nas três principais bases de legislação federal (em ordem alfabética: Câmara dos Deputados, Presidência da República e Senado Federal) com o texto da publicação do Diário Oficial da União. Os textos foram obtidos por consulta na Internet realizada em 1º de abril de 2011. Para a nossa surpresa, identificamos 52 dispositivos com imprecisões dos mais diversos tipos. Constatamos que nenhuma das bases de legislação federal possuía o texto idêntico ao da publicação oficial da Lei nº 4.320/1964. Ao ler este artigo, provavelmente, os erros aqui relatados já estarão corrigidos nas bases de legislação federal¹.

A Lei nº 4.320/1964 está vigente e “Estatui Normas Gerais de Direito Financeiro para elaboração e controle dos orçamentos e balanços da União, dos Estados, dos Municípios e do Distrito Federal”. A publicação oficial dessa norma foi veiculada pelo Diário Oficial da União na edição de 23 de março de 1964 e seu texto foi objeto de retificação no dia 9 de abril de 1964. O texto da promulgação de veto, publicado no dia 5 de maio de 1964, foi também objeto de retificação, que foi publicada no dia 3

¹ Para obter versões anteriores de uma página na *World Wide Web*, pode-se utilizar, no sítio <http://www.archive.org>, a funcionalidade “*WayBack Machine*”, sendo necessário apenas informar o endereço da página desejada.

de junho de 1964. A primeira retificação da norma foi pontual e modificou palavras da ementa, do preâmbulo e do *caput* do art. 14. A retificação do veto promulgado substituiu uma palavra do § 3º do art. 43.

A Tabela 1 apresenta um resumo dos erros encontrados no estudo de caso por base de dados de origem. Não identificaremos as bases enumeradas na tabela pois o objetivo do estudo não é mostrar qual base contém mais ou menos erros e, sim, mostrar que o problema ocorre em todas as bases e quais ações podem ser realizadas para atenuar ou, até mesmo, resolver o problema.

Os erros foram classificados nas seguintes categorias: Acentuação – ausência ou presença de acentos; Pontuação – ausência ou presença de sinais de pontuação; Capitulação – troca de maiúscula por minúscula ou vice-versa; Troca de Palavras – quando uma palavra é trocada por outra com diferente lema; Outros Erros – agrupa os demais casos, tais como mudança de gênero (masculino/feminino) ou número (singular/plural) das palavras.

Tabela 1 – Resumo dos Erros por Tipo e Base de Legislação (Lei 4.320/1964)

Tipo de Erro	Base 1	Base 2	Base 3
Acentuação	6	6	8
Pontuação	7	16	17
Capitulação	3	4	4
Troca de Palavra	8	14	12
Outros Erros	11	16	15
Total	35	56	56

A Tabela 2 apresenta alguns casos de imprecisão no texto da lei do estudo de caso, em que se observou a substituição de uma palavra por outra com sentido diverso da palavra do texto original.

Não é exagero afirmar que, em plena época de transição do suporte papel para o digital, vivemos uma situação parecida com a da Idade Média, quando os textos de manuscritos produzidos por copistas possuíam erros gerados pelo processo de

Tabela 2. Quadro resumo de troca de palavras na Lei nº 4.320/1964 (grifo nosso)

Dispositivo	Imagem DOU	Substituição de	Por	B1	B2	B3
art. 12, § 2º	manutenção	manutenção	manifestação	✓	!	!
art. 12, § 6º	lei especial anterior.	especial	especialmente	!	!	!
art. 36, § único	correm a conta de	correm	sorvem	✓	!	!
art. 49, caput	para efeito do	efeito	feito	!	!	!
art. 54, caput	da obrigação de	obrigação	observação	!	!	!
art. 61, caput	a especificação, e a	especificação	representação	✓	!	!
art. 72, caput	vinculadas a fundos	fundos	turnos	✓	!	!
art. 105, § 5º	nos paragrafos, anteriores e que, mediata ou indiretamente, possam vir a afetar o patrimônio.	mediata	imediate	✓	!	!
art. 107, caput	parafiscais da União	parafiscais	para fiscais	✓	✓	!

Legenda: ✓ Correto ! Errado

transcrição, gerando múltiplas versões incompatíveis entre si. Dom Paulo Evaristo Arns (2007, p. 59), na obra “A Técnica do Livro segundo São Jerônimo”, nos relata que:

“Os erros dos copistas eram tão frequentes que se receava sua presença em toda parte. (...) Uma cultura superficial está na origem de muitas correções. Quando o copista não entende, acrescenta, muda, para que a narrativa fique clara, isto é, errada.”

Para nosso espanto, essa “correção” por parte do “copista” também foi constatada na transcrição do § 3º do art. 105 da Lei nº

4.320/1964 para o meio digital. Pela análise de imagem do exemplar do Diário Oficial da União (Fig. 1) é possível perceber que o texto do § 3º está truncado na transição entre a segunda e a terceira linha.

O truncamento no texto deve ter motivado o copista a intervir. A Tabela 3 apresenta o resultado da comparação automática do texto do DOU com os textos das bases de legislação federal. As diferenças entre a versão oficial e as demais são destacadas utilizando-se itálico no texto acrescentado e tachado no texto removido. A análise é feita palavra a palavra. Numa eventual troca de

Figura 1. §§ 3º e 4º do art. 105 da Lei nº 4.320/1964 no DOU de 23/03/1964

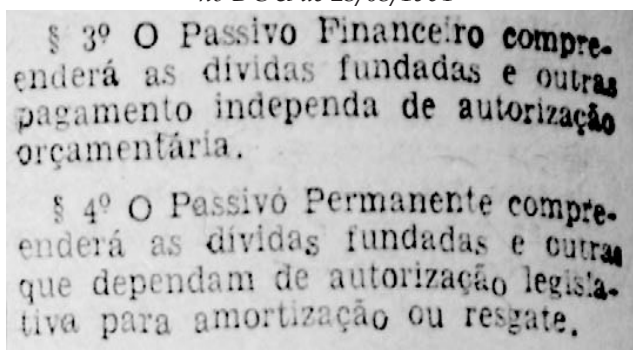





Tabela 3. Comparativo das versões do § 3º do art. 105 da Lei nº 4.320/1964

Art. 105, § 3º		DOU	O Passivo Financeiro compreenderá as dívidas fundadas e outras pagamento independa de autorização orçamentária.
		B1	O Passivo Financeiro compreenderá as dívidas fundadas e outras cujo pagamento independa de autorização orçamentária.
		B2	O Passivo Financeiro compreenderá as dívidas fundadas e outras, cujo pagamento independa de autorização orçamentária.
		B3	O Passivo Financeiro compreenderá as dívidas fundadas e outrosoutras pagamento independa de autorização orçamentária.

palavra, a palavra inserida é apresentada em itálico seguida da palavra removida apresentada em tachado.

Pela análise da Tabela 3, percebemos que o texto do § 3º do art. 105 possui três versões e que, por motivos diferentes, nenhuma delas coincide com a oficial. No texto da Base 1, notamos a ocorrência da palavra “cujo” incluída pelo copista; na versão do Base 2, a palavra “cujo” também ocorre, sendo precedida de uma vírgula. E, na versão da Base 3, não ocorre a palavra “cujo”, mas a palavra “outras” que consta do original foi substituída pela palavra “outros”.

Na busca de uma explicação para o texto truncado, seguimos a sugestão de Jerônimo a Rufino, relatada por Arns (2007, p. 178) da seguinte forma:

“Jerônimo indica a Rufino o meio de verificar sua afirmação sobre a autenticidade de uma carta: por que não recorrer aos arquivos e pedir o original, por que não ir à casa do remetente?”.

Recorremos à Secretaria de Arquivo do Senado Federal e obtivemos o texto do au-

tógrafo do Senado Federal do Projeto de Lei da Câmara nº 38 de 1952, matéria origem da Lei nº 4.320/1964. A intenção do legislador tornou-se clara, como podemos observar na Fig. 2. Em algum momento do processo legislativo, ocorreu a troca da expressão “os compromissos exigíveis cujo” pela expressão “as dívidas fundadas e outras”.

Para ter segurança de que o texto do § 3º do art. 105 do autógrafo retornado à Câmara dos Deputados não foi modificado, recorremos ao Centro de Documentação e Informação da Câmara dos Deputados e verificamos que o Substitutivo do Senado Federal foi aprovado com exceção do art. 22 (Emenda nº 22). A proposição que possuía 116 artigos no texto do substitutivo do SF passou a ser articulada em 115 artigos na versão final enviada à sanção presidencial. O art. 105 não foi objeto de alteração na última etapa do processo legislativo. No entanto, ao verificarmos o autógrafo enviado à sanção presidencial, percebemos que a adulteração do § 3º do art. 105 já estava presente (Fig. 3). Comparando as Figuras 1 e 3, percebemos que a publicação do Diário Oficial da União

Figura 2. Trecho do Autógrafo do PLC 38/1952

§ 3º - O Passivo Financeiro compreenderá os compromissos exigíveis cujo pagamento independa de autorização orçamentária.

§ 4º - O Passivo Permanente compreenderá as dívidas fundadas e outras, que dependam de autorização legislativa para amortização ou resgate.

Figura 3. Trecho do autógrafo da Câmara dos Deputados com a adulteração (grifo nosso)

§ 3º O Passivo Financeiro compreenderá as dívidas fundadas e outras pagamento independa de autorização orçamentária.

§ 4º O Passivo Permanente compreenderá as dívidas fundadas e outras que dependam de autorização legislativa para amortização ou resgate.

foi fiel ao texto do autógrafo. Podemos notar ainda que a segunda linha dos §§ 3º e 4º possui a mesma sequência de caracteres.

Então, a adulteração do texto publicado no DOU foi gerada de forma acidental por falha humana no momento do autógrafo. Arns (2007, p.172) relata a preocupação de Jerônimo no problema da adulteração dos escritos e divide a responsabilidade pela adulteração entre autores e copistas:

“Jerônimo, o tradutor da Bíblia, se vê forçado a cuidar da questão, e tenta descobrir suas causas.

Adulteração acidental

O longo encadeamento de cópias com os erros sucessivos devido aos escribas está na base da adulteração. Como o *exemplar* nem sempre é corrigido, os copistas só podem agravar ainda mais os erros”.

Enquanto na Idade Média os erros se propagavam de forma lenta, devido aos custos da realização das cópias, com a invenção da imprensa, um erro passou a ser multiplicado pela quantidade de exemplares de uma tiragem de uma edição. Na era da internet, como os custos de cópia e publicação foram reduzidos de forma drástica, os erros fluem na velocidade de um clique de *mouse*.

A Tabela 4 apresenta a quantidade de ocorrências das versões do § 3º do art. 105 da Lei nº 4.320/1964 que foram verificadas no dia 11 de abril de 2011 nos sítios da internet e também em sítios do governo na internet. Para realizar a contagem da disseminação das versões, utilizamos o principal buscador genérico disponível: o Google (www.google.com). Como o Google não considera a pontuação na indexação

Tabela 4. Disseminação das versões do § 3º do art. 105 da Lei nº 4.320/1964

Versões do § 3º do art. 105 da Lei nº 4 320/1964	Quantidade de ocorrências	
	Todos os sítios	Sítios gov.br
O Passivo Financeiro compreenderá as dívidas fundadas e outras pagamento independa de autorização orçamentária. (DOU)	73	56
O Passivo Financeiro compreenderá as dívidas fundadas e outras cujo pagamento independa de autorização orçamentária. (Base 1 e Base 2)	274	54
O Passivo Financeiro compreenderá as dívidas fundadas e outros pagamento independa de autorização orçamentária. (Base 3)	374	36
O Passivo Financeiro compreenderá os compromissos exigíveis cujo pagamento independa de autorização orçamentária. (versão do autógrafo)	95	5

dos documentos, as versões das Bases 1 e 2 foram quantificadas conjuntamente.

Interessante notar que a “versão do autógrafo” possui disseminação na internet maior que a “versão do DOU”. No entanto, em sítios governamentais, a versão do autógrafo é pouco difundida. Uma pesquisa adicional revelou que a “versão do autógrafo” foi disseminada pela doutrina, conforme explicação de nota no texto da Lei nº 4320/1964 disponível no Portal da Secretaria de Orçamento Federal do Ministério do Planejamento Orçamento e Gestão no endereço https://www.portalsof.planejamento.gov.br/bib/legislacao/Lei_4320.pdf, nos termos a seguir:

“O trexo [sic] original ‘... as dívidas fundadas e outras, ...’ tem sido substituído por ‘...os compromissos exigíveis, cujo...’ em diversas publicações de grande credibilidade como a ‘Revista das Finanças Públicas, Edição Suplementar, de Outubro de 1967’, a constante da terceira edição do texto atualizado até janeiro/1995 da Lei nº 4.320/64, supervisionado pela ABOP/RS e ‘A Lei 4.320 Comentada’, de J. Teixeira Mqchado [sic] Jr. E Heraldo da Costa Reis” (BRASIL, [2011]).

O caso citado não foi o único de intervenção do copista quando em dificuldades na transcrição do texto truncado ou faltando letras no DOU. Identificamos ainda dois outros casos de erro no texto publicado no DOU nessa mesma lei. O texto de norma jurídica publicado com erros no DOU, de acordo com conceitos da Filologia, é considerado um texto autêntico, mas não genuíno, pois não respeitou o desejo do autor. A retificação corrige a deficiência e torna o texto genuíno à medida que aproxima o texto oficial do texto pretendido pelo autor.

Razões ou causas para tantos erros

O filólogo Segismundo Spina, na obra “Introdução à Edótica” editada originalmente em 1977, considera que as alterações do texto original, no processo

de transcrição, são geradas por razões independentes da ferramenta utilizada no trabalho de transcrição, isto é, aplicam-se tanto aos copistas da Idade Média quanto aos linotipistas e, seguramente, podemos acrescentar, aos digitadores da atualidade. Ele nos ensina:

“É a atenção o suporte psicológico da boa ou má transcrição de um manuscrito ou texto impresso. O cansaço do ato de copiar, as interrupções do trabalho pelas causas mais diversas, a memória de quem lê os segmentos de texto para transcrevê-los, os deslizos de leitura ocasionados pelo movimento dos olhos que se deslocam do original para cópia e desta para o original, etc., etc., são geralmente os fatores mais comuns a intervirem na atenção do copista ou do linotipista, e conseqüentemente no processo de reprodução de um texto. (...) Ambos, linotipista e copista, não lêem palavra por palavra, mas segmentos fráscicos com sentido, e na reprodução sucede transcreverem com fidelidade a parte inicial do segmento, mas com erros a porção final porque a memória enfraqueceu. (...) Os chamados ‘erros homotelêuticos’, por exemplo, são também explicáveis por descuidos da atenção: ao retornar os olhos para o modelo, freqüentemente sucede que o copista salte uma linha ou mais, porque a palavra deixada se repete logo adiante e a transcrição recomeça a partir daí. Esse erro ocorre comumente quando palavras iguais ou semelhantes aparecem nos finais das linhas” (SPINA, 1977, p. 112).

Em dois outros trechos, o mesmo autor (SPINA, 1977, p. 115,117) relata as espécies dos erros acidentais e oferece uma classificação para os erros:

“Os erros acidentais são, entretanto, mais numerosos: grafias e sons equívocos, confusão de palavras semelhantes, transposição de letras e

de palavras, omissão de palavras ou frases, salto de linhas ou de versos, confusão ocasionada por parônimos, supressão de sílabas nas palavras (*haplografia*), ou sua repetição (*ditografia*), são as incorreções mais freqüentes nas cópias manuscritas ou nas composições tipográficas.

Os erros podem ser estudados a) quanto à sua *responsabilidade* (erros do autor, do copista, gráfico ou revisor, erros do editor); b) quanto à sua *natureza* (erros voluntários ou fraudes, involuntários, inevitáveis, correções intencionais); c) quanto ao *tipo* (omissão, confusão de letras, saltos, transposições, etc.) e d) quanto ao seu *condicionamento psicológico* (que compreende todos os aspectos da patologia da atenção)."

A seguir, apresentaremos a metodologia utilizada no estudo de caso. Trata-se de uma combinação da aplicação de técnicas da Edótica auxiliadas pela informática.

4. Metodologia do estudo de caso

A Edótica é uma disciplina da Filologia que cuida da busca da genuinidade do texto. Segismundo Spina (1977) nos apresenta o objetivo da Edótica e as principais fases do processo de apuração do texto:

"O objetivo da Edótica é um só: estabelecer um texto que se avizinha o mais possível do original. A esse processo de retorno, em que se tenta restaurar a forma original da obra com o auxílio das normas da crítica textual, se dá o nome de *constitutio textus*.

A constituição de um texto, observada a doutrina de Karl Lachmann – o fundador da moderna crítica textual –, passa por uma série de operações que se repartem por três fases mais ou menos distintas: a da RECENSIO (*Recensare*), a da ESTEMÁTICA (*Originem detegere*) e a da EMENDATIO

(*Emendare*). É através desta seqüência de operações que se chega à primeira etapa da edição crítica: a apuração do texto. (...)

A *recensio* consiste na pesquisa e coleta de todo o material existente de uma obra, manuscrito ou impresso (p. 88). Classificados todos os testemunhos da tradição direta (apógrafos, arquétipos, subarquétipos, antígrafos, etc.), parte o editor para a segunda fase da crítica textual, a que Lachmann denomina '*originem detegere*', isto é, '*revelar a ascendência*'. O editor procede, então, à apresentação genealógica dos testemunhos: como os mss. se filiam entre si e como se operou a sua transmissão (p. 94).

Ao conjunto de operações que visam à correção do texto se dá o nome de *Emendatio*" (p. 109).

Na apuração do texto que serviu de base para a comparação entre as versões dos textos das bases de dados de legislação, passamos pelas três etapas, que serão apresentadas nas três subseções seguintes.

4.1. Recensio

Nessa etapa, coletamos, inicialmente, os seguintes textos:

a) Texto da Lei nº 4.320/1964 (Publicação Original), disponibilizado pela base de legislação da Câmara dos Deputados em formato HTML no dia 1º de abril de 2011;

b) Texto da Lei nº 4.320/1964 (Publicação Original), disponibilizado pela base de legislação do Senado Federal em formato HTML no dia 1º de abril de 2011;

c) Texto da Lei nº 4.320/1964 (Texto Multivigente), disponibilizado pela base de legislação da Presidência da República em formato HTML no dia 1º de abril de 2011;

d) Cópia (imagem digital) do Diário Oficial da União do dia 23 de março de 1964 que publicou a Lei nº 4.320/1964;

e) Cópia (imagem digital) do Autógrafo do PLC 38/1952 enviado pelo Senado Federal à Câmara dos Deputados; e

f) Cópia (imagem digital) do Autógrafo do PL 201/1950 enviado pela Câmara dos Deputados à sanção presidencial.

Antes de partimos para a análise dos “erros comuns” às versões textuais, realizamos um procedimento de normalização com o objetivo de preparar os textos para o processamento pelo computador. Essa normalização consistiu basicamente de: a) padronização nos rótulos dos dispositivos; b) retirada dos caracteres repetidos, de espaço em branco, tabulação ou quebras de parágrafos; c) retirada das interpolações no texto da base da Presidência da República devido às alterações da norma e à promulgação do veto que foi parcialmente derrubado; d) conversão do formato HTML para o formato LexML realizada por um programa *Parser* desenvolvido pelo PRODASEN.

O formato LexML é uma linguagem de marcação derivada do XML (eXtensible Markup Language) que foi criada pelo Projeto LexML (<http://projeto.lexml.gov.br>) a partir de experiências bem sucedidas do exterior, com o objetivo de estruturar o texto de normas jurídicas brasileiras. Trata-se de um formato aberto, não proprietário, em que os caracteres que formam os dispositivos de uma lei são delimitados por marcadores. Dessa forma, o computador consegue “entender” quais símbolos formam um inciso, uma alínea, um parágrafo ou um *caput*. Uma vez convertido para o formato LexML, a cadeia de símbolos pode ser facilmente convertida para todos os outros formatos de consumo humano, tais como: HTML, PDF, EPUB, DOC, ODT, MP3 (síntese de voz), etc. Essa estruturação do texto é importante, pois a comparação é realizada dispositivo a dispositivo. A comparação entre os textos foi realizada por um programa desenvolvido pelo PRODASEN que utiliza o algoritmo JNDiff desenvolvido na Universidade de Bolonha e utilizado pelo Senado italiano.

Ao realizarmos a primeira comparação dos três textos coletados das bases de legislação, constatamos a existência dos “erros

comuns”, indicador forte de parentesco entre as versões, decorrente do processo histórico de formação das bases de legislação federal, como já mencionado na introdução deste estudo de caso. No processo de comparação dos textos, denominado colação das edições, normalmente se escolhe um texto para ser o “texto base” de comparação com os demais. A comparação automática realizada apenas entre cópias de texto de uma mesma família pode ser falha, pois os textos derivados carregam consigo os erros do texto que originou a família. Por esse motivo, decidimos acrescentar à nossa análise um novo texto, o da Lei nº 4.320/1964, que foi composto de forma independente por outro copista.

Para atender a essa necessidade de um novo texto, poderíamos solicitar a redigitação do texto do DOU. No entanto, para simplificar o processo, decidimos obter o texto desejado a partir do serviço de assinatura de bases de dados de informação jurídica que a Biblioteca do Senado Federal oferece aos servidores do Senado Federal na sua rede interna (*intranet*). Verificamos que, apesar de esse novo texto não possuir os “erros comuns” à família de textos, não estava livre de outros erros. Constatamos, assim, que o problema de precisão de textos legislativos em bases informatizadas ocorre também em base de legislação oferecida por empresas privadas às instituições governamentais.

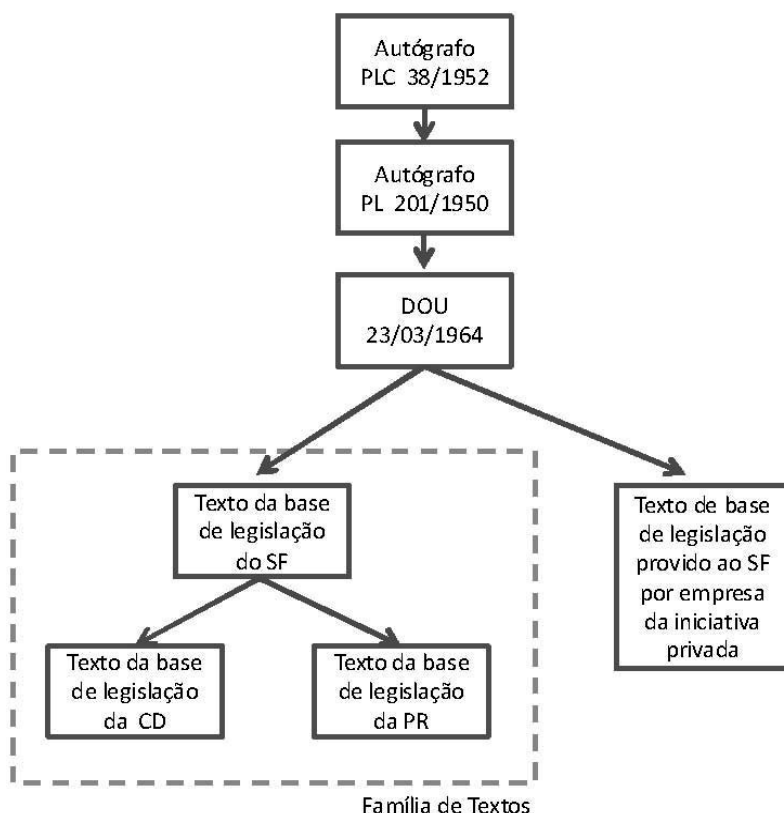
O texto de origem externa foi replicado para ser o “texto base” da comparação automática. Na fase de *Emendatio*, o “texto base” será corrigido inúmeras vezes com o objetivo de compor o texto desejado, denominado arquétipo. No nosso estudo, o arquétipo deve ser idêntico ao texto do DOU.

Nesse ponto, foi possível definir a árvore genealógica dos textos, apresentada na próxima seção.

4.2. Estemática

Nessa etapa, revelamos a ascendência dos textos coletados na etapa anterior. A Figura 4 apresenta a árvore genealógica dos

Figura 4. Árvore genealógica (estema)



textos, também conhecida como estema. Não foram considerados no estudo de caso nem representados no diagrama os textos das duas retificações publicadas no DOU e o texto da promulgação de veto, pois, ao considerarmos como arquétipo apenas o texto original do DOU, os textos posteriores passam a ser desnecessários para a análise. Um estudo mais completo deve considerar todos os textos oficiais.

4.3. Emendatio

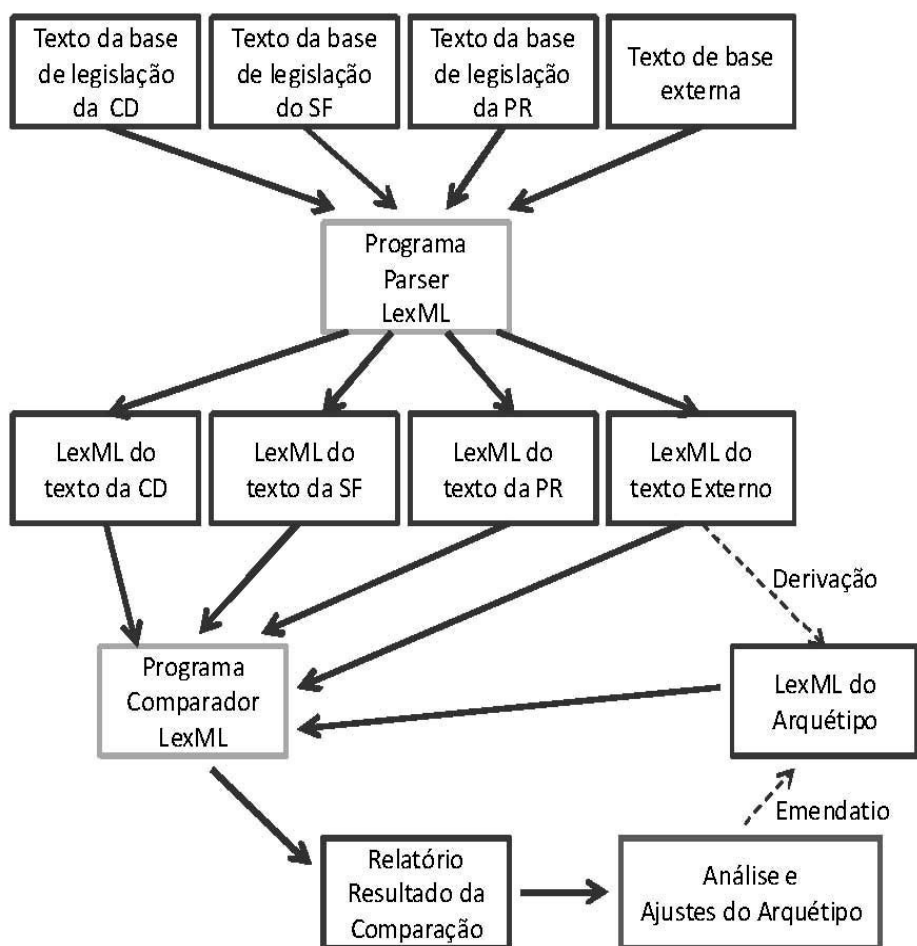
Nessa etapa, baseado nos resultados da comparação entre as diversas versões, identificamos os erros e, com a análise das imagens do DOU, e dos autógrafos, efetuamos as alterações no arquétipo para que ele se aproxime do texto do DOU.

A Figura 5 representa a preparação dos textos com a conversão pelo *Parser*; o pro-

cesso de derivação do arquétipo, o processo de comparação e, finalmente, o processo iterativo de análise e ajustes do arquétipo (*emendatio*). O processo de comparação e análise é realizado quantas vezes forem necessárias até que o analista considere que o arquétipo é suficientemente próximo do original do DOU. É interessante observar que, na primeira iteração, nenhum erro foi encontrado no texto que serviu de base para o arquétipo, pois esse foi derivado diretamente daquele por cópia integral. Nas iterações seguintes, com a identificação e modificações do arquétipo, os erros do texto da base externa começam a aparecer.

Esse processo pode ser realizado com qualquer quantidade de textos. Quanto mais textos de copistas diferentes alimentarem o sistema, maiores serão as chances de identificação dos erros.

Figura 5. Processo de Preparação, Comparação e Emendatio



5. Considerações finais e recomendações

Inexiste uma solução única para o problema enunciado neste estudo de caso. O problema da falta de precisão nos textos da legislação federal, muito provavelmente, também deve ocorrer nos acervos de legislação estadual, distrital e municipal. Edilene Passos (2009) constata a inexistência de um esforço coordenado na disponibilização da informação jurídica para o cidadão:

“No âmbito federal, três órgãos públicos oferecem o controle da informação jurídica normativa: o Senado Federal, a Presidência da República e

a Câmara dos Deputados. Os estados e o Distrito Federal seguiram o exemplo e passaram a disponibilizar na Internet bases de dados com a compilação da legislação estadual e distrital. Entretanto, devido à inexistência de recomendações internacionais ou nacionais sobre as funções e os pontos de acesso, cada unidade da Federação desenvolveu isoladamente sua base de dados, cada uma com suas qualidades e vantagens, mas basicamente diferentes umas das outras.”

Outro aspecto a considerar, com reflexo direto nas bases de legislação, é que *alguns vícios devem ser corrigidos antes da criação da*

norma. Melhorias no processo legislativo serão refletidas de forma positiva no ordenamento jurídico do Brasil.

Atualmente, em toda a cadeia do processo legislativo, o computador é utilizado de uma forma muito parecida com “máquina de datilografar”. Na iniciativa, a proposição é submetida como um arquivo codificado em um formato de um programa editor de textos comercial. A partir daí, o texto passa a ser migrado de um documento para outro, utilizando a conhecida operação “copia e cola” ou a cópia de arquivos, até que, finalmente, seja impresso em papel para o autógrafo. Para minimizar erros de transcrição do autógrafo na Presidência da República, envia-se também, junto com o autógrafo em papel, um arquivo digital do Congresso Nacional para a Presidência da República. Toda a manipulação dos textos é realizada de forma artesanal. *Essa cadeia pode se tornar melhor e mais segura na medida em que o computador passe a ser utilizado de forma mais inteligente.*

Por exemplo, várias regras da técnica legislativa, conforme definido pela Lei Complementar nº 95/1998, podem ser verificadas de forma automática por um programa que entenda as regras de articulação. Infelizmente, é fácil encontrar textos de leis que não seguem as regras de articulação de dispositivos da referida lei complementar. O recente Código Civil de 2002 utiliza “SUBTÍTULO” como dispositivo para agrupar artigos, elemento este não previsto no rol de agrupadores que considera apenas: Parte, Livro, Título, Capítulo, Seção e Subseção. *Este e outros tipos de erros formais podem ser facilmente evitados com a validação automática da estrutura de documentos textuais utilizando a tecnologia desenvolvida pelo LexML.*

Nesse contexto, apresentamos as seguintes recomendações:

Recomendação 1: Criação de uma base de informação legislativa derivada das bases existentes utilizando a metodologia apresentada neste estudo de caso.

Justificativa: A integração de esforços entre os gestores de informação de base de dados legislativa poderá minimizar a incidência de erros nos textos ao mesmo tempo em que compartilha os resultados sem retrabalho. O trabalho de “análise e ajustes do arquétipo” pode ser realizado de forma distribuída. Por exemplo, os gestores das bases de legislação das casas legislativas federais poderiam analisar as leis, as leis complementares e as emendas constitucionais, enquanto a Presidência da República poderia analisar os decretos. Todos receberiam os relatórios com a listagem dos erros identificados. O objetivo da base compartilhada é representar o texto tal como foi publicado no DOU com as retificações. A princípio, os textos derivados por compilação, anotação ou outro processo não devem fazer parte do escopo dessa base unificada. Dessa forma, todos os erros identificados serão corrigidos em todas as “bases fontes” de forma consistente.

Recomendação 2: Disponibilização da base unificada em formato aberto para toda a sociedade.

Justificativa: O texto preciso, correto e confiável é de interesse de todos. Não só os cidadãos, mas todos os segmentos da sociedade poderão se beneficiar do resultado desse esforço conjunto.

A ideia da disponibilização das informações jurídicas do país em formato XML não é inédita. A Inglaterra disponibilizou em 2010, no endereço <http://legislation.gov.uk>, todas as normas jurídicas no formato XML. Além de permitir a consulta do texto vigente de uma norma para qualquer data no portal, qualquer cidadão ou empresa pode realizar o *download* de todo o conjunto de arquivos.

Recomendação 3: Criação de um sistema que integre os sistemas existentes na Câmara dos Deputados, no Senado Federal, na Presidência da República e na Imprensa Nacional com o objetivo de permitir o intercâmbio de documentos eletrônicos com assinatura digital.

Justificativa: Instituições do Poder Judiciário já se adaptaram ao processo eletrônico motivadas principalmente pela grande quantidade de processos que tramita na justiça. Atualmente, já ocorre o intercâmbio de documentos eletrônicos com assinatura digital com certificação entre Tribunais Superiores, Tribunais Regionais e Tribunais de Justiça. A Lei nº 11.419/2006 permitiu a desmaterialização dos autos judiciais, a prática de atos processuais eletrônicos e a publicação oficial do Diário da Justiça Eletrônico. Trata-se de um grande avanço que contribui para a velocidade e segurança da prestação jurisdicional. O processo legislativo federal pode se beneficiar da experiência do processo judicial eletrônico.

O uso de um sistema informatizado para integrar toda a cadeia de valor adicionado do processo legislativo não é inédita no mundo. O projeto e-Law, da Áustria, em pleno funcionamento desde 2004, abrange todo o ciclo do processo legislativo: desde a iniciativa até a publicação da lei promulgada em um repositório digital autêntico. Essa iniciativa, a exemplo do LexML Brasil, também utiliza a tecnologia XML para estruturar o texto das normas.

Recomendação 4: Criação de uma base oficial de legislação disponível na internet apresentando a informação em múltiplos formatos.

Justificativa: A publicação do Diário Oficial da União no formato PDF com assinatura eletrônica pela Imprensa Nacional já representou um grande avanço no acesso à informação oficial, além de representar também uma grande economia de papel. No entanto, diferentemente do formato XML, o formato PDF dificulta o reaproveitamento da informação, pois foi projetado primordialmente para a apresentação da informação. A adoção de formatos textuais abertos baseados na tecnologia XML, como é o caso do LexML, pode facilitar a disseminação da informação jurídica oficial por múltiplos canais e formatos de expressão.

A exemplo do que ocorreu com o processo judicial eletrônico, faz-se necessária a criação de um marco normativo para permitir a modernização do processo legislativo federal e a instituição de uma base de dados de legislação com o caráter “oficial” que compreenda toda a produção normativa brasileira.

No caso da Áustria, foi necessária a promulgação de uma emenda constitucional para que a base de legislação do projeto austríaco e-Law se tornasse oficial. A publicação “*The e-Law Project in Austria*”, disponível em http://www.parlament.gv.at/ZUSD/PDF/2006-04-18_Publikation-Englisch.pdf, informa que a economia anual advinda da implantação do projeto no Parlamento superou a quantia de um milhão de Euros. Como a versão do periódico oficial parou de ser impressa, foram economizadas cerca de 60 toneladas de papel ao ano.

Recomendação 5: Criação de um modelo de requisitos para a gestão da informação jurídica.

Justificativa: As bases de dados de informação jurídica das instituições são construídas de forma independente. Inexiste um modelo que sirva de parâmetro para o desenvolvimento de sistemas informatizados que permita representar os textos das normas com todas as suas variantes e complementos. Além do texto original e do texto multivigente (texto compilado com todas as versões de um dispositivo), faz-se necessária a oferta também do texto vigente de uma norma para uma determinada data, pois o fato se regula pela norma vigente na data em que ele ocorreu. Apenas um sistema informatizado que aplique as modernas técnicas de manipulação textual conseguirá controlar a grande quantidade de textos derivados.

A criação de um modelo de requisitos para a gestão da informação jurídica permitirá uniformizar os esforços e a terminologia. O cidadão passará a ler notas de compilação mais uniformes de acordo com os requisitos de compilação.

Essa recomendação já começou a ser implementada pelo Comitê Gestor do Projeto LexML (<http://www.lexml.gov>), que constituiu um grupo de trabalho para conceber esse modelo de requisitos. O grupo de trabalho possui representantes dos seguintes órgãos: Senado Federal, Câmara dos Deputados, Advocacia-Geral da União, Câmara Legislativa do Distrito Federal, Receita Federal do Brasil e Superior Tribunal de Justiça. Os integrantes formam uma equipe multidisciplinar com larga experiência na gestão de informações jurídicas nas suas instituições.

Recomendação 6: Criação de uma etapa no processo legislativo de verificação da genuinidade do texto oficial da norma jurídica.

Justificativa: Como apresentamos anteriormente neste estudo de caso, um texto publicado no periódico oficial com erros é um texto autêntico, mas não genuíno. Faz-se necessária a definição de uma etapa no regimento interno das casas legislativas federais, estaduais e municipais para, ao final do processo legislativo, verificar se o texto publicado no periódico oficial está livre de erros.

Referências

ARNS, Paulo Evaristo. *A técnica do livro segundo São Jerônimo*. 2.ed. rev. e ampl. São Paulo: Cosac Naify, 2007.

BRASIL. Lei nº 4.320, de 17 de março de 1964. Estatui Normas Gerais de Direito Financeiro para elaboração e controle dos orçamentos e balanços da União, dos Estados, dos Municípios e do Distrito Federal. In: PUBLICAÇÃO consolidada da Lei 4.320, de 17 de março de 1964, determinada pelo art. 12 da Lei 9.528, de 10 de dezembro de 1997. Brasília: Secretaria de Orçamento Federal, [2011]. Disponível em: <https://www.portalsof.planejamento.gov.br/bib/legislacao/Lei_4320.pdf>. Acesso em: 15 abr. 2011.

KNIGHT, Shirlee-ann; BURN, Janice. Developing a Framework for Assessing Information Quality on the World Wide Web. *Informing Science: the International Journal of an Emerging Transdiscipline*, Santa Rosa, California, v. 8, p. 159-172, 2005. Disponível em: <<http://www.inform.nu/Articles/Vol8/v8p159-172Knig.pdf>>. Acesso em: 15 abr. 2011.

PASSOS, Edilenice. Requisitos recomendáveis para as bases de legislação estadual e distrital. *Senatus*, Brasília, v. 7, n. 2, p. 16-23, dez. 2009.

SPINA, Segismundo. *Introdução à Edótica: crítica textual*. São Paulo: Ars Poetica: Edusp, 1977.