

ANÁLISE EVOLUTIVA DA ESTRUTURA DE NOTAS FISCAIS ELETRÔNICAS: UM ESTUDO EM MICROEMPRESAS

Jeancarlo Campos Leão¹, Maria dos Anjos da Cruz dos Santos²

Resumo. A Nota Fiscal Eletrônica (NF-e) é um documento baseado em XML e de existência apenas digital, com o intuito de documentar, para fins fiscais, uma operação de circulação de mercadorias ou de prestação de serviços. Além do XML que representa a própria NF-e, existe ainda o XML que define um esquema de dados para cada versão. Com as alterações nas estruturas dos esquemas das NF-e, surge o questionamento sobre qual o impacto que estas alterações têm em seu uso. O objetivo deste trabalho é analisar a evolução estrutural da NF-e e de seus esquemas. Em prática, os esquemas não são totalmente aproveitados.

Palavras-chave: Nota Fiscal Eletrônica, Extração de Dados, Sumarização, Esquemas XML.

Introdução

XML foi projetada para descrever e armazenar os dados em uma única estrutura flexível. Esta descrição deve seguir a um esquema pré-definido que estabelece as possibilidades de conteúdo que podem aparecer no documento e a relação hierárquica entre eles. Neste esquema, pode ainda ser definido se um elemento é vazio ou pode incluir texto, quais os tipos de dados de seu conteúdo e os valores padrão contido, na prática a grande maioria dos documentos XML não possui esquema (MORO; VAGENA; TSOTRAS, 2008).

Uma das aplicações tanto do XML quanto dos seus esquemas (XML Schema Definition - XSD) são nas Notas Fiscais Eletrônicas (NF-e). A NF-e é um documento XML de existência apenas digital, emitido e armazenado eletronicamente, com o intuito de documentar, para fins fiscais, uma operação de circulação de mercadorias ou uma prestação de serviços, ocorrida entre as partes (FEDERAL, 2015).

Com as alterações nas estruturas dos esquemas das NF-e, surgem os questionamentos sobre a dimensão destas alterações e qual o impacto em seu uso? Em resposta a estes questionamentos, são apresentadas as hipóteses de que: (1) o modelo de conteúdo de NF-e apresenta aspectos fortes e peculiares a este domínio de aplicação; (2) Em prática, grande parte das notas fiscais não são desenvolvidas com total proveito dos esquemas propostos assim como dos elementos previstos por estes esquemas.

¹ Docente do IFNMG, Campus Araçuaí. Curso de Análise e Desenvolvimento de Sistemas. Email: jeancarlo.leao@ifnmg.edu.br

² Acadêmica do curso de Bacharelado em Administração do IFNMG, Campus Araçuaí. Email: mariacruzmaria@gmail.com

O objetivo deste trabalho é caracterizar e analisar a evolução estrutural da NF-e. Este objetivo é motivado pela larga utilização do formato XML em especial em Notas Fiscais Eletrônicas. Além disto, poucos estudos publicados com esta abordagem de uma aplicação tão específica de XML.

Trabalhos Relacionados

Mirella, Vagena e Tsostras (2008) definem os sumários estruturais e as apresentam possibilidades de recuperação de informação sobre eles em especial sobre arquivos esquemas XML (XSD). Grijzenhout e Systems (2010) propõem métodos de avaliação da qualidade de arquivos XML em geral. Laender, Moro et al (2009) fornecem uma visão geral sobre os arquivos XSD disponíveis na Web, identificando quais construções XSD são as mais e as menos utilizadas.

Material e Métodos

A metodologia se baseou na elaboração de processo e ferramentas para extração de dados sobre coleções de NF-e a partir da manipulação de seu conteúdo XML assim como dos seus esquemas XSD e a validação entre estes. Neste processo foram seguidas as etapas: obtenção de arquivos, validação, sumarização e visualização. Foram utilizadas NF-e de empresas voluntárias. Os esquemas (XSD), obtidos diretamente do repositório público da Receita Federal na Web (SRF, 2014) correspondentes a 4 versões: 1.10, 2.00, 3.00 e 3.10 do período de 2006 a 2015 totalizando 52 arquivos. No período de atuação com NF-e, as empresas utilizaram sistemas diversos para armazenamento e transporte dos arquivos XML: clientes de e-mail, sistema ERP da empresa, sistema de gestão eletrônica de documentos (GED). Ao final desta etapa, 4611 arquivos foram obtidos. Após a correção, deduplicação de arquivos e validação das NF-e com o XSD totalizaram 2651 arquivos de NF-e selecionados para os experimentos.

Resultados e Discussão

Na validação e sumarização dos arquivos XML, foram reportadas 804 NF-es não validadas por nenhum dos XSDs. Esta validação foi feita entre todas os esquemas disponíveis na coleção (52) e cada instância de NF-e (1716) totalizando 89232 processos de validação dentro dos quais houveram 912 positivas (NF-e válidas por seu XSD de layout de compartilhamento) e 88.320 exceções. Através dos logs gerados no processo de validação, foram obtidos os totais da tabela 01.

Versão NF-e	Elementos Fixos (Sempre presentes)	Elementos Distintos	Elementos Novos	Elementos Removidos
1.10	105	159	0	0
2.00	97	212	54	1
3.10	101	204	11	19

Tabela 05 – Sumário geral das NF-es.

Na tabela 01 a estrutura das NF-e é considerada individualmente para cada versão e de forma comparativa. Foram verificados elementos fixos, ou seja, presentes em toda nota fiscal; distintos, que representa a diversidade de elementos que podem ocorrer em uma determinada versão; novos, que foram concebidos naquela versão e; elementos depreciados, os quais foram utilizados apenas em versões anteriores. Dos 159 elementos propostos na versão 1.10 da NF-e, apenas 93 permaneceram até a versão 3.10. Neste intervalo de mudanças na estrutura da NF-e, 224 elementos distintos foram propostos, dentre os quais apenas 204 foram mantidos na versão final. Apenas o XSD de processamento de leiaute de compartilhamento de versão correspondente à NF-e conseguiu validá-las. Os outros 49 arquivos XSD geraram ao menos uma exceção.

Conclusões

O objetivo deste trabalho foi analisar a evolução estrutural da NF-e e conflitar as propostas de esquemas com notas fiscais eletrônicas reais. A medida em que seus esquemas são alterados, o leiaute dos XML das NF-e também devem ser. Considerando apenas os XML de NF-e bem formados, foram identificadas falhas na estruturação com base na validação com seus esquemas exceto no leiaute de compartilhamento de sua versão correspondente. Uma restrição do esquema que prevê isto é a especificação de um valor esperado para a versão do XML da NF-e. Estes resultados favorecem trabalhos futuros para a mineração de texto/dados com base em aspectos específicos das NF-e. Com uma amostragem mais rica em perfis diversos de empresas é possível ainda obter uma maior precisão sobre os resultados obtidos visto que os elementos de cada versão da NF-e podem ser subutilizados por alguns sistemas de automação.

Referências

FEDERAL, R. Manual de Integração do Contribuinte. Disponível em: <<http://www.nfe.fazenda.gov.br/>> Acesso em: 1 jul. 2015.

GRIJZENHOUT, S.; SYSTEMS, B. I. Quality of the Xml Web Draft. p. 1–78, 2010.

LAENDER, A. H. F. et al. An X-Ray on Web-Available XML Schemas. v. 38, n. 1, p. 37–42, 2009.

MORO, M.; VAGENA, Z.; TSOTRAS, V. XML Structural Summaries. Proceedings of the VLDB Endowment, v. 1, p. 1524–1525, 2008.

SRF., S. D. R. F. –. NF-e Nota Fiscal Eletrônica: Esquemas XML. Disponível em: <<http://www.nfe.fazenda.gov.br/>>. Acesso em: 20 maio. 2015.

Agradecimentos

Ao IFNMG - Instituto Federal do Norte de Minas Gerais pela condição de bolsista do servidor Jeancarlo Campos Leão através do PBQS - Programa de Bolsas para Qualificação de Servidores. Agradecemos também à empresa que foi receptiva a nós pesquisadores e à ciência.