

Extrator de dados de Notas Fiscais Eletrônicas

Higor Moreira¹, Ronaldo Serpa da Rosa¹, Edimar Manica^{1*}

¹Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS) - *Campus* Ibirubá. Ibirubá, RS, Brasil.

Uma Cooperativa de Ibirubá verificou a possibilidade do IFRS – *Campus* Ibirubá desenvolver uma solução tecnológica para minimizar o esforço manual necessário para a importação das Notas Fiscais Eletrônicas de fornecedores para seu sistema interno. Esse esforço é necessário porque o código e a descrição do produto no cadastro dos fornecedores são diferentes das especificações de produtos presentes no software utilizado pela empresa. Dentre as diferenças encontram-se variação na representação do conteúdo, omissão de determinadas informações, uso de acrônimos e abreviações. Nesse contexto, está sendo desenvolvido um projeto de pesquisa que tem como objetivo identificar as funções de similaridade textuais mais eficazes para encontrar especificações de produtos oriundas de sistemas diferentes que se referem ao mesmo produto. Para realizar essa identificação é necessário avaliar as funções de similaridade no contexto da empresa. Dessa forma, é fundamental que a empresa forneça exemplos de notas fiscais reais. No entanto, como essas informações são confidenciais para a empresa, uma vez que inclui informações sobre preços e tributação, foi necessário desenvolver um software que extraísse das notas fiscais apenas as informações requeridas para o projeto. Este resumo descreve o desenvolvimento deste extrator. O desenvolvimento do software foi realizado em quatro etapas: levantamento de requisitos, estudos das tecnologias, implementação e testes. O levantamento de requisitos foi realizado através da análise de documentos e de uma entrevista com a equipe contábil da empresa. O estudo das tecnologias consistiu em estudar os padrões dos arquivos XML e a biblioteca Java W3C.DOM. A implementação foi efetuada através da linguagem de programação Java, utilizando o ambiente de desenvolvimento NetBeans IDE 8.4 e a biblioteca Java W3C.DOM. Os testes foram realizados em duas fases. Na primeira fase, foram testadas vinte e uma notas fiscais eletrônicas de dez empresas diferentes obtidas de pessoas que realizaram compras pela internet. O software obteve 100% de precisão uma vez que conseguiu extrair os dados desejados em todas as notas. Na segunda fase, o software foi disponibilizado para a empresa que o testou com oitocentas notas fiscais reais. Nessa fase, o software também obteve 100% de precisão. Com isso, conclui-se que o extrator atende ao objetivo proposto.

Palavras-chave: Nota fiscal eletrônica. XML. Extrator de dados.

Trabalho executado com recursos do EDITAL IFRS Nº 25/2019 – PROBIC/PROBITI/IFRS/FAPERGS da Pró-Reitoria de Pesquisa