

UNIVERSIDADE SAGRADO CORAÇÃO

CAIO AUGUSTINHO DO NASCIMENTO

**APLICAÇÃO DAS TÉCNICAS XPATH E XQUERY
PARA A ORGANIZAÇÃO DOS ARQUIVOS XML
DA NOTA FISCAL ELETRÔNICA**

BAURU
2014

CAIO AUGUSTINHO DO NASCIMENTO

**APLICAÇÃO DAS TÉCNICAS XPATH E XQUERY
PARA A ORGANIZAÇÃO DOS ARQUIVOS XML
DA NOTA FISCAL ELETRÔNICA**

Trabalho de Conclusão de Curso apresentado ao Centro de Ciências Exatas e Sociais Aplicadas como parte dos requisitos para obtenção do título de bacharel em Ciência da Computação, sob orientação do Prof. Me. Patrick Pedreira Silva.

BAURU
2014

Nascimento, Caio Augustinho do.

N244a

Aplicação das técnicas XPath e XQuery para a organização dos arquivos XML da nota fiscal eletrônica / Caio Augustinho do Nascimento. -- 2014.

67f. : il.

Orientador: Prof. Me. Patrick Pedreira Silva.

Trabalho de Conclusão de Curso (Graduação em Ciência da Computação) – Universidade do Sagrado Coração – Bauru – SP.

1. XPath. 2. XQuery. 3. XML. 4. NFe. I. Silva, Patrick Pedreira. II. Título.

CAIO AUGUSTINHO DO NASCIMENTO

**APLICAÇÃO DAS TÉCNICAS XPATH E XQUERY PARA A ORGANIZAÇÃO
DOS ARQUIVOS XML DA NOTA FISCAL ELETRÔNICA**

Trabalho de Conclusão de Curso apresentado ao Centro de Ciências Exatas e Sociais Aplicadas da Universidade Sagrado Coração como parte dos requisitos para obtenção do título de bacharel em Ciência da Computação sob orientação do Prof. Me. Patrick Pedreira Silva.

Banca examinadora:

Me. Patrick Pedreira Silva
Universidade Sagrado Coração

Me. Henrique Pachioni Martins
Universidade Sagrado Coração

Dr. Elvio Gilberto da Silva
Universidade Sagrado Coração

Bauru, 02 de Dezembro de 2014.

Dedico este trabalho aos meus familiares,
amigos e a todos os envolvidos.

AGRADECIMENTOS

A Deus, pelo dom da vida e aos meus pais que sempre ensinou o caminho correto me dando total segurança, carinho, responsabilidade e muito amor.

Ao meu orientador Patrick Pedreira Silva, por me aceitar como orientando, com total atenção e incentivo para realização desse trabalho.

E a todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

RESUMO

Essa pesquisa teve a contribuição e o estudo do desenvolvimento de um software com uma base de dados que permite que empresas que manipulem notas fiscais eletrônicas possam, automaticamente, incluir, editar e excluir dados rapidamente, mostrando também a utilização das técnicas XPath e XQuery no processo de recebimento e organização de arquivos XML de Nota Fiscal Eletrônica. Utilizando o software desenvolvido na linguagem Delphi para realizar testes como comparar as duas técnicas (XPath e a XQuery), ressaltando aspectos relacionados sobretudo à velocidade de processamento. Os resultados obtidos evidenciam que a técnica XPath é a mais rápida, porém, talvez, menos eficaz em comparação com a XQuery, pois esta última permite a posterior manipulação das informações sobre os arquivos XML contidas no banco de dados.

Palavras-chave: XPath. XQuery. XML. NFe.

ABSTRACT

This research was contributed and the study of the development of software with a database that allows companies to manipulate electronic invoices can automatically add, edit and delete data quickly, also showing the use of XPath and XQuery techniques and process of receiving and organizing XML file Electronic Invoice. Using software developed in Delphi language to perform tests like comparing the two techniques (XPath and XQuery), highlighting aspects in particular the processing speed. The results show that the XPath technique is faster, but perhaps less efficient compared to XQuery, since the latter allows further handling of the information contained in the XML files database.

Keywords: XPath, XQuery, XML, NFe.

LISTA DE FIGURAS E ILUSTRAÇÕES

Figura 1 – Técnica XPath.....	41
Figura 2 – Técnica XQuery.....	25
Figura 3 – Modelo do Danf-e.....	31
Figura 4 – Leitor do certificado digital do Serasa	32
Figura 5 – Tela principal do Delphi.....	38
Figura 6 – Tela principal IBexpert.....	40
Figura 7 – Diagrama Jude.....	41
Figura 8 – Diagrama Classe.....	42
Figura 9 – Tela inicial do software.....	43
Figura 10 – Menu arquivos.....	44
Figura 11 – Configurar parametros do E-mail	45
Figura 12 – Tela de logs do sistema	46
Figura 13 – Tela de verificação pasta e servidor.....	47
Figura 14 – Tela de download E-mail.....	48
Figura 15 – Tela do icone analise XPath.....	49
Figura 16 – Tela do XPath.....	50
Figura 17 – Tela do icone analise XQuery	51
Figura 18 – Tela do XQuery	52
Figura 19 – Cronometro analise manual	54
Figura 20 – Tela de teste com XPath	55
Figura 21 – Tela de teste com XPath 2	56
Figura 22 – Tela de teste com XPath 3	56
Figura 23 – Tela de teste com XPath 4	57
Figura 24 – Tela de teste com XQuery.....	58
Figura 25 – Tela de teste com XQuery 2.....	59
Figura 26 – Tela de teste com XQuery 3.....	60
Figura 27 – Tela de teste com XQuery 4.....	61
Figura 28 – Tela de conclusão dos testes.....	61

LISTA DE ABREVIATURAS E SIGLAS

API	Application Programming Interface
APIs	Application Programming Interface
CPU	Graphics Processor Unit
DTD	Document Type Definition
DNS	Domain Name Service
ENAT	Encontro Nacional de Administradores Tributários
ECP	Extended Versatile Disk
FTP	File Transfer Protocol
GML	Game Maker Language
HTML	HiperText Markup Language
HTTP	HyperText Transfer Protocol
HTTPS	HyperText Transfer Protocol Secure
IBM	International Business Machines
MathML	Mathematical Markup Language
NCL	Nested Context Language
NFE	Nota Fiscal Eletrônica
POP3	Post Office Protocol
RDF	Resource Description Framework
RSS	Rich Site Summary
SDMX	Statistical Data and Metadata exchange
SGML	Standard Generalized Markup Language
SMIL	Synchronized Multimedia Integration Language
SPED	Sistema Público de Escrituração Digital
SVG	Scalable Vector Graphics
URL	Uniform Resource Locator
W3C	World Wide Web Consortium
WEB	World Wide Web
WAP	Wireless Application Protocol
WWW	World Wide Web
XBRL	extensible Business Reporting Language
XHTML	extensible Hypertext Markup Language

XML	extensible Markup Language
XSIL	Extensible Scientific Interchange Language
XSL	Extensible Stylesheet Language

SUMÁRIO

1 INTRODUÇÃO	13
2 ORGANIZAÇÃO DO TRABALHO	16
3 OBJETIVOS.....	17
3.1 Objetivo geral.....	17
3.2 Objetivos específicos	17
4 REVISÃO DA LITERATURA	18
4.1 Conceitos da XML.....	18
4.2 Breve Histórico sobre XML.....	18
4.3 Evolução da XML	19
4.4 Características da XML.....	20
4.4.1 Sintaxe da XML.....	21
4.4.2 Regras da XML.....	21
4.4.3 Estrutura da XML	22
4.4.4 Estilo da Linguagem XML	22
4.5 XML no dia a dia	22
4.6 Técnica XPath.....	23
4.7 Técnica XQuery	24
4.8 Definição da NF-e	25
4.9 Breve histórico sobre NF-e.....	25
4.9.1 Objetivo da NF-e.....	26
4.9.2 Organização da NF-e.....	28
4.9.3 Benefícios da NF-e	28
4.10 Validação da NF-e pela Sefaz.....	29
4.11 Problemas com a validação da NF-e	29
4.12 Consulta da NF-e.....	29
4.13 Cancelamento da NF-e	30
4.14 Danf-e	30
4.15 Certificado Digital	31

5 ENGENHARIA DE SOFTWARE.....	33
6 METODOLOGIA	35
6.1 Desenvolvimento do software.....	37
6.1.1 Delphi.....	37
6.1.2 Bando de dados.....	38
6.1.3 IBExpert.....	39
6.1.4 Jude Community.....	40
6.1.5 Modelagem.....	41
6.1.6 Interface.....	42
7 RESULTADOS.....	53
8 CONSIDERAÇÕES FINAIS.....	64
REFERÊNCIAS.....	65

1 INTRODUÇÃO

De acordo com Moraes (2003), o surgimento da WEB teve sua estrutura pautada na construção de documentos de textos e imagens, puramente informativos, sem interação imediata com os usuários. Eram informações em texto, som e imagem estáticos, apenas criados e disponibilizados para acesso. A falta de estruturação desses documentos pode, na maior parte das vezes, dificultar o acesso às suas informações, principalmente por meio de ferramentas computacionais. Diante deste fato, alternativas, surgiram, dentre as quais destaca a linguagem XML.

XML é uma recomendação para publicação, combinação e intercâmbio de documentos multimídia, desenvolvido pelo consórcio W3C. A tecnologia XML permite a criação de aplicações portáteis, distribuídas e completamente independentes de plataforma. Esta tecnologia se tornou ideal para o desenvolvimento de aplicações que exploram o consumo e a geração de informações por diversas classes de clientes. (PEREIRA, 2009).

Pereira (2009) afirma que a linguagem XML vem sendo usada cada dia mais pelas indústrias nos mais diversos segmentos como, no setor de vendas, contábil e informática. Por esses e outros motivos, XML é a tecnologia de maior sucesso para a Internet.

Antigamente o processo de venda era algo simples, onde o fornecedor emitia a Nota Fiscal, encaminhava a via do cliente, e uma cópia para o contador proceder aos registros contábeis e fiscais. Após esse procedimento, o cliente apenas guardava uma via da Nota Fiscal, e isso era o suficiente para finalizar o processo. Com a implantação do sistema de NF-e (“Nota Fiscal Eletrônica”), novos itens surgiram, porém, sem a devida divulgação de todas as etapas a serem seguidas.

Com o avanço da nota fiscal eletrônica (NF-e), ficou muito mais rápido, ágil e prático para as empresas disponibilizarem os arquivos de Nota Fiscal eletrônica (NF-e) para seus clientes através do site da empresa ou por e-mail.

Com base em estudos, muitas empresas ainda não estão familiarizadas com as novas normas implantadas para a Emissão das NF-e (“Notas Fiscais Eletrônicas”). A confusão vai desde a emissão até o recebimento dos documentos, e o arquivamento dos arquivos XML.

O XML é o padrão de documento utilizado pelo órgão responsável pela implantação do sistema de NF-e do Governo. O layout é padronizado em todo o território nacional, respeitando os conceitos antigos de escrituração fiscal e adicionando melhorias visando maior controle de tais transações. (ANDERSON, 2001).

O que muitas empresas ainda não entenderam é que a XML é baseado na própria Nota fiscal eletrônica e a leitura desse arquivo deve ser feita diretamente no site da secretaria da fazenda ou com programas específicos.

Entretanto, mesmo com todo esse avanço tecnológico, muitas empresas ainda encontram dificuldades para gerenciar esses arquivos, existindo assim a necessidade de automatizar o processo de organização desses documentos para não ser autuado pelo fiscal da receita e nem pelas auditorias locais.

O XML fornece a descrição, o armazenamento, o formato da transmissão, e todo conteúdo de uma nota fiscal eletrônica. A pesquisa de uma NF-e é bem ampla, podendo ser realizada uma pesquisa pela data de recebimento, número do CNPJ, número da nota fiscal, nome do fornecedor, endereço, cidade e etc. A pesquisa é sempre solicitada pelo fiscal da receita, e quase sempre a solicitação é por data do recebimento. (PORTAL NFE, 2011).

Dado o exposto, a explosão da utilização de XML em trabalhos de pesquisa para a Web, junto com a tecnologia móvel para acesso universal à informação motivou o desenvolvimento deste trabalho, visando melhorar a praticidade em empresas, auxiliando com total tecnologia e aprimoramento da linguagem XML.

O XML é uma linguagem de marcação, o que significa que os dados são sempre marcados por (tags¹).

A flexibilidade da especificação dos dados permite que XML seja utilizada no desenvolvimento de aplicações em diversos contextos. XML é consistentemente utilizada em aplicações de serviços Web (Web services) como o formato padrão de feeds RSS. (DEITEL, 2001).

¹ Estruturas de linguagem de marcação que consistem em breves instruções, tendo uma marca de início e outra de fim.

Uma estrutura de dados comum em XML pode melhorar a integração entre participantes de transações de comércio eletrônico, permitindo a troca de mensagens sobre compra e venda de produtos eficientemente.

Anderson (2001) diz que encapsulamento de dados XML é propício para o processamento de características avançadas (tais como privacidade) quando se acessa dados através da Web ou via aplicações desktop.

A XML oferece uma abordagem padrão para descrição, captura, processamento e publicação de informações. A linguagem XML é reconhecida como uma linguagem padrão para representação de informações tanto no meio acadêmico quanto no comercial, e tornou-se uma tecnologia bem-sucedida e onipresente para a troca de dados Web.

Para a manipulação de documentos XML é necessário a existência de um analisador (parser) que suporte todas as funcionalidades necessárias para percorrer as estruturas destes documentos, permitindo o acesso aos seus elementos e atributos. (MORAES, 2003).

Segundo Pereira (2009) basicamente é possível estabelecer regras que permitam a transformação de um documento XML em qualquer outro documento XML. Processadores XML são as principais ferramentas para a manipulação de documentos XML. Na verdade, estes processadores são APIs utilizadas para facilitar a leitura, criação e atualização de documentos no formato XML. A partir daí, métodos de busca eficientes em documentos podem ser diretamente aplicados em dados estruturados com XML nas técnicas XPath e XQuery.

A técnica XPath é um conjunto de regras de sintaxe que permite definir em partes documento XML, usa uma expressão de caminho para poder localizar nós chamados de tags em documentos XML. (W3C, 2010).

Já a XQuery é uma técnica criada a partir da XPath e é ideal para realizar consultas em vários tipos de documentos XML, possui uma grande estrutura que não é muito diferente da SQL, mesmo assim ela consegue gerenciar árvores na linguagem XML. (W3C, 2010).

2 ORGANIZAÇÃO DO TRABALHO

O Capítulo 1 apresenta a introdução do trabalho.

O Capítulo 2 apresenta a organização deste trabalho.

Já no Capítulo 3 os objetivos gerais e específicos são representados.

No Capítulo 4 apresenta-se a revisão da literatura.

Já o Capítulo 5 contém a engenharia de software, tendo como princípio exibir suas regras e características.

O Capítulo 6 contém a metodologia explicando o desenvolvimento do objetivo proposto.

No Capítulo 7 estão descritos a análise e os resultados obtidos na realização deste trabalho.

Por fim, o Capítulo 8 contém o fechamento deste estudo através das considerações finais.

3 OBJETIVOS

3.1 Objetivo geral

Desenvolver um software aplicando as técnicas XPath e XQuery para organizar o recebimento das notas fiscais eletrônicas em uma empresa, tendo por base arquivos XML visando economizar o tempo de seus colaboradores e evitar falhas humanas.

3.2 Objetivos específicos

- Proporcionar conhecimentos sobre as linguagem XML e as técnicas XPath e XQuery;
- Analisar os processos e aplicações da linguagem XML, detalhando o seu propósito principal e obter seus pontos positivos e negativos na organização das notas fiscais eletrônicas;
- Fazer o levantamento em livros, Internet e periódicos sobre as falhas de organização das notas fiscais eletrônicas sem a utilização da tecnologia;
- Criar uma base de dados para armazenamento dos documentos;
- Modelar a arquitetura do software proposto;
- Desenvolver um software que utiliza as técnicas XPath e XQuery para auxiliar na organização dos arquivos XML.
- Por fim realizar testes no software, a fim de aferir a sua eficiência para a tarefa proposta.

4 REVISÃO DA LITERATURA

4.1 Conceitos da XML

O XML foi desenvolvido pelo XML Working Group, que é conhecido como SGML com isso teve um grande patrocínio do www no ano de 1996, foi também presidido por Jon Bosak, proprietário da Sun Microsystems. Na data de fevereiro em 1998 conseguiu uma tecnologia recomendada por todas as pessoas, é sempre um conjunto de muitas regras para a definição semântica, que organiza em algumas partes identificáveis. É uma linguagem que define várias sintaxes para serem colocadas na criação de outras linguagens de marcação para um ou mais domínio, com uma estrutura semântica própria. (HAROLD, 1999).

4.2 Breve Histórico sobre XML

Quando foi criada, pensou-se que o propósito da XML seria de substituição da HTML. A história da XML foi iniciada nos anos 70, quando a IBM decidiu inventar a linguagem GML (General Markup Language – linguagem de marcação genérica). Assim, surgiu com a necessidade da empresa em armazenar uma grande quantidade de informações. A IBM e a GML conseguiriam classificar e processar rapidamente todos os documentos. A organização internacional para padronização teve em 1986 que trabalhar para padronizar esta linguagem. Neste mesmo momento, foi desenvolvida a SGML, uma Linguagem Padrão de Marcações Genéricas. (BERGLUND, 2004).

Em 1989, foi criada a HTML, uma linguagem de marcação que serve para criação de páginas de sites. A HTML foi derivada da SGML que utiliza-se um conjunto de tags, entre os símbolos < >, informando o que irá determinar a função de cada elemento. Em 1994, surge uma entidade chamada W3C (World Wide Web Consortium - Consórcio da WWW. E com isso decidiu colocar em ordem o HTML, e de formalizar suas regras, para que fosse um padrão. Mesmo assim, o HTML não cumpriu tudo o que propunha e planejava para a Internet,

crescendo de uma maneira descontrolada e desordenada. (MACORATTI, 2014).

A HTML está em constante evolução, porém, ela não consegue suprir tudo que as aplicações necessitam. Surge, em 1996, a XML que foi proposta com o objetivo de suprir as limitações da HTML, complementando-a. Ao ser publicada como uma recomendação do W3C (World Wide Web Consortium), em 1998, depois revisada em 2000, a XML rapidamente incorporou ferramentas que a colocam em um nível muito acima da simples publicação de documentos na Internet. (DEITEL, 2001).

4.3 Evolução da XML

Mesmo com sua criação em 1996, a XML vem crescendo gradativamente se tornando uma linguagem padrão de representação e intercambio de dados por ser utilizado em meios eletrônicos. Algumas pessoas acham que em um breve futuro essa linguagem substituirá o HTML que é a linguagem mais utilizada na Internet. (KAMANNA, 2003).

Com muito dinamismo, as informações contidas na Internet, já amplamente estudadas, ganham nova importância: a necessidade da informação, em outras palavras, torna-se obrigatório ter mecanismos para a representação e manipulação de um documento que sofre muitas alterações durante os anos.

Pereira (2009) destaca que hoje em dia existem muitas variedades de proposta que abordam essas questões, mais poucas que envolvem a XML, com a quantidade de plataforma existente no mundo da internet, precisa-se de aplicações nas quais tratam a informação com modelos de dados tradicionais e estruturados. Um exemplo é a edição de documentos, onde temos diversas versões e muita necessidade de alterar o conteúdo.

Segundo Castro (2002) existem muitas aplicações com o propósito de deixar mais acessível o histórico de alterações de documentos, sendo obrigados a realizar alterações às quais o documento é organizado, deixando mais simplificado e livre para total acesso do público.

O jeito com que os documentos XML são administrados está relacionado à organização efetiva e a representação do acesso a dados em formato de que as bases estão relacionadas. (SCHEMA, 2002).

4.4 Características da XML

Segundo Macoratti (2014), XML (Extensible Markup Language) é uma das linguagens que permite construir documentos legíveis para seres humanos, e que pode ser facilmente interpretada pelas máquinas, ou seja, a XML possui muitos marcadores que são mais fáceis e flexíveis do que em outras linguagens.

Uma pessoa um pode iniciar um novo documento XML com qualquer ferramenta de texto normal, podendo descrever todos os dados de um texto e também podendo mostrar qualquer tipo de dados. Além do mais, links podem conter muitas informações validas sobre os recursos referentes à ele, ou seja, devem desempenhar uma busca de informação na internet, e pode estar sendo cada dia mais usada, seja para construir os arquivos de configuração ou seja para os dados entre aplicações na internet. Além disso, existem várias outras configurações para a linguagem XML, sempre com o ideal de trocar os dados. (BERGLUND, 2004).

Afirmando que reconhecendo todos os fatos, os produtores de sistemas estão implementando a XML dentro de suas aplicações para poder ganhar em desempenho e funcionalidade na internet. Além disso, está cada dia sendo cada vez mais utilizada na construção de algumas bases de dados. Alguns estudiosos acreditam que no em um breve futuro a grande expansão da internet, a XML se tornará uma grande linguagem única para representar os dados. Logo, todas as aplicações poderão ser capazes de se interligar, uma vez que elas deverão entender os vocabulários e/ou marcações de outros documentos produzidos por outras aplicações. (DEITEL, 2001).

O XML é sempre capaz de descrever alguns tipos de dados e com facilidade de compartilhamento de informações através da Internet está entre linguagens baseadas em XHTML, RDF, SDMX , SMIL, MathML, NCL, XBRL, XSIL e SVG. (PEREIRA, 2009).

Com a insatisfação dos formatos existentes, o World Wide Web Consortium (W3C) começou a desenvolver entre 1985 e 1995 em uma linguagem que combinasse a flexibilidade da SGML com a simplicidade da HTML. O início do projeto era criar uma linguagem que pudesse ser lida por softwares e implantar-se com as outras linguagens, considerado o melhor formato para a criação de documentos com dados organizados. (W3, 2009).

Os princípios importantes da XML são a criação de arquivos para validação de estrutura (Chamados DTDs), interligação de bancos de dados, separação do conteúdo da formatação, possibilidade de criação de tags sem limitação, concentração na estrutura da informação, e não na sua aparência, simplicidade e Legibilidade, tanto para humanos quanto para computadores. (PEREIRA, 2009).

4.4.1 Sintaxe da XML

Segundo Hendler (2001), um documento XML pode ter muitas marcações para melhor representar uma estrutura de informações. Essa estrutura é onde ficam as informações do “teste”. Sempre tem uma maneira hierárquica de estruturar um documento, XML é semelhante a uma árvore, sendo os documentos que possuem elementos únicos de raiz e que contenha outros elementos chamados de filhos.

4.4.2 Regras da XML

Os documentos XML são totalmente formados o que significa que se devem seguir regras como, iniciar com uma letra ou caractere sublinhado “_”, podem ter qualquer tamanho, os nomes dos atributos seguem a mesma regra para nomes de elementos, valores dos atributos devem estar entre aspas, etc.

Deitel (2001) afirma que todos os documentos XML podem ser visualizados como árvores podendo possuir um elemento raiz, transformando as tabulações em espaço em branco, o que pode ser removido ou alterado.

4.4.3 Estrutura da XML

Um documento XML pode inserir outros documentos XML, esse processo é chamado de esquema ou validador, lembrando que o documento só tem esse nome se ele for da gramática ou esquema válido. (MORAES, 2003).

A XML deve ser sempre feito em blocos de construção, que nada mais é do que unidades de texto que pode ser de vários caracteres ou ter nível maior de complexidade.

4.4.4 Estilo da Linguagem XML

Diferentemente da linguagem HTML as informações de características da XML são formalizadas, impressos ou visualizados por navegadores com uma exibição na folha de estilo, isso mostra o conteúdo de um documento sendo um conjunto de ações que podem ser realizadas em partes identificadas, com isso é possível identificar facilmente as informações dentro de um arquivo XML. (PEREIRA, 2009).

4.5 XML no dia a dia

A XML é muito utilizado hoje em dia, pois sua versatilidade de oferecer uma estrutura de texto simples por meio de tags e de fácil localização dos dados, com suporte a várias arquiteturas permitindo mais informações codificada em linguagem humana que já mais foi oferecida por outra linguagem.

Alguns dos propósitos da XML são: auxiliar os sistemas de informação no compartilhamento de dados (especialmente via internet), codificar documentos e inserir seriais nos dados comparando o texto com o de outras linguagens baseadas em serialização. (Pereira 2009 p.68).

A linguagem XML está sendo muito utilizada nas empresas e no dia a dia das pessoas, uma grande aplicação da XML é na NF-e, pois muitas empresas utilizam arquivos XML para controlar seus meios de entrada e saída de produtos.

A maior parte dos editores txt modernos oferece recursos para destaque de XML.

4.6 Técnica XPath

No entendimento de Macoratti (2010), a XPath permite o endereçamento de subconjuntos em um único documento, portanto, permite estabelecer um caminho utilizado para percorrer um documento, sintaxe de XPath é muito parecida com a usada em sistema de arquivos, ou outras linguagens hierárquicas.

Segundo o W3C (2010), a técnica XPath permite construir expressões que recorrem e processam um documento XML de modo parecido a uma expressão regular. Pode também computar valores como números, valores Booleanos e cadeias de caracteres. O XPath foi definido no World Wide Web Consortium (W3C), e funciona de maneira parecida como se fosse uma árvore de diretórios. O resultado de cada expressão XPath é um conjunto de elementos especificados pelo caminho e tem como base os seguintes termos:

- É uma sintaxe para definir partes de um documento XML.
- Usa caminhos para definir elementos XML.
- Define uma biblioteca de funções padrão.

O XPath foi liberado com uma recomendação W3C em 16 de novembro de 1999 como uma técnica para endereçar partes de um documento XML, foi desenvolvido para ser usado por XSLT, XPointer e outros softwares XML parsing. (MICORATTI, 2014).

XPath, é parte importante da família de tecnologias XML, já que ela provê a possibilidade de selecionar e filtrar dados contidos em um ou mais documentos XML. Quando trabalhamos com XML em geral, vários são os casos de uso que surgem aos quais podemos aplicar expressões XPath para realizar pesquisas sobre esses documentos. A sintaxe adotada pelo XPath é bastante intuitiva, já que ela baseia-se em caminhos, como os caminhos de um sistema de arquivos. (Vasquez 2014 p.232).

A Figura 1 mostra a estrutura de código da técnica XPath e como ela é associada às informações do software.

Figura 1 – Técnica XPath.

```
XML document:
<?xml version="1.0" encoding="utf-8"?>
<wikimedia>
  <projects>
    <project name="Wikipedia" launch="2001-01-05">
      <editions>
        <edition language="English">en.wikipedia.org</edition>
        <edition language="German">de.wikipedia.org</edition>
        <edition language="French">fr.wikipedia.org</edition>
        <edition language="Polish">pl.wikipedia.org</edition>
      </editions>
    </project>
    <project name="Wiktionary" launch="2002-12-12">
      <editions>
        <edition language="English">en.wiktionary.org</edition>
        <edition language="French">fr.wiktionary.org</edition>
        <edition language="Vietnamese">vi.wiktionary.org</edition>
        <edition language="Trukish">tr.wiktionary.org</edition>
      </editions>
    </project>
  </projects>
</wikimedia>
```

Fonte: XPath Macoratti (2014).

4.7 Técnica XQuery

XQuery é uma técnica só de leitura (read-only), e foi projetada para fazer consultas em coleções de dados em XML, pode retornar uma instância do modelo de dados XQuery, mas não pode modificar uma instância existente.

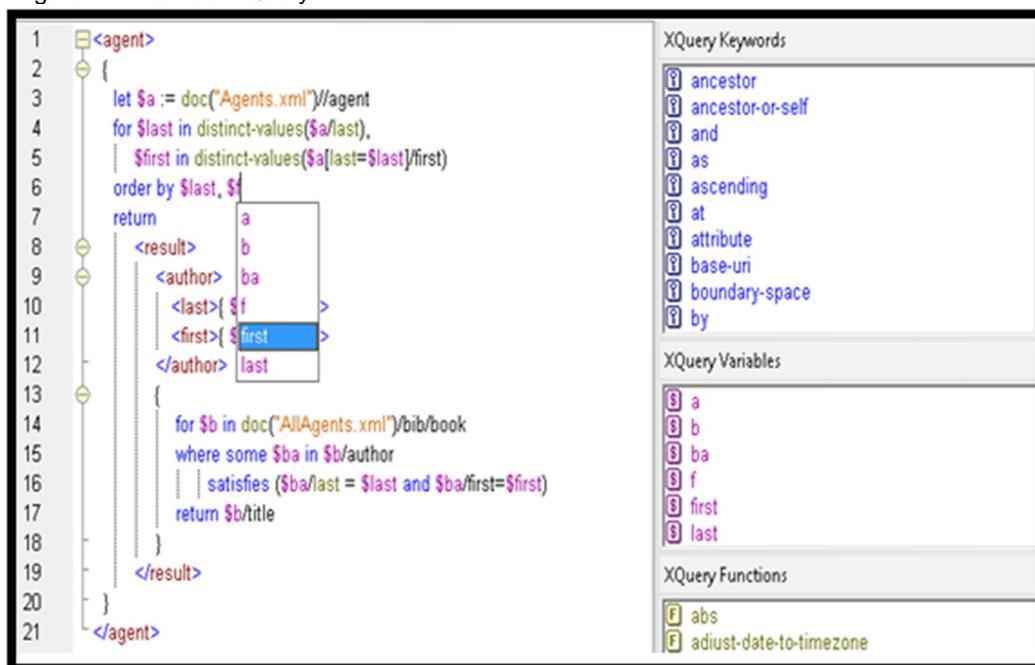
Todas as expressões XQuery podem ser classificadas em, expressões que efetuam atualização ou expressões que não efetuam atualização. É possível proporcionar uma ótima flexibilidade para exportar os dados reais e virtuais de documentos sobre a World Wide Web, pois proporciona a interação entre o mundo Web e Banco de Dados conforme ilustra a Figura 2. (DEBUGGER, 2010).

Para consultar uma instância XML armazenada em uma variável ou coluna de tipo xml, use os Métodos de tipo de dados xml. Por exemplo, é

possível declarar uma variável de tipo xml e consultá-la usando o método query() do tipo de dados xml. (Vasquez 2014 p.98).

A Figura 2 mostra um editor da estrutura de código da técnica XQuery e como ela é embutida nas informações do software, observa-se que a sua estrutura é parecida com as outras técnicas.

Figura 2 – Técnica XQuery.



Fonte: Xquery Debugger (2014).

4.8 Definição da NF-e

A Nota Fiscal Eletrônica (NF-e) é um documento que existe apenas em forma digital, obrigatório na compra e venda de qualquer tipo de mercadoria, após a sua emissão o seu armazenamento, é recomendável, para que fiscais ou auditores possam ter controle de uma operação de circulação de mercadorias ou de uma prestação de serviços.

4.9 Breve histórico sobre NF-e

Segundo o PORTAL NFE (2014), a nota fiscal eletrônica teve sua criação através de uma parceria com o ENCAT e a Receita Federal do Brasil. É um projeto inicial de três outros projetos que se interligam através de um

programa do governo federal chamado SPED (Sistema Público de Escrituração Digital da Receita Federal). A primeira fase do projeto principal da Nota Fiscal Eletrônica teve seu início em meados do ano de 2005, Tendo as secretarias da fazenda em várias grandes empresas que hoje recebem a NF-e referente aos seguintes estados: Goiás, Rio Grande do Sul, São Paulo, Bahia, Maranhão e Santa Catarina.

Com o objetivo de atuar e atender o disposto Constitucional; foi realizado, em julho de 2004 na cidade de Salvador, o I ENAT - Encontro Nacional de Administradores Tributários e com isso reunindo o Secretário da Receita Federal, além disso, teve um encontro que buscou algumas soluções conjuntas nas três partes de Governo que promovessem maior integração administrativa, padronização com melhor qualidade nas informações.

O mesmo aponta que em agosto de 2005, no evento do II ENAT - Encontro Nacional de Administradores Tributários, que foi em São Paulo, teve como objetivo o desenvolvimento e implantar o Sistema Público de Escrituração Digital e a Nota Fiscal Eletrônica.

O Sped fiscal faz parte do Projeto de Modernização da Administração Tributária e Aduaneira (PMATA) que tem o objetivo a implantação de novos procedimentos que apoiam os sistemas de informação integrados na tecnologia ou na infraestrutura.

A NF-e tem como objetivo atender todas as normas aplicadas no projeto da SEFAZ e tem sua validade jurídica garantida pela assinatura digital do emitente e recebimento pelo fisco.

4.9.1 Objetivo da NF-e

Em uma época onde o mundo tecnológico e a informação podem atravessar continentes em questão de segundos, a obrigatoriedade da NF-e traz grande avanço ao país e a todos, deixando todos os processos burocráticos menos custosos para as empresas.

De acordo com o Portal NFE (2010), a NF-e é um documento que existe exclusivamente em forma digital, com o objetivo de arquivar uma atividade de circulação de uma mercadoria ou qualquer tipo de prestação de serviço, sua

verificação de validade é garantida pelo fisco através de uma assinatura digital do emissor e entrada para a Secretaria da Fazenda.

Atualmente a grande maioria das empresas que recebem as Notas Fiscais Eletrônicas (NF-e) em arquivos em formato XML. Mesmo que esse processo já aconteça frequentemente, muitas empresas não sabem que autorizar a entrada de mercadorias sem a XML pode se tornar um enorme problema fiscal no futuro, com fiscais e auditores da receita.

O Portal NFE (2010) expõe que uma nota fiscal eletrônica é uma maneira mais simplificada, toda a mercadoria que uma empresa recebe ou dá entrada, é sempre obrigada a ter a NF-e em formato XML arquivada e validada. Muitas empresas arquivam o DANF-e (Documento Auxiliar a Nota Fiscal Eletrônica) e simplesmente deixa de lado o arquivo XML que o fornecedor envia, desta maneira pode acabar dando entrada em mercadorias que não poderiam ser aceitas.

O procedimento mais eficaz é, antes de receber qualquer tipo de mercadoria, sempre exigir que seu fornecedor envie a XML da NF-e referente à compra. Após ter recebido é com total obrigação da empresa conferir, junto da Fazenda, se aquela nota foi autorizada ou rejeitada, com a assinatura digital deste documento validado e se o arquivo está correto e não foi corrompido no ato da emissão.

Após a data de cancelamento, é sempre recomendado verificar novamente a validade do documento, pois pode ter sido cancelado pelo emissor. Após o recebimento da XML e com o devido tratamento, a empresa está pronta para receber a mercadoria acompanhada pelo DANFE, que possui as informações e deveria ter seu conteúdo correspondido com a XML que a empresa recebeu. Com estas verificações se torna complexas, envolvem muitos códigos enormes e obrigam recursos financeiros e de mão de obra. Já existe no mercado muitas soluções automatizam este processo e ajudam com a integração necessária com sistemas de gestão da empresa.

No Portal NFE (2010) ressalta que se tiver outras opções adotadas por empresas mais espertas, é uma verificação detalhada dos seus arquivos. Com isso pode ser feita através de sistemas especiais que busca as informações de compras de uma empresa com a base em arquivos XML, identificando assim eventuais mercadorias recebidas sem o respectivo arquivo XML da NF-e.

Sempre os responsáveis pelas áreas fiscal ou financeira devem estar espertos para que estes documentos estejam devidamente organizados e validados, se isso não ocorrer estão mantendo uma terrível falha em suas atividades, comprometendo o desempenho.

É muito importante lembrar que não fazer verificações é deixar de cumprir com a lei e correr o risco de pesadas multas junto ao Fisco em uma fiscalização. Então gerir, validar, organizar os arquivos XML, por meio de um fornecedor de TI especializado é o caminho mais seguro que sua empresa pode escolher. (PORTAL NFE, 2012).

4.9.2 Organização da NF-e

Segundo o Portal NFE (2010), as notas fiscais eletrônicas são arquivos XML estruturados, recebidos ou enviados para outras pessoas, que deve ser guardado em algum local seguro por pelo menos 5 anos. Fazer backup desses documentos em algum local seguro nem é um grande problema, já que hoje em dia tem os mais diversos sistemas de armazenamento como os CD/DVD, pen drives, HD externo e outros.

O problema aparece em caso de uma fiscalização, um vez que, nestes casos, um fiscal irá solicitar determinados arquivos XML, que deverão ser localizados em pouco tempo. Para resolver esse problema, que pode acarretar inclusive em multas a empresa fiscalizada, os colaboradores organizam esses arquivos, prestando muita atenção e tendo cuidado para que possam encontrá-los facilmente, assim organizando por data ou CNPJ, facilitando na busca quando solicitado.

4.9.3 Benefícios da NF-e

Segundo o Portal NFE (2010), muitos benefícios foram apresentados após a implantação da NF-e como redução de custo de impressão dos documentos fiscais, apesar de ainda ter muitos gastos com impressões deve-se notar que a economia gerada de cada arquivo XML recebido da NF-e reduzindo muito papel. Ainda é emitido um comprovante de cada nota fiscal recebida gerando um gasto, mas com certeza, muito menor do que antes.

Afim de reduzir o armazenamento de documentos fiscais, hoje todos e qualquer documento impresso deve ser guardado para a apresentação ao fisco pelo prazo decadencial, envolvendo também, não apenas partes físicas, mas também dados nos servidores.

4.10 Validação da NF-e pela Sefaz

Quando a Sefaz recebe a nota fiscal eletrônica, é feita uma verificação da validação do mesmo, observando aspectos, tais como: integridade, assinatura digital do emitente, regularidade dos campos, emissor autorizado, regularidades fiscais, base de dados na fazenda. Feita essa verificação e não constatado nenhum erro durante esse procedimento, a nota será recebida, organizada, armazenada, reestruturada por meio da Sefaz, Após isso, será gerado um protocolo e será liberada a consulta para ser realizada via internet pelo cliente, facilitando o controle, O Danf-e tem um código de barras padrão com um número de chave única”. (SEFAZ, 2010).

4.11 Problemas com a validação da NF-e

Pode-se apontar como os principais problemas com a validação da NF-e: interrupção na comunicação, rejeição ou não autorização de uso.

A interrupção da comunicação ocorrerá sempre por algum problema na comunicação de dados, sendo assim se não houver o recebimento do arquivo pela Sefaz, deverá se realizar uma nova transmissão do documento.

Quando o problema é a rejeição na validação de recepção, não será armazenado na base de dados de consulta da Sefaz, sendo possível a pessoa corrigir o problema.

O problema de não autorização do uso ocorrerá quando houver quaisquer problemas com o emissor ou destinatário da NF-e.

4.12 Consulta da NF-e

Segundo o Portal NFE (2014), todas as Nf-e podem ser consultadas no site da Sefaz para verificar sua validade de origem e emitente, a consulta é

feita através de uma chave de acesso da nota fiscal eletrônica que fica exposto no Danf-e, onde pode ser lida ou digitada pelo código de barras.

O sistema interno da empresa também pode consultar na Sefaz, e assim verificar a validade e autenticidade da nota, não sendo necessário entrar no site para verificar a consulta.

4.13 Cancelamento da NF-e

De acordo com o Portal NFE (2010), o cancelamento da NF-e deve estar dentro do prazo (máximo 24 horas da sua emissão), após este prazo, o software permitirá o cancelamento até 480 horas da data de autorização de uso da NFe, mais a empresa fica sujeita a multas conforme Decreto 45.490/2000 artigo 527 Inciso IV item z1, do RICMS/SP.

Se o colaborador tiver que cancelar a NFe após esse prazo, é necessário realizar um processo administrativo em algum Posto Fiscal.

Uma nota fiscal eletrônica somente deverá ser cancelada se caso o seu uso tenha sido autorizado e também que não tenha ocorrido nenhum fator ou também que não ocorreu a saída da mercadoria da empresa.

Segundo a Sefaz (2012), para se cancelar uma NF-e o responsável tem a função de solicitar um chamado específico gerando um arquivo XML. A aparência de um arquivo de solicitação de cancelamento deve ser consultada no site da Sefaz ou no Manual de Integração do Contribuinte. Sempre que o status da NF-e cancelada estiver como disponível poderá ser consultada no site da Secretaria da Fazenda do Estado da empresa ou no site nacional da Nota Fiscal Eletrônica (www.nfe.fazenda.gov.br).

4.14 Danf-e

Danf-e nada mais é do que a abreviação de um documento que auxilia a NF-e e tem como função deixar o transporte do produto mais seguro. Antes de existir o Danf-e era emitida uma via da Nota Fiscal que acompanhava a mercadoria. Agora ele auxilia a consulta da NF-e por todos os postos fiscais e pelo comprador da mercadoria, que podem verificar se aquele corresponde mesmo à uma Nota Fiscal eletrônica válida. (PORTAL NFE, 2011).

Segundo Portal NFE (2010), a Danf-e sempre deve ser impressa e deve também acompanhar a mercadoria, ele também deve possuir a codificação da chave de acesso da NF-e. Se a Danf-e tiver muitas informações e não couber em uma única folha, deverá imprimir todas as folhas, permitindo que tenha toda a informação possível do produto. Ele não vale igual a um documento fiscal e nem pode ser substituído no lugar de uma nota fiscal, mais existe um modelo pré-definido do Danf-e, porém cada empresa tem a liberdade de fazer o seu próprio layout, conforme ilustra a Figura 3.

Figura 3 – Modelo do Danf-e.

 Impulso LTDA Rua São Paulo, 21 Bucarein Joinville SC		DANFE Documento auxiliar da Nota Fiscal Eletrônica Saída: 1 1 Entrada: 2 RI N: 701 RI SÉRIE :1		CONTROLE DO FISCO 																													
NATUREZA DA OPERAÇÃO IMPULSO EQUIPAMENTOS PARA GASTRONOMIA																																	
REGISTRO ESTADUAL		REGISTRO ESTADUAL SUBSTITUÍDO		CNPJ																													
250607760				83123141000200																													
CHAVE DE ACESSO DA NF-E CONSULTA NO SITE: WWW.NFE.FAZENDA.GOV.BR 42081183123141000200550010000007010046403276																																	
DESTINATÁRIO/REMETENTE																																	
NOME/RAZÃO SOCIAL				CNPJ																													
????????????? DA SILVA				83123141000200																													
ENDEREÇO				CÉP																													
RUA DO CENTRO, 987				87500000																													
MUNICÍPIO		UF		REGISTRO ESTADUAL																													
UMUARAMA		SP		492526267118																													
DATA DE EMISSÃO 25/11/2008																																	
DATA DE SAÍDA/ENTRADA 25/11/2008																																	
HORA DE SAÍDA																																	
FATURA																																	
Número Cód. Veb. Valor																																	
CÁLCULO DO IMPOSTO																																	
BASE DE CÁLCULO DE ICMS		VALOR DO ICMS		VALOR TOTAL DOS PRODUTOS																													
980.00		909.00		100.00																													
VALOR DA FRETE		VALOR DO DESPES.		VALOR TOTAL DA NOTA																													
20.00		15.00		1.00																													
BASE DE CÁLC. DE ICMS SUBSTITUIÇÃO		VALOR DO ICMS SUBSTITUIÇÃO		VALOR TOTAL DA NOTA																													
980.00		678.00		1.00																													
TRANSPORTADOR/VOLUMES TRANSPORTADOS																																	
RAZÃO SOCIAL		FRETE POR QUANT.		CÓDIGO ANV																													
MGA TRANSP		1		123456																													
ENDEREÇO		MUNICÍPIO		UF																													
RUA TRANSP		MARINGÁ		PR																													
QUANTIDADE		ESPECIE		MARCA																													
1		ESPECIE		MARCA																													
QUANTIDADE		ESPECIE		MARCA																													
1		ESPECIE		MARCA																													
DADOS DO PRODUTO/SERVIÇO																																	
<table border="1"> <thead> <tr> <th>QTD PRODUTO</th> <th>DESCRIÇÃO DO PRODUTO/SERVIÇO</th> <th>NCM/SH</th> <th>CE</th> <th>2-DIG</th> <th>4-DIG</th> <th>CTE</th> <th>VLR UNIT</th> <th>VLR TOTAL</th> <th>ICMS</th> <th>VLR ICMS</th> <th>VLR IPI</th> <th>VLR OUTROS</th> <th>VLR TOTAL</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>...</td> </tr> </tbody> </table>						QTD PRODUTO	DESCRIÇÃO DO PRODUTO/SERVIÇO	NCM/SH	CE	2-DIG	4-DIG	CTE	VLR UNIT	VLR TOTAL	ICMS	VLR ICMS	VLR IPI	VLR OUTROS	VLR TOTAL	1
QTD PRODUTO	DESCRIÇÃO DO PRODUTO/SERVIÇO	NCM/SH	CE	2-DIG	4-DIG	CTE	VLR UNIT	VLR TOTAL	ICMS	VLR ICMS	VLR IPI	VLR OUTROS	VLR TOTAL																				
1																				

Fonte: Portal NFE. (2014).

4.15 Certificado Digital

O certificado digital é um método criado para validar o acesso de uma pessoa na internet ou softwares, podendo ser por meio de um software, onde se encontram informações necessárias para o acesso, ou por meio de um cartão de identificação que é colocando em uma leitora, e liberado o acesso.

Certificado Digital é uma assinatura com validade jurídica que garante proteção às transações eletrônicas e outros serviços via internet, permitindo que pessoas e empresas se identifiquem e assinem digitalmente de qualquer lugar do mundo com mais segurança e agilidade. (Serasa, 2014).

Sendo assim, o certificado digital é um meio para proteger o acesso a um sistema ou site, para que pessoas não autorizadas não possam entrar ou roubar dados.

A Figura 4 mostra um dos tipos de certificado digital existentes em nossa sociedade, que serve para acessar áreas restritas de sites e sistemas de governos.

Figura 4 – Leitor do certificado digital do Serasa.



Fonte: Boletim Informativo Serasa Experian. (2014).

5 ENGENHARIA DE SOFTWARE

Engenharia de software é uma área do conhecimento da computação voltada para a especificação, desenvolvimento e manutenção de sistemas de software, aplicando tecnologias e práticas de gerência de projetos, objetivando organização, produtividade e qualidade.

Quando um software é bem-sucedido e atende às necessidades dos usuários, opera perfeitamente durante um longo período, é fácil de modificar, mais fácil ainda, de utilizar, ele é realmente capaz de mudar as coisas para melhor. (PRESSMAN 2012 p.142).

A construção de um software possui estágios e sempre com a ideia inicial formada na mente do idealizador de forma ainda meio obscura, onde este identifica necessidades que precisam ser supridas por algum advento tecnológico. Talvez por falta de conhecimento técnico, este não seja capaz de especificar soluções, e métodos que atendam às necessidades, deve sempre aprovar uma primeira ideia para solução das suas necessidades.

Segundo Vasquez (2014), a metodologia clássica da Engenharia de Software está associada de forma geral a quatro fases de desenvolvimento e manutenção:

- a) Requisitos de Software: Descrição de todos os serviços, operações e funções que o software terá;
- b) designer de Software: Definição da arquitetura, os componentes, interfaces e outras características que o software irá possuir;
- c) construção de Software: Todo o processo de criação, através de codificação, unidades de teste, integração e debugging, utilizando-se dos requisitos e designer.
- d) testes de Software: Verificação de comportamento do programa.
- e) manutenção de Software: Atualizações, alterações e correções do software após ter sido finalizado.

Em resumo, a Engenharia de Software realiza uma análise do problema e define os métodos, ferramentas e procedimentos que serão empregados para

o desenvolvimento de um software, abrangendo também a fase de testes do programa e sua manutenção. (VELASQUEZ, 2011).

Teste é uma forma de verificação de conformidade e qualidade do produto, uma vez que após o desenvolvimento é muito importante realizar uma comparação detalhada de tudo que foi realizado. (MORELL, 1987).

Testes em sistemas e em documentações reduzem os riscos da ocorrência de defeitos do software no ambiente de produção onde são encontrados pelo cliente, contribuindo assim para a qualidade dos sistemas, pois, quanto mais cedo os defeitos forem encontrados antes da implantação do sistema o custo de correção é menor em relação ao encontrado da fase de produção. (Almeida 2010 p89).

Segundo Rezende (2006), um software tem a vida útil de 5 anos, pois nunca está por acabado, sempre têm manutenções, correções e melhorias para serem feitas. Podem existir vários tipos de ciclos de vidas em um software. Não existe uma forma única e padrão para a resolução dos problemas de se desenvolver um software, o que existe são vários modelos, estudos e métodos que abrangem todas as fases de desenvolvimento, que juntos e bem utilizados vão fazer a engenharia de software.

Existem poucos sistemas que fogem à essas regras, que nunca morrem, por estarem sempre a nível operacional, como Folhas de pagamento, Contabilidade, Contas a pagar, pois esses sistemas estão baseados em outros sistemas que foram formados e formatados a muito tempo e não devem mudar tão facilmente.

6 METODOLOGIA

As pesquisas exploratórias têm como propósito proporcionar ao investigador maior familiaridade com o problema, objetivando torná-lo mais explícito ou construir hipóteses. Uma pesquisa de cunho exploratório tende a ser bastante flexível, pois leva em consideração os mais variados aspectos relativos ao problema estudado. De modo geral, pesquisas realizadas com propósitos acadêmicos, pelo menos inicialmente, assumem esse caráter exploratório, pois neste momento é pouco provável que o pesquisador tenha uma definição clara do que irá investigar. Dessa forma, este projeto é inicialmente uma pesquisa exploratória, pois visa investigar métodos de construção de um software que permitisse agilizar o processo de organização de notas fiscais eletrônicas, utilizando teorias associadas às técnicas XPath e XQuery na organização de notas fiscais eletrônicas.

Para construção de um software dessa natureza pode-se observar uma série de etapas que devem ser seguidas até a validação do protótipo. Um bom projeto de sistema computacional, conforme recomendações da área de Engenharia de Software, deve seguir etapas bem definidas, sobretudo no que diz respeito ao planejamento e posterior desenvolvimento. Ter associada ao processo de desenvolvimento uma metodologia é fator decisivo para o sucesso de um sistema.

Um projeto pode ser visto como uma atividade bem definida cujo objetivo é produzir resultados pré-estabelecidos. O processo de desenvolvimento da ferramenta proposta iniciou-se com a definição da estrutura do projeto e a definição detalhada de metas e cronogramas. Um ponto importante no estabelecimento do projeto do sistema foi a decisão sobre que informação relevante deve ser mostrada ao usuário final. Assim, optou-se por uma interface simples e com um conjunto de funcionalidades bem específico e focado na resolução da tarefa proposta. O levantamento e escolha de ferramentas de apoio ao desenvolvimento é também uma questão importante uma vez que essas ferramentas devem ser utilizadas como facilitadoras na construção de protótipos. Na seção seguinte serão detalhadas todas as ferramentas utilizadas no processo de desenvolvimento. Cabe destacar que por

se tratar de uma ferramenta desenvolvida para lidar com um domínio particular associado processamento de arquivos XML de notas fiscais, um sistema deste tipo deve ser criado somente em um contexto de problemas bem definidos. O processo de criação dos algoritmos de processamento das notas fiscais foi realizado levando em consideração as características das técnicas XPath e XQuery.

A primeira etapa do desenvolvimento deste trabalho consistiu em um estudo das tecnologias disponíveis para processamento de XML. Após terem sido recolhidos todos os conhecimentos e informações relacionadas para a execução do trabalho, foi iniciada a fase da modelagem e desenvolvimento do protótipo. Este desenvolvimento consistiu principalmente na implementação das técnicas escolhidas (XPath e XQuery), da definição de um banco de dados firebird, pois ele é free e tem componentes nativos do delphi para melhor desempenho com a ferramenta e do acesso e processamento dos arquivos XML.

A criação do protótipo correspondeu ao estágio de desenvolvimento inicial, onde o valor do sistema foi avaliado, considerando testes iniciais executados pelo próprio autor deste trabalho. Estes testes iniciais foram úteis para verificar as funcionalidades básicas e os requisitos de funcionamento para o sistema. Os resultados obtidos dessa avaliação inicial do protótipo orientaram o estágio de desenvolvimento, permitindo o aprimoramento da interface e dos algoritmos associados às linguagem de processamento do XML.

A última etapa foi o estágio operacional, quando o sistema, em sua versão final, foi avaliado com relação a sua utilização na prática para que não haja falhas, verificando seus resultados frente à atividade para a qual ele foi desenvolvido. Para isso, conjuntos de dados (arquivos XML) com diferentes configurações foram criados e submetidos à ferramenta. Assim, foi possível fazer medir de modo quantitativo o desempenho do sistema sob a ótica das técnicas XPath e XQuery. Todas essas etapas de desenvolvimento e avaliação da ferramenta implementada serão discutidas nas seções subsequentes.

6.1 Desenvolvimento do software

Neste trabalho foram utilizadas várias ferramentas para a criação do software sendo elas: JUDE Community (que foi utilizado para realizar um diagrama de classe), a linguagem Delphi (como linguagem de programação), o banco de dados Firebird com o IBExpert (como administrador do banco de dados).

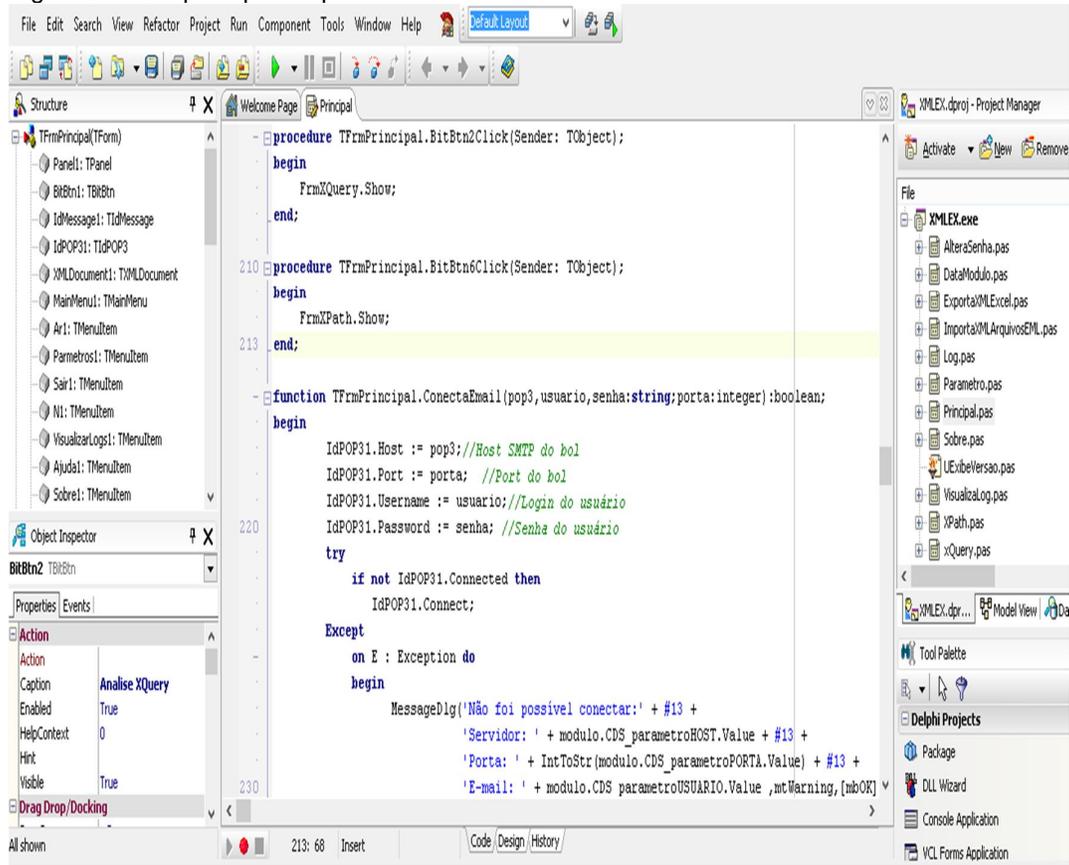
As ferramentas utilizadas, bem como a modelagem do software, estão detalhadas nas subseções seguintes.

6.1.1 Delphi

O ambiente de desenvolvimento foi o Borland® Delphi 7 é um ambiente integrado para desenvolvimento de software, produzido e mantido pela CodeGear, antigamente conhecida como Borland Software Corporation. A linguagem utilizada pelo Delphi é o Object Pascal, derivada da linguagem de programação estruturada Pascal (CANTÙ, 2003).

Uma grande vantagem de utilizar o Delphi, é o seu suporte a vários bancos de dados, principalmente utilizando os componentes dbExpress. Além de possibilitar o desenvolvimento de software estruturado em multicamadas. A Figura 5 apresenta a janela principal da IDE Delphi.

Figura 5 – Tela principal Delphi.



Fonte: Elaborado pelo autor (2014).

6.1.2 Banco de dados

O banco de dados Firebird é gratuito, tem muitos recursos, exemplos de uso e fóruns de listas de discussões encontrados na Internet. O projeto Firebird é aberto à comunidade para participação no desenvolvimento do código desse banco de dados, correção de erros, documentação, desenvolvimento de ferramentas, dentre outros.

Optou-se pelo Firebird no desenvolvimento do software, uma vez que o mesmo se integra facilmente ao Delphi, sendo uma das principais características de um sistema multicamadas, com acesso a banco de dados por meio dos componentes dbExpress. Esse banco de dados permite acesso de várias formas diferentes e acesso aos vários bancos de dados sem alteração na codificação do sistema. A única preocupação quanto a isso se

deve a conversão de procedures, views, triggers e tabelas de um banco de dados para outro.

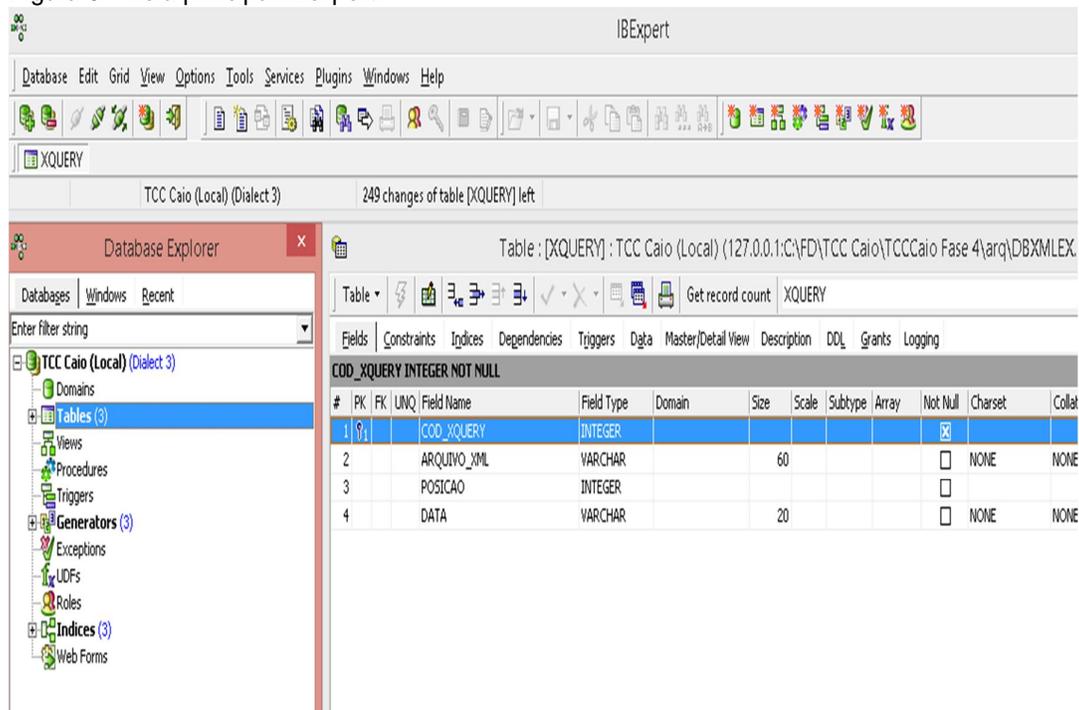
6.1.3 IBexpert

O IBExpert (IBEXPERT, 2011) é uma ferramenta para administração de bancos de dados Interbase e Firebird. Essa ferramenta permite criar e gerenciar usuários, tabelas, triggers, procedures, views, etc. Para utilizar o IBExpert é necessário registrar o banco de dados. Após esse registro é possível realizar as operações necessárias no banco de dados. O IBExpert possui uma versão gratuita para testes.

O IBExpert é um sistema gerenciador de banco de dados relacional (SGBDR), SQL (Structured Query Language) Cliente/Servidor que pode ser executado em uma diversidade de plataformas de sistemas operacionais servidores e clientes (CANTU, 2010).

Com o IBExpert é possível analisar dados, copiar objetos de bancos de dados, utilizar ferramentas de SQL, comparar bancos de dados ou tabelas de bancos de dados, bem como a opção de Database designer, para criar tabelas. A Figura apresenta a tela principal da ferramenta IBExpert. A área à direita é destinada à edição e os demais elementos estão na área à esquerda e na barra superior de ícones. A Figura 6 apresenta a janela inicial do IBExpert.

Figura 6 – Tela principal IBExpert.



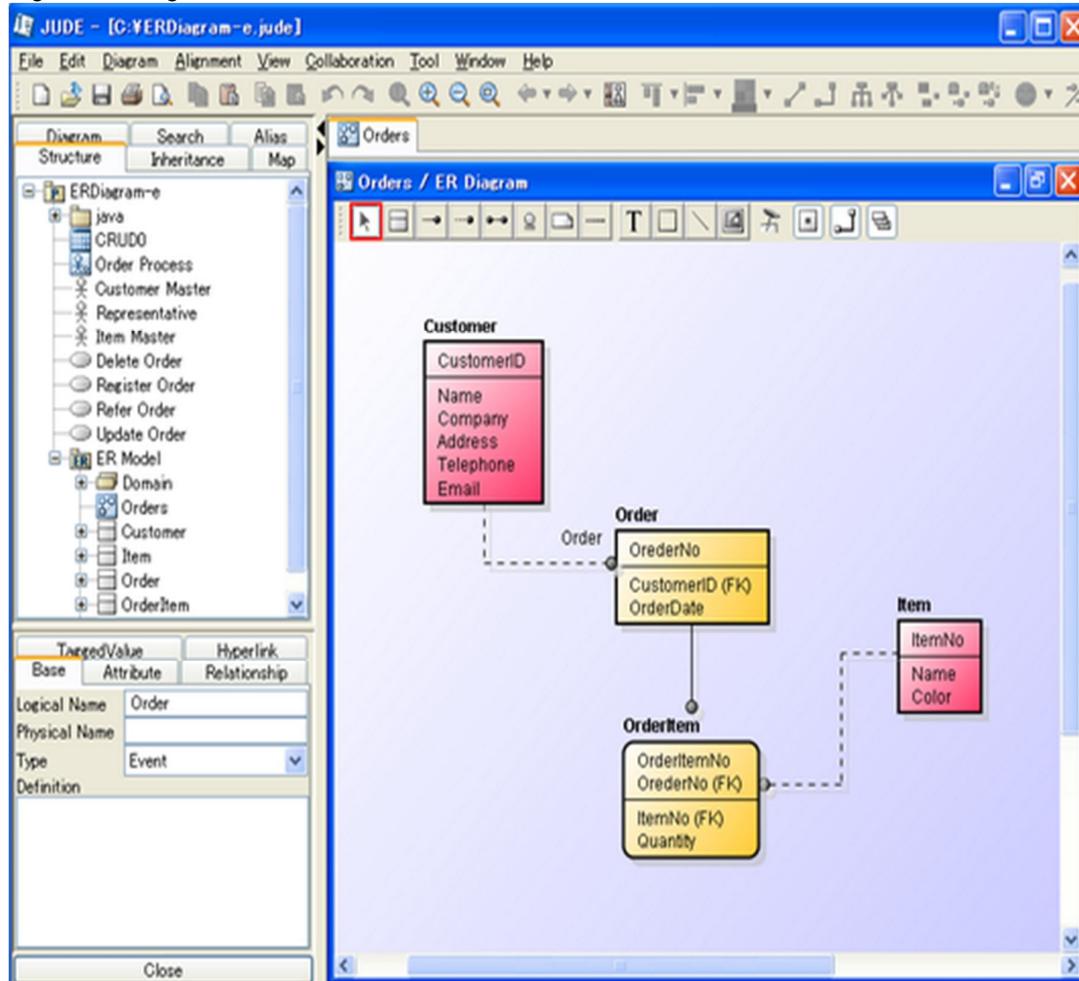
Fonte: Elaborado pelo autor (2014).

6.1.4 Jude Community

É um software desenvolvido na plataforma Java, o que garante sua portabilidade e seu funcionamento em qualquer plataforma que possui uma máquina virtual Java. É muito utilizado para a modelagem de dados, facilitando a criação do software e nas manutenções futuras. O Jude possibilita, a java reverse que é útil na importação de código java para criar um modelo e java forward que garante a geração de código fonte java a partir de um modelo desenhado.

Hoje em dia o Jude Community mudou de nome para o Astah Community, que tornou-se o editor gratuito mais utilizado pela sua grandiosidade de recursos oferecidos e por sua praticidade em elaborar diagramas completos. O Astah vem com a promessa de superar as expectativas dos tantos que utilizaram o Jude. (SALSOTTO, 2010). A Figura 7 mostra um exemplo de diagrama feito com o jude community.

Figura 7 – Diagrama Jude.

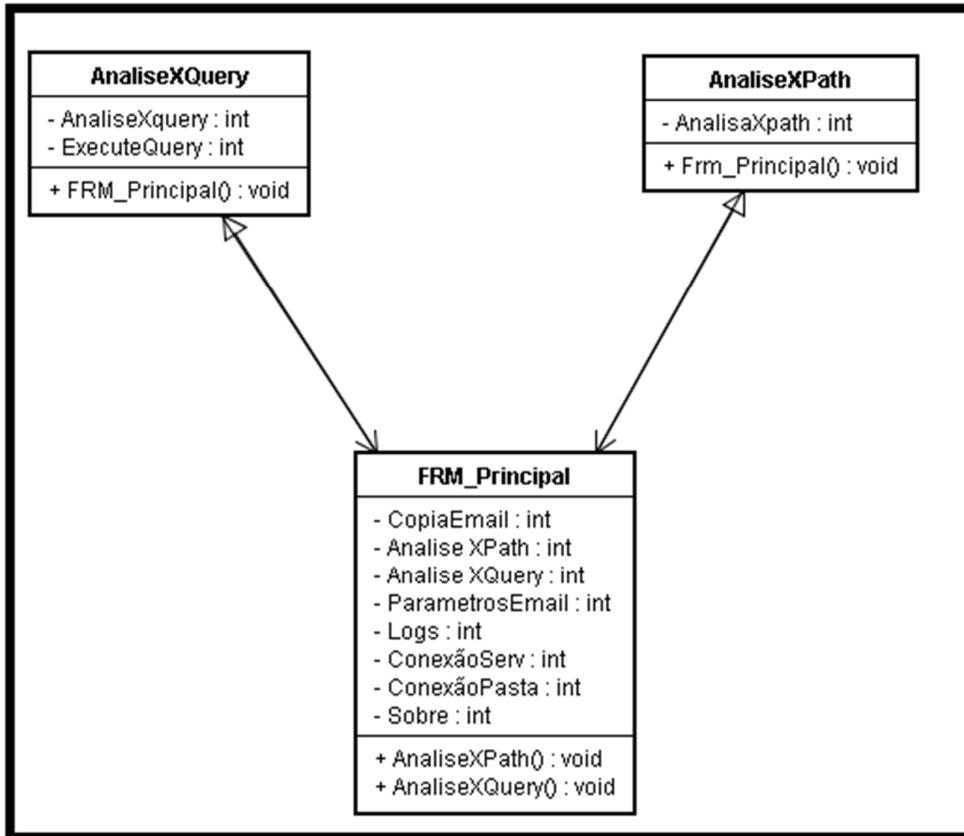


Fonte: Astah Users Change (2009).

6.1.5 Modelagem (diagramas)

Para o desenvolvimento do software foi necessário a criação de um diagrama de classe utilizando o software Jude. A Figura 8 demonstra o diagrama de classe criado com todos os atributos e métodos definidos para o desenvolvimento do software. Esse diagrama contém toda a estrutura do software, como análise XQuery, análise XPath e o form principal interligando as duas técnicas.

Figura 8 – Diagrama de classe.



Fonte: Elaborado pelo autor (2014).

6.1.6 Interface

A parte gráfica e lógica de um sistema é algo que os desenvolvedores devem fazer melhor. Pois através é dela que o usuário irá interagir com as funcionalidades do sistema. (MORAES, 2003). A seguir encontram-se as telas utilizadas no software desenvolvido e suas respectivas funcionalidades.

A Figura 9 apresenta o Form principal do software desenvolvido, tendo várias opções para se escolher, como configurar servidor de e-mail, verificar a conexão com e-mail e a pasta criada, iniciar copiar do e-mail para a pasta, utilizar as técnicas XPath, XQuery entre outras.

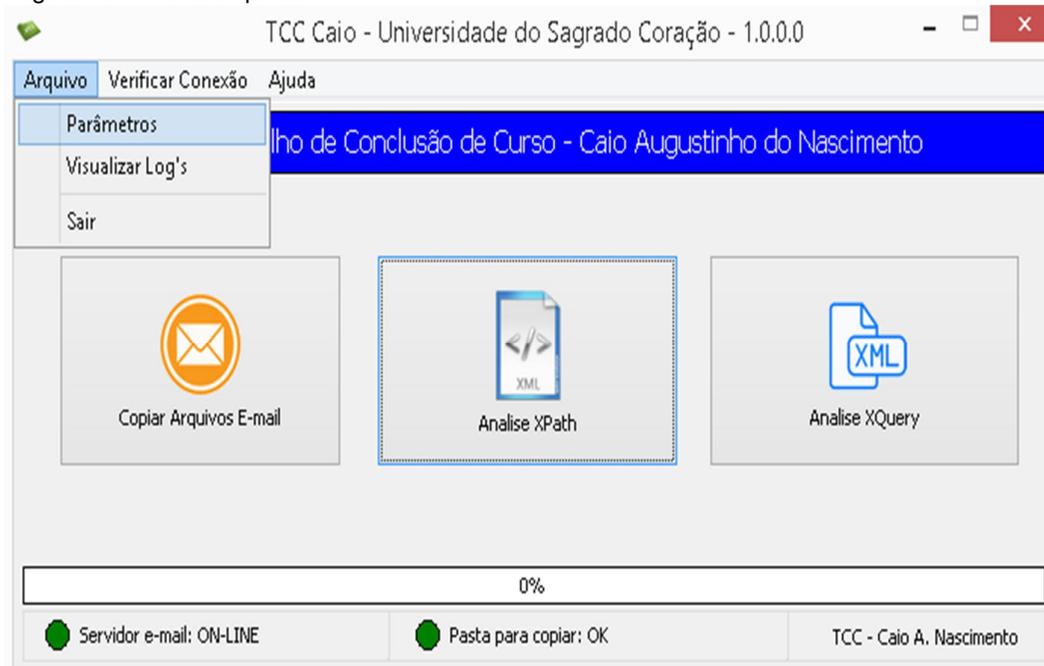
Figura 9 – Tela inicial do software desenvolvido.



Fonte: Elaborada pelo autor.

Já a Figura 10 mostra algumas opções dos itens menu arquivo, tipo os parâmetros de configuração do e-mail e caminho da pasta de armazenamento dos arquivos XML e também a visualização de log's gerados pela execução do XPath e XQuery.

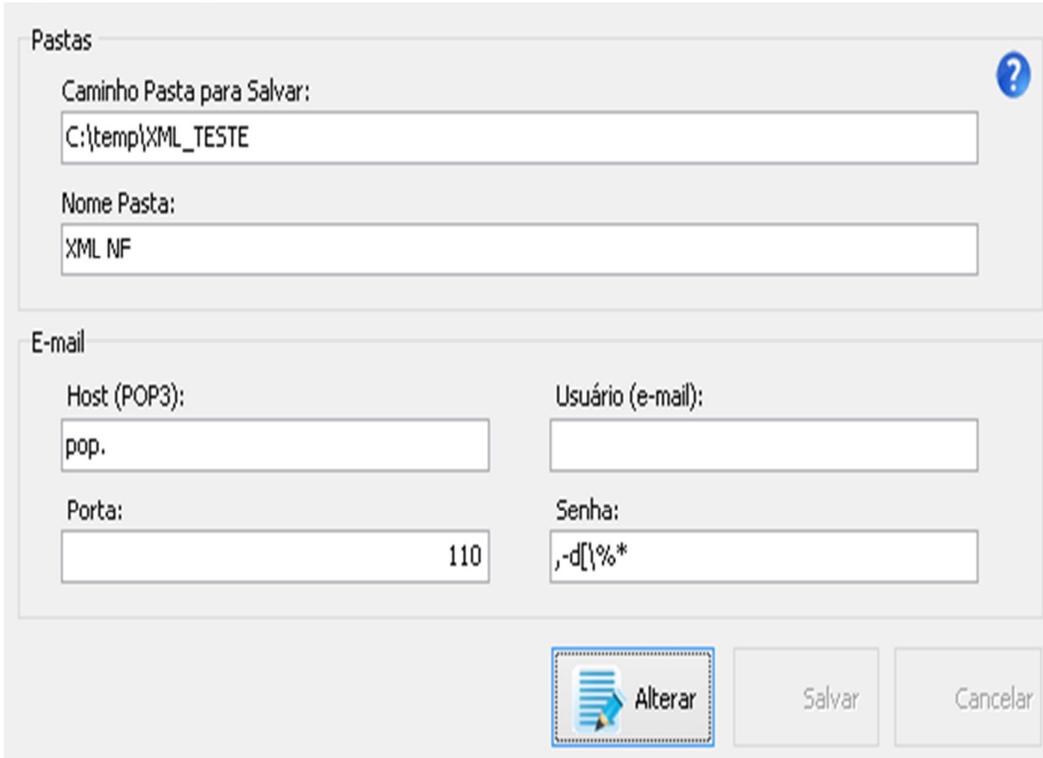
Figura 10 – Menu arquivos.



Fonte: Elaborado pelo autor.

Na Figura 11 é mostrada a tela de configuração do e-mail e a pasta de armazenamento dos XML. Pode-se perceber que os campos para a configuração são definidos manualmente em toda a tela, deixando, assim, para que o usuário configure da melhor forma possível.

Figura 11 – Configurar parâmetros do e-mail.



The image shows a configuration window with two main sections: "Pastas" and "E-mail".

Pastas

- Caminho Pasta para Salvar:** C:\temp\XML_TESTE
- Nome Pasta:** XML NF

E-mail

- Host (POP3):** pop.
- Usuário (e-mail):** (empty)
- Porta:** 110
- Senha:** ,-d[\ % *

Buttons at the bottom: Alterar (highlighted with a blue dashed border), Salvar, and Cancelar.

Fonte: Elaborado pelo autor.

Na Figura 12 é mostrada a tela de logs do sistema, deixando bem claro ao usuário a localização do arquivo XML ou se for o caso, a informação sobre problemas no processamento.

Figura 12 – Tela de logs do sistema.

Pesquisa por Período

De: Ate:

Status
 OK Erros
 Avisos Todos

Tipo
 Ação
 Período

Descrição do Log	Usuário	Computador
Movido o arquivo pasta origem: C:\temp\XML_TESTE\15140604322706000120570010000004051100992074.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\24120808645423000170550000000082391000927451.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\24120808645423000170550000000082421000931314.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\24120808645423000170550000000082451000936350.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\24120808645423000170550000000082461000937087.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\26120824441206000115550110000298061981673722.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\26120824441206000115550110000300101981641784.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\26120824441206000115550110000297341981680620.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\27120812285276000142550100000938601722179760.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\29120640482184000198550010000004451000004450.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\26140624441206000115550110000411451980129333.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\26140624441206000115550110000410891980137106.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\26140624441206000115550110000412831980111171.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\26140602733866000137570020000134331441477820.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\29140601034009000771570000000114671339971805.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\29140607405710000140550010000015761000015760.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\29140613504675000110550060000017341954038510.xml pasta destino: C:\Caio_	CAIO	CAIO
Movido o arquivo pasta origem: C:\temp\XML_TESTE\29140613504675000110550060000017321792064975.xml pasta destino: C:\Caio_	CAIO	CAIO

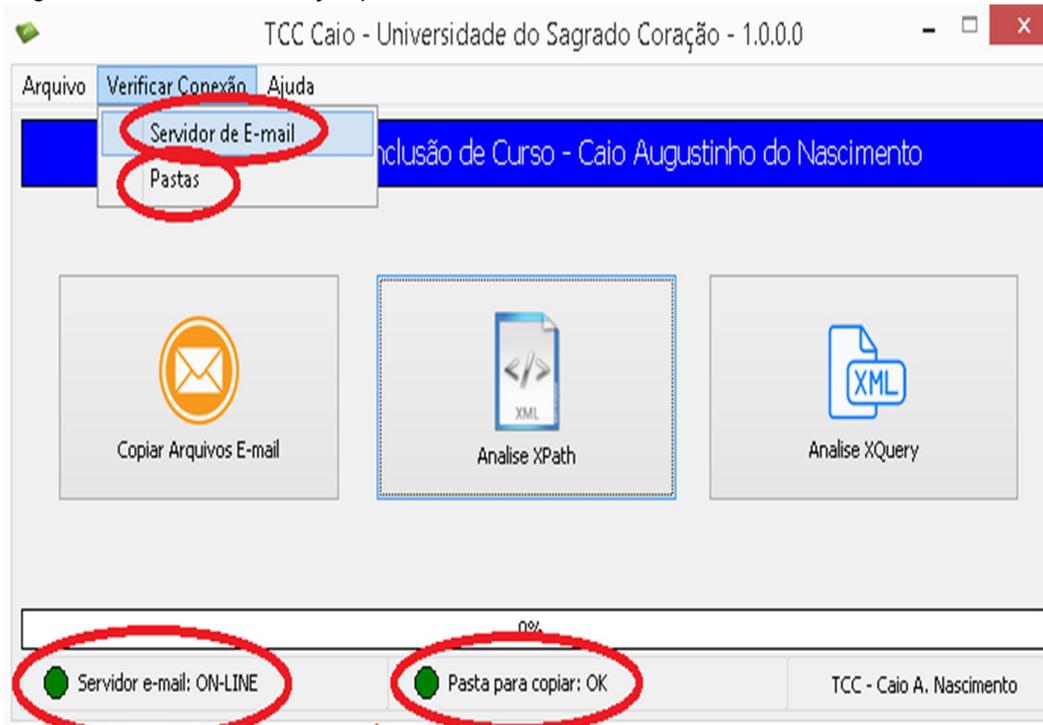
OK Avisos Erros

Logs encontrados: 5665

Fonte: Elaborado pelo autor.

Na figura 13 é mostrada a opção de verificação de conexão com o e-mail e com a pasta criada. Pode-se observar que essa tela, mesmo sendo simples, é bem útil e essencial, pois por ela o usuário pode verificar se a configuração realizada para acesso ao e-mail está certa ou não e também o local onde foi criada a pasta em que será copiado o XML.

Figura 13 – Tela de verificação pasta e servidor.



Fonte: Elaborado pelo autor.

Na Figura 14 é mostrada a opção de conexão com o servidor de e-mail. Ou seja, esse botão tem a função de se conectar ao e-mail e copiar todos os XML presentes para uma pasta desejada, conforme mostra a figura anterior.

Figura 14 – Tela de download do e-mail.



Fonte: Elaborado pelo autor.

Após realizado o download dos e-mails é possível utilizar as duas técnicas propostas nesse trabalho. A Figura 15 apresenta a técnica XPath que tem como característica a organização simples dos arquivos.

Figura 15 – Tela do ícone análise XPath.



Fonte: Elaborado pelo autor (2014).

Depois de clicado em Análise XPath é aberta uma outra tela onde é possível organizar todos os XML contidos dentro da pasta criada por data de recebimento. A Figura 16 mostra várias opções como data de início, data de fim, hora, tempo de execução e, por fim, tem um link para ir direto até a pasta com os XML organizados.

Figura 16 – Tela do XPath.

INFORMAÇÕES

Data Início	:	//
Data Fim	:	//
Hora Início	:	00:00:00
Hora Fim	:	00:00:00
Tempo de Execução	:	00:00:00
Pasta que contem os XML's	:	C:\temp\XML_TESTE

Iniciar processo de análise XPath

Processo XPath

0%

Fonte: Elaborada pelo autor.

Na Figura 17 é mostrado o botão “Análise XQuery”, ao pressioná-lo é aberto as configuração de organização utilizando a técnica XQuery.

Figura 17 – Tela do ícone análise XQuery.



Fonte: Elaborado pelo autor.

Após a execução é aberta a tela ilustrada na Figura 18, que exibe a data de início, data de fim, hora, tempo de execução e o caminho da pasta com os arquivos serão organizados, utilizando a técnica XQuery. Destaca-se que essa técnica é mais completa do que a XPath, com isso é capaz de fazer outras coisas como utilizar um comando SQL para trazer informações contidas dentro de um XML. Todas as informações extraídas do documento ficam salvas em um banco de dados, podendo ser acessada do software, manualmente ou até mesmo remotamente.

Figura 18 – Tela do XQuery.

Arquivo XML	Data	Posição

Query:

INFORMAÇÕES

Data Início	:	//
Data Fim	:	//
Hora Início	:	00:00:00
Hora Fim	:	00:00:00
Tempo de Execução	:	00:00:00
Pasta que contem os XML's	:	C:\temp\XML_TESTE


Iniciar processo de analise xQuery

Processo XPath

0%

Fonte: Elaborada pelo autor.

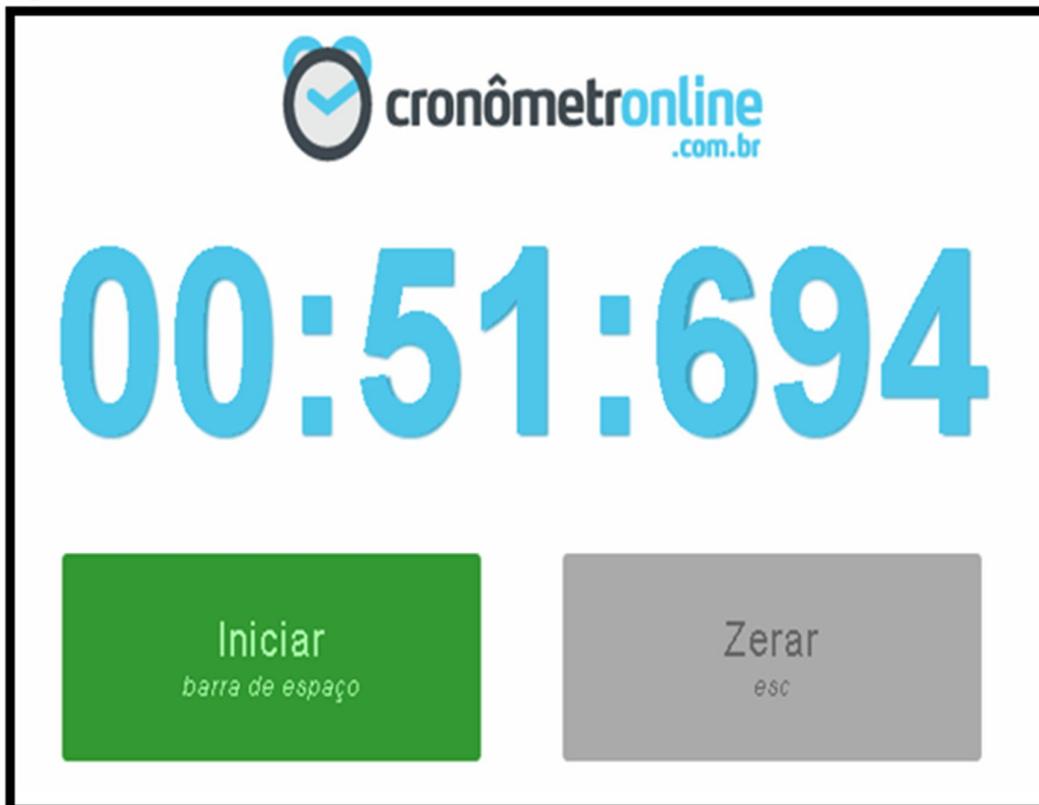
7 RESULTADOS

Com o intuito de avaliar as técnicas implementadas foram realizados 4 testes de eficiência com diferentes configurações. Dentre os aspectos testados foram considerados, por exemplo, os diferentes servidores de e-mails: Gmail, Hotmail, Icloud, Ig e outros. Também foi testado o número de e-mails cujo download foi realizado com sucesso pelo software. As técnicas XPath e XQuery foram testadas, considerando conjuntos com diferentes quantidades de arquivos XML.

O intuito principal dos testes foi, portanto, verificar se uma ferramenta que utiliza as técnicas XPath e XQuery poderia aumentar a eficiência da tarefa de organização de notas fiscais eletrônicas, conforme a hipótese inicial. Além disso, buscou-se verificar em que grau essa eficiência seria supostamente aumentada. Para que essa eficiência pudesse ser medida de alguma forma, inicialmente foi estabelecido como baseline o tempo médio de execução desta tarefa de recebimento e organização de arquivos dispendido por uma pessoa que faz tudo isso manualmente.

Uma pessoa fazendo o trabalho manualmente de copiar um arquivo XML da nota fiscal eletrônica que tem como padrão suas informações e número de campos, recebendo como anexo dentro de e-mail e salvá-lo dentro de uma pasta criada em um determinado local do computador, iria demorar em média de 00:51:694hs, isso realizado em 1 teste, conforme cronometro do site www.cronometroonline.com.br, mostrado na Figura 19.

Figura 19 – Cronometro da análise manual.



Fonte: Cronometro Online.

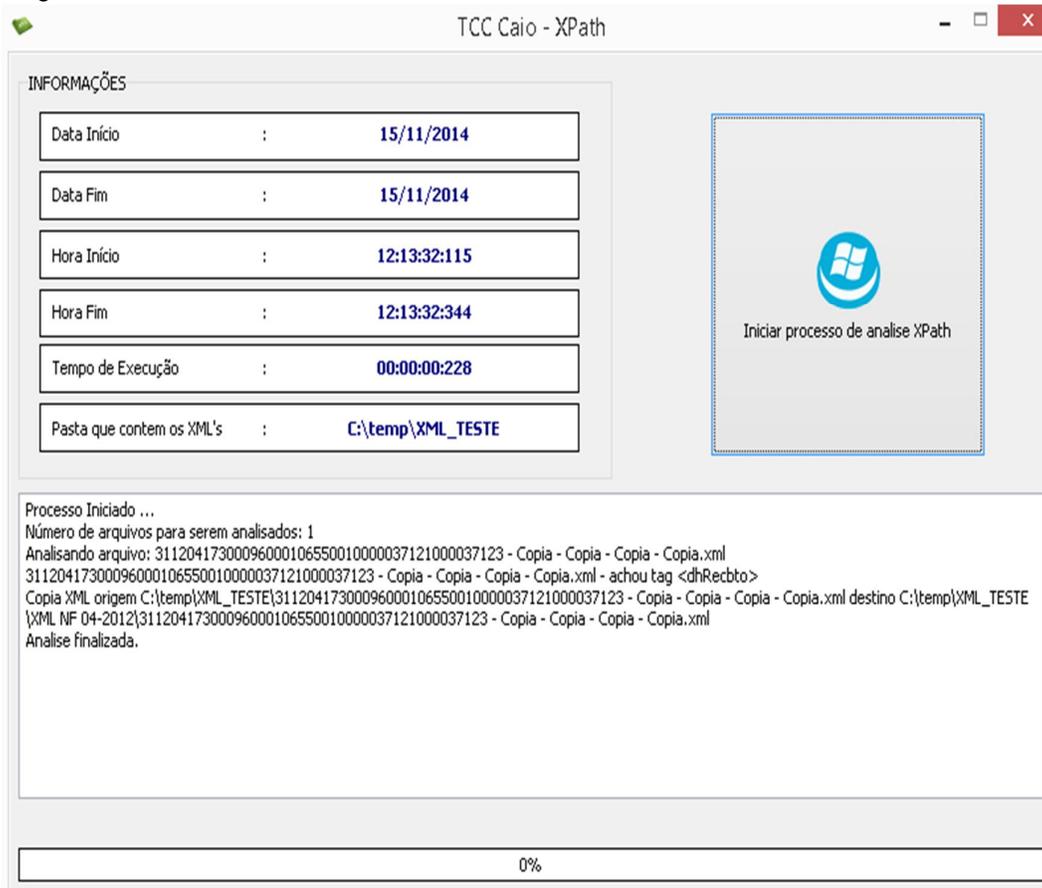
Com este tempo, foi possível estimar o tempo que uma pessoa levaria para realizar esta mesma tarefa com diferentes quantidades de arquivos. Por exemplo, esse mesmo procedimento realizado manualmente por uma pessoa com 100 arquivos XML iria demorar aproximadamente 86:15 horas, com 1.000 arquivos o tempo seria de 86.000 horas e com 10.000 demoraria aproximadamente 860.000 h.

Um funcionário de uma empresa iria demorar em torno de 35 dias para organizar 10.000 arquivos XML recebidos, isso trabalhando 24 horas por dia, o que é algo impraticável, mostrando que a execução da tarefa certamente demoraria muito mais tempo. Essa mesma tarefa, considerando turnos de 8 horas por dia de trabalho, exigiria aproximadamente 11 meses dedicados somente a esta organização de arquivos.

Fazendo esse mesmo teste com o software desenvolvido, considerando o processamento de um único arquivo XML, utilizando a técnica XPath como parâmetro e tempos reais gastos no processamento, observa-se

que são necessários apenas 0.228 segundos, conforme mostra a Figura 20 a seguir. Claramente, existe um ganho muito grande em termos de tempo, ou seja, se considerarmos um único arquivo o processamento via software consumiu apenas cerca de 0,45% do tempo gasto na execução manual. Em outras palavras, o software permitiu a redução do tempo de processamento em cerca de 99,55%.

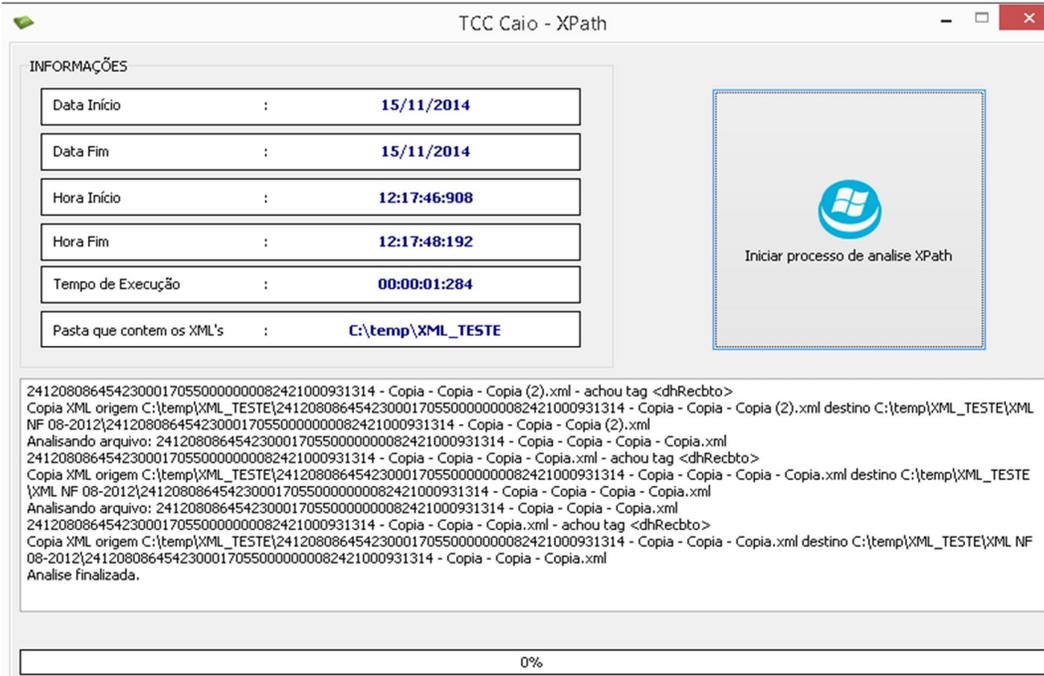
Figura 20 – Tela de teste com XPath.



Fonte: Elaborado pelo autor.

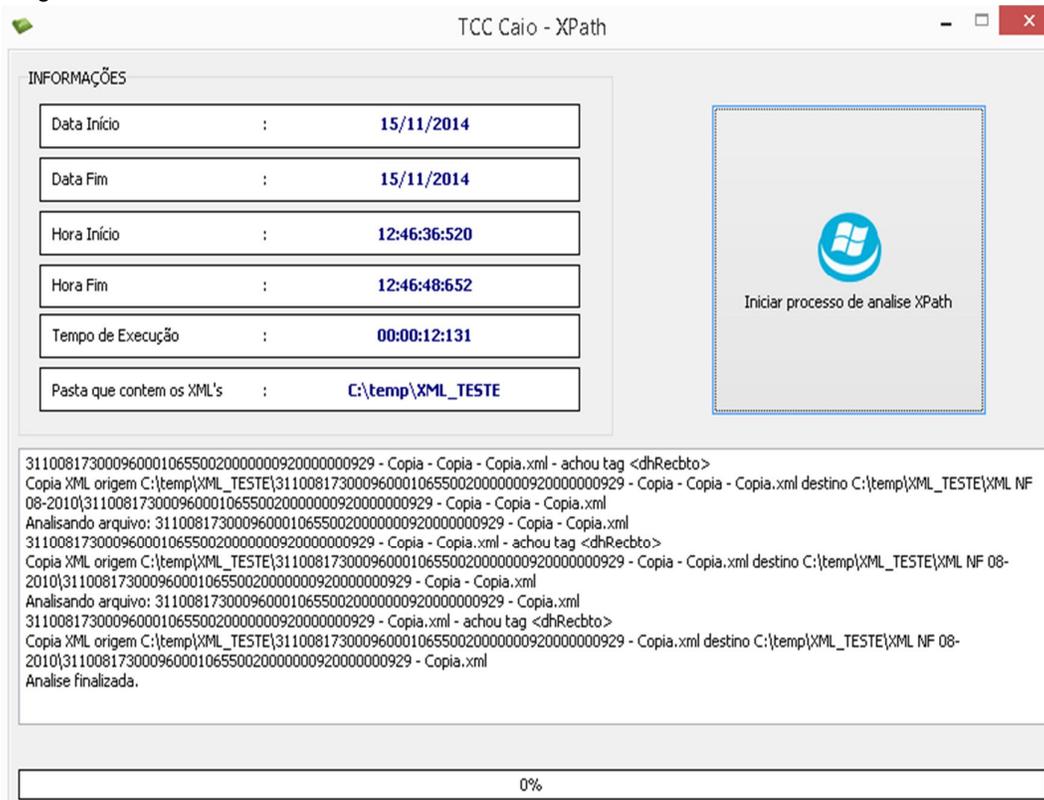
Considerando um conjunto de 100 arquivos XML e a técnica XPath, o tempo gasto foi de 01.284 s conforme mostra a Figura 21. A figura 22 mostra o teste para um conjunto de 1.000 arquivos (tempo de 12.131 s).

Figura 21 – Tela de teste com XPath 2.



Fonte: Elaborada pelo autor.

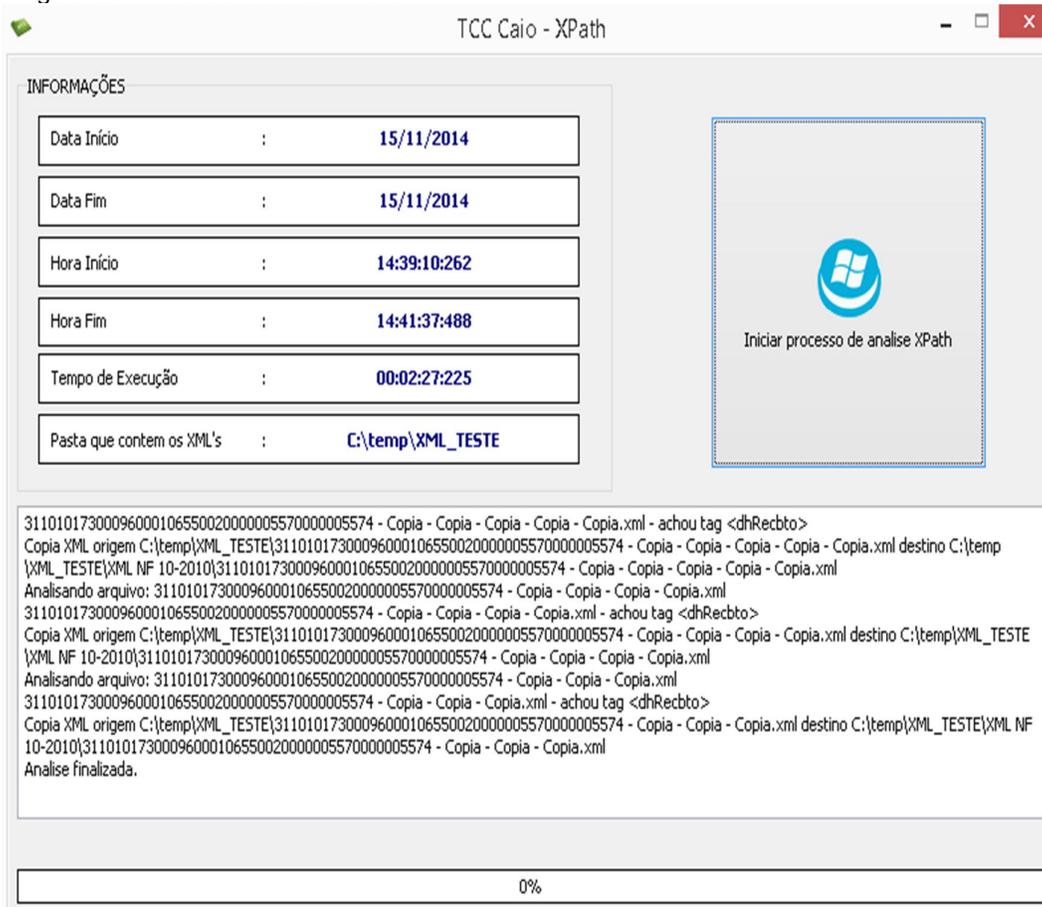
Figura 22 – Tela de teste com XPath 3.



Fonte: Elaborado pelo autor.

Utilizando como base para o teste um conjunto de 10.000 arquivos XML o tempo total gasto foi de 147.225 s, conforme ilustra a Figura 23.

Figura 23 – Tela de teste com XPath 4.



Fonte: Elaborada pelo autor.

Os mesmos teste foram refeitos utilizando a técnica XQuery, que é mais completa, uma vez que salva as informações dos arquivos XML em um banco de dados, tornando possível utilizar comandos SQL para buscar informações. A Figura 24 a seguir mostra o teste realizado em 1 arquivo XML, cujo tempo total gasto foi de 0.042 s.

Figura 24 – Tela de teste com XQuery.

TCC Caio - xQuery

Arquivo XML	Data	Posição
▶ 15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151

Query:

INFORMAÇÕES

Data Início	:	15/11/2014
Data Fim	:	15/11/2014
Hora Início	:	14:49:01:705
Hora Fim	:	14:49:01:747
Tempo de Execução	:	00:00:00:042
Pasta que contem os XML's	:	C:\temp\XML_TESTE



Iniciar processo de análise xQuery

0%

Fonte: Elaborado pelo autor.

A Figura 25 mostra a utilização do XQuery com 100 arquivos. Observa-se que o tempo total foi de 1.458 s.

Figura 25 – Tela de teste XQuery 2.

TCC Caio - xQuery

Arquivo XML	Data	Posição
15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151
15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151
15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151
15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151
15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151
15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151
15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151
15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151
15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151
15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151
15140604322706000120570010000004051100992074 - Copia - Copia	06-2014	7151

Query:

INFORMAÇÕES

Data Início	:	15/11/2014
Data Fim	:	15/11/2014
Hora Início	:	14:52:51:839
Hora Fim	:	14:52:53:297
Tempo de Execução	:	00:00:01:458
Pasta que contem os XML's	:	C:\temp\XML_TESTE



Iniciar processo de analise xQuery

0%

Fonte: Elaborado pelo autor.

Com 1.000 arquivos o tempo total é de 14.683 s (conforme ilustra a Figura 26), e aqui já é possível notar uma diferença entre o tempo de execução das duas técnicas, já que XPath gastou 12.131 s..

Figura 26 – Tela de teste XQuery 3.

TCC Caio - xQuery

Arquivo XML	Data	Posição
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia.xml	10-2009	20503
35091072381189000625550010001416746261866978.xml	10-2009	20503

Query:

INFORMAÇÕES

Data Início	:	15/11/2014
Data Fim	:	15/11/2014
Hora Início	:	14:55:05:438
Hora Fim	:	14:55:20:121
Tempo de Execução	:	00:00:14:683
Pasta que contem os XML's	:	C:\temp\XML_TESTE



Iniciar processo de análise xQuery

0%

Fonte: Elaborada pelo autor.

Por fim o teste com 10.000 arquivos XML consome um tempo total de 148.201 s utilizando a técnica XQuery, conforme Figura 27.

Figura 27 – Tela de teste XQuery 4.

TCC Caio - xQuery

Arquivo XML	Data	Posição
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia - Copia	10-2009	20503
35091072381189000625550010001416746261866978 - Copia.xml	10-2009	20503
35091072381189000625550010001416746261866978.xml	10-2009	20503

Query:

INFORMAÇÕES

Data Início	:	15/11/2014
Data Fim	:	15/11/2014
Hora Início	:	15:02:55:834
Hora Fim	:	15:05:24:035
Tempo de Execução	:	00:02:28:201
Pasta que contem os XML's	:	C:\temp\XML_TESTE

0%

Fonte: Elaborada pelo autor.

Tendo como base o teste realizado com 10.000 arquivos XML, fica extremamente evidente a diferença entre os métodos manuais e os automáticos (sejam eles usando as técnicas XPath ou XQuery). Manualmente uma pessoa iria demorar 860.000 horas para terminar a tarefa, utilizando o software proposto com a técnica XPath o tempo foi de 00:02:27:225 h e no XQuery 00:02:28:201 h. A Figura 28 mostra esse resultado.

Figura 28 – Tela conclusão teste.

	Manual - Aproximadamente	XPath	XQuery
1 - XML	00:51:694	00:00:00:228	00:00:00:042
100 - XML	86:15:000	00:00:01:284	00:00:01:458
1.000 - XML	86.000hs	00:00:12:131	00:00:14:683
10.000 - XML	860.000hs	00:02:27:225	00:02:28:201

Fonte: Elaborado pelo autor.

Pelos dados da Figura 28 é interessante notar que as evidências apontam que a técnica XQuery começa mais rápido e vai ficando mais lenta em comparação com a técnica XPath. Com esses testes de comparação, concluímos que a técnica XPath é a mais rápida, porém, talvez, menos eficaz em comparação com a XQuery, pois esta última permite a posterior manipulação das informações sobre os arquivos XML contidas no banco de dados.

8 CONSIDERAÇÕES FINAIS

Essa pesquisa mostrou a utilização das técnicas XPath e XQuery no processo de recebimento e organização de arquivos XML de Nota Fiscal Eletrônica por meio do software desenvolvido. Sendo de grande importância ressaltar que a tecnologia veio sem dúvida nenhuma contribuir para a manipulação eficiente de dados.

Foi realizada considerando os mais variados aspectos relativos ao tema estudado, seguindo os princípios acadêmicos e definindo uma pesquisa de caráter exploratório, realizando pesquisas de coleta de dados na internet, nos livros, e artigos científicos, sobre as técnicas XPath, XQuery e os problemas enfrentados pela sociedade com a nota fiscal eletrônica (GIL, 2010).

Esta investigação teve como uma de suas contribuições o desenvolvimento de um software com uma base de dados que permite que empresas que manipulem notas fiscais eletrônicas possam, automaticamente, incluir, editar e excluir dados rapidamente, realizando assim essas atividades com muito menos falhas e de modo muito mais eficiente. O software que foi desenvolvido oferece a função de extrair o arquivo XML anexado a um e-mail, oferecendo a possibilidade de recuperar a informação sobre a data de emissão da nota fiscal, sendo possível organizar todos os arquivos de modo a ajudar a controlar o recebimento da Nota Fiscal Eletrônica, automatizando o processo de organização e diminuindo as possíveis falhas existentes.

Além disso, outra contribuição foi uma comparação entre duas técnicas (XPath e a XQuery), ressaltando, aspectos relacionados sobretudo à velocidade de processamento.

A atual pesquisa vem a contribuir para uma melhor compreensão do processo de recebimento da nota fiscal eletrônica e automatização do seu recebimento, facilitando no dia a dia de muitas pessoas e empresas que não podem comprar software avançados que fazem esse procedimento.

Os conhecimentos obtidos através do desenvolvimento desta tese podem ser consideravelmente ampliados futuramente para a implementação no software como escolher automaticamente qual técnica será utilizada pela

quantidade de arquivos XML da NF-e e também exportar os dados do logs para mandar a outras pessoas que não tem acesso ao software.

REFERENCIAS

- ANDERSON; R. **Professional XML**. Ciência Moderna LTDA, 2001. P. 12
- BERGLUND, A; BOAG, S; CHAMBERLIN, D; FERNÁNDEZ, **XML Path Language (XPath) 2.0**. <http://www.w3.org/TR/2004/WD-xpath20-20041029/>. Acesso em: 01 mai. 2014.
- BERNERS-LEE, T; HENDLER, J; LASSILA, O. **The SemanticWeb**. Scientific American. Mai 2001.
- BOLETIM INFORMATIVO SERASA EXPERIAN, 2014 em **Serasa Experian: a gente trabalha para você crescer**. Disponível em: <http://www.serasaexperian.com.br/serasaexperian/publicacoes/bis/2010/118/bis_00356.htm> Acesso em: 29 abr. 2014.
- CASTRO, R; **Suporte Nativo a XML no SGBD Oracle 9i**. Recife: UFPE, Centro de informática, 2002.
- CANTÙ, M. Dominando o Delphi 7 **"A Bíblia"**. São Paulo: Makron Books, 2003.
- CANTU, C. H. Get to know Firebird in 2 minutes (Conheça o Firebird em dois minutos). Disponível em: <<http://www.firebirdnews.org/docs/f>>. Acesso em: 26 out. 2010.
- DEITEL, H.M, et al. **XML - Como Programar**. Porto Alegre: Bookman, 2001. P. 32
- DEITEL, H; DEITEL, M; PAUL, J. XML - **How to Program**. Prentice-Hall, Inc., 2000. P. 67
- EVANS, A.; KAMANNA, A.; MUELLER, J.; **XML e ASP.NET**. Rio de Janeiro: Editora Ciência Moderna, 2003.
- GIL, A. C. Como Elaborar Projetos de Pesquisa. 5. Ed. São Paulo: Atlas, 2010.
- HAROLD, E.; **The XML Bible**. IDG Books, 2ª edition, 1999.
- KAY, M. **XSL Transformations (XSLT) Version 2.0**. <http://www.w3.org/TR/2004/WD-xslt20-20041105/>, Acesso em: 02 abr. 2014.
- IBEXPERT. IBExpert developer studio. Disponível em: <<http://www.ibexpert.com/>>. Acesso em: 2 abr. 2011.
- JUDE/COMMUNITY. Disponível em <<http://jude.change-vision.com/jude-web/product/community.html>>. Acesso em: 20 abr. 2010
- MORAES, L. **XML e Java**. Instituto de Informática - UFG, 2003.

MACORATTI, J, **XML TUTORIAL XPATH**.

http://www.macoratti.net/vb_xpath.htm. Acesso em: 25 mai. 2014.

NOTA FISCAL ELETRÔNICA, 2014 em **Wikipédia: a enciclopédia livre**.

Disponível em: <http://pt.wikipedia.org/wiki/Nota_fiscal_eletr%C3%B4nica>

Acesso em: 15 abr. 2014.

NFE. **NOTA FISCAL ELETRÔNICA**. Disponível em:<

<http://www.nfe.fazenda.gov.br/portal/principal.aspx> >. Acesso em: 08 abr. 2014.

NFE 2010. **Governo do Estado de São Paulo: Secretaria da Fazenda**.

Disponível em <http://www.nfp.fazenda.sp.gov.br/> Acesso em: 08 abr. 2014.

NFE 2011. **NF-Eletrônica**. Disponível em <

https://www.nfeletronica.com.br/info_contrib/index.asp> Acesso em: 07 abr. 2014.

NATANYA; M. **Xml Black Book** – Solução e Poder. Makron, Mai 2010.

PEREIRA. A. P.; **O que é XML**. Disponível em

<http://www.tecmundo.com.br/programacao/1762-o-que-e-xml-.htm> em: 16 maio. 2014.

PORTAL NFE, 2014 em **Tudo em Foco**. Disponível em <

<http://www.tudoemfoco.com.br/portal-nfe.html>> Acesso em: 27 abr. 2014.

PRESSMAN, R. Engenharia de software. Rio de Janeiro: Mc Graw-Hill, 2005.

REZENDE, D. A. **Engenharia de software e sistema de informação**. 3d Ed. São Paulo: Brasport, 2006.

ROUSE, M, **XML Definition**. <http://searchsoa.techtarget.com/definition/XML>.

Acesso em: 01 mai. 2014.

VELASQUEZ, C. E. L. **Modelo de Engenharia de Software para o Desenvolvimento de Jogos e Simulações Interactivas**, 2010. Dissertação (Mestrado em Engenharia Informática) - Universidade Fernando Pessoa.

W3.; **XML Technology 2009**. Disponível em <http://www.w3org/standards/xml>

Acesso em: 07 abr. 2014.

W3Schools 2010. **XPath XML** Disponível em <http://www.w3schools.com/XPath/>

Acesso em: 01 abr. 2014.

W3 1999. **XML Path Language (XPath) Version 1.0**. Disponível em

<http://www.w3.org/TR/xpath/> Acesso em: 15 abr. 2014.

W3School 2010. **XML Tutorial**. Disponível em

<http://www.w3schools.com/xml/default.ASP> Acesso em: 10 abr. 2014.

W3 2010. **XQuery 1.0: An XML Query Language**. Disponível em <http://www.w3.org/TR/xquery/>> Acesso em: 30 abr. 2014.

XPATH, 2014 em **Wikipédia: a enciclopédia livre**. Disponível em <http://pt.wikipedia.org/wiki/XPath>> Acesso em: 10 abr. 2014.

XQUERY, 2010. **Wikipédia: a enciclopédia livre**. Disponível em: <http://pt.wikipedia.org/wiki/XQuery>> Acesso em: 01 abr. 2014.

XQUERY DEBUGGER, 2010. **Oxygen/>: xml editor**. Disponível em:< http://www.oxygenxml.com/xml_editor/xquery_debugger.html> Acesso em 23 abr. 2014.

Aplicação das Técnicas XPath e XQuery para a Organização dos Arquivos XML da Nota Fiscal Eletrônica.

Caio A. Nascimento, Patrick P. Silva, Henrique P. Martins, Elvio G. da Silva

Departamento de Ciências Exatas e Sociais Aplicadas - Universidade Sagrado Coração
(USC) – Bauru – SP – Brasil

caioaugustinho@ig.com.br, patrick.silva@usc.br, henmartins@gmail.com,
egsilva@usc.br

***Abstract.** This research shows the use of the techniques and XPath XQuery and receiving process and organization of XML files Electronic Invoice. Using software developed in Delphi language to carry out tests to compare the two techniques (XPath and XQuery), highlighting aspects related mainly to processing speed. The results show that the XPath language is faster, but perhaps less efficient compared to XQuery, since the latter allows further handling of the information contained in the XML files database*

***Resumo.** Essa pesquisa mostra a utilização das técnicas Xpath e Xquery e no processo de recebimento e organização de arquivos XML de Nota Fiscal Eletrônica. Utilizando um software desenvolvido na linguagem Delphi para realizar testes comparar as duas técnicas (XPath e a XQuery), ressaltando, aspectos relacionados sobretudo à velocidade de processamento. Os resultados obtidos evidenciam que a técnica XPath é a mais rápida, porém, talvez, menos eficaz em comparação com a XQuery, pois esta última permite a posterior manipulação das informações sobre os arquivos XML contidas no banco de dados...*

1. Introdução

Antigamente o processo de venda era algo simples, onde o fornecedor emitia a Nota Fiscal, encaminhava a via do cliente, e uma cópia para o contador proceder aos registros contábeis e fiscais. Após esse procedimento, o cliente apenas guardava uma via da Nota Fiscal, e isso era o suficiente para finalizar o processo. Com a implantação do sistema de NF-e (“Nota Fiscal Eletrônica”), novos itens surgiram.

Com o avanço da nota fiscal eletrônica (NF-e), ficou muito mais rápido, ágil e prático para as empresas disponibilizarem os arquivos de **Nota Fiscal eletrônica (NF-e)** para seus clientes através do site da empresa ou por e-mail.

O XML é o padrão de documento utilizado pelo órgão responsável pela implantação do sistema de NF-e do Governo. O layout é padronizado em todo o território nacional, respeitando os conceitos antigos de escrituração fiscal e adicionando melhorias visando maior controle de tais transações. (ANDERSON, 2001).

O XML fornece a descrição, o armazenamento, o formato da transmissão, e todo conteúdo de uma nota fiscal eletrônica. A pesquisa de uma NF-e é bem ampla, podendo ser realizada uma pesquisa pela data de recebimento, número do CNPJ, número da nota fiscal, nome do fornecedor, endereço, cidade e etc. A pesquisa é sempre solicitada pelo

fiscal da receita, e quase sempre a solicitação é por data do recebimento. (PORTAL NFE, 2011).

A técnica XPath é um conjunto de regras de sintaxe que permite definir em partes documento XML, usa uma expressão de caminho para poder localizar nós chamados de tags em documentos XML. (W3C, 2010).

Já a XQuery é uma técnica criada a partir da XPath e é ideal para realizar consultas em vários tipos de documentos XML, possui uma grande estrutura que não é muito diferente da SQL, mesmo assim ela consegue gerenciar árvores na linguagem XML. (W3C, 2010).

Diante deste contexto o objetivo deste trabalho foi desenvolver um software aplicando as técnicas XPath e XQuery para organizar o recebimento das notas fiscais eletrônicas em uma empresa, tendo por base arquivos XML visando economizar o tempo de seus colaboradores e evitar falhas humanas.

2. Referencial Teórico

O XML foi desenvolvido pelo XML Working Group, que é conhecido como SGML com isso teve um grande patrocínio do www no ano de 1996, foi também presidido por Jon Bosak, proprietário da Sun Microsystems.

Quando foi criada, pensou-se que o propósito da XML seria de substituição da HTML. A história da XML foi iniciada nos anos 70, quando a IBM decidiu inventar a linguagem GML (General Markup Language – linguagem de marcação genérica). Assim, surgiu com a necessidade da empresa em armazenar uma grande quantidade de informações.

Mesmo com sua criação em 1996, a XML vem crescendo gradativamente se tornando uma linguagem padrão de representação e intercambio de dados por ser utilizado em meios eletrônicos

Segundo Macoratti (2014), XML (Extensible Markup Language) é uma das linguagens que permite construir documentos legíveis para seres humanos, e que pode ser facilmente interpretada pelas máquinas, ou seja, a XML possui muitos marcadores que são mais fáceis e flexíveis do que em outras linguagens.

Segundo o W3C (2010), a técnica XPath permite construir expressões que recorrem e processam um documento XML de modo parecido a uma expressão regular. Pode também computar valores como números, valores Booleanos e cadeias de caracteres.

- É uma sintaxe para definir partes de um documento XML.
- Usa caminhos para definir elementos XML.
- Define uma biblioteca de funções padrão.

Já o XQuery é uma técnica só de leitura (read-only), e foi projetada para fazer consultas em coleções de dados em XML, pode retornar uma instância do modelo de dados XQuery, mas não pode modificar uma instância existente. Todas as expressões XQuery podem ser classificadas em, expressões que efetuam atualização ou expressões que não efetuam atualização.

3. Metodologia

Este projeto é uma pesquisa exploratória, pois visa investigar métodos de construção de um software que permite agilizar o processo de organização de notas fiscais eletrônicas, utilizando teorias associadas às técnicas XPath e XQuery na organização de notas fiscais eletrônicas.

Para construção de um software dessa natureza pode-se observar uma série de etapas que devem ser seguidas até a validação do protótipo. Um bom projeto de sistema computacional, conforme recomendações da área de Engenharia de Software, deve seguir etapas bem definidas, sobretudo no que diz respeito ao planejamento e posterior desenvolvimento. Ter associada ao processo de desenvolvimento uma metodologia é fator decisivo para o sucesso de um sistema.

Um projeto pode ser visto como uma atividade bem definida cujo objetivo é produzir resultados pré-estabelecidos. O processo de desenvolvimento da ferramenta proposta iniciou-se com a definição da estrutura do projeto e a definição detalhada de metas e cronogramas. Um ponto importante no estabelecimento do projeto do sistema foi a decisão sobre que informação relevante deve ser mostrada ao usuário final, neste caso, profissionais que lidam diretamente com notas fiscais eletrônicas. Assim, optou-se por uma interface simples e com um conjunto de funcionalidades bem específico e focado na resolução da tarefa proposta. O levantamento e escolha de ferramentas de apoio ao desenvolvimento é também uma questão importante uma vez que essas ferramentas devem ser utilizadas como facilitadoras na construção de protótipos. Na seção seguinte serão detalhadas todas as ferramentas utilizadas no processo de desenvolvimento. Cabe destacar que por se tratar de uma ferramenta desenvolvida para lidar com um domínio particular associado processamento de arquivos XML de notas fiscais, um sistema deste tipo deve ser criado somente em um contexto de problemas bem definidos. O processo de criação dos algoritmos de processamento das notas fiscais foi realizado levando em consideração as características das técnicas XPath e XQuery.

A primeira etapa do desenvolvimento deste trabalho consistiu em um estudo das tecnologias disponíveis para processamento de XML. Após terem sido recolhidos todos os conhecimentos e informações relacionadas para a execução do trabalho, foi iniciada a fase de desenvolvimento do protótipo. Este desenvolvimento consistiu principalmente na implementação das técnicas escolhidas (XPath e XQuery), da definição de um banco de dados e do acesso e processamento dos arquivos XML.

A criação do protótipo correspondeu ao estágio de desenvolvimento inicial, onde o valor do sistema foi avaliado, considerando testes iniciais executados pelo próprio autor deste trabalho. Estes testes iniciais foram úteis para verificar as funcionalidades básicas e os requisitos de funcionamento para o sistema. Os resultados obtidos dessa avaliação inicial do protótipo orientaram o estágio de desenvolvimento, permitindo o aprimoramento da interface e dos algoritmos associados às linguagem de processamento do XML.

A última etapa foi o estágio operacional, quando o sistema, em sua versão final, foi avaliado com relação a sua utilização na prática, verificando seus resultados frente à atividade para a qual ele foi desenvolvido. Para isso, conjuntos de dados (arquivos XML) com diferentes configurações foram criados e submetidos à ferramenta. Assim, foi possível fazer medir de modo quantitativo o desempenho do sistema sob a ótica das técnicas XPath e XQuery.

4. Resultado

Com o intuito de avaliar as técnicas implementadas foram realizados vários testes com diferentes configurações. Dentre os aspectos testados foram considerados, por exemplo, os diferentes servidores de e-mails: Gmail, Hotmail, Icloud, Ig e outros. Também foi testado o número de e-mails cujo download foi realizado com sucesso pelo software. As técnicas XPath e XQuery foram testadas, considerando conjuntos com diferentes quantidades de arquivos XML.

O intuito principal dos testes foi, portanto, verificar se uma ferramenta que utiliza as técnicas XPath e XQuery poderia aumentar a eficiência da tarefa de organização de notas fiscais eletrônicas, conforme a hipótese inicial. Além disso, buscou-se verificar em que grau essa eficiência seria supostamente aumentada. Para que essa eficiência pudesse ser medida de alguma forma, inicialmente foi estabelecido como baseline o tempo médio de execução desta tarefa de recebimento e organização de arquivos dispendido por uma pessoa que faz tudo isso manualmente.

Uma pessoa fazendo o trabalho manualmente de copiar um arquivo XML da nota fiscal eletrônica recebido como anexo dentro de e-mail e salvá-lo dentro de uma pasta criada em um determinado local do computador, iria demorar em média de 00:51:694hs, conforme cronometro do site www.cronometroonline.com.br.

Com este tempo, foi possível estimar o tempo que uma pessoa levaria para realizar esta mesma tarefa com diferentes quantidades de arquivos. Por exemplo, esse mesmo procedimento realizado manualmente por uma pessoa com 100 arquivos XML iria demorar aproximadamente 86:15 horas, com 1.000 arquivos o tempo seria de 86.000 horas e com 10.000 demoraria aproximadamente 860.000 h.

Um funcionário de uma empresa iria demorar em torno de 35 dias para organizar 10.000 arquivos XML recebidos, isso trabalhando 24 horas por dia, o que é algo impraticável, mostrando que a execução da tarefa certamente demoraria muito mais tempo. Essa mesma tarefa, considerando turnos de 8 horas por dia de trabalho, exigiria aproximadamente 11 meses dedicados somente a esta organização de arquivos.

Fazendo esse mesmo teste com o software desenvolvido, considerando o processamento de um único arquivo XML, utilizando a técnica XPath como parâmetro e tempos reais gastos no processamento, observa-se que são necessários apenas 0.228 segundos. Claramente, existe um ganho muito grande em termos de tempo, ou seja, se considerarmos um único arquivo o processamento via software consumiu apenas cerca de 0,45% do tempo gasto na execução manual. Em outras palavras, o software permitiu a redução do tempo de processamento em cerca de 99,55%.

Considerando um conjunto de 100 arquivos XML e a técnica XPath, o tempo gasto foi de 01.284 s e o teste para um conjunto de 1.000 arquivos (tempo de 12.131 s. Utilizando como base para o teste um conjunto de 10.000 arquivos XML o tempo total gasto foi de 147.225 s

Os mesmos teste foram refeitos utilizando a técnica XQuery, que é mais completa, uma vez que salva as informações dos arquivos XML em um banco de dados, tornando possível utilizar comandos SQL para buscar informações. O teste realizado em 1 arquivo XML, cujo tempo total gasto foi de 0.042 s. A mostra a utilização do XQuery com 100 arquivos. Observa-se que o tempo total foi de 1.458 s, Com 1.000 arquivos o tempo total é de 14.683 s e com 10.000 arquivos XML consome um tempo total de 148.201 s.

Tendo como base o teste realizado com 10.000 arquivos XML, fica extremamente evidente a diferença entre os métodos manuais e os automáticos (sejam eles usando as técnicas XPath ou XQuery). Manualmente uma pessoa iria demorar 860.000 horas para terminar a tarefa, utilizando o software proposto com a XPath o tempo foi de 00:02:27:225 h e no XQuery 00:02:28:201h, conforme Figura 1.

	Manual - Aproximadamente	XPath	XQuery
1 - XML	00:51:694	00:00:00:228	00:00:00:042
100 - XML	86:15:000	00:00:01:284	00:00:01:458
1.000 - XML	86.000hs	00:00:12:131	00:00:14:683
10.000 - XML	860.000hs	00:02:27:225	00:02:28:201

Figura 1. Tela conclusão testes

Pelos dados é interessante notar que as evidências apontam que técnica XQuery começa mais rápido e vai ficando mais lenta em comparação com a técnica XPath. Com esses testes de comparação, concluímos que a XPath é a mais rápida, porém, talvez, menos eficaz em comparação com a XQuery, pois esta última permite a posterior manipulação das informações sobre os arquivos XML contidas no banco de dados.

5. Considerações Finais

Essa pesquisa mostrou a utilização das técnicas Xpath e Xquery no processo de recebimento e organização de arquivos XML de Nota Fiscal Eletrônica por meio do software desenvolvido. Sendo de grande importância ressaltar que a tecnologia veio sem dúvida nenhuma contribuir para a manipulação eficiente de dados.

Essa pesquisa foi realizada considerando os mais variados aspectos relativos ao tema estudado, seguindo os princípios acadêmicos e definindo uma pesquisa de caráter exploratório, realizando pesquisas de coleta de dados na internet, nos livros, e artigos científicos, sobre as técnicas XPath, XQuery e os problemas enfrentados pela sociedade com a nota fiscal eletrônica (GIL, 2010).

Esta investigação teve como uma de suas contribuições o desenvolvimento de um software com uma base de dados que permite que empresas que manipulem notas fiscais eletrônicas possam, automaticamente, incluir, editar e excluir dados rapidamente, realizando assim essas atividades com muito menos falhas e de modo muito mais eficiente. O software que foi desenvolvido oferece a função de extrair o arquivo XML anexado a um e-mail, oferecendo a possibilidade de recuperar a informação sobre a data de emissão da nota fiscal, sendo possível organizar todos os arquivos de modo a ajudar a controlar o recebimento da Nota Fiscal Eletrônica, automatizando o processo de organização e diminuindo as possíveis falhas existentes.

Além disso, outra contribuição foi uma comparação entre duas técnicas (XPath e a XQuery), ressaltando, aspectos relacionados sobretudo à velocidade de processamento.

A atual pesquisa vem a contribuir para uma melhor compreensão do processo de recebimento da nota fiscal eletrônica e automatização do seu recebimento, facilitando no dia a dia de muitas pessoas e empresas que não podem comprar software avançados que fazem esse procedimento.

6. Referências

ANDERSON; R. Professional XML. Ciência Moderna LTDA, 2001. P. 12

NFE. NOTA FISCAL ELETRÔNICA. Disponível em:<
<http://www.nfe.fazenda.gov.br/portal/principal.aspx> >. Acesso em: 08 abr. 2014.

NFE 2011. NF-Eletrônica. Disponível em <
https://www.nfeletronica.com.br/info_contrib/index.asp> Acesso em: 07 abr.
2014.

GIL, A. C. Como Elaborar Projetos de Pesquisa. 5. Ed. São Paulo: Atlas, 2010.

W3Schools 2010. XPath XML Disponível em <http://www.w3schools.com/XPath/>
Acesso em: 01 abr. 2014.