

WXDC: Uma Ferramenta para Imposição e Validação de Restrições de Integridade baseadas na Linguagem XDCL

Stevan D. Costa¹, Alexandre T. Lazzaretti¹, Anubis G. de M. Rossetto¹

¹Instituto Federal de Educação, Ciência e Tecnologia Sul-Rio-Grandense (IFSUL) – Campus Passo Fundo – 99.064-440 – Passo Fundo – RS – Brazil

stevandanielicosta@gmail.com, {alexandre.lazzaretti, anubis.rossetto}@passofundo.ifsul.edu.br

Abstract. This paper presents the web tool WXDC, conceived as an auxiliary mechanism for XML, in order to verify and validate XML documents. Through an intuitive interface, the tool enables the generation of the document XDC with the restrictions to be imposed on the XML document, and validate the document using the parser XDCL. With the use of the tool can to easiear the validation of XML documents, when there is need to implement integrity constraints domain in the documents.

Resumo. Este trabalho apresenta a ferramenta WXDC, concebida como um mecanismo auxiliar a metalinguagem de marcação XML, com o objetivo de validar e verificar documentos XML. Por meio de uma interface intuitiva, a ferramenta possibilita a geração do documento XDC com as restrições a serem impostas sobre o documento XML, além de validar o documento utilizando o parser XDCL. Com o uso da ferramenta é possível agilizar a validação de documentos XML, quando há necessidade de aplicar restrições de integridade de domínio aos documentos.

1. Introdução

A tecnologia XML vem a cada dia sendo mais utilizada nas trocas de dados e em muitos casos pode haver necessidade de armazenar os dados de documentos XML em algum local, por exemplo, um banco de dados. Entretanto, no momento de se adicionar esses dados em um banco de dados, podem ocorrer muitos erros devido à inconsistência nos dados presentes no documento, pois o documento XML é escrito livremente pelo usuário, contando apenas com algumas restrições imposta por DTDs (*Data Type Definition*) e XSD (*XML Schema Definition*) e deixando de lado outras restrições de integridade, como por exemplo, de domínio. (LAZZARETTI, 2005)

As restrições de domínio são restrições que são utilizadas para garantir que os dados inseridos em uma determinada coluna, pertençam a um determinado domínio, ou seja, que os valores ali contidos sejam válidos (SILVA, 2000).

Visando a melhoria da segurança e integridade dos dados em documentos XML, foi desenvolvido por Lazzaretti (2005) um controle para tratar das restrições de domínio que não são suportadas pelas DTDs e XSDs. Esse controle é chamado XDC (*XML Domain Constraints*) e é responsável por analisar documentos XML com base em um documento XDC, que foi gerado a partir de uma linguagem chamada XDCL (*XML Domain Constraint Language*), sendo que é neste documento que estarão às restrições de integridade que o usuário deseja confrontar com o documento XML.

Com o desenvolvimento deste controle, tem-se a necessidade de implementar uma ferramenta gráfica para fazer a geração desses documentos XDCs de forma visual, abstraindo do usuário o conhecimento prévio da linguagem XDCL. Além do objetivo de

abstração, tem-se o propósito de agilizar o processo de validação dos documentos, assim como trazer uma ferramenta para auxiliar na correção de muitos dos erros encontrados em documentos XML. Portanto, este artigo vem apresentar a *interface* gráfica desenvolvida para suprir esta necessidade, e também facilitar e agilizar a validação dos documentos XML, garantindo maior consistência nos dados.

Na literatura alguns trabalhos relacionados foram encontrados, como por exemplo, a ferramenta “XML Validator Online” (SCRIBE’S, 2009), porém todos os trabalhos realizam a validação do documento somente com base nos *schemas XML*. O trabalho aqui apresentado se diferencia pelo fato de validar as restrições de domínio necessárias para garantir maior segurança e integridade dos dados que estão contidos no documento XML.

Este artigo está organizado da seguinte forma: a seção 2 se refere aos conteúdos necessários para o entendimento de documentos XML e restrições de integridade; a seção 3 é destinada aos resultados obtidos no desenvolvimento da *interface* e também as discussões levantadas; por fim, têm-se as considerações finais e as referências utilizadas.

2. Materiais e Métodos

2.1 . XML

O documento XML é escrito em arquivos textos normais, com o propósito de estruturar, armazenar e transmitir dados pela internet ou também para que esses dados possam ser armazenados em algum lugar específico, como por exemplo, um banco de dados. XML é uma metalinguagem de marcação, ou seja, todos os dados contidos neste documento estarão delimitados por um elemento inicial e um elemento final, de maneira que essas marcas qualificam e delimitam o texto contido nesse intervalo (W3C, 2013).

De acordo com a W3C, os elementos definidos em um documento XML seguem uma estrutura hierárquica, e o documento possui duas formas de validações: a validação de documento bem formado e a de documento válido. Um documento bem-formado é aquele que está de acordo com as regras estabelecidas no padrão XML.

Para verificar se a estrutura do documento é válida existem os esquemas XML, que, segundo a W3C um *schema XML* é uma linguagem para expressar restrições sobre o documento XML, sendo que os principais esquemas são a DTD e o XSD. A partir destes esquemas é possível avaliar se o documento está bem formado e é válido.

2.2. Restrições de Integridade

“ Uma restrição de integridade é uma expressão *booleana* que está associada a algum banco de dados e precisar ser avaliada o tempo todo como true.” (DATE, 2003).

As restrições de integridade se referem à precisão ou a correção dos dados contidos em um banco de dados, portanto elas garantem que quando o banco de dados sofra danos no momento que são feitas as alterações. (DATE, 2003, SILBERCHATZ *et al*, 1999) e segundo Heuser (1998), existem 4 categorias de restrições de integridade: a Integridade de vazio, Integridade de Chave, Integridade referencial, e a Integridade de domínio, que é a que falaremos a seguir.

Conforme Lazzaretti (2005) as restrições de integridade de domínio são as mais comuns de serem encontrada em um banco de dados, pois elas são verificadas sempre

que um novo dado é inserido no banco de dados. Elas são definidas em 5 tipos: restrições de atributo, restrições de tipo, restrições de tupla, restrições de banco de dados e restrições de transição de estados.

2.3. Controle XDC

O controle XDC vem a ser um recurso adicional à tecnologia XML existente, a fim de validar algumas restrições importantes e que não são suportadas pelos *schemas* XML. Para que o controle XDC seja usado, primeiramente deverá ser criado um documento texto com a extensão XDC e salvo, para que quando os documentos XML ou os bancos de dados nativos XML forem usá-los, possam fazer a referência ao arquivo. A validação das restrições é feita pelo *parser* XDCL, ele valida as expressões XDCL com as linhas contidas no documento XML, retornando após a validação uma mensagem de validação correta ou a indicação de que alguma inconsistência foi encontrada, esta validação é feita através do uso da linguagem XDCL. (LAZZARETTI, 2005)

O documento XDC é um documento no formato XML, portanto, ele deve estar bem-formado e ser válido, para que seja possível o uso desse documento para realizar as validações. Este documento possui vários elementos que definem o nome da restrição a serem impostas, as condições e os elementos que serão testados, as ações que serão efetuadas no momento em que uma inconsistência for encontrada.

3. Resultados

Com base nos estudos realizados e considerando as necessidades encontradas, foi realizado o desenvolvimento de uma ferramenta que possibilita ao usuário definir, através da linguagem XDCL, as restrições de integridade e fazer a verificação de seus documentos XML conforme as restrições definidas.

Diversas tecnologias foram utilizadas para o desenvolvimento da ferramenta, sendo que estas estão divididas em *cliente* e *servidor*, sendo *cliente* aquelas interpretada pelo navegador e *servidor* aquelas com execução no lado servidor. No lado *cliente* foram utilizadas as tecnologias JavaScript, Biblioteca Jquery, AJAX, Biblioteca XMLTree e no lado *servidor* foram utilizadas as tecnologias Java, Java Server Pages (JSP), Servlets, DOM e SAX.

3.1. Arquitetura da Ferramenta

Em relação à arquitetura da ferramenta, foi possível definir os requisitos funcionais que a ferramenta deve atender, os quais são:

- Carregar e mostrar documento XML existente;
- Definir restrições;
- Gerar documentos XDC a partir das restrições definidas;
- Carregar documento XDC.
- Validar estrutura de documento XDC;
- Validar e gerar XML aplicando XDC.

No Diagrama de Atividades (Figura 1), estão definidas as ações que o usuário deve seguir para fazer o processo completo de verificação de seu documento XML, que compreende desde a geração do documento XDC até a validação.

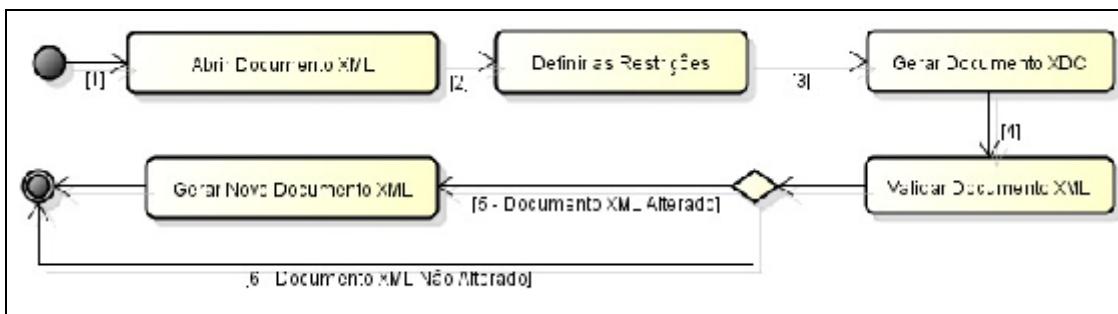


Figura 33. Diagrama de Atividade – Processo Completo

Conforme as atividades, primeiramente o usuário deve carregar o documento XML na ferramenta, após isso, define as restrições de integridade de domínio necessárias para validar o documento XML. Somente após a definição de todas as restrições, segue para a etapa seguinte que é a geração do documento XDC. Como ação opcional, é possível verificar se está bem-formado e válido, caso não esteja bem-formado é recomendado que os erros sejam corrigidos para que não ocorram problemas no momento da validação.

Após o usuário passar por todas estas etapas, segue para a etapa de validação do documento XML, com base no documento XDC, conforme mostrado no fluxo de controle 4. A ferramenta faz a validação e retorna mensagens informando o que foi feito. Caso o documento XML tenha sofrido alterações, será disponibilizado para download, este processo está sendo representado no fluxo de controle 5. Caso o documento não tenha passado por alterações, serão apenas mostradas as mensagens definidas nas restrições (fluxo de controle 6), terminando assim todo o processo de validação do documento.

Ainda é possível a utilização da ferramenta somente para a validação do documento XML, caso o usuário já tenha um documento XDC com as suas restrições, este processo é chamado processo parcial, pois o usuário não gera o documento XDC na própria ferramenta. Neste processo é necessário somente carregar os dois documentos, XML e XDC, e fazer a validação, de modo que a ferramenta dará como resultado as mensagens informadas no documento XDC, e caso ocorra alguma alteração no documento XML será disponibilizado um novo documento XML em seu estado consistente.

3.2. Funcionamento da Ferramenta

Após acessar o sistema, o usuário deve como primeiro passo, carregar o documento XML que deseja impor a (s) restrição (ões). A ferramenta possibilita a visualização do documento XML no formato de árvore com recursos para expandir e comprimir os elementos

Na sequência, vem a etapa de definição das restrições, onde deve informar o nome da restrição, o caminho *XPATH* (XPATH, 2013), os dois elementos que deseja comparar, quais os tipos dos elementos, o modo de comparação (condição) e as ações que a ferramenta deve realizar no momento que o documento for verificado. A Figura 2 apresenta a tela para definição das restrições.

The screenshot shows a software interface for defining restrictions. The main title is "Name Restriction" followed by "VerificaDatas". Below this, there are several configuration fields:

- "Path Element": A dropdown menu labeled "Select Path" with the value "/dados/prestadores/consultas/consulta".
- "Condition Type (Element 1)": A dropdown menu labeled "Element Date".
- "Element 1": A dropdown menu labeled "data_lancamento".
- "Operação": A dropdown menu labeled "Less than".
- "Condition Type (Element 2)": A dropdown menu labeled "Element Date".
- "Element 2": A dropdown menu labeled "data_realizacao".
- "Action": A group of checkboxes with "Msg." checked and "Ins.", "Del.", "Ren.", "Upd." unchecked.
- "Message": A text input field containing "Datas Inconsistente, verifique!" followed by a "+" button.

At the bottom left is a large "ADD Restriction" button.

Figura 34. Página para definição das restrições

Ao clicar no botão chamado "Select Path" é apresentada uma nova página com o documento XML já carregado, de forma a facilitar a seleção do elemento a ser validado. Na definição das restrições o usuário pode indicar uma ou mais ações a serem realizadas. Estas ações são de 5 tipos:

- *Message*: Mensagem informativa;
- *Insert*: Insere um novo elemento após o pai do mesmo;
- *Delete*: Remove um elemento;
- *Rename*: Renomeia um elemento;
- *Update*: Atualiza o valor de um elemento;

Ao concluir o processo clicando no botão "ADD Restriction", a restrição é incluída em uma lista de restrições que estará na sessão do usuário. Assim, é possível que o usuário inclua diversas restrições em um único documento, pois o documento será gerado com base nesta lista de restrições (Figura 3). O usuário visualizará uma lista das suas restrições com possibilidade de excluí-las. Ao término da geração do documento XDC, é solicitado ao usuário que faça o *download* do documento gerado, para o uso posterior.

The screenshot shows a software interface for managing XML documents. At the top, there's a toolbar with checkboxes for 'Msg.' (checked), 'Ins.' (checked), 'Del.' (checked), 'Ren.' (unchecked), and 'Upd.' (unchecked). Below the toolbar, there are two message boxes: 'Message' containing 'Datas Inconsistentes, verifique!' and 'Message' containing 'Valores Não Conferem.'. Underneath these, there are sections for 'Element Father Insert' and 'New Element'. In 'Element Father Insert', there's a dropdown menu set to 'nome_instituicao' with a value 'ercereco_instituicao' and a note 'Value Averida Brasi, 000'. There are also '+' and '-' buttons. In 'Element Delete', there are dropdown menus for 'consulta' and 'nome_convenio', each with a corresponding '+' and '-' button. At the bottom left is a 'ADD Restriction' button.

Figura 35. Inclusão de múltiplas ações

Ao passo que o documento XDC foi gerado, pode-se passar para a etapa de validação do documento, para verificar se o documento está bem-formado. Este processo é feito com base em um *schema XML* próprio da linguagem XDCL. Após estas etapas, é possível fazer a validação do documento XML com base no documento XDC.

A validação dar-se-á pelo *parser* XDCL, implementado por Lazzaretti (2005), que irá validar as condições e executar as ações definidas no documento XDC. O algoritmo do *parser* XDCL é definido conforme um conjunto de ações as quais se aplicam recursivamente sobre o documento XML. Após o término da validação do documento, a ferramenta retornará algumas mensagens de alerta para que o usuário tome conhecimento do que ocorreu durante o processo de validação. Caso tenha sido definida alguma ação que modificou o documento XML original, será disponibilizado o *download* do documento com as modificações feitas pelas ações que o usuário definiu.

3.3. Estudo de Caso

Para demonstrar o uso da ferramenta é apresentado a seguir um estudo de caso com base num documento XML (Figura 5), que será utilizado como exemplo de restrição que será imposta pelo documento XDC (Figura 6) gerado pela ferramenta.

O documento XDC, Figura 6, possui a restrição que será avaliada. Neste caso, será verificado se o elemento `nome` é diferente do elemento `nome_responsável`, caso seja diferente, irá executar a ação de *Update* no elemento `nome_responsável`, informando que o valor correto deste elemento é "Stevan Danielli Costa".

```
<?xml version="1.0" encoding="UTF-8"?>
<dados>
    <instituicoes>124</instituicoes>
    <nome_instituicao>Hosp. Prontodinicas Ltda</nome_instituicao>
    <prestadores>
        <prestador>1452</prestador>
        <nome_prestador>Hospital SP Ltda</nome_prestador>
        <responsavel>
            <nome>Stevan Danielli Costa</nome>
            <endereco>Rua Tuiuti, 177</endereco>
            <telefone>99645188</telefone>
        </responsavel>
        <consultas>
            <nome_responsavel>Paulo Freitas</nome_responsavel>
            <consulta>
                <autorizacao>H13371</autorizacao>
                <paciente>14570</paciente>
                <nome_paciente>Adalgisa Severo</nome_paciente>
                <convenio>78</convenio>
                <nome_convenio>Saude Brasil Individual</nome_convenio>
                <medico>200</medico>
                <nome_medico>Adao Soares</nome_medico>
                <data_realizacao>12/10/2004</data_realizacao>
                <data_lancamento>12/11/2004</data_lancamento>
                <valor_banco>27.00</valor_banco>
                <medico_auto>Adao Soares</medico_auto>
                <valor_pago>27.00</valor_pago>
            </consulta>
        </consultas>
    </prestadores>
</dados>
```

Figura 36. XML de entrada

```
<?xml version="1.0" encoding="UTF-8"?>
<xdc1_constraints>
    <xdc1_constraint xdc1_name="pasta_responsavel">
        <xdc1_on>/dados/prestadores/*</xdc1_on>
        <xdc1_statements>
            <xdc1_set_conditions>
                <xdc1_condition xdc1_operator="Alt;&gt;">
                    <xdc1_operand1 type_condition="Element5t">
                        nome
                    </xdc1_operand1>
                    <xdc1_operand2 type_condition="Element5t">
                        nome_responsavel
                    </xdc1_operand2>
                </xdc1_condition>
            </xdc1_set_conditions>
            <xdc1_actions>
                <xdc1_update>
                    <update name_element="nome_responsavel"
                           type_value="Constant">
                        Stevan Danielli Costa
                    </update>
                </xdc1_update>
            </xdc1_actions>
        </xdc1_statements>
    </xdc1_constraint>
</xdc1_constraints>
```

Figura 37. Documento XDC gerado para validação

4. Conclusão

O artigo apresentou a estrutura da ferramenta WXDC que tem como propósito a validação de imposição de restrições de integridade de domínio de documentos XML. A validação do documento XML ocorre por meio de um documento XDC e da linguagem XDCL. A ferramenta faz a geração dos documentos XDC, abstraindo do usuário conhecimentos específicos sobre a linguagem XDCL, além de fazer a validação das restrições impostas e retornar um documento XML consistente para o usuário.

Com o desenvolvimento da ferramenta e sua utilização no estudo de caso foi possível impor restrições de integridade de domínio em documentos XML, não

suportadas por esquemas XML. A ferramenta possibilitou realizar o processo de validação de forma simples e prática por meio da interface disponibilizada, alcançando assim seu objetivo.

Como trabalhos futuros pretende-se aprimorar alguns aspectos da ferramenta como: *i)* incorporação de restrições referentes aos atributos, *ii)* inclusão da opção de comparação com base em dois documentos XML, e após isso disponibilizar a ferramenta para uso.

Referências

- Date, C.J. *Introdução a Sistemas de Banco de Dados*. 9. Reimpressão. Tradução de Daniel Vieira. Rio De Janeiro: Elsevier, 2003.
- Heuser, Carlos Alberto. *Projeto de Banco de Dados*. 4. ed. Cidade: Ed.SagraLuzzato, 1998. Disponível em: <http://www.julianoribeiro.com.br/troca/banco_de_dados/material_der.pdf> Último acesso em: 14 Ago. 2013.
- Lazzaretti, Alexandre Tagliari. *XDC: Uma Proposta de Controle de Restrições de Integridade de Domínio em Documento XML*. Dissertação. Universidade Federal de Santa Catarina Programa de Pós-Graduação em Ciências da Computação, Florianópolis, Fevereiro de 2005. Disponível em: <http://usuarios.upf.br/~lazzaretti/downloads/DI_LAZZA.pdf> Último acesso em: 12 Jul 2013
- Scribe'S N.E.W Studio. XML Validator Online [Internet], 2009. Disponível em: <<http://xmlvalidator.new-studio.org>>.
- Silberchatz, Abraham; KORTH, Henry. F; SUDARSHAN, S. *Sistemas de Banco de Dados*. Ed. Makron Books, 1999. 3a Edição
- Silva, Marinaldo Nunes Da. *Um Sistema de Controle de Integridade para Modelo de Dados Aberto*. Disponível em: <http://docs.computacao.ufcg.edu.br/posgraduacao/dissertacoes/2000/Dissertacao_MarinaldoNunesdaSilva.pdf> Último acesso em: 23 Jul 2013.
- W3C. Disponível em <<http://www.w3.org/standards/xml/schema>>. Último acesso em 29 Jul 2013.
- Xpath Language. Disponível em <<http://www.w3.org/standards/techs/xpath#completed>>. Último acesso em 19 Set 2013