

**UMA LINGUAGEM PARA MODELAGEM
CONCEITUAL EM XBRL**

VAGNER CLEMENTINO

UMA LINGUAGEM PARA MODELAGEM
CONCEITUAL EM XBRL

Proposta de dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

ORIENTADOR: RODOLFO F. RESENDE

Belo Horizonte

Maio de 2015

Lista de Figuras

1.1	O uso da XBRL no mundo	3
2.1	Fluxo de dados financeiros em uma organização. Adaptado de (Bergeron [2004]).	5
2.2	Diferentes interfaces para a interoperabilidade entre sistemas. Adaptado de (Bergeron [2004]).	6
2.3	Melhoria no fluxo de informação com XBRL. Adaptado de (Hoffman & Strand [2001]).	7

Sumário

Lista de Figuras	v
1 Introdução	1
2 Justificativa	5
3 Revisão da Literatura	9
4 Metodologia	11
4.1 Revisão Sistemática	11
4.2 Desenvolvimento da Linguagem	12
4.3 Ferramenta de Modelagem	12
4.4 Avaliação	12
Referências Bibliográficas	13
A Cronograma	15

Capítulo 1

Introdução

No cenário atual, onde os sistemas de informação tem se tornado grandes e complexos, vemos a existência de um ecossistema de sistemas, também conhecidos como *System of System* (SoS) Nakagawa et al. [2013]. Organizações de grande porte, como os governos nacionais, precisam projetar sistemas de sistema ao invés vez de sistemas isolados a fim de enfrentar desafios tais como: (i) colaboração entre organizações financiadas e geridas de forma independente; (ii) migração para um ambiente orientado a serviços (SOA¹); (iii) processos de testes e verificação da conformidade para sistemas de sistemas. Neste contexto, surge a necessidade do desenvolvimento de *abordagens, técnicas e tecnologias* para a interação e evolução dos SoS.

No tocante a interoperabilidade de dados, diversos padrões vêm sendo propostos. Mais recentemente o formato JSON Crockford [2006] vêm crescendo em popularidade, contudo, o intercâmbio de dados é realizado primordialmente através da XML (Extensible Markup Language²) e seus derivados. Na área medica, o padrão HL7 V3 message³ vêm sendo largamente adotado para troca de mensagem entre sistemas médicos Andriakopoulos & Belsis [2013]. No contexto dos Sistemas de Informação Geográficas (SIG) a GML (Geography Markup Language⁴) consolidou-se com o principal instrumento para interoperabilidade de dados geográficos Lu et al. [2007]. Na área financeira diversas linguagens de marcação vêm sendo utilizadas para o intercâmbio de informações na Internet: *Open Financial Exchange*⁵ (OFX), *Eletronic Business Using XML*⁶ (ebXML),

¹Service Oriented Architecture

²<http://www.w3.org/XML/>

³<http://www.hl7.org/>

⁴<http://www.opengeospatial.org/standards/gml>

⁵<http://www.ofx.net/>

⁶<http://www.ebxml.org>

*Financial Information eXchange*⁷ (FIX), *Market Data Definition Language*⁸, dentre outras da Silva et al. [2006].

Com o advento da Web as linguagens de marcação cresceram em importância sobretudo devido a necessidade de se adicionar significado a informações sendo transferidas. Contudo, o padrão HTML (Hypertext Markup Language) foi desenvolvido com objetivo de descrever *como* a informação deve ser apresentada e não possui qualquer compromisso com o significado da informação. Nos meados da década de 1980, a Organização para Padronização Internacional (ISO) propôs uma metalinguagem padrão a fim de etiquetar informações com conteúdo semântico. Esta linguagem foi denominada como *Standard Generalized Markup Language* (SGML) Smith & Stutely [1988].

A linguagem SGML, embora fosse capaz de definir diferentes tipos de marcação, a sua flexibilidade trouxe como preço a complexidade. O conceito era correto, todavia, havia a necessidade de ser mais simples. Com este objetivo em mente, um pequeno grupo de trabalho e um maior número de interessados começou a trabalhar em meados dos anos 1990 em um subconjunto de SGML conhecido como *Extensible Markup Language* (XML). A primeira versão foi publicada em 1996 e, dois anos mais tarde, o World Wide Web Consortium⁹ (W3C) publicou uma versão revisada Fawcett et al. [2012].

Conforme exposto, a XML foi especificada a partir da SGML, na tentativa de se resolver as limitações da HTML e da SGML. Neste sentido, um documento em XML pode ser publicado na Web, interpretado por pessoas ou processado por aplicações. Apropriando-se desta características da XML, diversas linguagens vêm sendo propostas com a finalidade de troca de informações.

A XBRL (*eXtensible Business Reporting Language*) é uma linguagem para divulgação e intercâmbio de informações financeiras baseada em XML (da Silva et al. [2006]). O padrão vem sendo adotado por diversas instituições e empresas em todo mundo com o suporte de um consórcio global¹⁰ com mais de 650 membros que incentivam a criação de jurisdições locais. Atualmente o consórcio conta com 24 jurisdições, sendo que em países como Estados Unidos, Grã-Bretanha e Austrália, a XBRL já é a linguagem oficial para entrega de relatórios à órgãos de governo. A figura 1.1 exibe os países que estão promovendo a adoção da XBRL.

Os estudos para definição da XBRL iniciaram em 1998 nos Estados Unidos pelo contador Charles Hoffman com apoio *Institute of Certified Public Accountants*

⁷<https://fixspec.com/>

⁸<http://www.mddl.org>

⁹<http://www.w3.org>

¹⁰www.xbrl.org

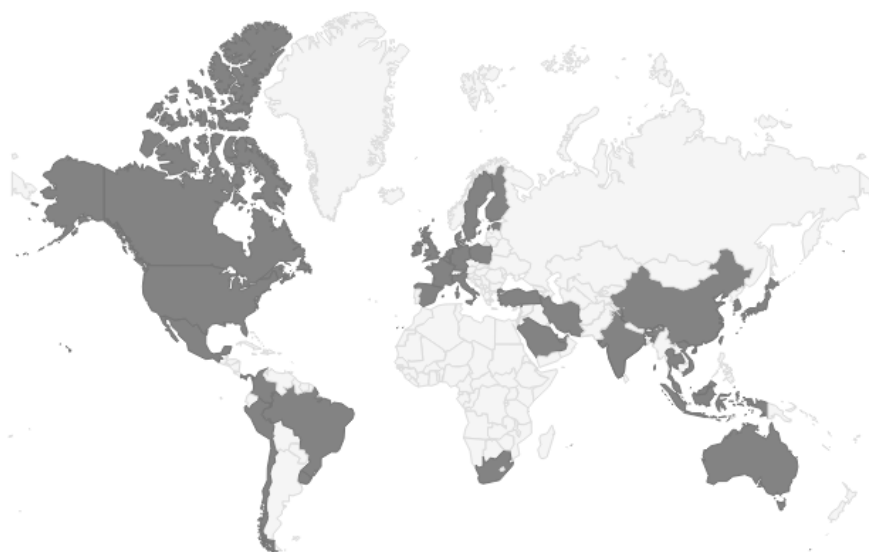


Figura 1.1. O uso da XBRL no mundo

(AICPA). O objetivo era utilizar a XML padronizar a divulgação de informações financeiras em formato eletrônico. A primeira versão do padrão, a XBRL 1.0, foi lançada em 2000, sendo a versão 2.0 lançada em dezembro de 2001. No mês de dezembro de 2003, foi lançada a versão 2.1 (Hoffman [2006]), corrigindo algumas deficiências detectadas na versão anterior. A versão 2.1 se mantém como a versão mais atual e estável da XBRL.

A linguagem XBRL define a estrutura básica dos documentos de instância, que são aqueles que portam os dados, e possibilita ainda a especificação de taxonomias que podem ser criadas para acomodar particularidades de cada organização por meio da introdução de novos elementos, denominados conceitos. Neste sentido, a linguagem possui elementos que facilitam a sua extensão e, por consequência, a adoção em diversos contextos.

Apesar de sua crescente adoção, falta à XBRL uma notação que facilite a sua modelagem e a comunicação entre os diferentes *stakeholders*. A necessidade de um modelo abstrato para a XBRL foi manifestada pelo próprio *XBRL Consortium*. Em 2010 (International [2010]), o consórcio declarou que a criação de um modelo conceitual é uma das seis iniciativas que darão suporte a contínua adoção do padrão. Com o objetivo de preencher esta lacuna propõe neste documento o desenvolvimento de uma linguagem conceitual para a XBRL. A seguir descreve-se como o documento está estruturado.

No Capítulo 2 discute-se as justificas para adoção da XBRL bem como da criação de uma linguagem conceitual. O Capítulo 3 apresenta-se a revisão da literatura no tocante a criação de ontologias em diversos campos dos conhecimento, especialmente

para área financeira e contábil. No Capítulo 4 é discutida a metodologia a ser aplicada. No Apêndice A é exibido o cronograma do trabalho.

Capítulo 2

Justificativa

Nas organizações as informações contábeis e financeiras são armazenadas em diversos formatos (planilhas eletrônicas, documentos de texto, bancos de dados relacionais e etc). Não obstante, se faz necessário a transformação destas informações em um formato único a fim de facilitar a sua recuperação bem como a sua transmissão para outros sistemas. O processo de transformação e redirecionamento da informação internamente na organização, entre a organização e suas filiais ou mesmo entre a empresa e os governos. Este fluxo de informação, especialmente para a geração de relatórios, é exibido na figura 2.1.

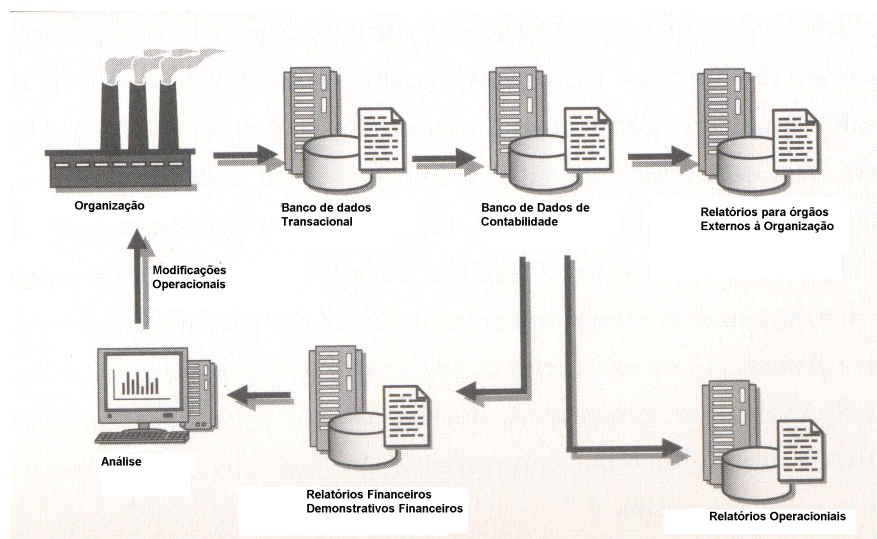


Figura 2.1. Fluxo de dados financeiros em uma organização. Adaptado de (Bergeron [2004]).

Se pensarmos em uma organização que necessitem receber informações de diversos locais, como por exemplo o governo federal de uma país que solicite a prestação de

contas de estados e municípios, onde cada ente possui seu próprio sistema para registro dos fatos financeiros. Neste contexto, haverá a necessidade de se criar diferentes interfaces para a conversão de formatos e padrões de contabilização. A figura 2.2 ilustra este cenário.

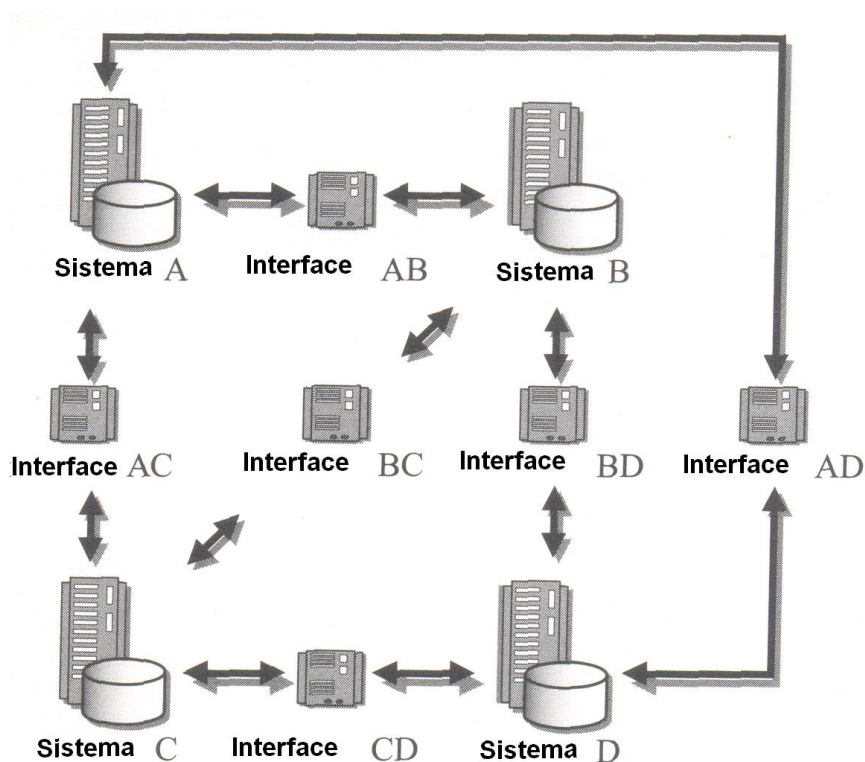


Figura 2.2. Diferentes interfaces para a interoperabilidade entre sistemas. Adaptado de (Bergeron [2004]).

A solução utilizada para minimizar estes problemas é adoção de uma linguagem de marcação que facilite o intercâmbio e apresentação na Internet, bem como proporcione o armazenamento em qualquer base de dados. Neste sentido a XBRL vêm sendo adotado como padrão em diversos países. O processo de troca de informações financeiras é simplificado pela XBRL tendo em vista que a transformação da informação original é realizada uma única vez para o formato XBRL. Posteriormente a informação poderá ser reutilizada e/ou distribuída automaticamente para diversos outros formatos. A figura 2.3 exibe a simplificação na troca de informação com a adoção do XBRL.

Apesar de sua importância e crescente adoção, o *XBRL Consortium* sentiu a necessidade de definir um modelo abstrato da XBRL (International [2010]) que facilitasse a compreensão da linguagem pelos profissionais da tecnologia da informação. Em 2012

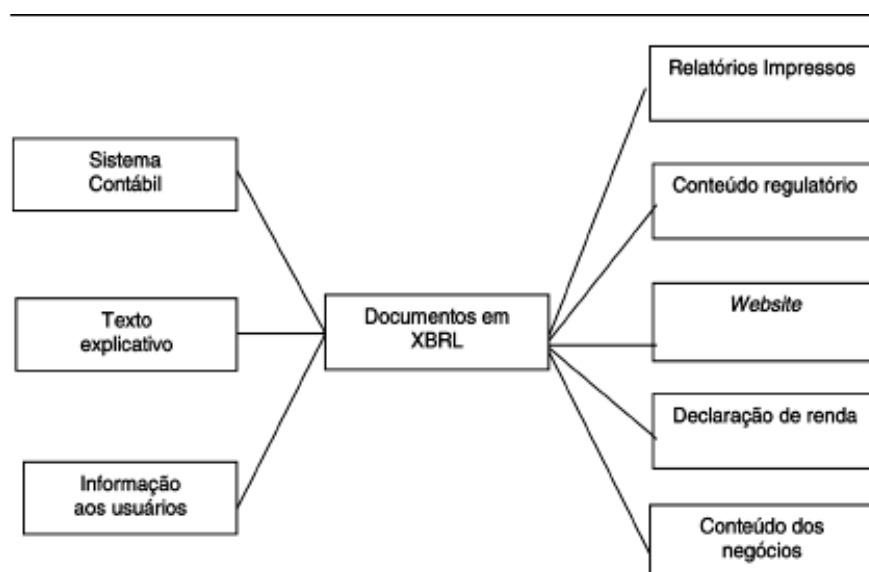


Figura 2.3. Melhoria no fluxo de informação com XBRL. Adaptado de (Hoffman & Strand [2001]).

foi proposto o *XBRL Abstract Model 2.0*¹ que consiste basicamente de uma versão estendida da UML com objetivo de captura aspectos semânticos da XBRL.

Um modelo abstrato deverá remover do seu escopo qualquer questão relativa à implementação do objeto modelado. Neste sentido, no caso de XBRL, o modelo abstrato deveria remover todas as referências a XML, o que não ocorreu no XBRL Abstract Model. Ademais, apesar da UML ser largamente utilizada entre os profissionais de Tecnologia da Informação, o seu uso para comunicação com os demais stakeholders muitas vezes não é o ideal (Peixoto et al. [2008]).

Neste contexto se faz necessário a proposição de linguagem conceitual para a XBRL com as seguintes características:

- Não contenha questões relativa ao XML
- Possibilite o desenho de aplicações que utilizem a XBRL
- Facilite a comunicação entre os diversos stakeholders envolvidos no domínio da XBRL.

No próximo capítulos iremos revisar a literatura relativo à definição de ontologias, especialmente no domínio contábil/financeiro.

¹<http://www.xbrl.org/Specification/abstractmodel-primary/PWD-2012-06-06/abstractmodel-primary-pwd-2012-06-06.html>

Capítulo 3

Revisão da Literatura

Bosch & Mathiak [2013] apresentam um estudo sobre a criação de ontologias com base em XML Schemas. Arquivos no formato XML Schemas são utilizados por linguagem como XBRL para definir os conceitos a serem utilizados em determinado domínio. Neste sentido, todas as informações localizadas em XML Schemas podem ser reutilizadas a fim de reduzir a necessidade de criar ontologias de domínio a partir do zero. Os autores argumentam que o tempo e esforço economizado poderiam ser utilizados para acrescentar informações semânticas específicas do domínio específico, tendo em vista que este tipo de detalhamento não existem nos XML Schemas.

Em Pierre [2008] há uma discussão sobre alguns dos problemas do uso de representações formais no domínio da contabilidade. O objetivo do autor era construir uma representação das informações financeiras de uma forma eficiente. Todavia, devido a própria maneira que a contabilidade é construída, basicamente de leis e normas, existem diversos problemas na representação dos elementos da mesma. Não obstante, Pierre considera a XBRL como uma boa alternativa para uma representação formal da contabilidade. O autor ressalta que a XBRL trouxe novos elementos para o processo de criação de relatórios financeiros: (i) facilidade na produção e publicação das informações contábeis; (ii) compartilhamento e possibilidade comparação partilha de informação; (iii) verificação e certificação da informação; (iv) ganho no processo de análise. Ele conclui que XBRL é bom padrão para armazenar informações, contudo, faz uma crítica ao padrão por não fornecer nenhuma formalização explícita de dados financeiros ou das normas contabilísticas.

Lupaşcu et al. [2010] realiza uma análise do REA Framework¹ como uma ontologia de contabilidade para sistemas de informação. Os elementos básicos do REA são recursos, eventos, agentes, fluxos, controle e *duality* (McCarthy [1982]). Os autores

¹<http://reatechnology.com/>

estendem o REA a fim de possibilitar a adição e compartilhamento do conhecimento.

Gailly et al. [2007] utilizaram a UML para representar graficamente ontologias baseadas no REA. Os autores utilizaram metodologias tradicionais para a criação de ontologias para desenvolver o modelo proposto e argumentam que REA é uma ontologia de domínio do negócio. Eles concluem que linguagens conceituais tais como UML possuem a riqueza necessária para representar REA componentes ontológicos. Todavia, alguns detalhes dos elementos representados não são capazes de serem explicitados.

Sugumaran & Storey [2002] discutem a criação, uso e gestão de ontologias para modelagem conceitual. Eles também propõe um processo para gestão de ontologias. Os objetivos dos trabalho incluem demonstrar a importância de ontologias na modelagem conceitual e propor um modelo baseada em heurística para a criação de ontologias em determinado domínio. Eles observam que a maioria das ontologias são criadas manualmente e que não existe uma metodologia largamente aceita para construção das mesmas. O seu método de criação ontologia proposto é composto por quatro etapas. O primeiro passo é a identificação de termos básicos do domínio. Isto inclui duas subetapas: a identificação dos termos mais frequentes e a identificação de sinônimos e termos relacionados. O segundo passo é a identificação dos relacionamentos. Existem três tipos de relações: relações entre termos básicos, relações entre ontologias, e as relações entre ontologias e sub-ontologias. O terceiro passo é a identificação de restrições e regras do domínio. O quarto passo é a identificação de restrições de nível superior. Estas restrições capturar o conhecimento de domínio. Os autores observam ainda que, enquanto a maioria das ontologias são construídos estaticamente, a maioria dos domínios, na verdade, evoluem com o tempo. O desenvolvimento de uma ontologia deve incluir esta evolução.

Capítulo 4

Metodologia

O processo de desenvolvimento deste trabalho pode ser dividido em quatro partes principais: *I - Revisão Sistemática da Literatura*; *II - Desenvolvimento da Linguagem*; *III - Construção da Ferramenta*; *IV - Avaliação*. Cada uma das etapas é detalhada nas próximas seções.

4.1 Revisão Sistemática

O primeiro passo no desenvolvimento do trabalho será revisar a literatura relativa a criação de ontologias, especialmente no domínio contábil/financeiro. Conforme proposto por Wohlin et al. [2012] será definido um protocolo de revisão com o objetivo de conduzir o processo de coleta de artigos. Este protocolo conterá, dentre outros tópicos, os critérios de seleção de trabalhos, estratégia para extração de dados; definição de questões de pesquisas, dentre outros. O protocolo será revisto com o orientador visando garantir suas consistência e validade.

Nesta etapa também será revisada a especificação da XBRL¹ com o objetivo de conseguir subsídios para a criação da ontologia. Conforme discutido em Lupaşcu et al. [2010], um dos passos para criação de uma ontologia efetiva é a identificação de termos básicos do domínio. Nesta tarefa poderá utilizado técnicas de recuperação da informação como índices invertidos (Baeza-Yates et al. [1999]).

¹<http://specifications.xbrl.org/specifications.html>

4.2 Desenvolvimento da Linguagem

Finalizado o levantamento teórico do trabalho, o próximo passo é definir os elementos da linguagem. Por se tratar de uma linguagem conceitual visando a facilitar modelagem e a comunicação da XBRL, nesta fase será construído os elementos gráficos da linguagem. Tais elementos irão se apropriar dos objetos de linguagens como a UML e a BPMN, contudo, fornecendo primitivas adequadas para a representação dos dados contábeis e financeiros.

4.3 Ferramenta de Modelagem

A adoção de uma nova notação será facilitada se existirem ferramentas que facilitem a sua criação, alteração e disponibilização. Neste sentido, faria pouco sentido definir uma linguagem conceitual para a XBRL sem, contudo, criar uma ferramenta que lhe desse apoio. Assim, uma das etapas deste trabalho será o desenvolvimento de uma ferramenta que possibilite a criação de diagrama no modelo proposto. Esta ferramenta será de suma importância na última etapa do trabalho, que trata da avaliação da linguagem proposta.

4.4 Avaliação

A fim de avaliar se a linguagem proposto atendeu aos seus requisitos, será realizadas uma bateria de testes com diferentes usuários. O baseline será o *XBRL Abstract Model*, onde pretende-se analisar fatores como capacidade de comunicação, expressividade, dentre outros aspectos das duas linguagens.

O experimentos de avaliação será composto de duas etapas. Em um primeiro momento, profissionais de tecnologia da informação e contabilidade, serão convidados a modelar uma taxonomia XBRL utilizando a linguagem proposta e o XBRL Abstract Model. Posteriormente, estes diagramas serão avaliados por outros profissionais no tocante as aspectos como expressividade. Estes dados serão tabulados para posterior avaliação.

Referências Bibliográficas

- Andrikopoulos, P. K. & Belsis, P. (2013). Towards effective organization of medical data. Em *Proceedings of the 17th Panhellenic Conference on Informatics*, PCI '13, pp. 305--310, New York, NY, USA. ACM.
- Baeza-Yates, R.; Ribeiro-Neto, B. et al. (1999). *Modern information retrieval*, volume 463. ACM press New York.
- Bergeron, B. (2004). *Essentials of XBRL: Financial Reporting in the 21st Century*. Essentials Series. Wiley. ISBN 9780471458081.
- Bosch, T. & Mathiak, B. (2013). How to accelerate the process of designing domain ontologies based on xml schemas. *Int. J. Metadata Semant. Ontologies*, 8(3):254--266. ISSN 1744-2621.
- Crockford, D. (2006). The application/json media type for javascript object notation (json). RFC 4627, IETF.
- da Silva, P. C.; da Silva, L. G. C. & Aquino jr, I. J. e. S. (2006). *XBRL Extensible Business Reporting Language - Conceitos e Aplicações*. Editora Ciência Moderna, 1ª edição. ISBN: 8573934999.
- Fawcett, J.; Ayers, D. & Quin, L. R. E. (2012). *Beginning XML*. Wrox Press Ltd., Birmingham, UK, UK, 5th edição. ISBN 1118162137, 9781118162132.
- Gailly, F.; Laurier, W. & Poels, G. (2007). Positioning rea as a business domain ontology. Em *Resource Event Agent-25 (REA-25) Conference, Newark, Delaware, USA*. Citeseer.
- Hoffman, C. (2006). *Financial Reporting Using XBRL: IFRS and US GAAP Edition*. Lulu.com, first edição. ISBN-10: 1411679792 ISBN-13: 978-1411679795.

- Hoffman, C. & Strand, C. (2001). *XBRL Essentials: A Nontechnical Introduction to EXtensible Business Reporting Language (XBRL), the Digital Language of Business Reporting*. American Institute of Certified Public Accountants. ISBN 9780870513534.
- International, X. (2010). Preserve. promote. participate. moving xbrl forward. Relatório técnico, XBRL International. Disponível em: <http://www.xbrl.org/2010Initiatives/xbrl2010initiatives.pdf>.
- Lu, C.-T.; Dos Santos, Raimundo F., J.; Sripada, L. & Kou, Y. (2007). Advances in gml for geospatial applications. *GeoInformatica*, 11(1):131–157. ISSN 1384-6175.
- Lupaşcu, A.; Lupaşcu, I. & Negoescu, G. (2010). The role of ontologies for designing accounting information systems. *The Annals of “Dunarea de Jos” University of Galati, Fascicle I, Economics and Applied Informatics. Years XVI no, 1*.
- McCarthy, W. E. (1982). The rea accounting model: A generalized framework for accounting systems in a shared data environment. *Accounting Review*, pp. 554--578.
- Nakagawa, E. Y.; Gonçalves, M.; Guessi, M.; Oliveira, L. B. R. & Oquendo, F. (2013). The state of the art and future perspectives in systems of systems software architectures. Em *Proceedings of the First International Workshop on Software Engineering for Systems-of-Systems*, SESoS '13, pp. 13--20, New York, NY, USA. ACM.
- Peixoto, D.; Batista, V.; Atayde, A.; Borges, E.; Resende, R. & Pádua, C. (2008). A comparison of bpmn and uml 2.0 activity diagrams. *VII Simposio Brasileiro de Qualidade de Software*, 56.
- Pierre, T. (2008). The use of a formal representation of accounting standards. *IJCSA*, 5(3b):93–116.
- Smith, J. & Stutely, R. (1988). *SGML: the user's guide to ISO 8879*. Ellis Horwood Series in Computers and their Applications. E. Horwood. ISBN 9780470211267.
- Sugumaran, V. & Storey, V. C. (2002). Ontologies for conceptual modeling: their creation, use, and management. *Data & knowledge engineering*, 42(3):251--271.
- Wohlin, C.; Runeson, P.; Höst, M.; Ohlsson, M. C.; Regnell, B. & Wesslén, A. (2012). *Experimentation in software engineering*. Springer Science & Business Media.

Anexo A

Cronograma

#	Atividade	Início (MM/AAAA)	Termino (MM/AAAA)
01	Revisão da literatura	07/2015	08/2015
02	Revisão da especificação XBRL	09/2015	09/2015
03	Definição da linguagem	10/2015	12/2015
04	Construção da ferramenta	01/2016	02/2016
05	Avaliação da linguagem	03/2016	03/2016
06	Redigir a dissertação	04/2015	05/2016
07	Defesa da dissertação	07/2016	07/2016