

(XOR) XBRL-OBJECT-RELATIONAL: A FRAMEWORK FOR EXTRACTION, TRANSFORMATION, LOADING AND AUDITING OF FINANCIAL DATA IN RELATIONAL BASES AND XBRL INSTANCES

Abstract: XBRL is a relatively new technology that has been widely used worldwide in the process of standardizing financial data. However, despite all its flexibility and possibilities, it is not yet fully explored in countries like Brazil, due to the complexity inherent in the technological adaptations necessary for its implementation. This work presents a framework whose main objective is to minimize the difficulties inherent to the Extraction, Transformation and Loading (ETL) process of data from relational databases, mapped and validated by XBRL taxonomies. The proposed framework consists of a specific programming language for the financial and XBRL domain, which allows the creation of algorithms and operations for financial inspection and audits on the mapped data.

Keywords: XBRL, ETL (Extract, Transform and Load), Financial Language, ramework

(XOR) XBRL-OBJETO-RELACIONAL: UM *FRAMEWORK* PARA EXTRAÇÃO, TRANSFORMAÇÃO, CARGA E AUDITORIA DE DADOS FINANCEIROS EM BASES RELACIONAIS E INSTÂNCIAS XBRL

Resumo: A XBRL é uma tecnologia relativamente nova que vem sendo amplamente utilizada em todo mundo no processo de padronização de dados financeiros. Todavia, apesar de toda a sua flexibilidade e possibilidades, ainda não é totalmente explorada em países como o Brasil, devido a complexidade inerente as adequações tecnológicas necessárias a sua implementação. Esse trabalho apresenta um *framework* cujo objetivo principal é minimizar as dificuldades inerentes ao processo de Extração, Transformação e Carga (*ETL*) de dados oriundos de bases relacionais, mapeados e validados por taxonomias XBRL. O *framework* proposto é constituído por uma linguagem de programação específica para o domínio financeiro e de XBRL, que permite a criação de algoritmos e operações para fiscalização e auditorias financeiras sobre os dados mapeados.

Palavras-chave: XBRL, *ETL* (Extração, Transformação e Carga), Linguagem Financeira, Framework

1. INTRODUÇÃO

A ampliação do uso de Tecnologias da Informação e Comunicação (TICs) nos últimos anos favoreceu a expansão de diversas atividades econômicas, permitindo maior transparência e atratividade de investimentos, principalmente pela facilidade na difusão e análise de informações (LUCIANO et. al., 2017).

A inovação tecnológica propiciou um mercado globalizado, onde eficiência e eficácia tornaram-se pilares fundamentais para a longevidade das empresas. Nesse contexto, as organizações passaram a sofrer pressão para agir de forma transparente, tornando-se necessário a divulgação constante de informações de seu desempenho social, ambiental e econômico (LUCIANO et. al., 2017). Entretanto, o processamento e a análise de informações de gestão, financeiras ou tributárias são complexas e onerosas, face ao volume de dados produzidos e a quantidade de bases de armazenamento existentes, e que, na maioria dos casos, não possuem o interfaceamento necessário que possibilite sua extração e transformação de forma que facilite a organização e disponibilização dessas informações. Ademais, há de se considerar a necessidade de um padrão que satisfaça todos os usuários que utilizam os dados, pois as sucessivas transformações podem acarretar distorções ou erros (MAUS et. al., 2009).

A XBRL (eXtensible Business Reporting Language), apesar de não ser uma tecnologia recente, ganhou visibilidade nos últimos anos por minimizar o impacto decorrente do intercâmbio de informações financeiras pelo mundo, tornando-se um padrão. Essa tecnologia proporcionou um grande avanço porque permite a criação de regras que validem e permeiem a legislação vigente. Em muitos países, incluindo o Brasil, a XBRL ganhou espaço, como exemplo podemos citar o projeto da Secretaria do Tesouro Nacional, que desenvolveu o projeto SICONFI (Sistemas de Informações Contábeis e Fiscais do Setor Público Brasileiro), cuja base de transmissão de dados é baseada em XBRL¹.

Entretanto, apesar da tecnologia XBRL se apresentar como possível solução na diminuição dos impactos inerentes as diferenças semânticas e de complexidade na interoperabilidade de dados financeiros, sua implementação ou utilização são escassas, principalmente em função da falta de capital humano capacitado para implementação e suporte ou pelo custo de sua implantação devido ao impacto tecnológico, conforme pode ser observado no estudo realizado por Farias (2014).

Assim, este trabalho propõem a construção de um *framework* XBRL-OBJETO-RELACIONAL que diminua sobremaneira a complexidade na manipulação de dados para exportação ou importação de informações em um padrão XBRL, podendo ser utilizada tanto como ferramenta ETL (*Extract, Transform, Load*) para importação/exportação a partir de arquivos TXT, XML ou banco de dados relacional, mapeados por um arquivo XML do *framework*, como para realização de trabalhos de fiscalização e auditoria através da manipulação de informações XBRL utilizando uma linguagem financeira incorporada ao *framework*.

Este trabalho é composto por quatro seções. A segunda seção é destinada à fundamentação teórica sobre a tecnologia XBRL. Em seguida, na terceira seção, é apresentada a proposta do *framework* a fim de facilitar o processo ETL, fiscalização e auditoria tanto por pequenas empresas, quanto por municípios e estados. Por fim, na quarta seção, são colocadas as conclusões deste trabalho seguidas das referências bibliográficas.

¹ <https://siconfi.tesouro.gov.br/siconfi/pages/public/conteudo/conteudo.jsf?id=584>

2. XBRL E AS DIFICULDADES DECORRENTES DE SUA UTILIZAÇÃO

Desenvolvida pelo contador Charles Hoffman, em 1998, a XBRL (*eXtensible Business Reporting Language*) foi concebida com a finalidade de padronizar o intercâmbio de informações financeiras, o que permite que os dados concebidos por essa tecnologia possam ser utilizadas por diversos aplicativos de software, minimizando as sucessivas transformações necessárias ao intercâmbio das informações (SILVA, 2006).

Um documento XBRL possui em sua estrutura elementos que são definidos pela interrelação de componentes, conforme pode ser observado abaixo (SILVA, 2006):

- Taxonomia: parte da XBRL que descreve como deve ser a estrutura, rótulos, relações, características e até a ordem de apresentação dos fatos. Também é um documento XML, derivado do XML Schema e Xlink, funcionando como uma biblioteca ou vocabulário para o documento de instância da XBRL;
- Documento de Instância: onde os dados financeiro ou contábeis estão dispostos e organizados segundo a especificação da XBRL *Instance Schema* e taxonomia;
- Documentos *linkbases*: expressam os relacionamentos entre elementos da taxonomia e os elementos identificadores com o documento XBRL.

A tecnologia XBRL é empregada por várias entidades públicas e privadas no mundo, não apenas pela possibilidade de padronização dos dados financeiros e contábeis, mas porque não é como a maioria das linguagens de marcação que possibilitam apenas a apresentação de relatórios estáticos, pois ela é dinâmica e interativa, oferecendo diversas possibilidades de análise em um mesmo documento, além de sua extensibilidade (MAUS, 2009).

Pelos benefícios proporcionados, a XBRL passou a ser pesquisada também no Brasil, na tentativa de ser incorporada pelo governo e empresas. Todavia, como apresentado por Farias (2014), existem entraves tecnológicos e de mão-de-obra que terminam inviabilizando sua utilização. Farias (2014) ressalta que a XBRL não têm desenvolvimento ou aplicação fácil, que apesar do alcance mundial, existem barreiras complexas que precisam ser dirimidas, das quais ele destaca: “a complexidade da XBRL; cultura para aceitação das mudanças; custo de implantação; desconhecimento dos benefícios a serem alcançados; escassez de capital humano com conhecimento sobre o tema; Impacto tecnológico; sistema jurídico do país; e processo de convencimento dos envolvidos;”.

3. FRAMEWORK XOR (XBRL-OBJETO-RELACIONAL)

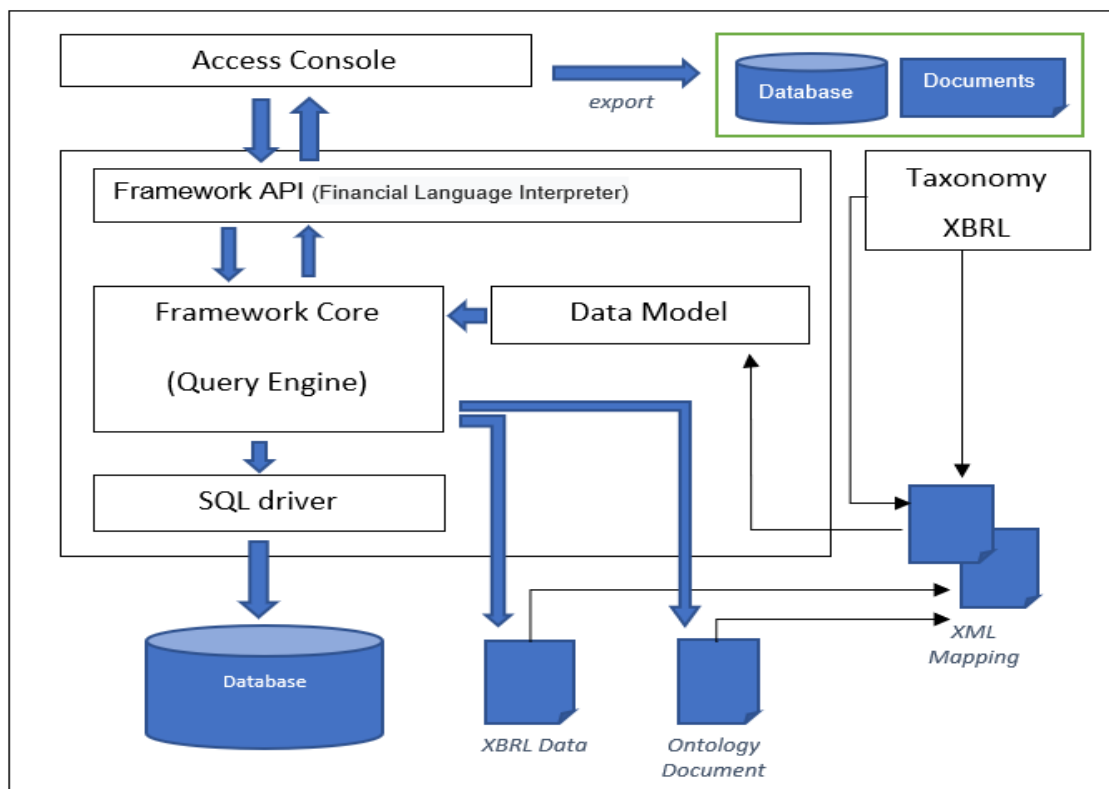
Este trabalho surgiu das necessidades apresentadas pelos estudos de Farias (2014) e outras, i.e. falta de capital humano capacitado para implementação e suporte, custo de implantação, impacto tecnológico, complexidade de XBRL, apesar da sua potencialidade comprovada, tendo como objetivo principal o desenvolvimento de um *framework open source* XBRL-OBJETO-RELACIONAL que facilite o processo de

ETL de dados de qualquer base relacional e ou arquivos XML/TXT para instâncias XBRL ou vice-versa, utilizando uma determinada taxonomia, conforme arquitetura ilustrada na Figura 1.

A proposta da arquitetura dar-se-á mediante mapeamentos em arquivos XML com os dados de conexão, as tabelas para estabelecer a relação entre os atributos e a taxonomia XBRL adotada. O *framework* deve possibilitar a conexão a diversas bases relacionais ao mesmo tempo e consultar informações via linguagem *LMDQL* (*Link-Based Multidimensional Query Language*), cujos conceitos podem ser obtidos no trabalho desenvolvido por Silva (2009), bem como através da linguagem financeira implementada – ou ambas ao mesmo tempo, incorporada ao *framework* para realização de operações, incluindo a criação de algoritmos que usarão a nomenclatura dos campos mapeados no arquivo XML.

Com base na Figura 1, é possível vislumbrar a arquitetura e funcionalidades do *framework*, as quais serão detalhadas a seguir.

Figura 1 – Arquitetura do *Framework*



Fonte: Autores

O *framework* será construído em linguagem Java e Java CC (Java Compiler Compiler) pela facilidade que possui tanto para a especificação da sintaxe da nova linguagem financeira quanto pela sua flexibilidade na execução em diversos sistemas operacionais. Nesse aspecto é importante salientar que uma nova linguagem está sendo especificada em virtude de que no contexto financeiro existe ainda uma dificuldade crescente no que tange a manipulação e tratamento de dados em instâncias XBRL que possibilite a profissionais como auditores ou analistas codificarem algoritmos de propósitos gerais para esse domínio.

O *Access Console* é o componente do *framework* onde comandos em SQL, que serão adaptados para manipular XML, ou através da linguagem financeira, poderão ser executados sobre os dados, bem como realizar os processos de exportação e importação baseados na taxonomia definida.

O *Framework API* é onde serão definidas ou parametrizadas as sintaxes das linguagens *LMDQL* e financeira, sendo também responsável pela interpretação dos comandos a serem executados, estabelecendo relação com as interfaces do *Framework Core*. A *LMDQL* é linguagem de consulta OLAP para documentos XML interligados por XLink desenvolvida por Silva (2009). O processo de mapeamento em XML pode ser observado através do exemplo apresentado na Figura 2, onde os atributos são mapeados para os campos de determinada tabela, bem como a conexão com o banco de dados – essas propriedades podem ser verificadas no atributo *class* do XML.

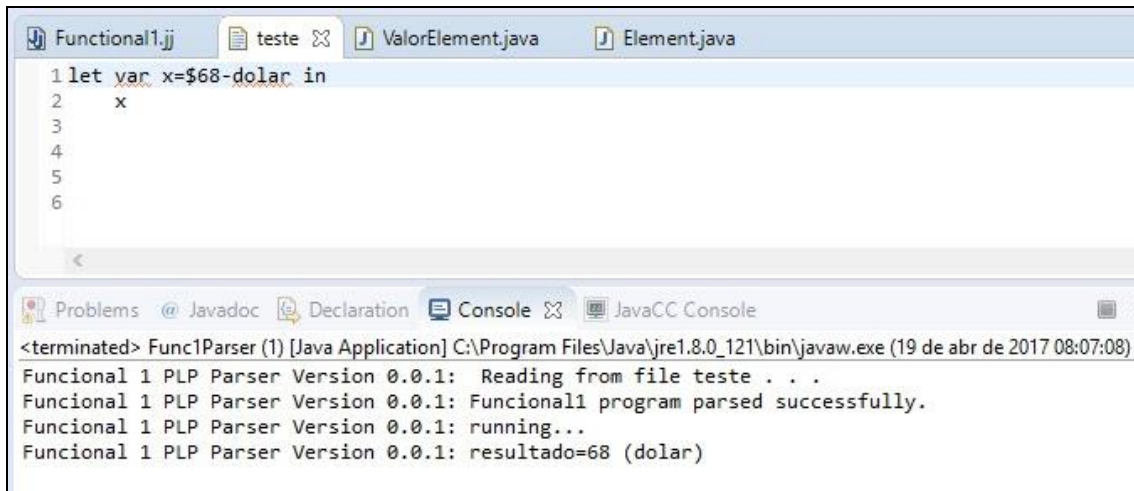
Figura 2 – Arquivo de Mapeamento (Exemplo)

```
<?xml version="1.0" encoding="utf-8" ?>
<xbrlframe-mapping xmlns="urn:nxbrlframe-mapping-1.0">
  <class name="PagamentosAbril" table="financeiro"
  conection="empresaDois">
    <id name="Codigo" column="codigo" type="Int32">
      <period>
        <period:startDate property name="start"
        column="data_inicio" type="Date" length="15" />
        <period:startEnd property name="end" column="data_fim"
        type="Date" length="15" />
      </period>
      <property name="totalOfReceivables" column="valor_total"
      type="Float" length="12" />
    </class>
  </xbrlframe-mapping>
```

Fonte: Adaptado de Bauer e King (2005)

A linguagem financeira, em desenvolvimento, foi estabelecida pela necessidade de analistas e auditores financeiros executarem algoritmos de cunho contábil ou financeiro, utilizando expressões próprias para essa finalidade. Assim, para essa nova linguagem serão definidos tipos e operações específicas desta área. Dentre os principais tipos já implementados podemos destacar: **Monetary Value (\$)**: variável de valor monetário, refere-se a uma determinada quantia financeira com a qual é possível descrever o valor exato e a respectiva moeda (exemplo mostrado na Figura 3); **Element Value (\$ele)**: variável de valor *elemento*, descreve o elemento financeiro (de contabilidade), a qual será processada no algoritmo. Isso propicia semântica financeira no desenvolvimento do código (exemplo ilustrado na Figura 4); **Schema Value (\$sch)**: variável de valor *Schema*, descreve um conjunto de *elements* que podem estar contidos no algoritmo. Serve para validação dos elementos financeiros descritos no código; **Link Expression (@)**: A expressão *Link* é uma expressão binária, a qual permite estabelecer hierarquia entre elementos financeiros (atributo comum em relatórios financeiros).

Figura 3 – Exemplo de *Monetary Value (\$)*

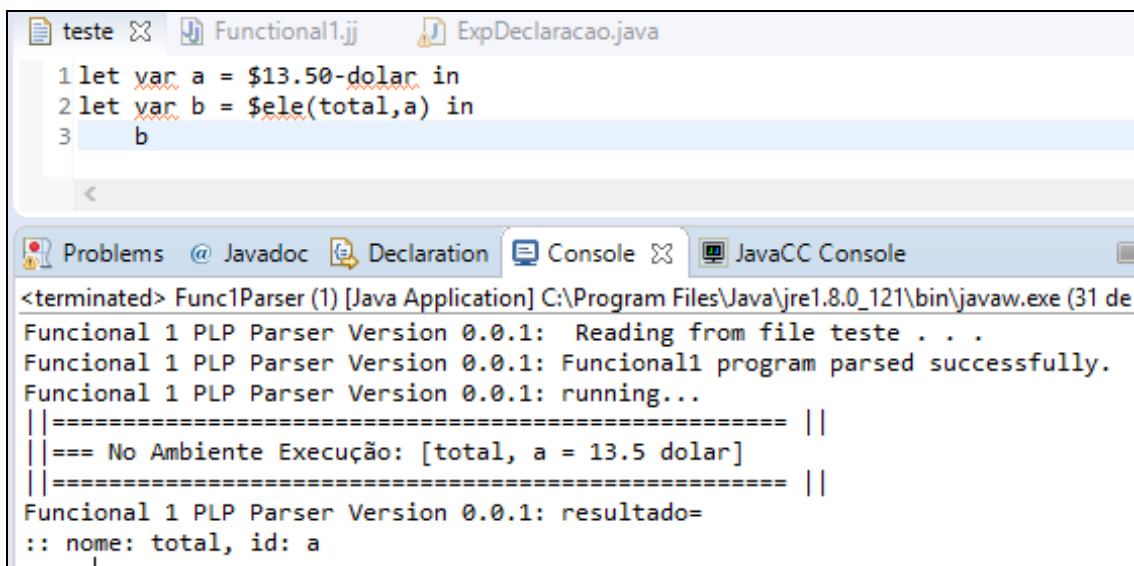


```
Functional1.jj  teste  ValorElement.java  Element.java
1 let var x=$68-dolar in
2   x
3
4
5
6

Problems  Javadoc  Declaration  Console  JavaCC Console
<terminated> Func1Parser (1) [Java Application] C:\Program Files\Java\jre1.8.0_121\bin\javaw.exe (19 de abr de 2017 08:07:08)
Funcional 1 PLP Parser Version 0.0.1: Reading from file teste . . .
Funcional 1 PLP Parser Version 0.0.1: Funcional1 program parsed successfully.
Funcional 1 PLP Parser Version 0.0.1: running...
Funcional 1 PLP Parser Version 0.0.1: resultado=68 (dolar)
```

Fonte: Autores

Figura 4 – Exemplo de *Element Value (\$ele)*



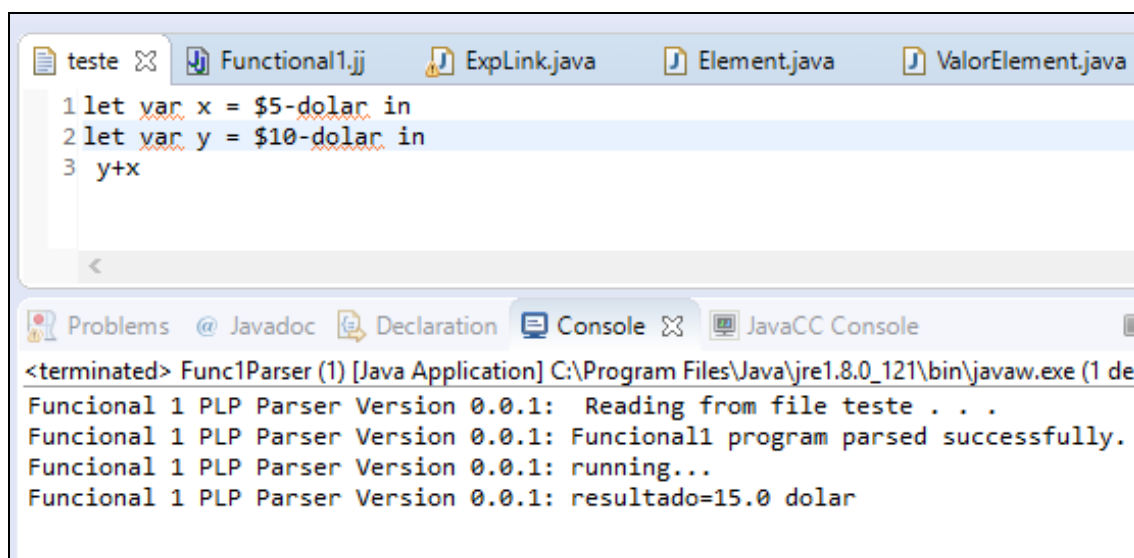
```
teste  Functional1.jj  ExpDeclaracao.java
1 let var a = $13.50-dolar in
2 let var b = $ele(total,a) in
3   b

Problems  Javadoc  Declaration  Console  JavaCC Console
<terminated> Func1Parser (1) [Java Application] C:\Program Files\Java\jre1.8.0_121\bin\javaw.exe (31 de
Funcional 1 PLP Parser Version 0.0.1: Reading from file teste . . .
Funcional 1 PLP Parser Version 0.0.1: Funcional1 program parsed successfully.
Funcional 1 PLP Parser Version 0.0.1: running...
||=====||
||=== No Ambiente Execução: [total, a = 13.5 dolar] ||
||=====||
Funcional 1 PLP Parser Version 0.0.1: resultado=
:: nome: total, id: a
```

Fonte: Autores

Além dos tipos já implementados e que podem ser observados nas Figuras 3 e 4, existem também alguns operadores em fase de implementação e teste, como o *MonetarySum Expression (+)*: que permite fazer soma entre valores monetários de mesma moeda); o *MonetarySubtraction (-)*: que permite fazer subtração entre valores monetários de mesma moeda; *MonetaryDivision (/)*: que permite fazer divisão entre valores monetários de mesma moeda; *MonetaryMultiplication (*)*: que permite fazer multiplicação entre valores monetários de mesma moeda. Na Figura 5 é possível identificar um exemplo da execução do operador *MonetarySum Expression (+)*.

Figura 5 – Exemplo de *MonetarySum Expression* (+)



```
1 let var x = $5-dolar in
2 let var y = $10-dolar in
3 y+x

<terminated> Func1Parser (1) [Java Application] C:\Program Files\Java\jre1.8.0_121\bin\javaw.exe (1 de
Funcional 1 PLP Parser Version 0.0.1: Reading from file teste . . .
Funcional 1 PLP Parser Version 0.0.1: Funcional1 program parsed successfully.
Funcional 1 PLP Parser Version 0.0.1: running...
Funcional 1 PLP Parser Version 0.0.1: resultado=15.0 dolar
```

Fonte: Autores

O bloco referente ao *Framework Core* é composto por interfaces que serão usadas pelos aplicativos do *framework*. As interfaces consideradas inicialmente e adaptadas de Bauer e King (2005) são: **Interface de Sessão**: responsável pelo armazenamento em cache ou armazenamento da coleção de objetos para a unidade de trabalho. Importante para o processo de criação e destruição de sessões de trabalho; **Interface de Configuração**: usado para especificar o local de mapeamento dos arquivos ou documentos de configuração do *framework*; **Interface Query e Criteria (Query Engine)**: responsável por executar consultas, inserções e atualizações nos bancos de dados e consultas em arquivos texto, XML ou XBRL baseadas nas classes criadas em **Data Model** e nos arquivos de mapeamento XML, que definem essa relação, além de controlar como isso é executado. Os comandos poderão ser escritos em SQL adaptado ou na linguagem financeira estabelecida para o *framework*. Já função *export*, ainda em referência a Figura 1, tem como objetivo exportar um conjunto de dados para uma instância XBRL conforme a definição da taxonomia ou para uma outra base de dados definida pelo operador e mapeada através de um arquivo *XML Mapping*.

Outrossim, a estrutura do *framework* foi pensada de forma a separar as principais funções permitindo sua extensibilidade ao longo do tempo. Por ser um software *open source* será disponibilizado para permear a utilização da tecnologia XBRL face as dificuldades apresentadas por Farias (2014) e pela ampla utilização de seus recursos no mercado financeiro mundial e governos, conforme explicitado nos trabalhos de Luciano e Silva (2016) e Maus et. al. (2009).

4. CONCLUSÃO E TRABALHOS FUTUROS

A construção do *framework* é um projeto que teve como finalidade geral criar mecanismos que visam facilitar o uso da tecnologia XBRL. O *framework* pode ter seu uso como ferramenta *ETL* ou para fins de fiscalização e auditoria, já que também permite a criação de algoritmos através de uma linguagem puramente financeira. As

dificuldades apresentadas por Farias (2014), onde os setores público e privado no Brasil apresentam dificuldades técnicas na adoção dessa tecnologia, já amplamente utilizada no mundo, impulsionaram esse projeto. A proposta é que o *framework* seja disponibilizado através da licença *open source*, tão logo o projeto seja concluído, permitindo sua utilização e colaboração de outros pesquisadores.

Como projeto futuro, se pretende incluir ontologias de cunho financeiro para a realização de mapeamento e validação de dados para os mais diversos formatos, bem como a possibilidade de criação de banco de dados baseado na ontologia, facilitando o processo de modelagem. Assim, espera-se que a proposta aqui apresentada possa contribuir para o processo de gestão e auditoria fiscal, ressalvadas as adaptações necessárias.

REFERÊNCIAS BIBLIOGRÁFICAS

Farinelli, F., Almeida, M. B. (2014). Interoperabilidade semântica em sistemas de informação de saúde por meio de ontologias formais e informais: um estudo da norma Openehr. In: CONFERÊNCIA BIREADIAL ISTECC – ACESSO ABERTO, PRESERVAÇÃO DIGITAL, INTEROPERABILIDADE, VISIBILIDADE E DADOS CIENTÍFICOS. 2014.

Farias, E. L. (2014). Desafios da Implantação do XBRL no Brasil: um estudo utilizando a teoria de difusão de inovações. Dissertação apresentada ao Departamento de Contabilidade e Atuária da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo – USP. Disponível em: <<https://teses.usp.br/teses/disponiveis/12/12136/tde-01122014-180454/publico/EduardoLopesFariasVC.pdf>>. Acessada em: 06/05/2020.

Luciano, G. J., Silva, P. C. da. (2008). Taxonomia XBRL - SPED: Simplificação e Auditoria no âmbito do Sistema Público de Escrituração Digital (SPED). iSys – Revista Brasileira de Sistemas de Informação. V. 11, Nº 04. 2008.

Mauss, C. V., Bleil, Vanti, C., A. A.. (2009). XBRL na Gestão Pública com Business Intelligence (BI). BASE – Revista de Administração e Contabilidade da Unisinos 6(1):5-18. janeiro/abril 2009. Disponível em: <<http://revistas.unisinos.br/index.php/base/article/viewFile/4811/2073>>. Acessado em: 04/05/2020.

Silva, P. C. da, Silva L. G. C., Júnior, I. J. de S. A. (2006). XBRL – eXtensible Business Reporting Language. 2006. Rio de Janeiro – RJ. Editora Ciência Moderna.

SILVA, P. C. da. (2009) LMDQL: Link-based and Multidimensional Query Language. In: DOLAP09 - ACM Twelfth International Workshop on Data Warehousing and OLAP, Hong Kong. ACM Twelfth International Workshop on Data Warehousing and OLAP, 2009.