

ANÁLISE E DESENVOLVIMENTO DE UMA APLICAÇÃO PARA INTEGRAÇÃO DE DADOS ENTRE SISTEMAS (TCC)¹

Felipe Santin²

Alexandre T. Lazzaretti³

RESUMO

A integração de dados de diferentes esquemas é uma etapa importante a ser considerada no momento em que existe uma troca de sistema informatizado em empresas. Em sua maioria, o desenvolvimento de códigos específicos para integração destes dados demandam custos financeiros e de pessoal. Neste sentido, este trabalho foi desenvolvido com o objetivo de gerar uma aplicação visando melhorar o processo de integração de dados entre esquemas. Assim, este trabalho mostra o desenvolvimento de uma aplicação que realiza a integração de dados de sistemas legados para novos sistemas, sua utilização é dinâmica, possibilitando a associação entre tabelas e seus respectivos campos conforme a demanda dos usuários, assim podendo realizar a migração conforme os esquemas disponibilizados. Além disto, facilitar futuras conversões, a ferramenta também armazena arquivos templates, com o mapeamento dos esquemas já realizados. Para teste da ferramenta, foi aplicado um estudo de caso de uma aplicação ERP.

Palavras-chave: ETL, ERP, Web, Genérico.

1 INTRODUÇÃO

No contexto de empresas de desenvolvimento de software, a atividade de integração de dados é um ponto a ser levado em consideração. O problema desta atividade consiste na combinação de dados existentes em diferentes fontes, com objetivo de prover ao usuário uma visão destes dados.

Um processo fundamental durante a implantação de um novo sistema informatizado é a utilização dos dados existentes nos sistemas legados. Para a realização desta tarefa, torna-se necessário desenvolver uma ferramenta específica para acesso ao sistema legado, e isto envolve analisar as estruturas deste sistema e

¹ Trabalho de Conclusão de Curso (TCC) apresentado ao Curso de Tecnologia em Sistemas para Internet do Instituto Federal Sul-rio-grandense, Câmpus Passo Fundo, como requisito parcial para a obtenção do título de Tecnólogo em Sistemas para Internet, na cidade de Passo Fundo, em 2017.

² Aluno do curso de Tecnologia de Sistemas para Internet no Instituto Federal de Educação, Ciência e Tecnologia Sul-rio-grandense de Passo Fundo (IFSul). E-mail: felipesantin123@hotmail.com.

³ Orientador, professor do Instituto Federal de Educação, Ciência e Tecnologia Sul-rio-grandense de Passo Fundo (IFSul). E-mail: alexandre.lazzaretti@passofundo.ifsul.edu.br.

criar um novo projeto somente para esta migração, demandando tempo e todos os custos relacionados ao desenvolvimento de um projeto. (Ribeiro & Oliveira, 2010). Por este motivo pretende-se desenvolver uma aplicação genérica que faça esta tarefa de migração de dados do sistema legado existente. Neste sentido, a integração de dados é uma tarefa importante [...] e difícil de ser resolvida. (Macura, 2014).

Analisando a importância da integração de dados, este trabalho mostra uma ferramenta web que possui a funcionalidade de extrair dados de sistemas legados, e gerar um novo conjunto de dados associado ao esquema do novo sistema a ser implantado.

Para realizar este desenvolvimento estudou-se conceitos relacionados a ETL (Extract Transform and Load), integração de dados e trabalhos relacionados. Nas pesquisas realizadas, foram encontrados poucos trabalhos relacionados a integração de dados entre sistemas ERP, e menos ainda trabalhos relacionados ao objetivo deste trabalho.

Um trabalho encontrado, foi o realizado por Ribeiro (Ribeiro, 2010), nele é mostrado um projeto para migração de dados para um sistema ERP específico. Neste projeto foi implementado o processo de migração em cerca de 29% das entidades do banco de dados do sistema. Mas como o software foi desenvolvido especificamente para o sistema de sua empresa, o mesmo acabou ficando restrito e assim impossibilitando uma conversão genérica o que é o objetivo da aplicação que está sendo apresentado neste artigo.

Este artigo está dividido da seguinte forma: (i) a segunda seção apresenta o referencial teórico, realizando uma apresentação ao leitor de alguns temas importantes, ou seja, processos utilizados e tecnologias; (ii) a seção que a sucede, é demonstrado a metodologia abordada para o desenvolvimento do trabalho e o caso de uso; (iii) na quarta seção estão dispostas as considerações finais e a premissas de trabalhos futuros.

2 REFERENCIAL TEÓRICO

Este capítulo tem por objetivo retomar o conceito de alguns temas importantes para o desenvolvimento desta aplicação, como: ETL, .NET, XML e Banco de Dados.

2.1 ETL

No processo de integração de dados entre sistemas uma das partes mais importante é o ETL, que refere-se à sigla de Extração, Transformação e Carga (Extract, Transform and Loader), sendo uma metodologia que se divide em três estágios: um para leitura e identificação dos dados, um para manipulação, conversão ou ajuste, e outro para preparação para inserção na tecnologia de destino (Oliveira, 2012). Estas etapas estão descritas a seguir.

2.1.1 Extração (*Extract*)

O primeiro passo é acessar os dados de diferentes fontes para poder extrair os dados necessários (Novak e Rabuzin, 2014).

2.1.2 Transformação (*Transform*)

Na obtenção de dados de fontes muitas vezes desconhecidas e que foram gerenciadas por sistemas de informação antigos e por falha de projeto ou sem a utilização de um sistema gerenciador de banco de dados adequado, que possui a função de propiciar um ambiente tanto conveniente quanto eficiente (sem falhas) para a recuperação e armazenamento das informações (Ribeiro & Oliveira apud Silberschatz et al., 1999), é comum encontrar problemas de integridade referencial ou inconsistências como datas inválidas, atributos obrigatórios não preenchidos, somatórios numéricos inconsistentes, falta de normalização e diversos outros problemas (Ribeiro & Oliveira, 2010).

Tendo estes problemas em mente esta etapa é a mais importante, pois é nela onde os dados coletados são verificados, limpos e conformados, estes dados

passavam por várias atividades a fim de melhorar a qualidade dos dados. (Novak e Rabuzin, 2014).

2.1.3 Carga (*Load*)

O processo de carga deverá gravar os dados extraídos, tratados e transformados nas etapas anteriores, no banco de dados de destino (Ribeiro & Oliveira, 2010).

Uma vez executadas as rotinas de transformação, se faz necessária a saída em linguagem SQL, através de comandos de inserção [...]. Quando o sistema de destino for um único SGBD, deverá executar um arquivo texto contendo os comandos (Oliveira, 2012) para concluir a integração de dados entre os sistemas.

2.2 .NET

É a uma plataforma de desenvolvimento da Microsoft que tem como foco principal o desenvolvimento de Serviços Web. Sendo que um serviço Web transcende ao que conhecemos como páginas dinâmicas, as quais podem ser acessadas a partir de um browser. A ideia central de um Web Service consiste em permitir que as aplicações, sejam elas da Web ou Desktop, ou ainda middleware, se comuniquem e troquem dados de forma simples e transparente, independente do sistema operacional ou da linguagem de programação (Lima, 2002).

2.3 XML

A tecnologia XML (*eXtensible Markup Language*) possui uma meta-linguagem de marcação que contém métodos de acesso que possibilitam a manipulação dos dados. Um documento XML é organizado de forma hierárquica, carrega a estrutura dos dados (tags) e os seus conteúdos (dados). Ele é composto por elementos simples (um exemplo: *nome* Figura 1, linha 4) ou elementos compostos (exemplo: *configuracao* Figura 1, linha 3), atributos (exemplo: *xmlns:xsi* Figura 1, linha 3),

instruções de processamento (exemplo: Figura 1, linha 1) e comentários (exemplo: Figura 1, linha 2) (W3C, 2012).

Figura 1. Arquivo de configuração gerado pelo sistema, contendo o nome da conversão, tipos e string de conexões dos bancos.

```

1. <?xml version="1.0" encoding="UTF-8"?>
2. <!--Exemplo de arquivo de configuração -->
3. <Configuracao xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
4.   <nome>Conversao Teste de caso</nome>
5.   <tipo0>F</tipo0>
6.   <string0>
7.     User=SYSDBA;Password=masterkey;Database=localhost:basedados; F
8.   </string0>
9.   <tipoD>F</tipoD>
10.  <stringD>
11.    User=SYSDBA;Password=masterkey;Database=localhost:basedados; Port
12.  </stringD>
13.  <ligacoes> ... </ligacoes>
14. </Configuracao>

```

Fonte: O autor

2.4 Banco de dados

É uma coleção de dados que contém as informações relevantes ao domínio da aplicação (Silberchatz et al., 2006).

Para Silberchatz *et al.* (2006) um modelo de dados tem por finalidade apoiar a estrutura do banco de dados, pois ele fornece ferramentas conceituais para descrever os dados, relações entre os dados, semântica dos dados e restrições de consistência. A seguir é descrito o modelo de dados relacional.

O modelo relacional representa o banco de dados como um coleção de relações (tabelas). Onde cada relação é representada por uma tabela de valores, onde cada linha da tabela representa uma coleção de valores de dados relacionados. Esse modelo possui um conjunto de conceitos associados, que são: relações, atributos (campos) e tuplas (registros) (Lazzaretti, 2012).

3 RESULTADOS

3.1 Metodologia de Funcionamento

A Figura 2, representa as funcionalidades propostas para o desenvolvimento da aplicação. Sendo elas:

- Possibilitar a ligação entre tabelas e seus respectivos campos;
- Converter os dados do banco de origem para o de destino
- Gerar arquivos com as ligações realizadas para facilitar futuras conversões (arquivos templates);
- Gerar os comando necessários para inserir os dados desejados no banco de destino.

A aplicação foi desenvolvida baseado na arquitetura MVC, sendo uma aplicação web e podendo ser acessada por qualquer navegador. Para utilizar o sistema desenvolvido a empresa deverá implantar a aplicação em um servidor que possua IIS (*Internet Information Services*), que é um servidor para aplicações desenvolvidas em .NET.

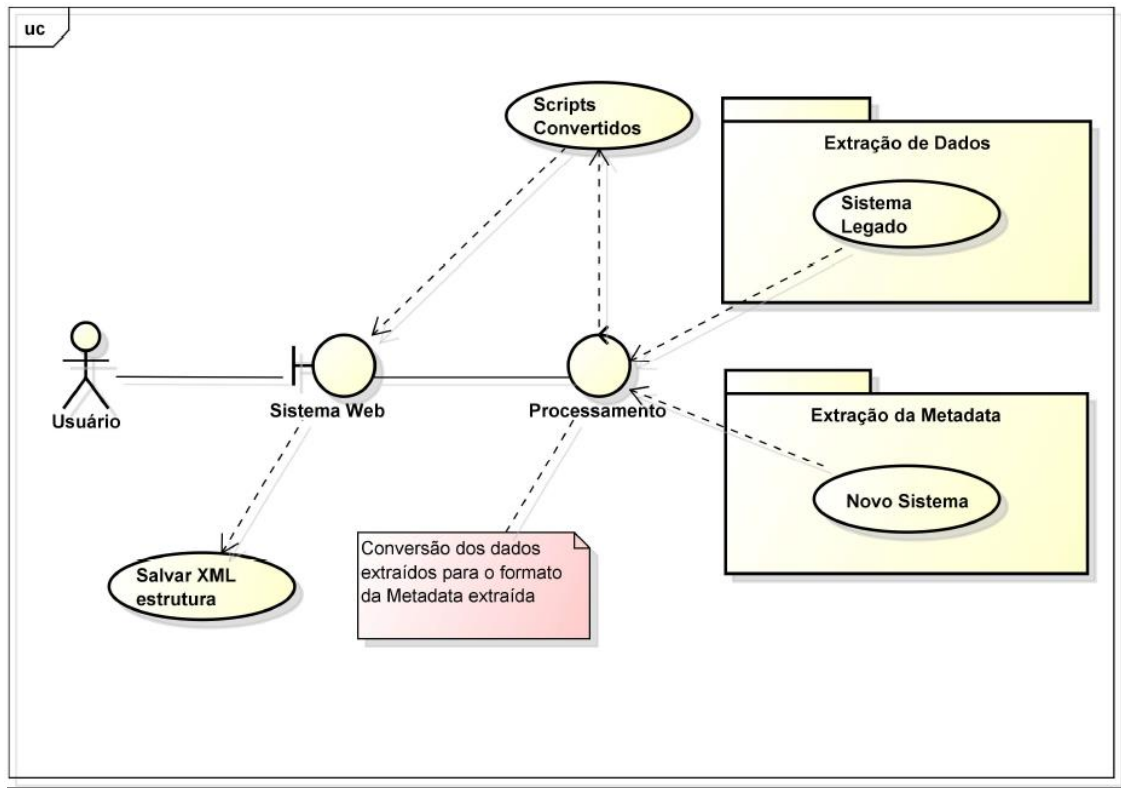
Ao utilizar o sistema implantado, o usuário informará uma string de conexão, permitindo ao sistema conectar-se aos bancos informados. Essa conexão é realizada através de bibliotecas externas que permitem a conexão com um tipo de banco de dados específico.

Após ocorrer a validação desta conexão o usuário poderá selecionar quais são as ligações entre tabelas e seus respectivos campos, que devem ser integradas com o banco de destino.

Após, o sistema irá realizar a extração dos dados do sistema legado realizando as transformações necessárias e ajustando-os conforme o esquema do banco de destino. Ao final, o sistema gera um arquivo com os comando necessários para inserir os dados na nova base de dados.

Além disso, o sistema possui uma opção para salvar os mapeamentos entre as tabelas e os campos. Para isto, é gerado um arquivo .xml, desta forma, permite futuras conversões de bancos com o mesmo esquema de dados.

Figura 2. Representação das funcionalidades que deverão ser desenvolvidas para esta aplicação.



Fonte: O autor

3.2 Estudo de Caso

No contexto de empresas de desenvolvimento de software, a atividade de integração de dados é um ponto a ser levado em consideração. O problema desta atividade consiste na combinação de dados existentes em diferentes fontes, com objetivo de prover ao usuário uma utilização.

Para testar as funcionalidades da ferramenta desenvolvida, foi implementado um estudo de caso relacionado a um banco de dados de um cliente da empresa Holma Software, que utiliza um banco de dados firebird e fez-se necessário a migração para uma nova base de dados firebird com uma estrutura diferente, que por sua vez é um sistema ERP (*Enterprise Resource Planning*).

Segundo Sallaberry (2009), os sistemas ERP tem como finalidade dar suporte a maior parte da empresa como suprimentos, manufatura, manutenção, administração financeira, contabilidade, entre outros, com um sistema de informação

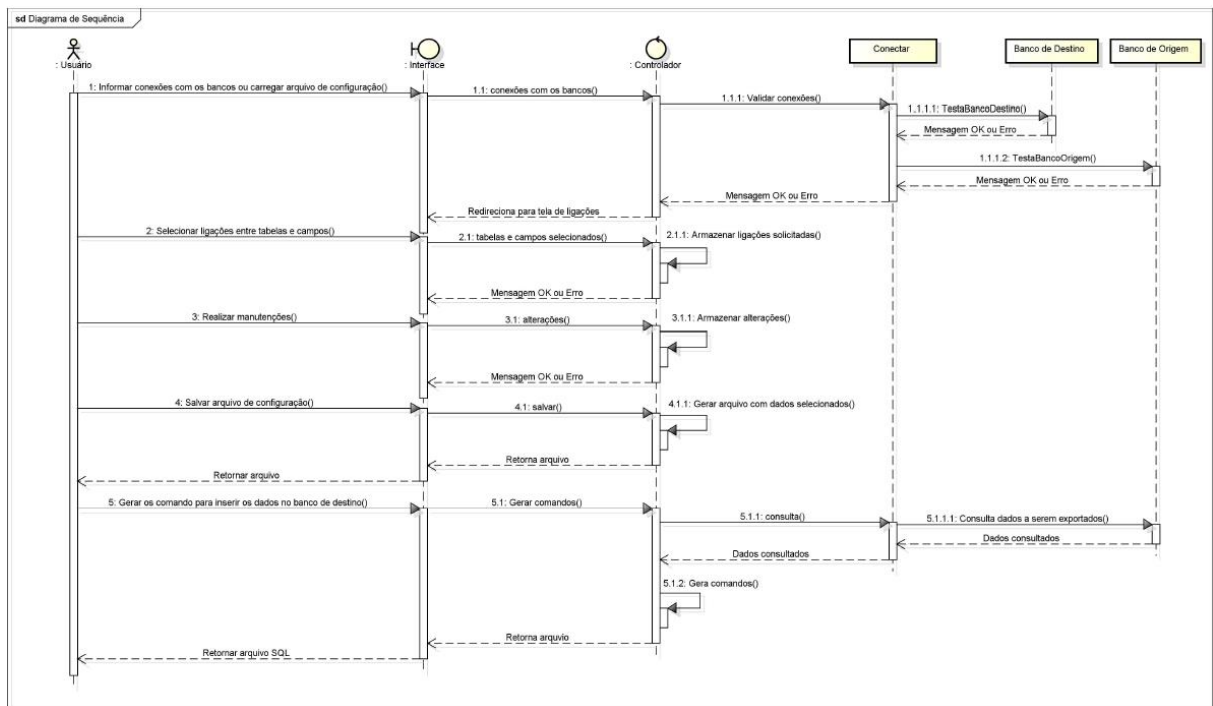
integrado, ou seja, um sistema que possua integração com todas estas área na empresa, facilitando a comunicação e troca de informações entre os funcionários.

Apesar do ERP ter suporte para várias áreas, a integração de dados proposta neste trabalho, tem foco na área de manufatura, focando na conversão de produtos, processos de produção e afins.

3.3 Aplicação do Estudo de Caso

Para realizar as tarefas disponibilizadas pelo sistema, deve-se seguir uma seqüência de tarefas, conforme demonstrada no Diagrama de Sequência na Figura 3.

Figura 3. Diagrama de seqüência.



Fonte: O autor

Na Figura 4, pode-se ver as opções da tela inicial do sistema. Em sua parte superior, existe uma opção para carregar uma configuração salva anteriormente, assim facilitando conversões de banco com esquemas de dados iguais.

Em seguida, é possível informar um nome para o arquivo que contém as regras de mapeamento. Sugere-se usar um nome significativo que corresponda ao mapeamento realizado.

Ainda na Figura 4, pode ser configurado de qual SGBD será realizada a conversão de dados e para qual SGBD será convertido. Nesta versão, pode-se escolher entre 2 tipos de SGBD's: Firebird e Postgres.

Referente a configuração dos SGBD's ao selecionar um dos dois tipos, o sistema traz automaticamente os dados de uma conexão padrão, como o servidor em que o banco está, a porta de conexão, nome e senha do banco de dados, permitindo alterar conforme as necessidades.

Figura 4. Parte superior da tela inicial do sistema desenvolvido.

The screenshot displays the upper portion of a web-based configuration interface. It is organized into several sections:

- Seleção de arquivo de configuração:** Features a 'Choose File' button and the text 'No file chosen'.
- Crie um arquivo de configuração:** Includes a text input field labeled 'Descrição'.
- Configuração do banco de origem:** Contains a dropdown menu currently set to 'Firebird', with a list of options including 'Firebird' and 'PostgreSQL'. Below the dropdown, a text field shows a connection string: 'User=SYSDBA;Password=masterkey;Database=localhost:basedados; Port=3050;Dialect=3'.
- Configuração do banco de destino:** Features a dropdown menu currently set to 'PostgreSQL'.
- Configuração de conexão do banco de destino:** A text field at the bottom displays: 'Server = 127.0.0.1; Port = 5432; Database = basedados; User Id = postgres; Password = '.

Fonte: O autor

A Figura 5, demonstra a área do sistema onde pode-se selecionar as tabelas do banco de dados do sistema de Origem. Ao selecionar a tabela, são listadas os campos da respectiva tabela selecionada.

Figura 5. Tela de ligação, banco de origem.

Banco de Dados de Origem

PESSOAS

Colunas	
BARR_COBR_PESS	Ligar
BARR_ENTR_PESS	Ligar
BARR_PESS	Ligar
CEP_COBR_PESS	Ligar
CEP_ENTR_PESS	Ligar
CEP_PESS	Ligar
CODG_CIDD	Ligar
CODG_CIDD_COBR	Ligar
CODG_CIDD_ENTR	Ligar
CODG_PESS	Ligar
CODG_TRNS	Ligar
CODG_VEND	Ligar
COMP_ENDR	Ligar
COMP_ENDR_COBR	Ligar
COMP_ENDR_ENTR	Ligar

Fonte: O autor

A Figura 6 demonstra a área do sistema onde pode ser selecionar as tabelas do banco de dados de Destino. Após realizada a seleção da tabela, são listadas as colunas da tabela selecionada.

Figura 6. Tela de ligação, banco de destino.

Banco de Dados de Destino

CLIFORN

Colunas	
ANTECIPADO	Ligar
ANTECIPUTILIZADO	Ligar
AREATOTAL	Ligar
AVALIACAO	Ligar
BAIRRO	Ligar
BAIROCOBRANCA	Ligar
CALCOMISSAO	Ligar
CALCULAIFI	Ligar
CAPITALSOCIAL	Ligar
CASAP	Ligar
CELULAR	Ligar
CEP	Ligar
CEPCOBRANCA	Ligar
CHAVE	Ligar
CLASSE	Ligar

Fonte: O autor

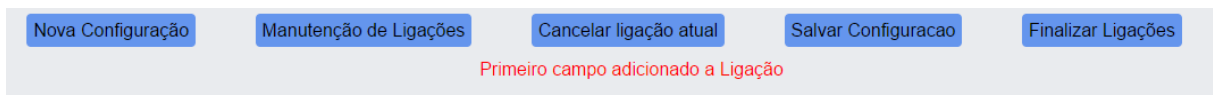
Após serem selecionadas as tabelas (origem e destino) e os respectivos campos, é necessária realizar o mapeamento (ligação) entre o que foi selecionado. A Figura 7, mostra as opções disponíveis na tela de ligação.

Na primeira opção pode-se criar uma nova configuração para o mapeamento. Na segunda opção, pode-se gerenciar as tabelas e colunas selecionadas, podendo remover ligações realizadas ou especificar uma condição, através da clausula “where”, a ser executada no momento da conversão.

O terceiro botão serve para cancelar a ligação de um campo que está sendo realizada no momento, ao clicar em uma coluna errada, poder removê-la e clicar na correta. O quarto botão serve para salvar um arquivo com as configurações realizadas, visando facilitar futuras conversões com estruturas iguais.

Por fim, na última opção é utilizada para gerar os comando necessários para inserir os dados no banco de destino. Abaixo dos botões, na tela, existe uma área onde serão mostradas as mensagens para o usuário.

Figura 7. Tela de ligação, parte inferior.



Fonte: O autor

A tela de manutenção de ligações pode ser dividida em duas partes. A primeira, conforme a Figura 8, mostra todas as tabelas ligadas e dá a possibilidade de remover completamente uma tabela e também adicionar a clausula where.

Figura 8. Tela de manutenção de ligação.

Tabela de Origem	Tabela de Destino	Where	
PESSOAS	CLIFORN	Adicionar	Remove
PRODUTOS_COMPOSICAO	COMPOSICAO	Adicionar	Remove

Fonte: O autor

Na segunda parte, conforme a Figura 9, pode-se visualizar todas as colunas que foram ligadas, assim tendo a opção de remover uma coluna específica.

Figura 9. Tela de manutenção de ligação.

Campos de Origem		
CODG_PESS	CODCLIFORN	Remover
NOME_PESS	DESCRCLIFORN	Remover
NOME_FANT	COMPLEMENTO	Remover
ENDR_PESS	ENDERECO	Remover
NUMR_ENDR	NRO	Remover
COMP_ENDR	COMPENDERECO	Remover

Fonte: O autor

A Figura 10, mostra a tela com o comando SQL gerado para a execução da integração. Nele, pode-se adicionar condições a serem impostas pela cláusula WHERE no momento de sua execução.

Esta opção foi adicionada ao sistema pois em uma importação o cliente pode solicitar somente alguns dados, como por exemplo somente os produtos que estavam ativos no sistema legado.

Figura 10. Tela para adicionar clausula Where em uma ligação.

```
select PRODUTOS_COMPOSICAO.CODG_PRDT, PRODUTOS_COMPOSICAO.CODG_ITEM,
PRODUTOS_COMPOSICAO.QUNT_PRDT, PRODUTOS_COMPOSICAO.QUNT_PRDT from
PRODUTOS_COMPOSICAO where
```

Cancelar Salvar

Fonte: O autor

Assim após realizar todos os processos necessários para uma migração e clicar no botão “Finalizar Ligações”, o sistema gerará um arquivo SQL contendo os comando para inserir estes dados no banco de destino, como pode ser visto no exemplo da Figura 11.

Figura 11. Parte do Arquivo com os comandos gerados.

```
--Inserts gerados para as seguintes tabelas: CLIFORN, COMPOSICAO

--CLIFORN
insert into CLIFORN(CODCLIFORN, DESCRCLIFORN, COMPLEMENTO, ENDEREÇO, NRO, COMPENDEREÇO, BAIRRO, CEP);
insert into CLIFORN(CODCLIFORN, DESCRCLIFORN, COMPLEMENTO, ENDEREÇO, NRO, COMPENDEREÇO, BAIRRO, CEP);
insert into CLIFORN(CODCLIFORN, DESCRCLIFORN, COMPLEMENTO, ENDEREÇO, NRO, COMPENDEREÇO, BAIRRO, CEP);
.
.
.
--COMPOSICAO
insert into COMPOSICAO(CODPRODUTO, CODCOMPONENTE, QTD, QTDLIQUIDA ) values(1393, 1446, 2, 2);
insert into COMPOSICAO(CODPRODUTO, CODCOMPONENTE, QTD, QTDLIQUIDA ) values(1393, 1214, 0.06, 0.06);
insert into COMPOSICAO(CODPRODUTO, CODCOMPONENTE, QTD, QTDLIQUIDA ) values(1393, 475, 0.8, 0.8);
```

Fonte: O autor

4 CONSIDERAÇÕES FINAIS

Este trabalho teve por objetivo desenvolver uma ferramenta para realizar a integração de dados entre bancos de dados, desta forma, mantendo os dados importantes para os usuários em seus novos sistemas.

Para este desenvolvimento foram utilizadas diversas tecnologias, como a plataforma de programação .NET, que foi utilizada para desenvolvimento da interface web, processamento e geração dos comandos necessários para migração dos dados.

Como principal diferencial, têm-se a funcionalidade de criação dos mapeamentos de forma dinâmica, realizando a leitura de tabelas e seus respectivos campos. Desta forma, possibilitando a integração de diferentes esquemas de dados.

Para a validação da ferramenta, foi aplicado um estudo de caso baseado num sistema ERP existente. Através dos resultados verificou-se que foi possível realizar a migração dos dados.

Para trabalhos futuros, em primeiro momento, pretende-se adicionar uma funcionalidade para realizar teste se o SQL gerado está funcionando corretamente.

Pretende-se também adicionar uma opção para poder selecionar várias tabelas em um mesmo SQL, assim podendo melhorar o nível de consulta que o sistema poderá realizar.

ABSTRACT

The integration of data from different schemas is an important step to be considered when there is an export of computerized systems in companies. In the most part, the development of specific codes for the integration of this data demands financial and people costs. In this sense, this work was developed with the purpose of generating an application to improve the process of data integration between different schemas. Thus, this work shows the development of an application that performs the integration of data from existing systems to other new systems, its use is dynamic, allowing the association between tables and their respective fields according of the necessity, thus being able to perform the migration according to the schemas available. In addition, to facilitate future conversions, the tool also stores template files, with the mapping of the schemes already performed. To test the tool, a case study of an ERP application was applied.

Keywords: ETL. ERP, Web, Generic.

REFERÊNCIAS

DEGAN, Joyce Otsuka Côrtes. **Integração de dados corporativos: uma proposta de arquitetura baseada em serviços de dados**. Instituto de Computação Universidade Estadual de Campinas. 2005.

LAZZARETTI, Alexandre T. **Revisão bibliográfica sobre banco de dados**. Universidade de passo fundo programa de pós-graduação em agronomia. Junho de 2012.

LIMA, Edwin. **C# e .net para desenvolvedores**. Rio de janeiro: campus, 2002

MACURA, Marek. **Integration of data from heterogeneous sources using etl technology**. Computer Science. 15 de fevereiro de 2014. Disponível em <http://dx.doi.org/10.7494/csci.2014.15.2.109>.

NOVAK, Matija; RABUZIN, Kornelije. **Prototype of a web etl tool**. (IJACSA) International journal of advanced computer science and application. vol. 5. no. 6. Varazdin, Croatia, 2014, p. 97-103.

OLIVEIRA, Cleber silva de. **Metodologias e estratégias de migração de dados**. Sinergia, São Paulo, v. 13, n. 3, p. 183-191, set./dez. 2012

RIBEIRO, Alexandre lima; OLIVEIRA, Edemar Costa. **Processos de implantação e migração de dados com utilização de etl para um erp comercial**. Universidade luterana do brasil. Curso de sistemas de informação. Campus canoas. Canoas, RS. 2010.

SILBERSCHATZ, Abraham; KORTH, Henry; SUDARSHAN S. **Sistema de Banco de Dados**. 5a Edição. Editora Elsevier. 2006.

SALLABERRY, Cícero Rodrigues. **Implementação de um sistema ERP em uma empresa construtora: impactos no processo de aquisição de materiais**. Universidade Federal do Rio Grande do Sul, Departamento de Engenharia Civil. Porto Alegre. Dezembro 2009.