



BUSINESS INTELLIGENCE – DESMISTIFICANDO A CAMADA ETL (EXTRAÇÃO, TRANSFORMAÇÃO E CARGA)

Fabio Casarotti JUNIOR¹
Juliene Aglio O. PARRÃO²
Paula Jurema Piloto LANGHI³

RESUMO: O presente trabalho terá como temas abordados o *Business Intelligence* ou Inteligência de Negócios que se remete a importância dos dados que geramos dia após dia e a camada ETL (Extração, Transformação e Carga) correspondente ao processo de extração dos dados de bases diversas para uma base única, o Data Warehouse, a qual será utilizada como fonte de análise. Serão analisados exemplos de informações preciosas que são possíveis de se extrair de fontes de dados com potencial para alavancar o negócio de uma empresa.

Palavras-chave: *Business Intelligence*. Extração. Transformação. Carga. Análise de Dados.

1 INTRODUÇÃO

Neste artigo serão apresentadas as etapas da camada ETL, a qual trata dos processos de extração dos dados de uma ou mais fontes de negócio, sequencialmente a transformação, onde podem ser aplicadas regras de negócio para padronização dos dados, tornando-os interpretáveis e fazendo com que durante a análise possam ser transformados em informações de nível gerencial, e por último a etapa da carga, a qual trata da injeção dos dados transformados em uma base de dados, que irá armazenar as informações podendo gerar futuramente relatórios de alto nível e decisões gerenciais.

¹ Discente do 6º ano do curso de Sistemas de Informação do Centro Universitário “Antônio Eufrásio de Toledo” de Presidente Prudente. E-mail: fabiojunior@toledoprudente.edu.br.

² Orientadora do Trabalho. Docente do curso de Arquitetura e Urbanismo no Centro Universitário Antônio Eufrásio de Toledo de Presidente Prudente. Doutora em Serviço Social pela PUC/SP, mestre em Políticas Sociais pela UEL/PR. E-mail: coord.social@toledoprudente.edu.br.

³ Docente do curso de Sistemas de Informação do Centro Universitário “Antônio Eufrásio de Toledo” de Presidente Prudente. Mestre em Ciências Cartográficas pela Universidade Estadual Paulista (UNESP) e Especialista em *Business Intelligence* pela Toledo Prudente Centro Universitário. E-mail: paula_langhi@toledoprudente.edu.br. Orientadora do trabalho.

O *Business Intelligence* pode ser entendido como um processo composto de várias etapas até chegar a sua fase final, que é a transformação dos dados em informações de alto nível para um negócio. Ter ferramentas de inteligência de negócios dentro de uma empresa possibilita a extração de informações, desde suas vendas em um período, até estatísticas acerca do melhor produto para vender em uma determinada data, incrementando um retorno maior de lucro.

Este artigo tem como objetivo mostrar a relevância dos dados em qualquer modelo de negócio e ainda demonstrar de forma clara e concisa o volume de dados que pode ser produzido diariamente, gerando toneladas de informações, as quais não passariam de bytes armazenados se não soubermos efetivamente o que pode ser interpretado através delas, gerando melhorias nos sistemas e nos negócios das empresas, pois, quando se obtém a informação certa, na hora correta, sobre determinado público ou produto, é possível criar novas expectativas para o negócio, aumentar a margem de lucro e implementar um sistema que possa trazer inteligência para todas as partes da empresa.

2 A IMPORTÂNCIA DOS CADASTROS DE INFORMAÇÕES EM SISTEMAS OLTP NAS EMPRESAS.

OLTP (*Online Transaction Processing*) é a sigla para sistemas de processamento de transações online. Esses sistemas são utilizados no processamento dos dados de rotina que são gerados diariamente através dos sistemas de informação da empresa e dão suporte às funções de execução do negócio organizacional. Por exemplo, quando se faz um saque em um caixa eletrônico é subtraído da conta corrente o valor sacado ou quando é feito um depósito é adicionado o valor do depósito na conta corrente, estes exemplos podem ser definidos como operações OLTP. Uma vez que essas operações são feitas, são registradas no banco de dados informações como quem, quando, onde, qual valor e a data que a operação foi realizada, esses dados são cruciais para análises de *Business Intelligence* futuras, quando qualquer dado é registrado de forma incorreta,

uma inconsistência surge na base de dados e o prejuízo para fixar um erro no banco de dados é altíssimo.

2.1 O TEMPO PERDIDO COM INCOSISTÊNCIAS NAS BASES DE DADOS

Toda a empresa tem a sua forma de trabalhar e seu modo de fazer negócios, empresas que utilizam sistemas de softwares necessitam de pessoas para manusear os processos das transações comerciais.

Durante o primeiro semestre de 2020 a orientadora deste artigo Paula J. Piloto Langhi disponibilizou durante a disciplina de Banco de Dados III, dados em planilhas de Excel, advindos de diversos sistemas de informação, para que a turma executasse trabalhos relacionados ao processo ETL. As planilhas da professora continham dados reais, porém tais informações como nome, endereço e telefone foram alteradas de forma aleatória para manter a integridade e o anonimato das pessoas nas bases de dados. Com as planilhas era possível enxergar claramente

	A	B	C	D	E	F
	PACIENTE	SEXO	DATA NASCIMENTO	LOGRADOURO	NUMERO	BAIRRO
3	MARIA BARBOSA DE SANTOS		23/07/50	R	0	JD NOVA EUROPA
4	JOSUÉ DE SANTOS		23/07/50	R	0	PRIMAVERA
10	ABDIAS BISPO OLIVEIRA		18/11/47	AL DAS PEROBAS	13	CENTRO
11	ABDIAS BISPO OLIVEIRA		18/11/47	AL DAS PEROBAS	13	CENTRO
12	JOSUÉ APARECIDO DOMINGOS		07/11/50	RUA FRANCISCO GOMES	86	JD PAULISTANO
14	ABDIAS ALVES		31/05/55	R JAPAO	200	VILA INDUSTRIAL
16	MARIO SACOMAM		20/10/41	AV DR IBRAIN NOBRE	400	PARQUE FURLANETO
17			05/07/40	R VALTER HADDAD	30	CENTRO
18	MARIANO FIGUEIREDO DA SILVA		26/08/78	JOAO CORAZZA	163	JARDIM DOS REIS
20	ADILSON DIAS FERREIRA		14/10/66	RUA ANGIUS	508	CENTRO
21	ANTONIO ALVES PIRES		22/05/87	OLIVEIRA	45	VILA OPERARIA
22	ADAUTO PERETTI FILHO		20/07/62	R JOAQUIM NABUCO	389	CENTRO
28	ADELINO CARNELOSSI		03/09/50	R DESB CEARA	159	BOSQUE
29	ADENIL DOS SANTOS COSTA		06/06/56	R FUAD MALULY	140	CENTRO
30			24/10/62	RUA DUQUE DE CAXIAS	360	CAMPO REDONDO
31	ADAIR TADEU VILLA REAL		24/10/62	RUA DUQUE DE CAXIAS	360	CAMPO REDONDO
32	ADAIR DE ALMEIDA		01/07/45	R JOAO BUENO DA COSTA	65	CENTRO
34	ADAIR ALVES RIBEIRO		25/11/69	RUA IPIRANGA	577	VL FORMOSA
36	ALADIA TAPIAS GERALDO		17/10/39	R JOAQUIM NABUCO	955	CENTRO
40	MARIANO LACERDA KALIL	MASCULINO	15/03/79	DONA ERMELINDA PEREIRA	314	JARDIM ESTORIL
42	ADAO TONI PIRONDI	MASCULINO	21/11/72	R TUIUTI	90	JARDIM CAICARA

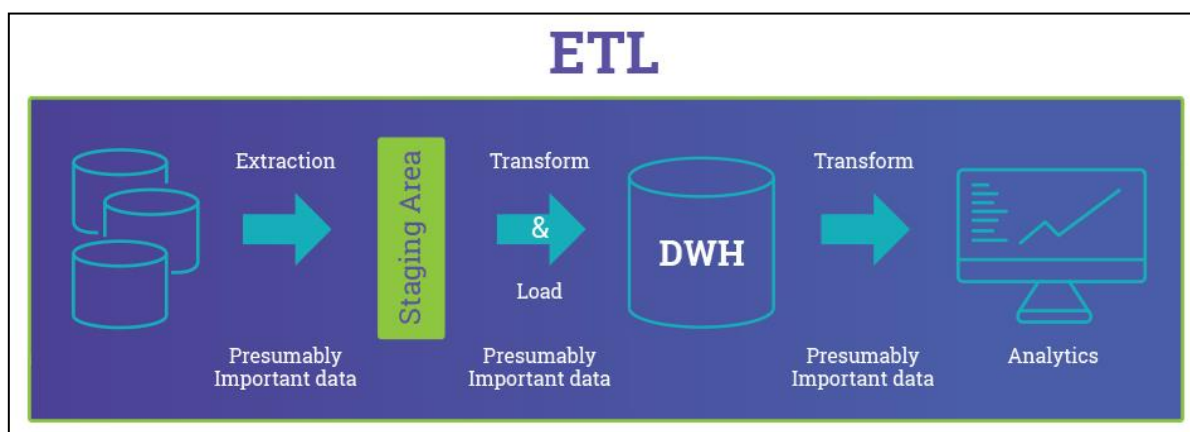
que erros de digitação eram mais comuns do que se poderia imaginar e a correção dos dados de forma manual, dentre os milhares de registros, seria inviável.

Figura 1 - Planilha de Excel 1. Tabela Clientes com divergência no campo sexo. Fonte: Paula J. Piloto Langhi (2020).

3 A GERAÇÃO DE INFORMAÇÃO PRECISA ATRAVÉS DO PROCESSO ETL

3.1 A CAMADA ETL

A camada ETL (*Extraction, Transform and Load*) corresponde ao processo de Extração, Transformação e Carga de dados e em geral é utilizada para dar carga em ambientes de *Data Warehouse*. Dentro da camada ETL são executados três importantes processos, sendo eles: o processo de extração de dados responsável pelo carregamento das informações vindas de diversas fontes de dados de uma empresa. Processo de transformação, no qual, é realizada a transformação dos dados aplicando regras de negócio para fim de padronização, correção, validação e filtragem das informações carregadas e por último o processo de carga, o qual carrega os novos dados para dentro do *Data Warehouse*, onde as



informações vão ficar disponíveis para análise e criação de relatórios gerenciais. A Camada ETL pode ser visualizada através da Figura 2.

Figura 2 – Representação gráfica da camada ETL. Fonte: Cetax (2016).

3.2 O PROCESSO DE EXTRAÇÃO DE DADOS

O processo de extração de dados pode ser realizado de diversas fontes tais como Planilhas de Eletrônicas, Documentos de texto, Banco de Dados, páginas de internet entre outras. A extração dos dados resulta em estabelecer uma fonte confiável de informações para que possam futuramente ser usadas na elaboração de relatórios complexos a fim de tomadas de decisões gerenciais que impactam diretamente o negócio das empresas.

Ter uma base de dados constante que fomenta o processo de extração de dados reincidirá na velocidade da extração tanto quanto na etapa final de entrega dos relatórios.

3.3 O PROCESSO DE TRANSFORMAÇÃO DOS DADOS

Sendo a segunda etapa do processo ETL, a transformação ocorre após os dados terem sido extraídos. Nesta etapa serão aplicadas regras de negócio nos dados para que se tornem informações de alto nível.

As regras da transformação visam padronizar os dados e fazer validações para que possam estar capacitadas no momento em que forem usados para relatórios gerenciais e tomadas de decisões.

3.3.1 TRANSFORMANDO OS DADOS PRIMITIVOS EM INFORMAÇÕES DE ALTO NÍVEL

As transformações dos dados podem variar de acordo com o negócio no qual os dados estão inseridos ou até mesmo padronização para que fiquem uniformes. Na planilha de Excel 1, mostrada na Figura 3, o campo sexo possui diversas referências para denominar o sexo do paciente seja ele feminino ou masculino, por meio de ferramentas de transformação uma padronização é aplicada nessa coluna para diminuir as opções de sexo existentes e deixando apenas duas. É possível criar uma regra de negócio de forma que, toda palavra que comece com a letra “F” seja substituído o valor atual por “Feminino”, da mesma forma para com “M”

porém resultando em “Masculino”, este é um exemplo de padronização de dados para que não fiquem redundantes as informações do campo sexo na base de dados ou gerem inconsistências em análises futuras.

	A	B	C	D
1	PACIENTE	SEXO	DATA_NASCIMENTO	LOGRADOURO
2	MARIA BARBOSA DE SANTOS		01/04/65	R MARIA THEREZA ALBERTI
3	MARIA BARBOSA DE SANTOS		23/07/50	R
4	JOSUÉ DE SANTOS		23/07/50	R
5	MARIA BARBOSA DE SANTOS		15/02/81	R MARIA THEREZA ALBERTI
6	MARIA BARBOSA DE SANTOS		15/02/81	R MARIA THEREZA ALBERTI
7	MARIA BARBOSA DE SANTOS		15/02/81	R MARIA THEREZA ALBERTI
8	ADRIANY DIAS BERGAMASCHI DE LA RUA		15/08/81	ENGENHEIRO ALFRED LIEMERT
9	ALAYDE CUSTODIO ERBELLA		06/03/57	RUA CAPITAO ALBERTO MENDES JR
10	ABDIAS BISPO OLIVEIRA		18/11/47	AL DAS PEROBAS
11	ABDIAS BISPO OLIVEIRA	MASCULINO	18/11/47	AL DAS PEROBAS
12	JOSUÉ APARECIDO DOMINGOS	MASCULINO	07/11/50	RUA FRANCISCO GOMES
13	ADELIA TYOKO SILVA YOKOSAWA	FEMININO	07/02/49	R ALBERTO S TANABE
14	ABDIAS ALVES	MASCULINO	31/05/55	R JAPAO
15	MARIA OLIVA GONINI ESTRELA	FEMININO	24/03/67	R PADRE DIONIZIO GONZALES

Figura 3 - Planilha de Excel 1. Tabela Clientes com divergências no campo sexo. Fonte: Paula J. Piloto Langhi (2020).

Além da padronização e regras de negócios que podem ser aplicadas aos dados, existe a validação que é possível usar para verificar os dados que estão faltando ou até mesmo os dados que não estão de acordo com o negócio. Usando novamente como exemplo, a planilha de Excel 1 (Figura 4), existem linhas que estão com alguns dados faltantes, pode-se determinar uma regra de transformação que faça a leitura individualmente de cada linha e para cada campo faça-se a verificação de campos nulos. As linhas que tiverem campos sem dados não prosseguirão para a etapa da carga, porém, estas linhas serão encaminhadas para uma base separada onde estarão sujeitas á investigação para descobrir a causa de sua falta de informações.

	A	B	C	D	E	F	G	H
	PACIENTE	SEXO	DATA_NASCIMENTO	LOGRADOURO	NUMERO	BAIRRO	MUNICIPIO	UF
3	MARIA BARBOSA DE SANTOS		23/07/50	R	0	JD NOVA EUROPA	CAMPO REDONDO	MT
4	JOSUÉ DE SANTOS		23/07/50	R	0	PRIMAVERA	CAMPO REDONDO	MT
10	ABDIAS BISPO OLIVEIRA		18/11/47	AL DAS PEROBAS	13	CENTRO	SAO TOME	PR
11	ABDIAS BISPO OLIVEIRA		18/11/47	AL DAS PEROBAS	13	CENTRO	SAO TOME	PR
12	JOSUÉ APARECIDO DOMINGOS		07/11/50	RUA FRANCISCO GOMES	86	JD PAULISTANO	NATAL	RN
14	ABDIAS ALVES		31/05/55	R JAPAO	200	VILA INDUSTRIAL	CAMPO REDONDO	MT
16	MARIO SACOMAM		20/10/41	AV DR IBRAIM NOBRE	400	PARQUE FURLANETO	NATAL	RN
17			05/07/40	R VALTER HADDAD	30	CENTRO	CAMPO REDONDO	MT
18	MARIANO FIGUEIREDO DA SILVA		26/08/78	JOAO COAZZA	163	JARDIM DOS REIS	NATAL	RN
20	ADILSON DIAS FERREIRA		14/10/66	RUA ANGIUS	508	CENTRO	NATAL	RN
21	ANTONIO ALVES PIRES		22/05/87	OLIVEIRA	45	VILA OPERARIA	NATAL	RN
22	ADAUTO PERETTI FILHO		20/07/62	R JOAQUIM NABUCO	389	CENTRO	NATAL	RN
28	ADELINO CARNELOSSI		03/09/50	R DESB CEARA	159	BOSQUE	NATAL	RN
29	ADELNIL DOS SANTOS COSTA		06/06/56	R FUAD MALLUY	140	CENTRO	MACAIBA	RN
30			24/10/62	RUA DUQUE DE CAXIAS	360	CAMPO REDONDO	CAMPO REDONDO	MT
31	ADAIR TADEU VILLA REAL		24/10/62	RUA DUQUE DE CAXIAS	360	CAMPO REDONDO	CAMPO REDONDO	MT
32	ADAIR DE ALMEIDA		01/07/45	R JOAO BUENO DA COSTA	65	CENTRO	CAMPO REDONDO	MT
34	ADAIR ALVES RIBEIRO		25/11/69	RUA IPIRANGA	577	VL FORMOSA	NATAL	RN
36	ALADIA TAPIAS GERALDO		17/10/39	R JOAQUIM NABUCO	955	CENTRO	CERRO CORA	MT
40	MARIANO LACERDA KALIL	MASCULINO	15/03/79	DONA ERMELINDA PEREIRA	314	JARDIM ESTORIL	NATAL	RN
42	ADAO TONI PIRODI	MASCULINO	21/11/72	R TUIUTI	90	JARDIM CAICARA	NATAL	RN
43	ABBADIA FRANCO CARDOSO	MASCULINO	20/12/36	R TTE NICOLAU MAFFEI	208	CENTRO	NATAL	RN
45	ADELINO DOS SANTOS COSTA	MASCULINO	06/06/56	R FUAD MALLUY	140	CENTRO	MACAIBA	RN
46	ALBINO OLIVA	MASCULINO	22/06/38	R EUCLIDES DA CUNHA	618	VILA PRQ BOA VISTA	NATAL	RN
47	JOSUÉ DOS SANTOS	MASCULINO	07/03/62	BRIGADEIRO TOBIAS	1675	JD PORTAL DO SOL	CAMPO REDONDO	MT
50	ADEMIR APARECIDO ALVES	MASCULINO	17/05/56	R RUI BARBOSA	291	CENTRO	SAO TOME	PR
53	ALCIDES PESSOA LOURENCO	MASCULINO	04/11/47	R FAGUNDES VARELLA	426	VL LESSA	NATAL	RN
54	ADELINO DOS SANTOS COSTA	MASCULINO	06/06/56	R FUAD MALLUY	140	CENTRO	MACAIBA	RN
55	ADELINO FERNANDES CARRION	MASCULINO	09/03/47	R 7 DE SETEMBRO	87	JARDIM AVIACAO	NATAL	RN
56	ADOLFO RODRIGUES DE ALMEIDA NETO	MASCULINO	28/04/84	R JERONIMO GARCIA DUARTE	86	VILA SANTA TEREZA	NATAL	RN
57	MARIOA DE FREITAS VIEIRA	MASCULINO	26/09/45	AV PRES VARGAS	574	JARDIM PAULISTA	ANGIUS	RN
58		MASCULINO	25/03/32	RUA MENDES DE JESUS	775	CENTRO	NATAL	RN

Figura 4 - Planilha de Excel 1. Tabela Clientes com linhas sem nome. Fonte: Paula J. Piloto Langhi (2020)

3.4 O PROCESSO DE CARGA: CARREGAMENTO DOS DADOS TRANSFORMADOS EM UM DATA WAREHOUSE OU COMO SUPLEMENTO EM OUTROS SISTEMAS

A última etapa do processo de ETL, a carga, baseia-se em fazer a inserção dos dados que já foram transformados para ambientes como *Data Warehouse* ou até mesmo outros sistemas de softwares que possam fazer uso desses dados e gerar relatórios de alto nível para tomadas de decisões gerenciais de um negócio. Com os dados prontos para uso é possível extrair as mais específicas informações dos negócios da empresa como quem são as pessoas que fizeram as melhores vendas e até mesmo dados de satisfação do cliente após a venda de um determinado produto.

3.4.1 CARREGANDO OS DADOS TRANSFORMADOS EM UM DATA WAREHOUSE

Data Warehouse é um conceito superior de tipo de banco de dados onde nele os dados são organizados para serem usados de forma inteligente integrando informações de todas as partes da empresa, criando um mapa de relações onde é possível enxergar os detalhes das transações comerciais para servirem de uso em relatórios de tomadas de decisões gerenciais. Dentre as características do *Data Warehouse* as que mais se destacam é a sua integração com os ambientes operacionais dos sistemas das empresas de forma a padronizar os dados provindos de diversas fontes. Outra característica importante é que, os dados podem ser minerados, tendo em conta que o *Data Warehouse* mantém os históricos dos dados promovendo um arquivo completo e facilmente acessível, devido á todas essas peculiaridades desse banco de dados, cria-se uma confiabilidade maior acerca das informações, uma vez que, as informações da empresa estão sendo padronizadas proporcionando uma fonte confiável e segura para extração dos relatórios e criação de painéis de monitoramento.

O processo ETL determina o tratamento dos dados de acordo com as regras e passos inseridos nas ferramentas utilizadas. O software executor dos passos do processo ETL realizará as tarefas programadas, porém o tempo é um fator decisivo até a finalização de cada passo, dependendo da quantidade de regras que estão sendo aplicadas aos dados, existe a dependência de em quanto tempo essa tarefa será realizada, surgindo assim o conceito de temporalidade.

Quando o processo ETL está sendo executado, recursos são consumidos da máquina ativa e das fontes de dados que enviam as informações para a ferramenta de ETL, desta forma a máquina abriga os dados para a extração e uma das possíveis fontes pode ser o banco de dados dos sistemas da empresa. O processo de extração executará comandos para selecionar os dados e em sequência enviar para a próxima fase da camada ETL. Ao mesmo tempo em que, este banco reúne as informações, outros sistemas, pessoas e processos também fazem acesso aos mesmos dados pedindo requisição de informações, podendo sobrecarregar as requisições ao banco causando lentidão nas suas operações e conseqüentemente, fazendo com que os sistemas dependentes deste mesmo banco tenham uma demora além do normal. Como o processo de ETL requer muito processamento dependendo da quantidade de dados a serem transformados, as rotinas para realização deste devem ser executadas em horários de menor pico de

uso do sistema sem causar lentidão nos outros sistemas que consomem o mesmo recurso do banco.

Bancos de dados extensos podem levar até mais de um dia para realizar as rotinas daquele exato período e enquanto a ferramenta realiza o processo de ETL, pilhas e pilhas de dados já estão sendo gerados entrando em uma fila de espera que irá consumir processamento e hardware excessivo da máquina que está fazendo o processamento.

Diante do exposto acima, é possível perceber que, escolher a temporalidade para a execução das rotinas ETL é essencial para o resultado com boa performance, além do que, a geração em massa de novos dados pode requerer que o processo ETL seja executado mais vezes, fazendo com que as análises tornem-se mais precisas em tempos mais hábeis.

4 PRINCIPAIS FERRAMENTAS UTILIZADAS NO PROCESSO ETL

Como já mencionado, os dados a serem extraídos e carregados em um *Data Warehouse*, podem estar contidos em diversas fontes, tais como blocos de notas, planilhas de Excel, arquivos XML, documentos Word, sistemas ERP, sistemas CRM entre outros. Devido às inerentes possibilidades de banco de dados, os softwares que trabalham com o processo de ETL demandam constante atualização para se adaptar às diferentes formas de banco de dados que existem, como também as futuras novas tecnologias emergentes.

Cada empresa possui seus sistemas e suas formas de gerar bases de dados, existem inúmeras ferramentas atuais que possibilitam a realização de um processo ETL. Uma empresa não precisa ter um sistema específico que funcione apenas com um software de ETL, o mercado de *Business Intelligence* é amplo e competitivo, oferece altas opções de software para cada necessidade desde as ferramentas sem custo até as ferramentas pagas.

Atualmente, estão entre as principais ferramentas que possibilitam o processo de ETL:

Pentaho Data Integration: Ferramenta gratuita com interface gráfica de simples entendimento para a realização das transformações ETL. O software é

disponibilizado pela empresa Hitachi. Algumas empresas que utilizam o Pentaho são: OLX, CERN (Organização Europeia para a Pesquisa Nuclear), Nasdaq e entre outras.

Skyvia: Software online para soluções ETL com diversas compatibilidades com outros softwares e possui um banco de dados próprio para armazenamento na nuvem, além de prometer pouco nível de conhecimento na ferramenta para começar a usar. A própria empresa Skyvia oferece alguns planos gratuitos de uso, porém com limitações.

Cognos Data Manager: Mais uma ferramenta online que utiliza Inteligência de negócios e faz uso do processo ETL é o Cognos Data Manager diretamente da IBM, onde o usuário envia as fontes de dados para a ferramenta e o software faz todo o processamento e traz as opções de listas inteligentes para fazer uso dos dados. Uma versão de experimento gratuita está disponibilizada pela IBM, caso contrário a empresa oferece planos de mensalidade da ferramenta. Destaca-se algumas empresas que utilizam o Cognos são: ECCO, FleetPride, Wimbledon e entre outras.

Hadoop: Software open-source gratuito que armazena diversos tipos de dados e executa aplicações no seu sistema de clusters em um grande nível de processamento e podendo executar diversas tarefas ao mesmo tempo. As empresas que fazem uso deste software são: MasterCard, World Wide Technology, Cerner e entre outras.

5 CONCLUSÃO

No início deste trabalho foi apresentado uma visão de sistemas de empresas contemplando formas de operação para registros de transações e armazenamento de dados junto com os possíveis erros comuns de gravação dos dados e a importância dos registros de informações para tratamentos de análises futuras. O processo da carga é a injeção dos dados provindos de diversas fontes de dados para dentro da ferramenta de ETL. A transformação explicou formas em que os dados podem ser padronizados e limpos com diferentes regras de negócio aplicáveis.

A etapa de carregamento abordou o uso do *Data Warehouse* como o banco de dados para onde as informações serão migradas e de onde serão retiradas as informações necessárias para a elaboração de relatórios, análises e mineração de dados com base no histórico das informações inseridas e ainda diferentes formas de temporalidade.

Atualmente quando pesquisamos por algum assunto é feito uma análise nessa pesquisa e futuramente somos bombardeados com toneladas de propagandas referentes á tal, faça um teste, abra seu navegador e digite “comprar mouse” em seguida abra paginas de algumas lojas que apareceram na pesquisa e depois feche todas, vá á qualquer site onde tenha propagandas como, por exemplo, um site de noticias e estará lá aquele objeto pesquisado dentro da propaganda te pedindo para comprar com apenas alguns cliques, até onde os registros dos seus dados podem te persuadir a fazer uma especifica compra? Ou apoiar algum candidato de politica? Ou até mesmo fazer sua opinião mudar em relação á algum assunto? Portanto os processos que envolvem tratamento dos dados como camada ETL para análises futuras de *Business Intelligence* são poderosos por agirem especificamente em um grupo seletivo de pessoas persuadindo-as com base nos seus dados para desenvolver os futuros produtos e projetos para melhor atender as demandas da sociedade.

REFERÊNCIAS BIBLIOGRÁFICAS

TURBAN, Efraim. **Business Intelligence – Um enfoque Gerencial**. Porto Alegre, 2009. Acesso em 27 de Maio de 2020, 110p.

Diego. O que significa OLTP e OLAP na prática? . **Canal Tech**, 2014. Disponível em: < <https://canaltech.com.br/business-intelligence/o-que-significa-oltp-e-olap-na-pratica/> >. Acesso em 17 de ago. de 2020.

Cetax. O QUE É ETL – EXTRACT TRANSFORM LOAD? . **Cetax**, 2016. Disponível em: < <https://www.cetax.com.br/blog/etl-extract-transform-load/> >. Acesso em 19 de ago. de 2020.

Da Redação. Veja como as empresas usam Hadoop para reforçar seus projetos de big data. **Computer World**, 2016. Disponível em: < <https://computerworld.com.br/2016/01/22/como-sete-empresas-usam-hadoop-para-reforcar-aplicacoes-de-big-data/> >. Acesso em 20 de ago. de 2020.

Helder. Entenda o que são Fontes de Dados. **Cultura Analítica**, 2018. Disponível em: < <https://culturaanalitica.com.br/o-que-sao-fontes-de-dados/> >. Acesso em 29 de Maio de 2020.

IBM Cognos Business Intelligence. **IBM**, 2020. Disponível em: <https://www.ibm.com/support/knowledgecenter/en/SSEP7J_10.2.2/com.ibm.swg.ba.cognos.cbi.doc/bi_qsg.htm >. Acesso em 20 de ago. de 2020.

Marcelo, Daniel e Fábio. Pentaho Data Integration - ETL em Software Livre. **Infoq**, 2017. Disponível em: < <https://www.infoq.com/br/articles/pentaho-pdi/> >. Acesso em 20 de ago. de 2020.

Leandro. Saiba quais empresas estão usando o Pentaho. **Know Solution**, 2016. Disponível em: < <https://www.knowsolution.com.br/saiba-quais-empresas-estao-usando-o-pentaho/> >. Acesso em 20 de ago. de 2020.

Hadoop. **SAS**, 2019. Disponível em: < https://www.sas.com/pt_br/insights/big-data/hadoop.html >. Acesso em 20 de ago. de 2020.

Data Integration. **Skyvia**, 2020. Disponível em: < <https://skyvia.com/data-integration/> >. Acesso em 20 de ago. de 2020.

15 Best ETL Tools In 2020 (A Complete Updated List). **Software Testing Help**, 2020. Disponível em: < <https://www.softwaretestinghelp.com/best-etl-tools/> >. Acesso em 03 de jun. de 2020.