

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE ELETRÔNICA
MBA EM GESTÃO DA TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO

DANILO FERREIRA DE FREITAS

**ANÁLISE DE FERRAMENTAS DE *BUSINESS INTELLIGENCE*: COMO
FERRAMENTAS *OPEN SOURCE* PODEM REDUZIR CUSTOS EM
NEGÓCIOS**

MONOGRAFIA

CURITIBA

2017

DANILO FERREIRA DE FREITAS

**ANÁLISE DE FERRAMENTAS DE *BUSINESS INTELLIGENCE*: COMO
FERRAMENTAS *OPEN SOURCE* PODEM REDUZIR CUSTOS EM
NEGÓCIOS**

Monografia apresentada como requisito parcial à obtenção do título de Especialista em Gestão da Tecnologia da Informação e Comunicação, do Departamento Acadêmico de Eletrônica, da Universidade Tecnológica Federal do Paraná.

Orientador: Prof. MSc. Luiz Pinheiro Junior

CURITIBA

2017



TERMO DE APROVAÇÃO

Título da monografia

Análise de Ferramentas de *Business Intelligence*: Como Ferramentas *Open Source* Podem Reduzir Custos em Negócios

Por

Danilo Ferreira de Freitas

Esta monografia foi apresentada às **19:30h** do dia **29/09/2017** como requisito parcial para a obtenção do título de Especialista no CURSO DE ESPECIALIZAÇÃO EM GESTÃO DE TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO, da Universidade Tecnológica Federal do Paraná, **Campus Curitiba**. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho:

1		Aprovado
2		Aprovado condicionado às correções Pós-banca, postagem da tarefa e liberação do Orientador.
3		Reprovado

Prof. Msc. Alexandre Jorge Miziara
UTFPR - Examinador

Prof. Msc. Luiz Pinheiro Junior
UTFPR – Orientador

Prof. Msc. Alexandre Jorge Miziara
UTFPR – Coordenador do Curso

Dedico este trabalho à minha esposa
Fernanda, pelo amor, apoio e
compreensão de sempre.

AGRADECIMENTOS

Agradeço a Deus por ter me guiado até aqui.

Também agradeço aos meus professores, que demonstraram dedicação e conhecimento ao longo do curso.

Gostaria de deixar registrado também, o reconhecimento à minha família, pois eles sempre foram o apoio para as conquistas em minha vida.

RESUMO

FREITAS, Danilo Ferreira de. **Análise de ferramentas de *Business Intelligence*: como ferramentas *open source* podem reduzir custos em negócios**. 2017. 58p. Monografia de Especialização em Gestão da Tecnologia da Informação e Comunicação - Universidade Tecnológica Federal do Paraná. Curitiba, 2017.

O objetivo deste trabalho é explorar como a utilização de softwares de Inteligência de Negócios (BI) *Open Source* pode contribuir nos custos de licenciamento para as organizações. Este objetivo destaca-se em opções de software livre para Inteligência de Negócios que podem eventualmente substituir ferramentas pagas em determinados contextos. Este objetivo se alcança através da demonstração das funcionalidades de algumas ferramentas escolhidas para análise, as quais serão vistas em detalhes e comparadas entre si. Para isso foi realizada uma revisão da literatura sobre inteligência dos negócios conceituando as principais ferramentas, seus contextos e uma explanação sobre *Big Data*. Foi utilizada uma metodologia qualitativa comparando-se três ferramentas *open source* de BI, sendo elas: Jaspersoft, Pentaho e SpagoBI. Na análise de dados foram analisados 16 critérios: versões, componentes, licenças de utilização, idiomas de documentação, desempenho, consultas ad hoc, arquitetura, plataforma, suporte técnico e documentação, agendamento, dashboards, exportação de relatórios, suporte a dispositivos móveis, análise preditiva, custo e carregamento de dados de diferentes fontes (Big Data). Conclui-se que as ferramentas *open source* citadas neste trabalho podem ser usadas, dentro de suas capacidades, para substituir ferramentas pagas de BI por apresentarem diversas funcionalidade compatíveis com as disponíveis no mercado hoje. Como sugestões futuras que sejam feitas mais análises comparativas utilizando-se os mesmos critérios, porém com ferramentas tanto *open source* quanto pagas.

Palavras-chave: *Business Intelligence*. *Open source*. Jaspersoft. Pentaho. SpagoBI.

ABSTRACT

FREITAS, Danilo Ferreira de. **Análise de ferramentas de *Business Intelligence*: como ferramentas *open source* podem reduzir custos em negócios**. 2017. 58p. Monografia de Especialização em Gestão da Tecnologia da Informação e Comunicação - Universidade Tecnológica Federal do Paraná. Curitiba, 2017.

The purpose of this paper is to explore how the use of Open Source Business Intelligence (BI) software can contribute to the licensing costs for organizations. This goal stands out in free software options for Business Intelligence that can eventually replace paid tools in certain contexts. This objective is reached through the demonstration of the functionalities of some tools chosen for analysis, which will be seen in detail and compared to each other. For this, a review of the literature on business intelligence was carried out, conceptualizing the main tools, their contexts and an explanation about Big Data. A qualitative methodology was used comparing three open source BI tools: Jaspersoft, Pentaho and SpagoBI. In the data analysis, 16 criteria were analyzed: versions, components, user licenses, documentation languages, performance, ad hoc queries, architecture, platform, technical support and documentation, scheduling, dashboards, report export, mobile device support, analysis predictive, cost and load data from different sources (Big Data). It is concluded that the open source tools mentioned in this work can be used, within their capabilities, to replace paid BI tools by presenting several functionalities compatible with those available in the market today. As future suggestions, it could be made more comparative analysis using the same criteria, but with both open source and paid tools.

Keywords: Business Intelligence. Open source. Jaspersoft. Pentaho. SpagoBI.

LISTA DE ILUSTRAÇÕES

Figura 1 – As cinco forças que formam a concorrência da indústria.	19
Figura 2 – Passos necessários para o data mining.	23
Figura 3 – Esquema estrela e cubo OLAP.	24
Figura 4 – Interface do JasperReports Server.	33
Figura 5 – Ambiente no qual o JasperReports é executado.	35
Figura 6 – Plataforma analítica do Pentaho.	38
Figura 7 – Dashboard Interativo Pentaho.	39
Figura 8 – Ferramenta de Integração de Dados do Pentaho.	40
Gráfico 1 – O BI provoca benefícios tangíveis e intangíveis substanciais.	20
Gráfico 2 – Onde o BI provoca resultados.	20
Gráfico 3 – Dados estruturados e não estruturados, repetitivos e não repetitivos. ...	25
Gráfico 4 – Big Data consiste em dados não estruturados.	26
Gráfico 5 – Aplicação de BI em Software Livre em percentual Departamental e Corporativo.	28
Gráfico 6 – BI em Software livre nas empresas por tamanho.	28
Quadro 1 – Comparação entre aplicações OLAP e OLTP.	22
Quadro 2 – Fontes oficiais de pesquisa das ferramentas	29
Quadro 3 – Principais diferenças entre as versões Jaspersoft Community Edition e Commercial Edition.	32
Quadro 4 - Descrição das versões das ferramentas de Inteligência de Negócios analisadas.	47
Quadro 5 – Critérios básicos: Características consideradas básicas para qualquer ferramenta de Inteligência de Negócios	48
Quadro 6 – Relatórios: Usabilidades dos relatórios e gráficos	50
Quadro 7 – Funcionalidades Web: Disponibilidade da Empresa via Web para suporte.	52
Quadro 8 – Simulação de Cenários: Análises feitas com a interação do usuário	53
Quadro 9 – Produto: Custos do produto no mercado	54
Quadro 10 – Ferramenta de Planejamento: Modo de distribuição dos relatórios.....	54

LISTA DE SIGLAS

BI	Business Intelligence (BI) é um termo que pode ser usado como inteligência de negócios. É um método para tomada de decisões inteligentes, mediante dados e informações.
DW	Data Warehouse, em português tem como significado armazém de dados e é variável com o tempo para auxiliar na tomada de decisões.
GNU	General Public License ou Licença Pública Geral GNU, é uma licença de software cujos softwares são distribuídos sob a mesma licença. É utilizada por projetos de software livre e de código aberto.
GUI	Refere-se à denominação "Graphical User Interface" ou Interface Gráfica do Usuário, que é em um modelo de interface do utilizador para a interação com dispositivos.
OEM	Sigla de "Original Equipment Manufacturer", ou "Fabricante Original do Equipamento". São produtos fabricados especialmente para grandes empresas, não destinados ao consumidor final.
OLAP	Online Analytical Processing é a capacidade de analisar grandes volumes de dados.
OLTP	Online Transaction Processing se refere aos sistemas transacionais.

SUMÁRIO

1 INTRODUÇÃO	13
1.1 JUSTIFICATIVA	14
1.2 PROBLEMA E OBJETIVOS.....	15
1.2.1 Objetivo Geral	15
1.2.2 Objetivos Específicos	16
1.3 ESTRUTURA DO TRABALHO	16
2 INTELIGÊNCIA DE NEGÓCIOS	18
2.1 TECNOLOGIAS PARA INTELIGÊNCIA DE NEGÓCIOS.....	21
2.2 'ETL': EXTRAÇÃO, TRANSFORMAÇÃO E CARGA DE DADOS	22
2.3 ESQUEMAS ESTRELA E CUBOS OLAP	23
2.4 <i>BIG DATA</i>	24
2.5 FERRAMENTAS GRATUITAS PARA INTELIGÊNCIA DE NEGÓCIOS	26
3 PROCEDIMENTOS METODOLÓGICOS	29
3.2 CRITÉRIO DE ESCOLHA DAS FERRAMENTAS	30
3.3 CRITÉRIOS DE COMPARAÇÃO ENTRE AS FERRAMENTAS	30
4 ANÁLISE COMPARATIVA DAS FERRAMENTAS	32
4.1 JASPERSOFT	32
4.1.1 Componentes do Jaspersoft.....	34
4.1.2 Referência de desempenho do servidor Jasper Reports.....	35
4.2 PENTAHO	36
4.2.1 Componentes do Pentaho	37
4.2.2 Big Data com Pentaho.....	38
4.2.3 Integração de dados com Pentaho.....	40
4.2.4 Performance com Pentaho.....	40
4.3 SPAGOBI	41
4.3.1 Diferenciais do SpagoBI	41
4.3.2 Componentes do SpagoBI	42
4.3.3 Arquitetura do SpagoBI	44
4.3.4 SpagoBI Analytics	45
4.4 ANÁLISE COMPARATIVA ENTRE AS FERRAMENTAS	46
4.4.1 Considerações sobre o Jaspersoft	55
4.4.2 Considerações sobre o Pentaho	55
4.4.3 Considerações sobre o SpagoBI.....	56
5 CONSIDERAÇÕES FINAIS	57
5.1 LIMITAÇÕES DA PESQUISA E TRABALHOS FUTUROS	58

1 INTRODUÇÃO

O conhecimento estratégico de negócio tem sido exigido constantemente pelo mercado, afim de que uma organização possa se destacar ou atingir suas metas de acordo com seus objetivos. Para isso, tem-se utilizado cada vez mais ferramentas que auxiliam na exploração de dados que são traduzidos em informações que geram conhecimentos até então implícitos, tanto sobre o próprio negócio, quanto do mercado e até mesmo de concorrentes.

Um exemplo que pode ser citado é o mostrado pela EXAME (2016) onde foi visto que quando empresas analisam o comportamento dos indivíduos na internet, estas podem oferecer produtos ou serviços personalizados. Para isso, dados são coletados para estruturar pesquisas e ações em diversos sites e os resultados servirão de informação útil para definir novas estratégias de mercado.

As plataformas de *Business Intelligence* vêm sendo cada vez mais usadas pela gestão de empresas de grande e médio porte. Essa tendência de usar sistemas que transformam um grande volume de dados em informações estratégicas e para apoiar decisões também está se intensificando na administração pública para acompanhar, dinamicamente, indicadores estratégicos como o de arrecadação.

Com isso, baseado nessa necessidade de mercado e também visando elaborar um material acadêmico de pesquisa com base em dados relevantes coletados de diversas fontes explicitadas nos procedimentos metodológicos, este trabalho tem como tema principal uma análise comparativa de ferramentas de Inteligência de Negócios, também conhecida pelo termo *Business Intelligence* (ou simplesmente BI)

Considerou-se a utilização de ferramentas de software livre ou *open source*, como forma de redução de custos na implementação de Inteligência de Negócios pois, como será visto, existem opções bastante robustas disponíveis para livre utilização e por outro lado a aquisição de uma licença de software deste tipo pode ser inviável em diversos casos, como em pequenas empresas ou onde há necessidade de corte de custos.

Diz-se software livre quando o código é aberto e se pode ler, auditar, realizar modificações e adaptações. Diversas empresas de tecnologia - Google, Facebook, Amazon, Twitter e Microsoft têm partes de sua estrutura baseadas em código aberto assim como diversas *startups* de tecnologia se beneficiam deste modelo de

desenvolvimento de software. Uma razão é a econômica: ao adotar soluções livres, não é necessário comprar programas prontos (com licenças muitas vezes de alto custo). Outra razão é o aproveitamento de programas já criados e que podem ser adaptados às necessidades da gestão.

Além disso, há razões de segurança e privacidade de dados pois quando um código é aberto, ele pode ser auditado minuciosamente e pode se ter a certeza que não há nenhum programa espião embutido pois se torna possível ler cada linha de código dele. Por outro lado, um ponto negativo que pode ser considerado é que as soluções em software livre requerem desenvolvimento, treinamento e adaptação dos utilizadores. Por isso, é um processo que pode ser custoso uma vez que muitas pessoas estão habituadas a utilizar soluções prontas vendidas comercialmente.

1.1 JUSTIFICATIVA

Os dados estatísticos e preditivos fornecidos por sistemas de apoio a decisão, permitem que uma organização tome decisões seguras. Porém, muitas organizações deixam de explorar informações cruciais que estão disponíveis nas bases de dados de seus sistemas de informações pelo fato de ser necessário dispor de tempo e investimento financeiro para organizar e transformar esses dados em informações úteis para a tomada de decisões e por não conhecerem ferramentas alternativas as pagas, material de estudo deste trabalho.

A partir do momento que as organizações passam a utilizar um sistema de forma livre ou em parte gratuito, o investimento financeiro pode diminuir, porém pode-se manter a qualidade da exploração de dados para a tomada de decisões. Isto, portanto, pode ser a solução para aquelas que não podem despende de muito investimento com ferramentas de Inteligência de Negócios, ou ainda, das que gostariam de reduzir custos, substituindo-se as ferramentas pagas utilizadas atualmente, por ferramentas livres ou parcialmente livres de custos com licença de software.

Com isso, este trabalho se justifica por fornecer um material de pesquisa que pode servir de apoio aos que buscam maiores informações sobre as ferramentas em software livre na área de Inteligência de Negócios, tanto para a área acadêmica quanto para a utilização prática em negócios.

Na área acadêmica este material poderá ser utilizado para obter de forma prática e resumida os benefícios que cada ferramenta pode trazer na geração de conhecimento a partir de diversas fontes de dados, como é o caso por exemplo do Big Data. Já para as organizações, este trabalho abrange de forma simples e resumida as vantagens de cada ferramenta estudada, e como elas podem ser incorporadas em seu portfólio de serviços para agregar valor a seus produtos e realizar melhores tomadas de decisões.

1.2 PROBLEMA E OBJETIVOS

O processo de integração e limpeza dos dados, muitas vezes, não é vista como prioridade pelos gestores por seu alto custo e difícil estimativa de retorno financeiro. Contudo, este processo torna confiável e assertiva a tomada de decisão em negócios.

Sendo assim, este estudo busca responder a seguinte pergunta de pesquisa:

Como a utilização de softwares de Inteligência de Negócios Open Source pode contribuir nos custos de licenciamento para as organizações?

Para responder essa pergunta de pesquisa, estruturou-se o objetivo geral e específicos a seguir.

1.2.1 Objetivo Geral

O objetivo geral deste trabalho é explorar como a utilização de softwares de Inteligência de Negócios *Open Source* pode contribuir nos custos de licenciamento para as organizações. Este objetivo destaca-se em opções de software livre para Inteligência de Negócios que podem eventualmente substituir ferramentas pagas em determinados contextos.

Este objetivo se alcança através da demonstração das funcionalidades de algumas ferramentas escolhidas para análise, as quais serão vistas em detalhes nos próximos capítulos e comparadas entre si.

1.2.2 Objetivos Específicos

O objetivo geral deste trabalho é explorar como a utilização de softwares de Inteligência de Negócios *Open Source* pode contribuir nos custos de licenciamento para as organizações. Neste sentido, detalham-se em objetivos específicos:

1. Identificar as ferramentas de Inteligência de Negócio *Open Source* do mercado;
2. Comparar três ferramentas de Inteligência de Negócio *Open Source*;
3. Compreender como as ferramentas de Inteligência de Negócio *Open Source* podem contribuir nos custos de licenciamento.

Além disso, este material servirá de apoio para futuras pesquisas na área acadêmica envolvendo softwares para produção de Inteligência de Negócios, e que sejam desenvolvidos em software livre ou ainda que possam ser comparados entre si, independentemente de serem ferramentas pagas ou não.

1.3 ESTRUTURA DO TRABALHO

Este trabalho está dividido em cinco capítulos, os quais contém as seguintes sessões descritas a seguir.

No primeiro capítulo é feita a introdução, com a justificativa do desenvolvimento do trabalho assim como o problema de pesquisa e os objetivos gerais e específicos.

Em seu segundo capítulo são apresentados conceitos e ferramentas de Inteligência de Negócios e sua importância do momento atual, com alguns pontos importantes para o entendimento de como podem ser utilizadas para a tomada de decisões. Também é apresentado o conceito do chamado processo ETL o qual inclui a extração, transformação e carga de dados. Também são vistos os conceitos de cubo de dados, estrelas OLAP, Big Data e ferramentas gratuitas para Inteligência de Negócios.

Já a partir do terceiro capítulo, são apresentados os procedimentos metodológicos para a pesquisa com as ferramentas escolhidas para este trabalho, apresentando os critérios de escolha e as formas de análise escolhidas para um comparativo entre as ferramentas de BI, em ordem alfabética: Jaspersoft, Pentaho e SpagoBI.

No quarto capítulo é feito um comparativo entre as ferramentas estudadas, apresentando-se as principais características de cada uma delas e utilizando-se critérios comuns a ferramentas de BI.

Por último, no quinto capítulo encerra-se o trabalho através de uma conclusão que cita quais ferramentas podem ser úteis de acordo com o perfil de cada organização ou usuário.

2 INTELIGÊNCIA DE NEGÓCIOS

Existe uma série de definições para Inteligência de Negócios, e uma das descrições gerais mais amplamente aceitas vem do Gartner: “um termo abrangente que inclui os aplicativos, infraestrutura e ferramentas e as melhores práticas que permitem o acesso e análise de informações para melhorar e otimizar decisões e desempenho.” (FORBES INSIGHTS, 2016)

Pode-se dizer que o conceito de BI já era usado pelos povos antigos, como a sociedade do Oriente Médio antigo que utilizava os princípios básicos do BI quando cruzavam informações obtidas junto à natureza em benefício de suas aldeias, como o comportamento das marés e os períodos de seca. Estas informações eram utilizadas para tomada de decisões importantes. (PRIMAK, 2008)

Philip Kotler (2015), em seu livro “Confronting Capitalism, Real Solutions for a Troubled Economic System” cita que Andy Grove, quando era o CEO da Intel, teve que lidar com a incerteza e com todos os tipos de ameaças no negócio de fabricação de chips de computadores. Seria preciso apenas um competidor ágil fabricar com um chip superior a um preço mais baixo para possivelmente derrubar a Intel. Então, a Intel teve que criar um sistema de alerta precoce que revelasse sinais de problemas iminentes, criar diferentes cenários de precauções, e planejar diferentes respostas aos diferentes cenários no caso deles ocorrerem.

Os negócios funcionam melhor quando se têm uma imagem clara do seu público-alvo, seus concorrentes e as principais forças demográficas, econômicas, sociais, tecnológicas e políticas que afetam o ambiente do negócio. Por outro lado, as empresas operam mal quando muitos desses elementos que afetam profundamente o desempenho do negócio são aleatórios e imprevisíveis (KOTLER, 2015).

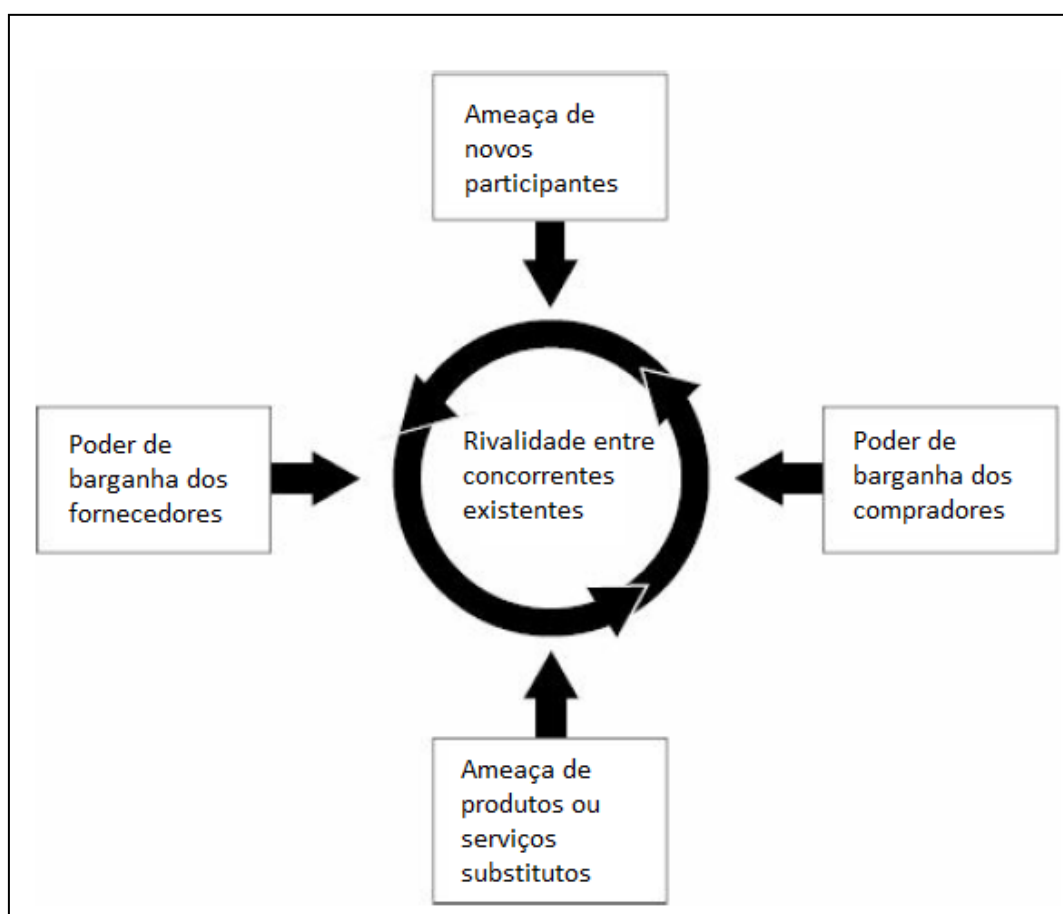
Ainda, segundo KOTLER (2015), a turbulência de mercado é agora a condição normal das indústrias, mercados e empresas. A turbulência é a nova normalidade, pontuada por intermitências de prosperidade e desaceleração. Além dos desafios cotidianos de lidar em uma arena competitiva, os líderes empresariais precisam reconhecer todos os pontos que desafiam seu planejamento de negócios.

A Tecnologia da Informação é o apoio que permite às organizações nutrirem-se de informações. Sendo assim, existe um grande potencial a ser descoberto pelas

empresas: a necessidade de obter informações do ambiente de negócios externo e incorporá-las ao processo de tomada de decisão. Esta descoberta está relacionada com a inteligência competitiva (POZZEBON *et al*, 1997).

Michael Porter (2008), diz que compreender as forças competitivas (**Figura 1**) e suas causas subjacentes revela as raízes da rentabilidade atual de uma indústria, ao mesmo tempo que fornece uma estrutura para antecipar e influenciar a concorrência (e a lucratividade) ao longo do tempo. Uma estrutura industrial saudável deve ser uma preocupação tanto competitiva para os estrategistas como a própria posição de sua empresa. Compreender a estrutura da indústria também é essencial para o posicionamento estratégico efetivo. Defender a força competitiva e moldá-los em favor da empresa são cruciais para a estratégia.

Figura 1 – As cinco forças que formam a concorrência da indústria.

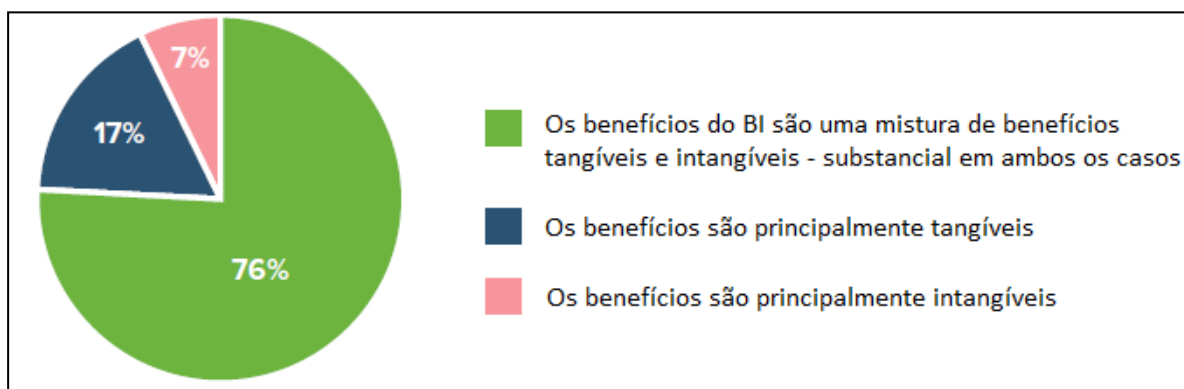


Fonte: Adaptado de PORTER (2008).

Em pesquisas conduzidas pela Forbes Insights (2016), as organizações informam que estão obtendo retornos inovadores dos investimentos em Inteligência

de Negócios (BI). As soluções são incorporadas em líderes da indústria em todo o mundo. Enquanto isso, com tantos mais dados e soluções de análise, os executivos e seus relatórios estão sendo transformados de simples gerentes para verdadeiros trabalhadores da informação. O sucesso desses programas está gerando benefícios substanciais, tangíveis e intangíveis, conforme o **Gráfico 1**.

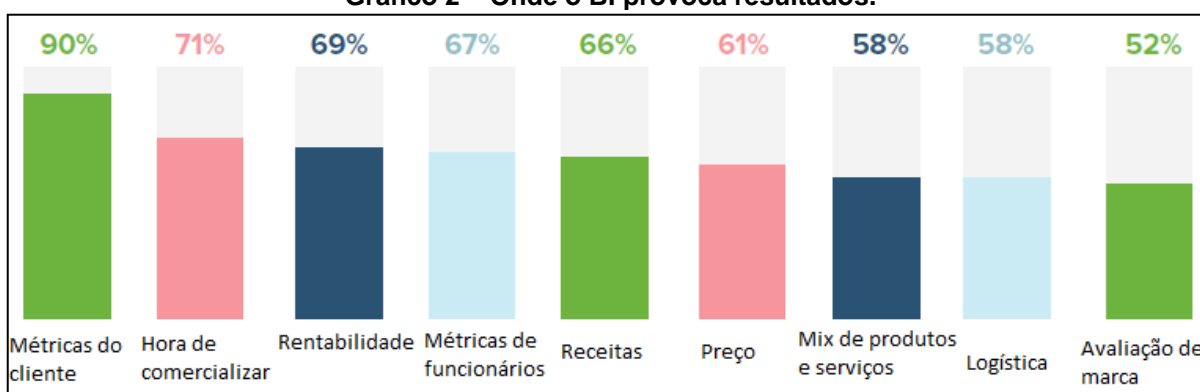
Gráfico 1 – O BI provoca benefícios tangíveis e intangíveis substanciais.



Fonte: Adaptado de FORBES INSIGHTS (2016).

Ainda em termos de benefícios do BI, eles abrangem uma ampla gama de atividades organizacionais e processos de negócios. Estes incluem métricas aprimoradas do cliente, tempo acelerado para o mercado, mixes de produtos e serviços mais fortes, valorização e reconhecimento de marca aprimorados e maior rentabilidade, conforme o **Gráfico 2**.

Gráfico 2 – Onde o BI provoca resultados.



Fonte: Adaptado de FORBES INSIGHTS (2016).

Sob um outro aspecto, nos dias atuais os governos podem utilizar BI até mesmo para obter inteligência que permita uma melhor compreensão e análise de

grupos terroristas e extremistas que utilizam a chamada *Dark Web*, a parte da *World Wide Web* utilizada para atingir objetivos ilícitos. Organizações terroristas, grupos extremistas, grupos de ódio e grupos de supremacia racial estão usando a Web para promover sua ideologia, facilitar as comunicações internas, atacar seus inimigos e realizar atividades criminosas, e estes geram dados que podem ser cruzados e analisados por ferramentas de BI. (CHEN, 2012)

2.1 TECNOLOGIAS PARA INTELIGÊNCIA DE NEGÓCIOS

De acordo com SANTOS e RAMOS (2006), os sistemas de *Business Intelligence* estão tradicionalmente associados a três tecnologias: *Data Warehouses (DW)*, *Online Analytical Processing (OLAP)* e *Data Mining*. Um *Data Warehouse*, de uma forma resumida, é um armazém de dados, analisados por uma ferramenta OLAP e/ou ferramentas de *Data Mining*.

Ainda segundo SANTOS e RAMOS (2006), estes sistemas devem ser vistos apenas como ferramentas de apoio que dentro das circunstâncias organizacionais e de gestão adequadas favorecem a inteligência, aprendizagem e criatividade organizacional. São necessárias então políticas adequadas de gestão de recursos humanos, gestão estratégica e operacional sob o risco destes sistemas não cumprirem os objetivos para que foram adotados.

As soluções de Inteligência de Negócios compõem-se por diversos elementos que tornam possível um ambiente de tomada de decisão. Elas podem combinar os componentes para cada situação, sendo que para a camada de apresentação dos dados (informações) utilizam-se as soluções OLAP. Os dados vindos dos modelos dimensionais são geralmente apresentados pelas soluções OLAP, podendo ser de maneira tabular ou gráfica ou ainda combinadas com técnicas de mineração de dados. (CECI, 2012)

O processo OLAP se diferencia do processo de transações online (OLTP), pois o OLTP é focado em processamento de transações repetitivas em grande quantidade e de manipulação simples. Já o OLAP envolve uma análise de muitos itens de dados em relacionamentos complexos, buscando padrões, tendências e exceções, conforme o Quadro 1. (PRASAD, 2007)

Quadro 1 – Comparação entre aplicações OLAP e OLTP.

Características	OLTP	OLAP
Tipo de banco de dados	Transacional	Analítico
Processamento de dados	Operacional	Análise de dados
Dimensões	Uma dimensão	Diferentes dimensões
Utilizadores	Pessoal do serviço operacional	Pessoal da gerência e estratégia
Tipo de dados	Detalhes das transações	Dados históricos e sumarizados
Tamanho do banco de dados	Menor se comparado ao OLAP por ser normalizado	Maior que o OLTP por apresentar redundância de dados

Fonte: Adaptado de PRASAD (2007).

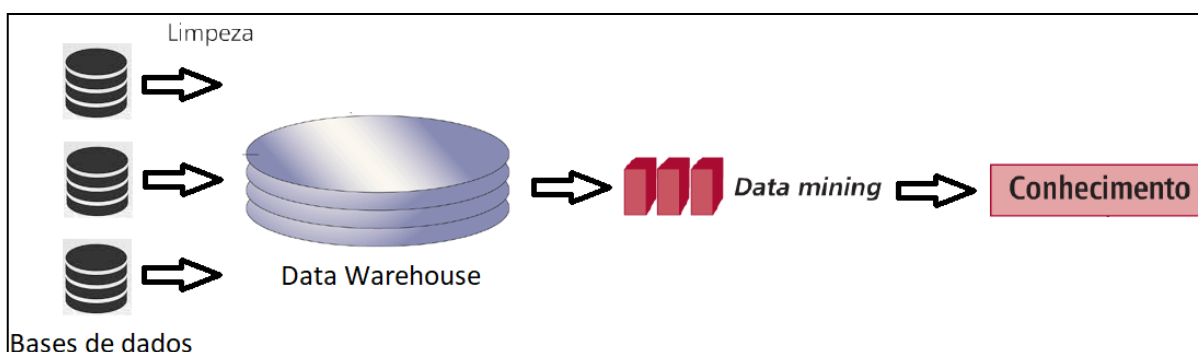
2.2 'ETL': EXTRAÇÃO, TRANSFORMAÇÃO E CARGA DE DADOS

Nos anos 1980, com o surgimento dos novos movimentos metodológicos, os dados atingiram uma espécie de estrelato, pela primeira vez. Foi a época do surgimento da administração de dados, da modelagem de dados, da engenharia da informação e da análise de dados. (BARBIERI, 2011)

E no final da década de 1980 e início da década de 1990, os bancos de dados de suporte à decisão começaram a ser implantados separadamente dos bancos de dados transacionais e, em vez disso, foram implantados no que se tornou conhecido como *data warehouses*. Este modelo de separação das cargas de dados foi impulsionado pela necessidade de informar, pesquisar e analisar quantidades cada vez maiores de dados com diferentes níveis de complexidade. (STACKOWIAK, 2007).

Os passos necessários para uma bem-sucedida mineração de dados têm seu início na fonte de dados, que passa por uma limpeza para a remoção de ruídos. Desse processo nascem os repositórios organizados, como o *data warehouse*. Desse repositório, podem-se selecionar algumas colunas para passar pelo processo de mineração, de forma interativa, até que apareçam valiosos padrões, conforme a **Figura 2**. (NAVEGA, 2002)

Figura 2 – Passos necessários para o data mining.



Fonte: Adaptado de NAVEGA (2002).

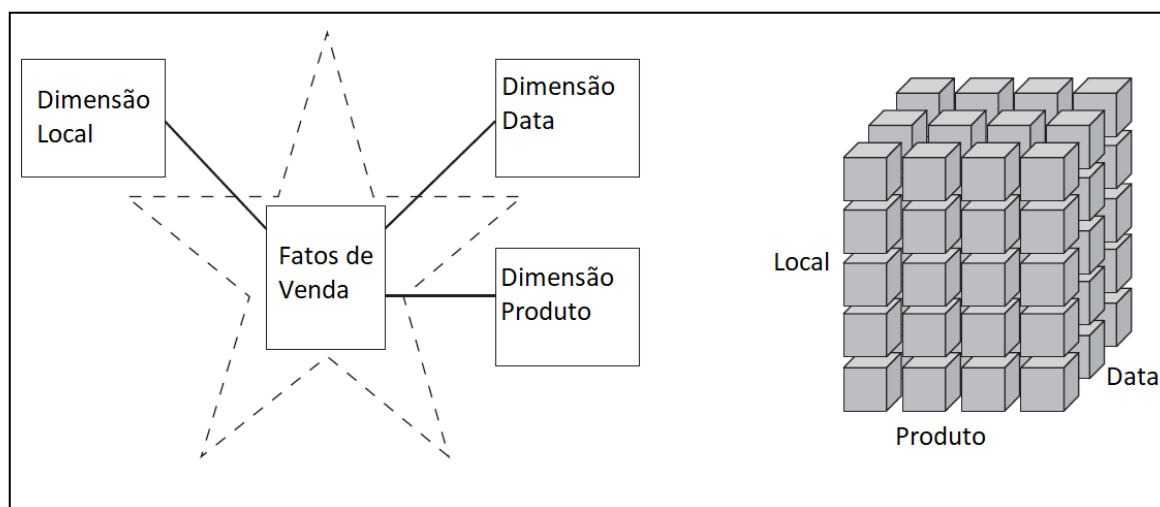
Por fim, conforme SÁ *et al* (2012), uma tendência para a Inteligência de Negócios é o uso de dispositivos móveis (smartphones, tablets, etc), onde os desenvolvedores de sistemas de BI devem estar atentos, uma vez que o cenário ideal será o desenvolvimento de aplicações centralizadas que consigam receber pedidos e introduzir informação para usuários móveis.

2.3 ESQUEMAS ESTRELA E CUBOS OLAP

Um dos recursos mais importantes de qualquer organização é a sua informação. Este recurso é quase sempre usado para duas finalidades: manutenção de registros operacionais e tomada de decisão analítica. Simplesmente falando, os sistemas operacionais são onde você coloca os dados e o sistema DW / BI é onde você tira os dados. (KIMBALL, R. e ROSS, M. (2013))

Os modelos dimensionais implementados em sistemas de gerenciamento de banco de dados relacionais são referidos como esquemas em estrela devido à sua semelhança com uma estrutura semelhante a uma estrela. Modelos dimensionais implementados em ambientes de banco de dados multidimensionais são referidos como cubos de processamento analítico online (OLAP), conforme ilustrado na **Figura 3**.

Figura 3 – Esquema estrela e cubo OLAP.



Fonte: Adaptado de KIMBALL, R. e ROSS, M. (2013).

2.4 BIG DATA

O *Big Data* apresenta três características básicas definidoras: volume (taxa de crescimento e taxa de dados), variedade (tipos de dados, como logs de sensores, microblogs - como Twitter e Facebook) e velocidade (a velocidade da fonte de fluxo de dados). Esses três termos são muito utilizados quando se trata de discussões Big Data e muitas vezes eles são chamados de "os 3 Vs" (ZIKOPOULOS e EATON, 2011). Porém além dos "3 Vs" há ainda outros "Vs" a se considerar quando se trata de Big Data, como veracidade e valor, totalizando assim "5 Vs".

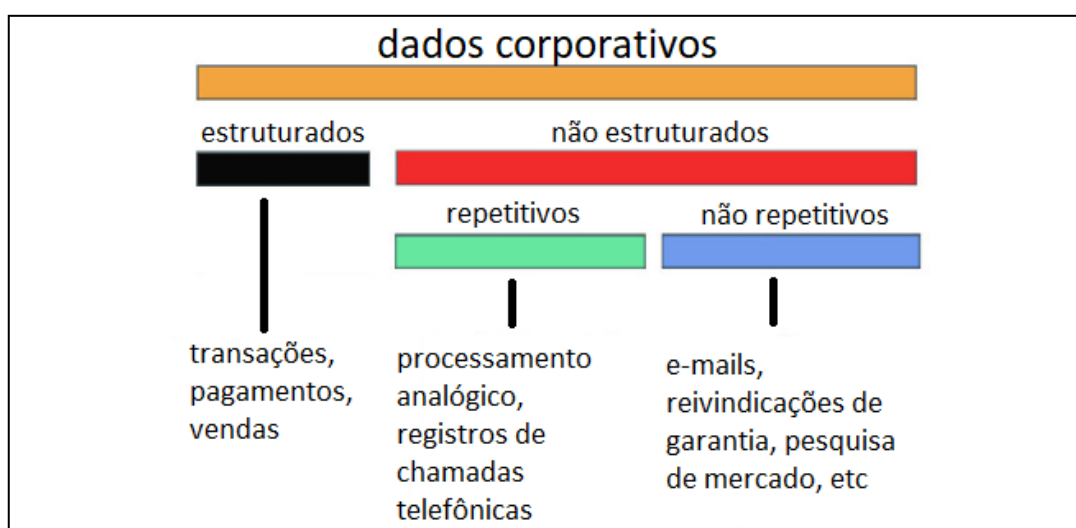
O modo como sua organização define sua estratégia e abordagem de dados – incluindo sua escolha de tecnologias de Big Data e nuvem – fará uma diferença essencial em sua capacidade de competir no futuro. (ORACLE, 2017)

Para Inmon e Linstedt (2014), considerando a totalidade dos dados encontrados na corporação, uma das maneiras de dividi-los é estruturados e não estruturados, conforme observado no **Gráfico 3**.

Os dados estruturados são bem definidos e possuem um formato previsível e periódico. Eles são gerenciados por um sistema de gerenciamento de banco de dados (SGBD) e consiste em registros, atributos, chaves e índices. Como regra, a maioria das unidades de dados no ambiente estruturado pode ser localizada de forma muito rápida e fácil.

Já os dados não estruturados, inversamente, são dados imprevisíveis e sem estrutura reconhecível. Em regra geral, os dados não estruturados são bastante desajeitados para o acesso, onde as sequências de dados longas devem ser pesquisadas sequencialmente para encontrar uma determinada unidade de dados. Existem muitas formas e variações de dados não estruturados. Talvez a forma mais comum de dados não estruturados seja o texto. (INMON W. H e LINSTEDT, Dan (2014))

Gráfico 3 – Dados estruturados e não estruturados, repetitivos e não repetitivos.

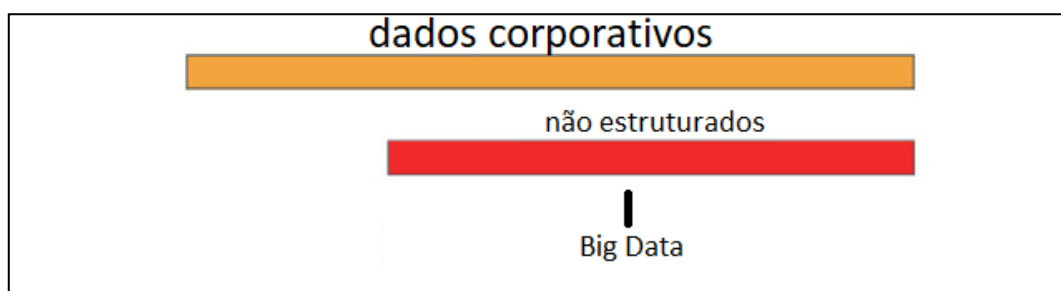


Fonte: Adaptado de INMON W. H. e LINSTEDT, Dan (2014).

Dados não estruturados repetitivos são dados que ocorrem muitas vezes, inclusive na mesma estrutura e na mesma forma de realização. A estrutura dos dados repetitivos parece exatamente igual ou substancialmente a mesma que o registro anterior. Não existe uma infraestrutura maciça e elaborada que gerencie o conteúdo de dados não estruturados repetitivos. Os dados não estruturados não-repetitivos são dados em que os registros são substancialmente diferentes um do outro.

É interessante notar que o *Big Data* consiste em dados não-estruturados repetitivos e não estruturados não-repetitivos na corporação, conforme observado na **Gráfico 4**. (INMON W. H e LINSTEDT, Dan (2014))

Gráfico 4 – Big Data consiste em dados não estruturados.



Fonte: Adaptado de INMON W. H e LINSTEDT, Dan (2014).

Os grandes dados são uma consequência do ambiente digital de hoje, que gera dados que fluem continuamente de todas as direções a uma velocidade e volume sem precedentes, e esses dados geralmente requerem limpeza. Eles são considerados "sujos" principalmente por causa da fusão de múltiplas fontes. (RATNER, 2011)

Não é fácil medir o volume total de dados armazenados eletronicamente, mas uma estimativa coloca o tamanho do "universo digital" em 4.4 zettabytes em 2013 e prevê um crescimento de dez vezes em 2020 para 44 zettabytes. Um zettabyte é equivalente a aproximadamente a um bilhão de terabytes. (WHITE, 2015)

Existe uma clara falta de profissionais com o conhecimento "profundo" necessário para gerenciar os V's de *Big Data*. Há também uma crescente demanda por indivíduos com o profundo conhecimento necessário para gerenciar as três "perspectivas" da tomada de decisões empresariais: análises descritivas, preditivas e prescritivas. (CHEN e CHIANG, 2012)

Segundo TOBACOWALA e GUPTA (2016), as empresas investiram milhões de dólares em grandes dados e análises, mas relatórios recentes sugerem que a maioria ainda não viu uma recompensa nesses investimentos. Isso confirma a necessidade de se observar outras maneiras de extrair dados como por exemplo, utilizando-se ferramentas menos custosas.

2.5 FERRAMENTAS GRATUITAS PARA INTELIGÊNCIA DE NEGÓCIOS

Conforme FELLER, Joseph *et al* (2005) os termos "software livre" e "software de código aberto" referem-se a produtos de software distribuídos em termos que permitem aos usuários:

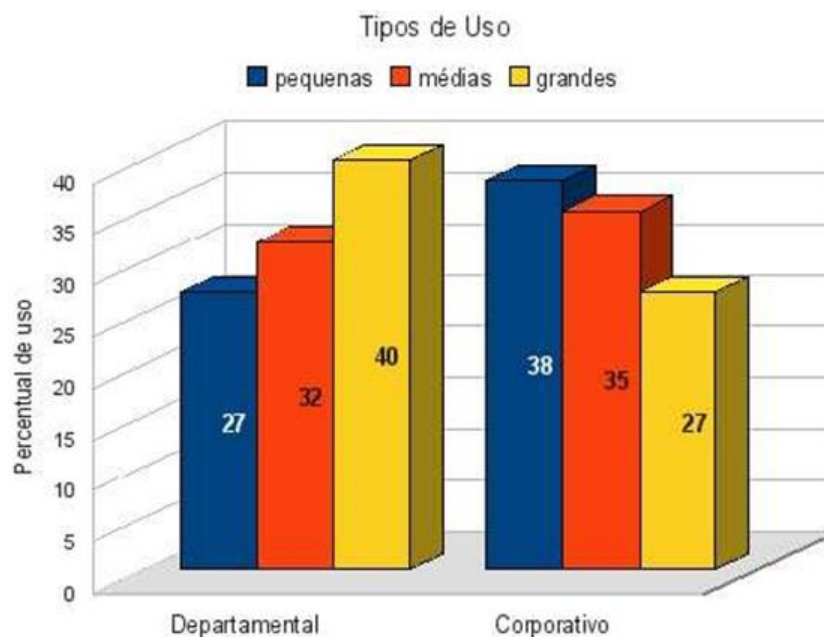
- Usar o software.
- Modificar o software.
- Redistribuir o software.

Qualquer que seja a forma de uso, não se exige que se pague o autor do software uma taxa, ou taxa para se realizar alguma das atividades listadas. Em geral, esses termos de distribuição também protegem o que o mundo da publicação chama o "direito moral" do autor do software a ser identificado como tal. Produtos como o sistema operacional GNU / Linux, o servidor da Web Apache, o navegador da Web Mozilla, a linguagem de programação PHP e a ferramenta OpenOffice são exemplos bem conhecidos desse tipo de software.

Em 2004, Jaspersoft e Pentaho foram fundados para oferecer conjuntos de BI completos que alavancaram as soluções de BI autônomas existentes. Desde então, essas empresas têm visto um forte crescimento e adoção de mercado, tornando as soluções de BI de código aberto uma alternativa séria para os fornecedores proprietários estabelecidos. (BOUMAN e DONGEN, 2009)

Foi feito um estudo por MADSEN (2009), que mostra uma proporção importante de soluções de BI departamentais nas empresas de grande porte. Empresas pequenas usam BI em software livre corporativamente em maior percentual que as grandes. Este estudo foi baseado em entrevistas com consultores e profissionais de TI entre julho e agosto de 2009, com uma amostra de 1.000 pessoas, embora nem todos os respondentes respondessem a todas as perguntas.

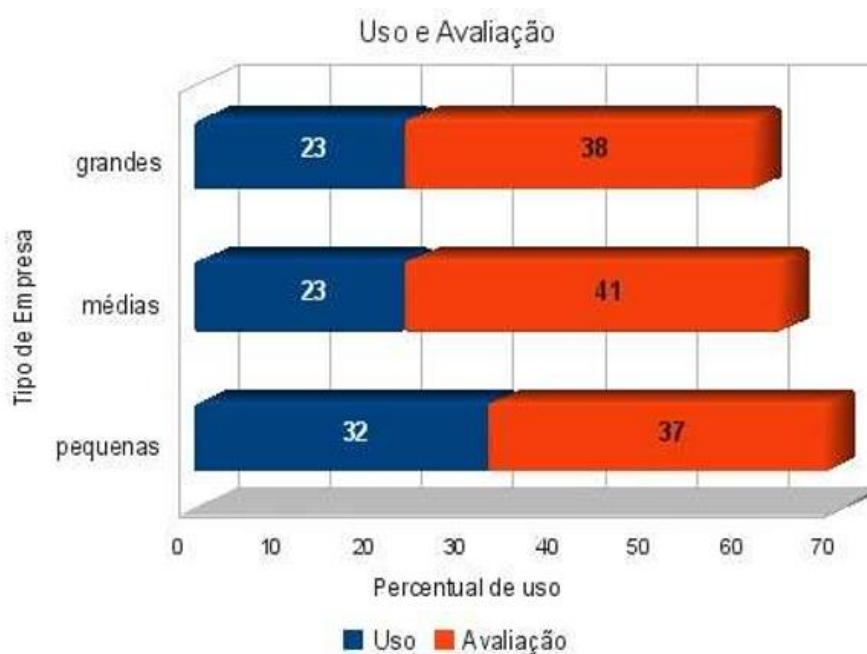
Gráfico 5 – Aplicação de BI em Software Livre em percentual Departamental e Corporativo.



Fonte: Adaptado de MADSEN (2009).

Empresas pequenas usam mais soluções BI baseadas em software livre do que as grandes, mas ao mesmo tempo vemos que as grandes têm um percentual maior de empresas que estão avaliando o uso, conforme o **Gráfico 6**.

Gráfico 6 – BI em Software livre nas empresas por tamanho.



Fonte: Adaptado de MADSEN (2009).

3 PROCEDIMENTOS METODOLÓGICOS

Este trabalho é um estudo de caso com dados secundários onde se selecionou três ferramentas de Inteligência de Negócios para análise de suas funcionalidades, características e qualidades. São elas: Pentaho, SpagoBI e Jaspersoft, que serão vistas em detalhes nos tópicos a seguir.

As ferramentas escolhidas são todas de código aberto, software livre e apresentam ao menos uma versão gratuita. Elas podem ser utilizadas por empresas, organizações e instituições de maneira geral para extração, organização e exploração de seus dados e metadados afim de explorar informações até então desconhecidas ou ignoradas pelas mesmas.

Quanto as ferramentas de inteligência de negócios licenciadas em software livre, foi feita uma análise de mercado em 2008 pelo Gartner Group que diz que somente o SpagoBI, Jaspersoft, Pentaho e Birt podem atender as necessidades de BI de forma corporativa. (LACY, Miguel Koren O'Brien de (2010))

3.1 FONTES DE PESQUISA

Para o estudo, foi realizada uma pesquisa bibliográfica, tendo como base autores que exploram o universo do Marketing, Inteligência de Negócios, *Big Data* e *Data Warehouse*.

Também se utilizou como fonte de pesquisa os sites das ferramentas citadas, para que então, com base nas informações fornecidas pelos fabricantes e desenvolvedores das ferramentas fosse possível fazer uma demonstração das características principais de cada uma, conforme a **Tabela 2**.

Quadro 2 – Fontes oficiais de pesquisa das ferramentas

	Jaspersoft	Pentaho	SpagoBI
Web site	https://www.jaspersoft.com	http://www.pentaho.com	http://www.spagobi.org
Período	Julho a Agosto de 2017	Julho a Setembro de 2017	Agosto de 2017
Conteúdo explorado no site	<i>Documentation</i>	<i>Product Overview</i>	<i>Product;</i> <i>BI Components;</i> <i>Architecture</i>

Fonte: A autoria própria.

3.2 CRITÉRIOS DE ESCOLHA DAS FERRAMENTAS

Com o objetivo de comparar as ferramentas de BI, foi feito um estudo detalhado destas ferramentas de BI *open source* citadas a fim de expor as principais características, funcionalidades, pontos fortes e fracos de cada uma.

Foram selecionadas as ferramentas utilizando os seguintes critérios de escolha:

1. Ferramentas bem-conceituadas pelo Gartner, e que fossem *open source*.
2. Ferramentas com casos de utilização no mercado.
3. Disponibilização de documentação de pesquisa em sites ou periódicos.
4. Possibilidade de efetuar testes práticos caso necessário.
5. O autor já possuir contato com a ferramenta, que é o caso da ferramenta Pentaho, e pretende aprofundar os seus conhecimentos na mesma.

A partir da demonstração das características de cada ferramenta pesquisada, fosse possível fazer um comparativo para auxiliar na escolha da melhor alternativa para uso, de acordo com o perfil da organização que irá utilizar a ferramenta.

3.3 CRITÉRIOS DE COMPARAÇÃO ENTRE AS FERRAMENTAS

Para a análise comparativa entre as ferramentas foram selecionados os seguintes critérios:

1. Diferenças entre as versões comerciais e gratuitas (caso haja alguma diferença);
2. Componentes da cada ferramenta.
3. Licenças de utilização das versões gratuitas.
4. Idiomas de documentação.
5. Desempenho: Avalia se a ferramenta tem uma boa performance ao processar consultas com um alto volume de dados.
6. Consultas *ad hoc* OLAP: Avalia se a ferramenta permite ao usuário ter a liberdade de definir consultas que acredita ser melhor em um dado contexto.

7. Arquitetura: Avalia se a solução implementa arquiteturas OLAP que possuem alta escalabilidade, como por exemplo, ROLAP ou HOLAP.
8. Plataforma: Avalia se a ferramenta pode ser executada nos sistemas operacionais mais difundidos, como Windows, Linux e UNIX.
9. Suporte Técnico/Documentação e fóruns: Avalia o nível de qualidade da documentação e suporte técnico oferecido pela ferramenta.
10. Agendamento: Avalia se a ferramenta permite o agendamento de relatórios.
11. *Dashboards*: Avalia se a ferramenta possibilita a criação de painéis.
12. Exportação para outros formatos: Avalia se a ferramenta dispõe de recursos de exportação para formatos como PDF, HTML e ODT/ODS (para permitir futuramente a integração com ferramentas livres).
13. Suporte a dispositivos móveis: Avalia se a ferramenta suporta o uso de dispositivos móveis como Android e IOs.
14. Análise preditiva e simulação "*What If*": Avalia se a ferramenta possui o recurso de simulação de cenários hipotéticos, exibindo-os através de gráficos e dados. Também se consegue prever tendências futuras e comportamento.
15. Custo: Avalia o custo para a compra da licença da ferramenta, caso exista uma versão paga.
16. Carregamento de dados de diferentes fontes (Big Data): Avalia a possibilidade de integração da solução com fontes de dados heterogêneas.

Ao fim da análise comparativa e da avaliação destes critérios, foram feitas considerações sobre cada ferramenta afim de auxiliar a melhor escolha de acordo com o perfil do utilizador.

4 ANÁLISE COMPARATIVA DAS FERRAMENTAS

Neste capítulo são apresentadas as ferramentas pesquisadas neste trabalho, começando em ordem alfabética pela ferramenta JasperSoft, em seguida, é apresentada a ferramenta Pentaho e enfim a ferramenta SpagoBI. Após isso, é realizada uma análise comparativa entre elas feito um comparativo entre as ferramentas estudadas.

4.1 JASPERSOFT

A ferramenta Jaspersoft está dividida em duas versões, a Jaspersoft *Community Edition* e a *Commercial Edition*, comercializada pela empresa TIBCO Software Inc., que adquiriu a Jaspersoft em 2014. As principais diferenças entre elas serão demonstradas no **Quadro 3** abaixo.

Quadro 3 – Principais diferenças entre as versões Jaspersoft Community Edition e Commercial Edition.

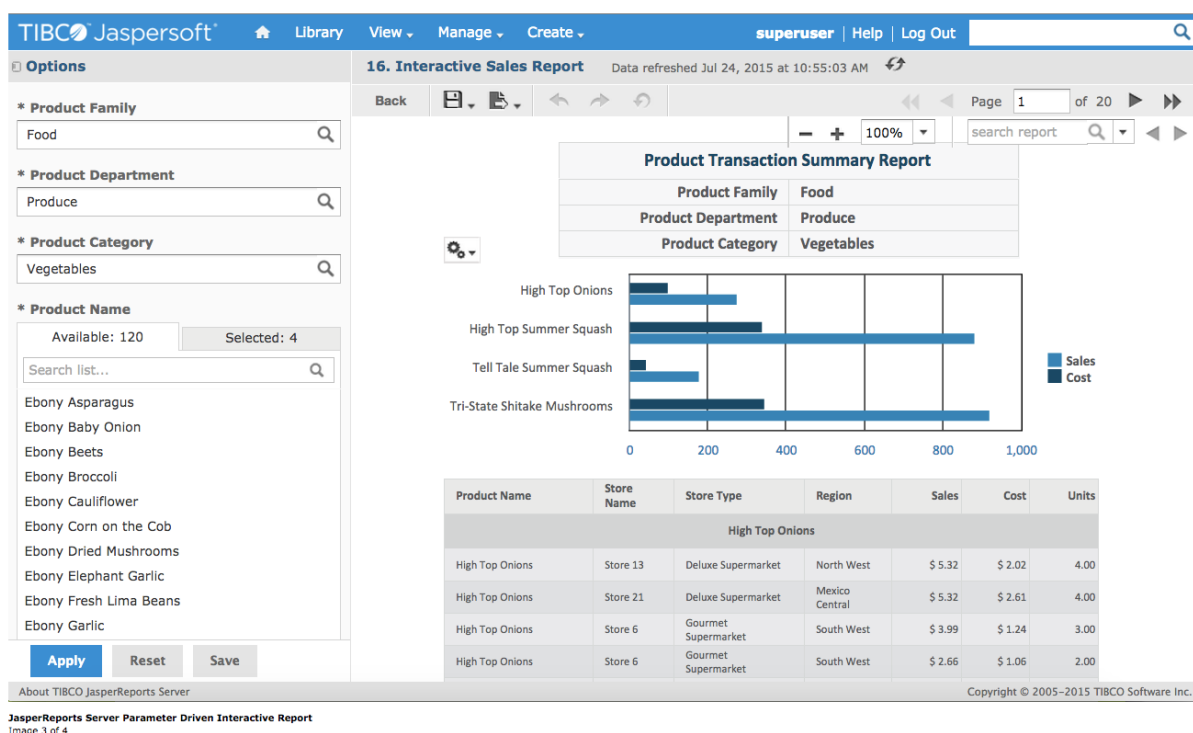
Versão	<i>Community Edition</i>	<i>Commercial Edition</i>
Acesso livre. Binários de software disponíveis publicamente e código fonte sem nenhum custo.	Sim	Não
Direitos de distribuição. Capacidade para OEM, incluir ou incorporar o software Jaspersoft em um produto, aplicativo ou serviço publicamente distribuído.	Sim, em termos de licenças de código aberto	Sim, nos termos do acordo.
Suporte de Plataforma Certificada. Sistema operacional proprietário, servidor de aplicativos ou banco de dados.	Não	Sim
Ciclos de liberação gerenciada. Aperfeiçoamentos de produtos em menos lançamentos.	Não	Sim
Garantias de suporte. Suporte garantido para aplicativos de produção de missão crítica e/ou lançamentos mais antigos.	Não	Sim
Assuntos legais. Garantias e indenização do produto.	Não	Sim

Funcionalidade avançada. Recursos de valor agregado, como relatórios baseados na Web, painéis interativos e APIs avançadas.	Não	Sim
---	-----	-----

Fonte: Adaptado de JASPERSOFT (2017).

O principal produto Jaspersoft é o JasperReports Server (**Figura 4**), uma aplicação web Java EE que fornece recursos avançados do servidor de relatórios, como agendamento de relatórios e permissões. Ele está disponível sob uma licença de código aberto para uso em conjunto com infraestrutura de código aberto, como o MySQL e o JBoss, ou uma licença comercial para implantações corporativas envolvendo bancos de dados comerciais e servidores de aplicativos.

Figura 4 – Interface do JasperReports Server.



Fonte: JASPER COMMUNITY (2017).

JasperReports é uma ferramenta de geração de relatórios Java de código aberto que pode escrever em vários destinos, como: tela, impressora, em PDF, HTML, Microsoft Excel, RTF, ODT, valores separados por vírgulas ou arquivos XML. Ele pode ser usado em aplicativos habilitados para Java, incluindo Java EE ou

aplicativos da Web, para gerar conteúdo dinâmico. Ele lê suas instruções a partir de um arquivo XML ou .jasper. (JASPERSOFT, 2017)

A JasperSoft fornece uma variedade de versões para seu software de BI, o JasperSoft BI Suite. A versão Community possui um ambiente de relatórios e gráficos com uma infraestrutura completa que fornece todo suporte necessário.

O gerador de gráficos suporta imagens, tabelas de referência cruzada e sub-relatórios para layouts de relatórios. O visualizador de relatório interativo é um browser baseado em um visualizador de relatório que fornece filtragem, ordenação e formatação.

Além disso, o JasperSoft BI possui um repositório centralizado que fornece uma infraestrutura completa para relatórios e armazenamentos de perfis de usuários, dashboards e visualização analítica. Outra funcionalidade interessante é o suporte para agendamento automático de relatórios e a versão Mobile BI que é suportada por aparelhos iPhone e Android.

A versão comercial da ferramenta adiciona mais *dashboards*, camada de metadados, análise em memória, capacidade de integração de informações e visualizações interativas.

4.1.1 Componentes do Jaspersoft

Entre os componentes de BI da ferramenta Jaspersoft, podem-se destacar:

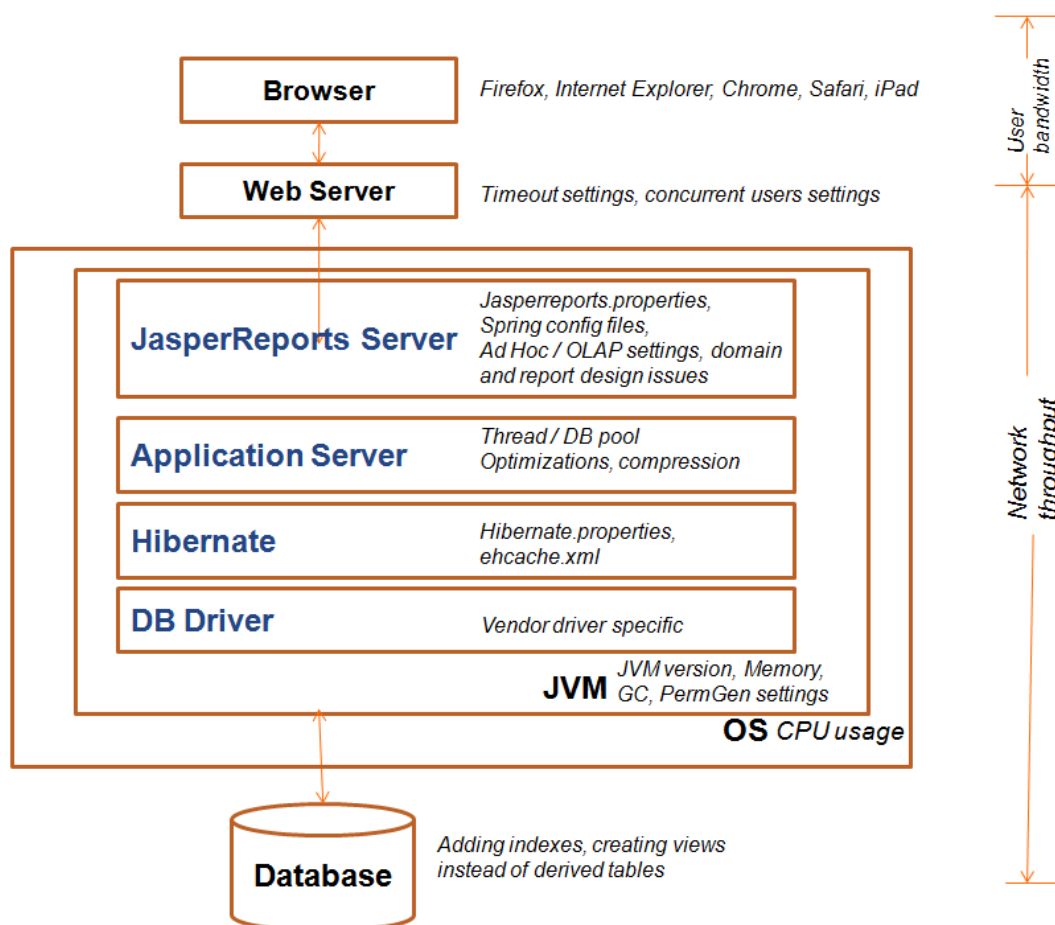
- **Jaspersoft BI Enterprise Edition:** É um servidor de BI completo que permite criar, implementar e gerenciar relatórios, painéis e visualizações de dados, sendo voltado para desenvolvedores, administradores de BI, usuários empresariais e analistas de dados.
- **Jaspersoft Studio Professional Edition:** É um designer de desktop usado para criar relatórios simples ou muito complexos e visualizações de dados. Tem como público-alvo os desenvolvedores de relatórios, desenvolvedores de aplicativos e especialistas em BI.
- **Jaspersoft ETL:** É uma ferramenta de integração de dados para integrar, limpar e criar perfil de dados. Os utilizadores são engenheiros de dados e especialistas em integração de dados.

4.1.2 Referência de desempenho do servidor Jasper Reports

A implementação de serviços de BI da Jaspersoft pode variar de um relatório simples de serviço de processador único em um banco de dados local de aplicativos, para um conjunto de Servidores JasperReports que se conectam a dezenas de várias fontes de dados que funcionam em tantos outros servidores.

Os problemas de desempenho podem variar de problemas com um Servidor JasperReports específico para problemas de transferência de rede para as muitas fontes de dados que o servidor pode querer consultar. Abaixo, na **Figura 5** está um diagrama que atende o ambiente no qual JasperReports Server é executado. Além do navegador do usuário, muitos ou todos esses componentes podem residir na mesma instância do servidor (física ou virtual).

Figura 5 – Ambiente no qual o JasperReports é executado.



Fonte: Adaptado de JASPER COMMUNITY (2017).

4.2 PENTAHO

A Pentaho é uma das principais soluções de *Business Intelligence* baseada em software livre do mundo e está no mercado há mais de 10 anos, sendo reconhecida mundialmente como ferramenta de BI e integração de dados. Muitas empresas e órgãos públicos vem adotando esta tecnologia por conta da sua franca evolução e custo. (OPENIN, 2017)

O Pentaho BI Suite é um software de código aberto onde o usuário é livre para usar e distribuir seus programas e até mesmo modificar seu código fonte gratuitamente. Pentaho está programado na linguagem de programação Java. (BOUMAN e DONGEN, 2009)

O Pentaho inclui um espectro de ferramentas para analisar, visualizar, explorar, informar e prever facilmente. Aberto, incorporável e extensível, o Pentaho é projetado para garantir que cada membro da equipe, desde desenvolvedores a usuários empresariais, possa converter dados em valor. (PENTAHO, 2017)

Ele é uma suíte de software ou conjunto de aplicativos para Inteligência de Negócios. Esta coleção de aplicativos possui funcionalidades relacionadas, compartilhando uma interface de usuário semelhante e a capacidade de trocar dados facilmente entre si.

Entre as **vantagens** deste modelo do tipo 'suíte de software' estão:

- Menos custo financeiro do que comprar pacotes individuais
- GUI idêntica ou muito parecida.
- Projetado para interagir um com o outro
- Ajuda a curva de aprendizado do usuário

Já as **desvantagens** deste modelo, pode-se dizer que são:

- Alguns recursos comprados nunca podem ser usados pelo usuário
- Pode levar uma quantidade significativa de espaço em disco, o que pode ser chamado de '*bloatware*', que é o nome dado a um conjunto de softwares indesejados.
- Provavelmente exigirá um esforço significativo para o desenvolvedor de software atualizar e manter toda a coleção de aplicativos ao longo do tempo.

A suíte Pentaho oferece suporte a relatórios, ferramentas de *data mining*, e atinge as necessidades da maioria das organizações pois se observa na suíte chamada 'Pentaho BI Suite Community Edition (CE)' as funcionalidades de ETL, OLAP, metadata, *data mining*, relatórios e *dashboards*. Uma variedade de soluções *open source* são reunidas para oferecer diversas funcionalidades, entre elas: Weka para *data mining*, Kettle para integração de dados, Mondrian para OLAP, e diversas outras para relatórios de BI, *dashboards*, OLAP, e até mesmo *big data*.

Essa plataforma também oferece um ambiente para construção de soluções de BI e inclui autenticação, uma *engine* para regras e *web services*. O servidor do Pentaho BI fornece gerenciamento de relatórios baseados na web, integração de aplicações e *workflow*. O Pentaho Report Designer, Reporting Engine e Reporting SDK suportam a criação de relatórios relacionais e analíticos com muitos formatos de saída e diferentes fontes de dados. O Pentaho também oferece versões com suporte, treinamentos e consultoria, assim como mais funcionalidades e customizações.

4.2.1 Componentes do Pentaho

A solução Pentaho atualmente consiste dos seguintes componentes:

- **Pentaho Data Integration:** também chamado de Kettle, é uma ferramenta para extração, transformação e carga de dados, o chamado ETL.
- **Pentaho Analysis Services:** também conhecido como Mondrian OLAP server, é uma ferramenta para análise de dados, também conhecido como processo OLAP.
- **Pentaho Reporting:** derivado do projeto JFreeReport.
- **Pentaho Data Mining:** derivado do projeto Weka, um conjunto de ferramentas relacionadas com a mineração de dados.
- **Pentaho Dashboard:** para painéis dinâmicos utilizados para análise de dados.
- **Pentaho for Apache Hadoop:** é uma plataforma de computação distribuída voltada para clusters e processamento de grandes massas de dados.

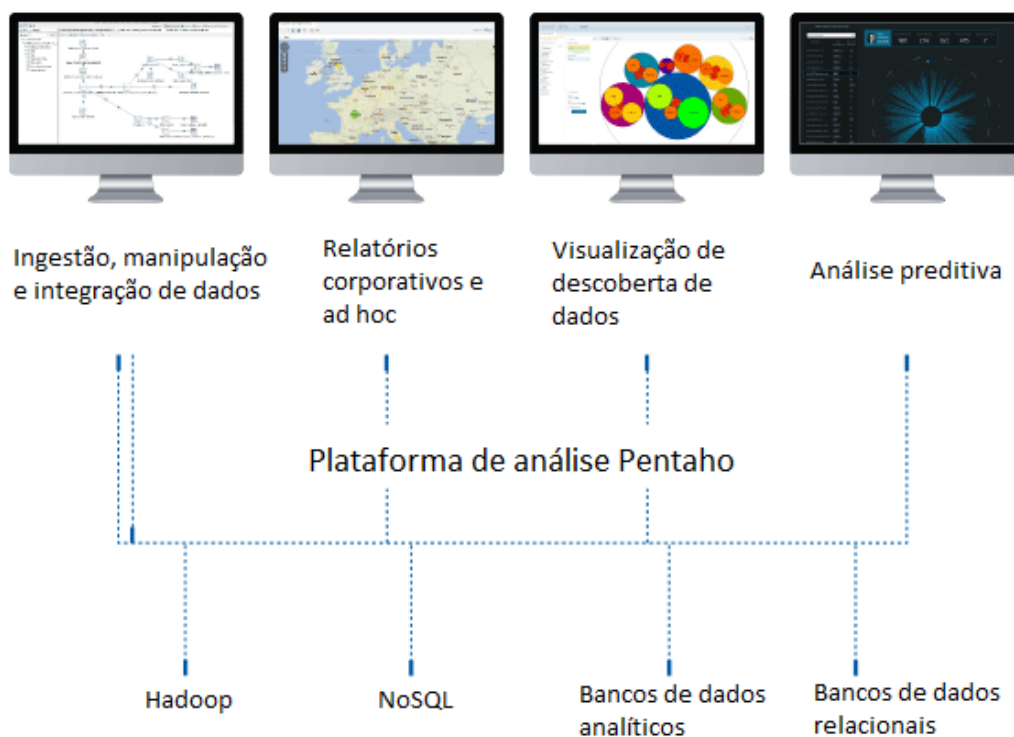
4.2.2 Big Data com Pentaho

O Pentaho fornece ferramentas de dados importantes para extrair, preparar e misturar dados, além das visualizações e análises. Com integração Hadoop e Spark para NoSQL, o Pentaho permite transformar grandes massas de dados em informações úteis para gerar conhecimento.

No Pentaho há uma plataforma de integração de dados e de análise de negócios (conforme **Figura 4**) fortemente acoplada que acelera a realização de valor a partir de dados. Pode-se destacar os recursos de acesso e integração de dados à visualização de dados e análises preditivas. Permissão aos usuários arquitetar grandes misturas de dados na fonte e transmiti-los diretamente para análises mais completas e precisas. Além disso pode-se alternar ou combinar mecanismos de processamento de dados com a execução no cluster para maximizar a capacidade de processamento existente.

Ainda é capaz de detectar dados em tempo real com acesso imediato a análises, incluindo gráficos, visualizações e relatórios, a partir de qualquer etapa na preparação de dados.

Figura 6 – Plataforma analítica do Pentaho.



Fonte: Adaptado de PENTAHO (2017).

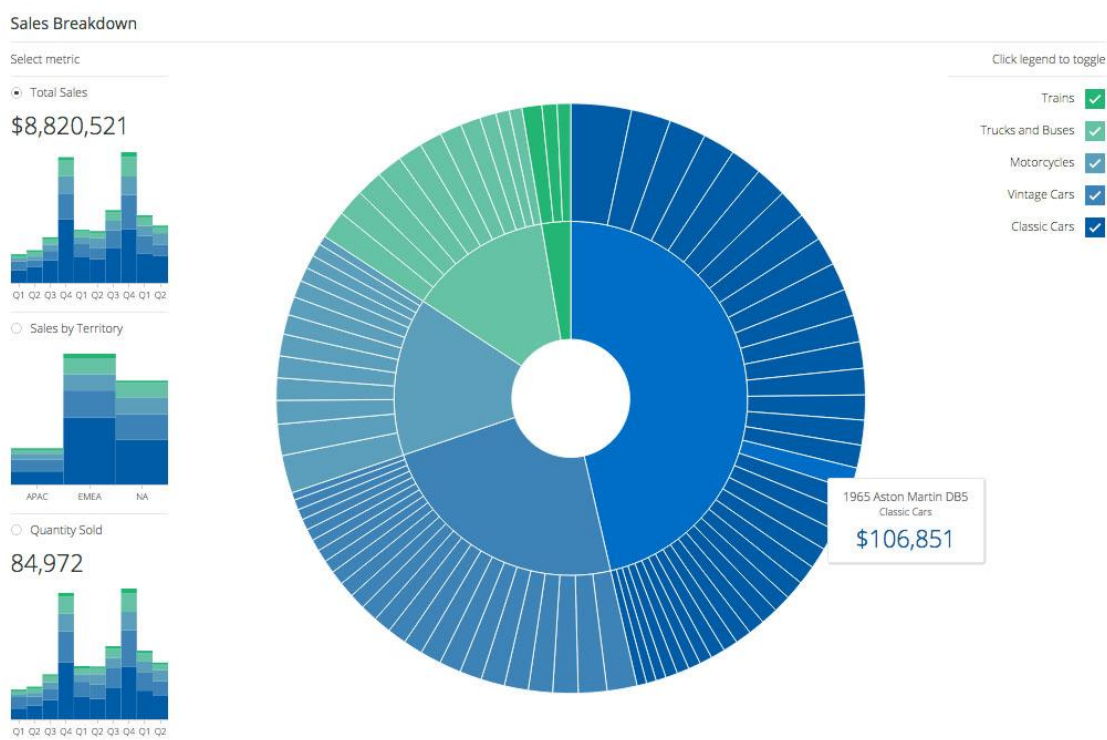
A ferramenta Pentaho possui conexões nativas e uma camada de Big Data adaptativa às últimas versões de grandes fornecedores de fontes de dados, como:

- Suporte para as distribuições Hadoop da Cloudera, Hortonworks, MapR e Amazon Web Services.
- Capacidade de acessar dados para preparação via SQL no Spark e orquestrar aplicativos existentes Spark em Scala, Java e Python.
- Integração com NoSQL incluindo MongoDB e Cassandra.
- Conectividade a bancos de dados analíticos, incluindo HPE Vertica, Amazon Redshift e SAP HANA.

A Pentaho permite visualizar, analisar e divulgar dados em várias dimensões em formatos de análise interativa, filtração, zoom, grades de calor, mapas geográficos, *sunbursts* (destaque de atributos para uma maior visão). É possível fazer armazenamento de dados em memória em escala extrema para análise em alta velocidade de grandes volumes de dados e relatórios interativos de autoatendimento para relatórios empresariais de alto volume.

Além disso gera *dashboards* de fonte de dados do tipo *big data*, conforme a **Figura 5**.

Figura 7 – Dashboard Interativo Pentaho.



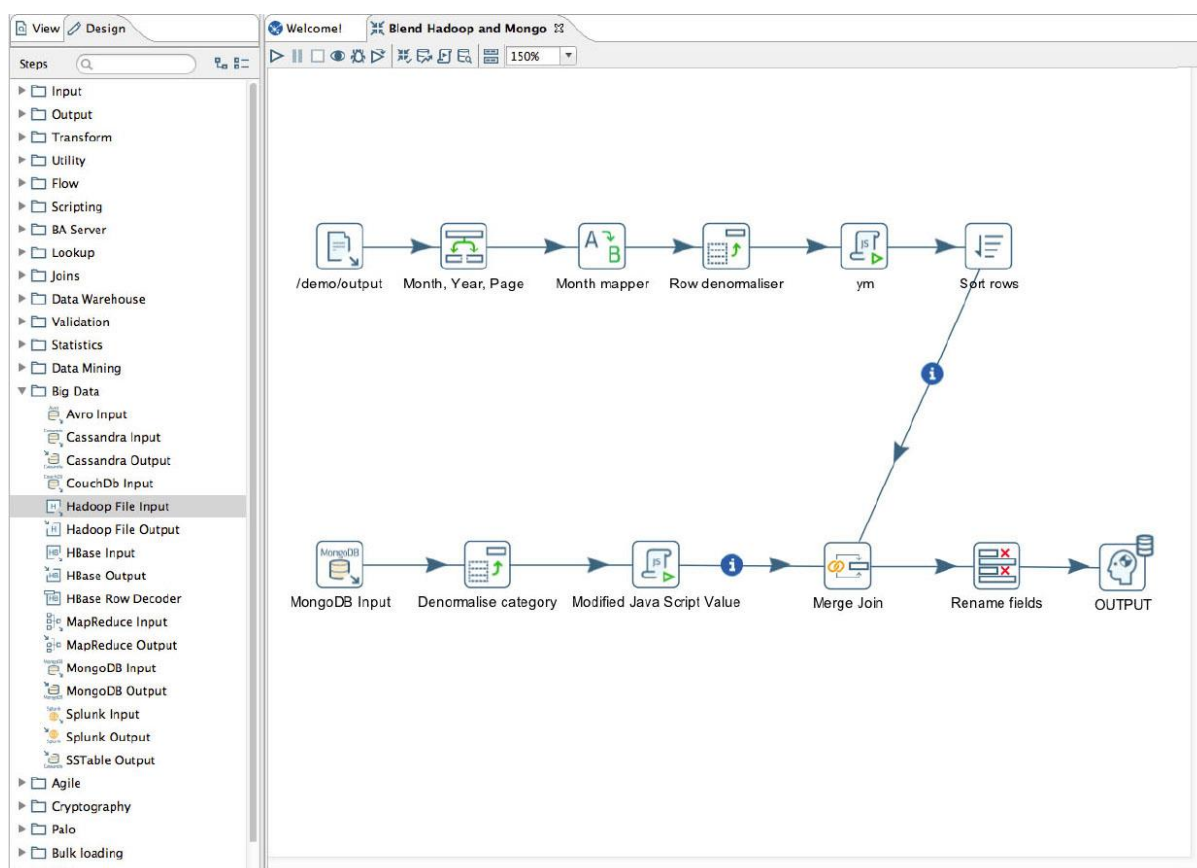
Fonte: PENTAHO (2017).

4.2.3 Integração de dados com Pentaho

O *Pentaho Data Integration* permite aos usuários ingerir, misturar, limpar e preparar diversos dados de diversas fontes com ferramentas visuais para eliminar a codificação e a complexidade.

Ele oferece integração de dados com funções de arrastar e soltar, acoplada a conectividade de dados que abrange arquivos planos e RDBMS para Hadoop e outras fontes, incluindo designer gráfico de extração-transformação (ETL) para simplificar a criação de pipelines de dados, conforme a **Figura 6**.

Figura 8 – Ferramenta de Integração de Dados do Pentaho.



Fonte: PENTAHO (2017).

4.2.4 Performance com Pentaho

O Pentaho Data Integration oferece recursos avançados de *clustering* e particionamento que permitem que as organizações dimensionem suas implantações de integração de dados.

Os clusters de integração de dados Pentaho são criados para aumentar o desempenho e o processamento das transformações de dados; em particular, eles são construídos para realizar clássicos "dividir e conquistar" o processamento de conjuntos de dados em paralelo. (PENTAHO PERFORMANCE, 2017).

4.3 SPAGOBI

A SpagoBI é considerada uma das melhores ferramentas de Inteligência de Negócios *open source* disponíveis atualmente pois possui uma grande coleção de softwares para BI, criando-se assim uma ampla capacidade para suportar análise de dados. Além das características essenciais que uma ferramenta de BI possui, a SpagoBI inclui outros domínios como *data mining* e BPM fazendo com que grandes empresas como FIAT e Gamebay já aderissem à SpagoBI como parte da sua solução de BI.

A SpagoBI é também a única solução de BI *open source* que a versão livre é a única versão disponível. Portanto, esta ferramenta destaca-se por ser completa e totalmente livre. A quantidade de ferramentas oferecidas é bastante diversificada e cada área é servida por um grande número de funcionalidades, entre elas: Relatórios, análises multidimensionais, gráficos, KPI, Relatório *ad-hoc*, *Location Intelligence*, *Data mining*, *Dashboard* e Console, colaboração, automação, ETL e Mobile.

É possível realizar treinamentos e obter suporte através de pagamento diretamente no site oficial da ferramenta, onde também é possível obter o software para download.

4.3.1 Diferenciais do SpagoBI

Um dos diferenciais do SpagoBI em relação a outras soluções de BI licenciadas em software livre é o apoio da empresa Engineering Ingegneria Informatica que apoia o desenvolvimento da ferramenta, porém tem o foco em consultoria. Ele é desenvolvido pela Engineering Group, uma das 10 principais empresas europeias de TI, com escritórios localizados na Itália, Bélgica, Líbano, Brasil e Argentina. (SPAGOWORLD, 2017)

Além disso, o SpagoBI é um sistema de BI completo em software livre que pode ser implementado e usado corporativamente sem pagamento de licenças. O termo 'uso corporativo' se refere a funções de escalabilidade, arquitetura, funções e segurança necessários em uma plataforma desse tipo. Geralmente os outros sistemas mantêm uma versão chamada de "Community" com funcionalidades reduzidas e uma versão "Enterprise" com pagamento anual de suporte, o que não se aplica no SpagoBI.

4.3.2 Componentes do SpagoBI

O SpagoBI é o único conjunto inteiramente *open source* de Inteligência de Negócios e abrange todas as áreas analíticas de projetos de BI, com temas e tecnologias inovadoras (SPAGOBI, 2017). O SpagoBI oferece uma ampla gama de ferramentas analíticas, conforme destacado a seguir:

- **Relatórios:** Realiza relatórios estruturados e exporte-os usando o formato mais adequado (HTML, PDF, XLS, XML, TXT, CSV, RTF).
- **Análise multidimensional (OLAP):** Explora dados em diferentes níveis de detalhes e de diferentes perspectivas, através de processos detalhados, drill-across, slice-and-dice, drill-through.
- **Gráficos:** Desenvolve gráficos prontos para uso de acordo com gráficos únicos (histogramas, gráficos de torta, gráficos de barras, gráficos de área, diagramas de dispersão, gráficos de linhas, gráficos de bolhas, gráficos de dispersão) e interativos (controles deslizantes temporais, adicionar / excluir séries).
- **KPIs:** O SpagoBI oferece um conjunto completo de ferramentas para criar, gerenciar, visualizar e navegar modelos de hierarquia KPI, através de diferentes métodos, regras de cálculo, limiares e regras de alarme.
- **Cockpits interativos:** Agrega análises diferentes em uma única visualização, definindo caminhos de navegação e explorando dados de forma dinâmica e gráfica.
- **Relatórios *ad hoc*:** Cria relatórios de várias folhas, incluindo tabelas, tabelas cruzadas e gráficos.

- **Localização inteligente:** Visualiza os dados em mapas (ou seja, catálogos de mapas estáticos ou serviços de mapeamento / recurso web) e interagem dinamicamente para obter visualizações instantâneas.
- **Query:** Um mecanismo QbE (*Query by Example*) faz a exploração e navegação de dados particularmente intuitiva e fácil, devido a uma interface totalmente gráfica e baseada na web. É possível guardar as consultas para uso futuro.
- **Mineração de dados:** Análise de dados avançada que lhe permite extrair conhecimento de grandes volumes de dados, para melhorar estratégias de tomada de decisão e de negócios.
- **Análise de rede:** Visualiza e interpreta as relações entre entidades através de visualizações especializadas. Uma entidade pode ser animada (por exemplo, usuários de redes sociais) ou inanimada (por exemplo, países, empresas, projetos).
- **ETL:** O SpagoBI integra o produto de código aberto TOS (Talend Open Studio), para carregar dados no *data warehouse* e gerenciá-los conforme sua conveniência.
- **Colaboração:** É possível criar dossiês de relatórios estruturados, e enriquecer a análise com notas pessoais e comentários postados pelos usuários. Em seguida, pode-se compartilhar através de um fluxo de trabalho colaborativo.
- **Automação de escritório:** Pode-se publicar documentos pessoais no ambiente de BI, integrando ferramentas comuns do Office (Open Office ou MS Office).
- **Gerenciamento de Masterdata:** Os usuários podem escrever de volta para o banco de dados e modificar dados de tabela através de uma interface de usuário intuitiva, cujo comportamento pode ser configurado por meio de parâmetros de configuração simples, usando modelos pré-definidos.
- **Processos externos:** Gerencia processos analíticos, que podem ser executados em segundo plano ou agendados para iniciar e parar em uma hora agendada.

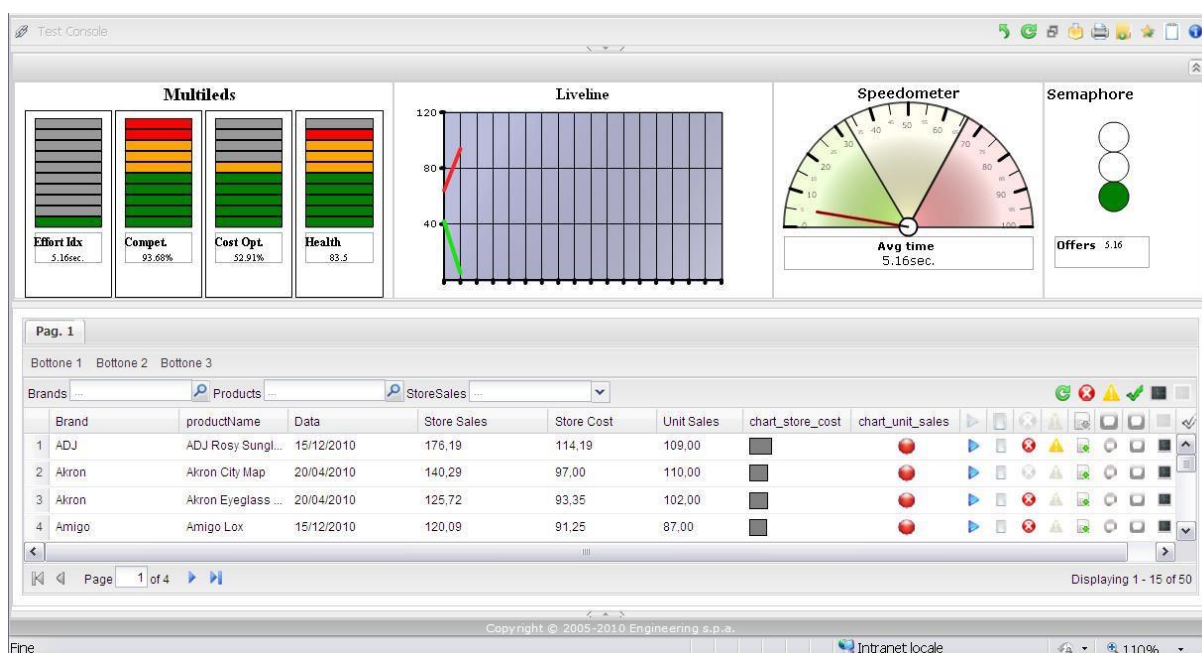
4.3.3 Arquitetura do SpagoBI

O SpagoBI fornece ferramentas para diversos tipos de usuários. Dos usuários empresariais e tomadores de decisão, para desenvolvedores e administradores. Ele é composto pelos seguintes módulos principais:

- Servidor SpagoBI, o núcleo da suíte, incluindo as ferramentas analíticas e os recursos
- SpagoBI Studio, o ambiente de desenvolvimento integrado
- SpagoBI Meta, o ambiente de metadados
- SpagoBI SDK, a camada de integração que permite usar o SpagoBI com ferramentas externas
- Aplicações SpagoBI, uma coleção de modelos analíticos verticais que são desenvolvidos usando o SpagoBI.

A lista completa de motores analíticos está disponível no Wiki do projeto em seu site oficial, assim como a lista completa de ambientes certificados. A interface do SpagoBI pode ser observada na **Figura 9**.

Figura 9 – Interface do SpagoBI



Fonte: LOGZ.IO (2017)

4.3.4 SpagoBI Analytics

SpagoBI *Analytics* são modelos analíticos dirigidos a mercados específicos (e-governo, gerenciamento de risco financeiro) ou domínios de negócios (por exemplo, inteligência competitiva, inteligência de serviços, geocomercialização). O *Analytics* aproveita todo o potencial do pacote SpagoBI para criar soluções personalizadas que atendam às necessidades dos usuários. Os modelos são:

- **Spago4Q** é a plataforma *open source* para medir, analisar e monitorar a qualidade dos produtos, processos e serviços.
- **O SpagoBI for Areas** é o modelo analítico da AREAS, o primeiro ERP italiano projetado e desenvolvido especificamente para o domínio da saúde. Fornece as diferentes áreas de aplicação de saúde com uma camada de inteligência alargada, capaz de suportar os processos gerenciais de decisão e os processos gerenciais administrativos e clínicos. A Inteligência Clínica oferece um apoio real para melhorar a eficiência do curso de cura dos pacientes, através da realização de um grande número de avaliações para identificar possíveis relações entre os vários eventos referentes à história clínica recente e passada dos pacientes.
- **O SpagoBI para Elisa** é o modelo analítico que suporta o processo local de gestão tributária e a recuperação dos montantes devidos à evasão fiscal. De acordo com as leis e regulamentos italianos, a cobrança de impostos locais é responsável pelos municípios locais. O ELISA apoia a gestão fiscal local e a descoberta de rendimentos de evasão fiscal através de análises cruzadas sobre os dados de propriedade dos municípios locais. Os dados podem ser analisados em várias fontes de dados, em diferentes níveis de detalhe através de várias visualizações analíticas predefinidas e fáceis de usar.
- **O SpagoBI da Adaxa** é o modelo analítico do Adaxa Suite, uma coleção de aplicativos de negócios de código aberto, abordando as principais áreas funcionais da maioria dos domínios de negócios. Ele fornece um alto nível de inteligência e relatórios em todas as fontes de dados da suíte, incluindo fontes de dados externas, oferecendo vistas

de negócios que seriam muito caras ou muito complexas para criar com sistemas tradicionais e proprietários. O Adaxa Suite e o SpagoBI são ferramentas complementares. O Adaxa Suite suporta o log de transações de várias fontes de negócios e a SpagoBI é capaz de identificar a forma e as tendências dos negócios confiando em suas capacidades analíticas. Os relatórios padrão fornecidos pelo aplicativo incluem: fluxo de pedidos, pedidos de volta, margem de vendas, desempenho do estoque, eficácia da estocagem, recuperações de fabricação, desempenho de entrega de fornecedores, vendas e pedidos de orçamento, meta mensal, pipeline.

- **O SpagoBI para o Pagoda** é o modelo analítico que fornece painéis de exibição prontos para uso para a Governança de Laboratório Clínico. Pagoda permite a coleta, integração e processamento de dados provenientes de laboratórios clínicos. Os dados são coletados e analisados para fornecer informações valiosas para monitorar, verificar e otimizar o fluxo de trabalho. O Pagoda permite que os gerentes de laboratório monitorem de forma fácil e autônoma dados históricos e em tempo real relativos a volumes, atividades, consumíveis, adequação, qualidade e epidemiologia.

Por fim, o SpagoBI é uma suíte de BI *open source* que inclui ferramentas de relatórios, gráficos e mineração de dados. É desenvolvido pelo *Open Source Competency Center of Engineering Group*, uma grande empresa italiana de software e serviços que também oferece serviços profissionais, como suporte ao usuário, manutenção, consultoria e treinamento.

4.4 ANÁLISE COMPARATIVA ENTRE AS FERRAMENTAS

Após a apresentação das ferramentas, onde foi possível obter uma visão geral de suas funcionalidades, foi realizada a comparação de fato entre elas. Para isso foram utilizados critérios comuns a qualquer ferramenta de Inteligência de Negócios para que a análise comparativa seja justa e eficaz.

Dentre diversas possibilidades de critérios, foram selecionados os julgados mais essenciais para uma ferramenta de BI. As fontes de informações utilizadas para inferir as avaliações foram baseadas em pesquisas nos sites das ferramentas, sites especializados, trabalhos acadêmicos, artigos publicados e em livros especializados no assunto.

Como forma de comparar as ferramentas, foram definidos alguns critérios de avaliação eleitos com base nas tecnologias disponibilizadas por cada uma. Estes critérios foram divididos em nove tabelas e enumerados para melhor leitura e avaliação.

Quadro 4 - Descrição das versões das ferramentas de Inteligência de Negócios analisadas.

	JASPERSOFT	PENTAHO	SPAGOBI
Edições e Versões atuais	<p>Edição comercial (60 dias de experimento)</p> <ul style="list-style-type: none"> • Jaspersoft BI Enterprise Edition v6.4 • Jaspersoft ETL v6.3 <p>Edição comunitária:</p> <ul style="list-style-type: none"> • JasperReports® Server v6.4.0 • JasperReports® Library v6.4.1 • Jaspersoft® Studio v6.4.0 • iReport Designer v5.6.0 • Jaspersoft® ETL v5.6.2 	<p>Edição comercial (30 dias de experimento)</p> <ul style="list-style-type: none"> • Pentaho Edition 7.1 <p>Edição comunitária</p> <ul style="list-style-type: none"> • Pentaho Community Edition 7.1 	<p>Edição única e totalmente gratuita:</p> <ul style="list-style-type: none"> • SpagoBI 5.2
Licenças das versões gratuitas	AGPLv3 (GNU Affero General Public License version 3)	GPLv2 (GNU General Public License Version 2)	LGPL (GNU Lesser General Public License)
Idiomas de documentação	Inglês	Inglês, Alemão e Francês	Francês, Espanhol, Inglês e Italiano

Fonte: Autoria própria.

Quadro 5 – Critérios básicos: Características consideradas básicas para qualquer ferramenta de Inteligência de Negócios

Critério	JASPERSOFT	PENTAHO	SPAGOBI
<p>1.1. Desempenho:</p> <p>Avalia se a ferramenta tem uma boa performance ao processar consultas com um alto volume de dados.</p>	<p>A implementação de serviços de BI da Jaspersoft pode variar de um relatório simples de serviço de CPU único em um banco de dados local, para um conjunto de Servidores JasperReports que se conectam a dezenas de fontes de dados. Para relatórios maiores, o JasperReports Server usa um virtualizador para evitar falta de memória. Ele usa o JRFileVirtualizer e está configurado para retroceder após 300 páginas.</p>	<p>Pentaho fornece uma solução escalável que atende organizações de todos os tamanhos. O Pentaho Business Analytics Server pode ser implantado em diferentes configurações, desde um único nó de servidor até um conjunto de nós distribuídos em vários servidores.</p>	<p>O SpagoBI fornece muitas capacidades analíticas graças a mais de vinte motores analíticos e operacionais. Toda a gama de recursos de inteligência de negócios é coberta por vários mecanismos para obter a melhor solução que atenda às necessidades específicas. (SPAGOWORLD, 2017)</p>
<p>1.2. Consultas ad hoc OLAP:</p> <p>Avalia se a ferramenta permite ao usuário ter a liberdade de definir consultas que acredita ser melhor em um dado contexto.</p>	<p>A funcionalidade Ad Hoc está disponível apenas para usuários da edição JasperReports Server Professional. Criado com o editor Ad Hoc, essas visualizações podem ser baseadas em conexões de clientes OLAP. Uma visualização Ad Hoc pode ser uma tabela, gráfico ou <i>crosstab</i> e é o ponto de entrada para operações de análise, e detalhamento. Pode-se salvar a visualização Ad Hoc como um relatório para editá-la no visualizador de relatório interativo, agendá-la ou adicioná-la a um painel de controle. (TIBCO, 2017)</p>	<p>Consultas ad hoc podem ser feitas através do módulo Pentaho Reporting, que contempla duas ferramentas, uma ferramenta de geração de relatórios, também conhecida como JFreeReport e outra para geração de metadados, a qual permite a criação Ad-Hoc de relatórios via web browser.</p>	<p>Self-Service BI é a forma para o usuário final que permite criar conjuntos de dados e análises. Esses recursos estão disponíveis no menu MyData e MyAnalysis. Pode-se selecionar o conjunto de dados para criar uma planilha ou um 'Qbe' (<i>QueryByExample</i>) que é um mecanismo que permite pesquisar uma fonte de dados relacional através de uma interface gráfica sem escrever uma única linha de sql.</p>

<p>1.3. Arquitetura:</p> <p>Avalia se a solução implementa arquiteturas OLAP que possuem alta escalabilidade, como por exemplo, ROLAP ou HOLAP.</p>	<p>Jaspersoft OLAP é baseado em um servidor ROLAP de código aberto chamado Mondrian. Ele permite analisar dados organizados em um modelo dimensional hierárquico, que tem cubos e dimensões como entidades primárias.</p>	<p>Utiliza o sistema OLAP Mondrian que consiste em quatro camadas: a camada de apresentação, a camada dimensional, a camada de estrela e a camada de armazenamento. Mondrian fornece uma API para aplicativos clientes para executar consultas.</p>	<p>SpagoBIJPivotEngine é o motor OLAP baseado no cliente OLAP JPivot. O motor vem com uma versão incorporada do servidor Mondrian.</p>
<p>1.4. Plataforma:</p> <p>Avalia se a ferramenta pode ser executada nos sistemas operacionais mais difundidos, como Windows, Linux e UNIX.</p>	<p>Jaspersoft suporta diversas plataformas, sendo elas: Windows, Linux e Mac OS X. Possui SDK: Apple iOS e Android. Jaspersoft é 100% Java, então é capaz de rodar em qualquer plataforma java de versão compatível.</p>	<p>Pentaho é desenvolvido em java e executa em diversas plataformas:</p> <p>Servidor:</p> <p>Microsoft Windows 2008 Server R2 & 2012 Server R2, CentOS 6 & 7, Red Hat Enterprise 6 & 7, Ubuntu Server 14.04 LTS & 16.04 LTS, SUSE Linux SLES 11 (SP3+)</p> <p>Desktop:</p> <p>Microsoft Windows 7, 8, & 10, Ubuntu Desktop 14.04 LTS & 16.04, OS X 10.11 & 10.12</p>	<p>Qualquer sistema operacional que suporta Java (JVM 1.5) como Windows ou Linux.</p>

<p>1.5. Suporte Técnico e Documentação:</p> <p>Avalia o nível de qualidade da documentação e suporte técnico oferecido pela ferramenta.</p>	<p>Um vasto material pode ser encontrado no site da comunidade em: <https://community.jaspersoft.com/documentation>.</p> <p>Também há suporte via fórum da comunidade em: <https://community.jaspersoft.com/answers></p> <p>Possui uma comunidade de mais de 250.000 membros registrados.</p>	<p>Uma boa documentação pode ser acessada através do site <https://help.pentaho.com/Documentation>.</p> <p>Também é possível participar do fórum da comunidade que presta suporte em: <http://forums.pentaho.com/></p> <p>A Pentaho mantém uma comunidade forte e ativa ao longo de 13 anos.</p>	<p>Muitos links não funcionam na parte de documentação do site oficial em: <http://www.spagobi.org/homepage/services/documentation/>.</p> <p>Existe um fórum de suporte em: <https://www.spagoworld.org/jforum/></p>
--	---	--	---

Fonte: Autoria própria.

Quadro 6 – Relatórios: Usabilidades dos relatórios e gráficos

Critério	JASPERSOFT	PENTAHO	SPAGOBI
<p>2.1. Agendamento:</p> <p>Avalia se a ferramenta permite o agendamento de relatórios.</p>	<p>A ferramenta possui suporte para agendamento automático de relatórios. É possível programar relatórios para distribuição por e-mail e armazenamento no repositório.</p>	<p>É possível usar a página Schedules do console do usuário para agendar um relatório para ser executado em intervalos regulares, em determinadas datas e horários, e com diferentes parâmetros. Também se pode configurar um relatório agendado para ser enviado automaticamente por e-mail.</p>	<p>É possível agendar relatórios e o formato de saída padrão do SpagoBI é HTML, podendo-se alterar o formato de saída para PDF e enviá-lo por e-mail.</p>

<p>2.2. Dashboards:</p> <p>Avalia se a ferramenta possibilita a criação de painéis.</p>	<p>Um Jaspersoft <i>Dashboard</i> exibe vários relatórios em uma única visualização integrada. Um painel de controle pode incluir outros painéis, controles de entrada para escolher os dados exibidos em um ou mais quadros e quadros personalizados que apontam para URLs para outros conteúdos. Ao combinar diferentes tipos de conteúdo relacionado, é possível criar painéis de leitura atraentes e ricos em dados que rapidamente transmitem tendências.</p>	<p>O Pentaho fornece ferramentas e serviços para criar painéis personalizados para suportar grandes visualizações de dados e projetos. É possível explorar demonstrações do painel para ver exemplos de como o Pentaho pode ser personalizada com base em objetivos comerciais específicos.</p>	<p>O SpagoBI usa um motor para criar painéis dinâmicos com o OpenLazslo. O formato do dados e projetos. É possível explorar demonstrações do painel para ver exemplos de como o Pentaho pode ser personalizada com base em objetivos comerciais específicos. Os valores podem ser recuperados de um conjunto de dados associado ao documento.</p>
<p>2.3 Exportação para outros formatos:</p> <p>Avalia se a ferramenta dispõe de recursos de exportação para formatos como PDF, HTML e ODT/ODS (para permitir futuramente a integração com ferramentas livres).</p>	<p>É possível usar JasperViewer para exibir o relatório antes de salvar nos formatos PDF e ODT. Também é possível exportar relatórios em formato HTML.</p>	<p>É possível selecionar o tipo de saída para a documentação gerada (PDF, HTML, DOC, Excel, CSV ou METADATA). Aceita como entrada Excel (XLS, XLSX), incluindo formato OpenOffice (ODS), mas não foi encontrada documentação que cita se exporta em formato ODS.</p>	<p>Jasper Server permite exportação de relatórios nos formatos de pdf, Html, xls, flash, doc, rtf e ods/odt.</p>

Quadro 7 – Funcionalidades Web: Disponibilidade da Empresa via Web para suporte.

Critério	JASPERSOFT	PENTAHO	SPAGOBI
<p>3.1 Suporte a dispositivos móveis:</p> <p>Avalia-se a ferramenta suporta o uso de dispositivos móveis como Android e IOs.</p>	<p>O servidor JasperReports fornece relatórios e análises que podem ser incorporados em uma aplicação web ou móvel, além de servir como um repositório de informações que podem ser entregues em tempo real ou agendado.</p>	<p>O Pentaho Mobile está disponível para permitir visualizar e editar relatórios do Analyzer ou rever os relatórios feitos pelo Dashboard Designer, Interactive Reporting e Report Designer através do iPad da Apple. O Pentaho Mobile é um aplicativo da Web disponível apenas com a versão Enterprise Edition do Pentaho, para permitir acesso aos usuários de Apple iPad. A Pentaho já não enviará o Pentaho Mobile com o produto na versão 7.1 e removerá o suporte para o Pentaho Mobile na 8.0.</p>	<p>A aplicação SpagoBIMobileEngine é uma aplicação web que precisa ser implantada em um servidor de aplicativos para os usuários acessarem e executarem documentos. Um aplicativo específico para dispositivos móveis, acessa o mesmo Servidor e interage diretamente com ele para produzir documentos interativos disponíveis em dispositivos móveis. Está disponível um aplicativo personalizado para Apple iPad, que instala um link para o motor móvel SpagoBI. É possível criar ou modificar documentos.</p>

Fonte: Autoria própria.

Quadro 8 – Simulação de Cenários: Análises feitas com a interação do usuário

Critério	JASPERSOFT	PENTAHO	SPAGOBI
<p>4.1. Análise preditiva e simulação “What If”:</p> <p>Avalia se a ferramenta possui o recurso de simulação de cenários hipotéticos, exibindo-os através de gráficos e dados. Também se consegue prever tendências futuras e comportamento.</p>	<p>Com o Jasper Server é possível executar relatórios complexos criados com Jasper Reports e por exemplo, preencher um parâmetro no relatório que permita especificar uma porcentagem do crescimento do Volume de Negócios e ver o que acontece com seu custo se a proporção de volume de negócios para custo permanecer igual.</p>	<p>O Pentaho Server contém sistemas e componentes que fornecem relatórios e análises com recursos de <i>slice-and-dice</i>, simulação “what if” e mineração de dados. O Pentaho faz a análise “what if”, fornecendo um relatório parametrizado incorporado em uma página contendo um “painel de controle” com controles para definir os parâmetros do relatório. Ele utiliza controles deslizantes usando a escala 0 - 100% para modificar o valor de um parâmetro por essa fração. Então ele permite que o usuário modifique as variáveis de entrada para observar a mudança resultante nas saídas.</p>	<p>O SpagoBI Whatlf Engine é o novo motor OLAP SpagoBI com capacidades de gravação de retorno, baseado no popular motor OLAP Mondrian. O “What If” é a capacidade de mudar os valores das células do cubo e ver como essas mudanças impactam no resto do cubo. Depois de alguma interação com o cubo, o usuário pode salvar os novos dados modificados, sem tocar no DW original.</p>

Fonte: Autoria própria.

Quadro 9 – Produto: Custos do produto no mercado

Critério	JASPERSOFT	PENTAHO	SPAGOBI
6.1. Custo: Avalia o custo para a compra da licença da ferramenta, caso exista uma versão paga;	Preço disponível apenas por orçamento entrando em contato diretamente com a empresa para detalhes de preços.	A Pentaho não publica seus preços corporativos em seu site, mas é possível obter uma cotação gratuita.	Spago BI não possui versão paga. No entanto é oferecido suporte ao usuário ao custo de € 6,000 anualmente. Valor disponível no site oficial < http://www.spagobi.org/homepage/services/professional-services/ > em Setembro de 2017.

Fonte: Autoria própria.

Quadro 10 – Ferramenta de Planejamento: Modo de distribuição dos relatórios

Critério	JASPERSOFT	PENTAHO	SPAGOBI
7.1. Carregamento de dados de diferentes fontes (Big Data): Avalia a possibilidade de integração da solução com fontes de dados heterogêneas.	O Jaspersoft Big Data Connectors fornece conectividade nativa para iReport, JasperReports e JasperReports Server para Hadoop e fontes de dados NoSQL. Os conectores fornecem as ferramentas necessárias para gerenciar suas necessidades de Big Data Analysis usando ETL e também para realizar pesquisas diretas e interativas a relatórios diretos de fontes de dados Hadoop e NoSQL.	O Pentaho fornece ferramentas de dados para extrair, preparar e misturar dados, além das visualizações e análises. Com integração Hadoop e Spark para NoSQL, o Pentaho permite transformar grandes massas de dados em informações úteis para gerar conhecimento.	O SpagoBI usa a plataforma de dados MapR. Permite criar análises, mesmo usando dados provenientes de seu <i>data lake</i> (repositório dentro da empresa, com todos os dados brutos disponíveis). Comumente usa-se o Hadoop para trabalhar com os Data Lakes. O Apache Drill é o mecanismo que permite a explorar dados em <i>Big data</i> com um SQL sem esquemas.

Fonte: Autoria própria.

4.4.1 Considerações sobre o Jaspersoft

Jaspersoft é uma plataforma de *Business Intelligence* (BI) com recursos de visualização, análise e relatórios de dados. Projetado para as pequenas e médias empresas, a Jaspersoft permite que os desenvolvedores de aplicativos incorporem painéis, relatórios e análises altamente interativos em suas aplicações.

Arquitetural e economicamente escalável, a Jaspersoft oferece às organizações de todos os tamanhos a capacidade de obter informações de múltiplas fontes de dados e tomar melhores decisões de negócios.

Usando um modelo comercial de código aberto e confiável por uma comunidade de mais de 250.000 membros registrados, a solução de software de código aberto da ferramenta foi baixada em mais de 14,5 milhões de vezes. Jaspersoft tem escritórios em todo o mundo.

Embora o relatório da Jaspersoft para a elaboração de relatórios seja comparativamente melhor do que o Pentaho Report Designer, as capacidades de painel do Pentaho em termos de funcionalidade são melhores. Isso ocorre já que a funcionalidade do Painel está presente apenas na edição Enterprise do Jaspersoft, enquanto que no Pentaho é acessível na edição básica.

Para resumir, o foco da Jaspersoft é mais sobre relatórios e análises e o foco da Pentaho na integração de dados, ETL e automação de fluxo de trabalho. Jaspersoft tem um forte foco em relatórios e análises. Possui uma melhor interface de usuário que o Pentaho e é mais fácil de usar, pois o Pentaho se concentra na integração de dados, ETL e automação de fluxo de trabalho.

4.4.2 Considerações sobre o Pentaho

Alguns pontos positivos do Pentaho são a gestão e distribuição de informes e painel de controle sobre plataforma de código aberto: otimização do intercâmbio de informação e a colaboração. Já os pontos negativos que podem ser destacados são a pouca documentação e necessidade de conhecimento técnico elevado.

Também pode-se destacar as capacidades técnicas do Pentaho como:

- Integração com diferentes fontes de dados e aplicativos;

- Capacidade de usar APIs, serviços web, modificar planilhas, regras de negócios e inclusive o código fonte;
- Flexibilidade nas opções de saída, podendo ser: PDF, HTML, Microsoft Excel, texto plano, etc;
- Ser multi-plataforma: Windows, Linux, Macintosh, Unix, etc.

Entre as organizações que adotam o Pentaho, uma das vantagens sentidas é o baixo tempo de integração e o custo de infraestrutura comparado por exemplo ao SAP BIA e outras ferramentas. Junto com isso, o suporte da comunidade disponível 24 horas por dia, com fóruns de suporte ativo. Sua visualização ilimitada e fontes de dados podem lidar com qualquer tipo de dados, juntamente com um bom conjunto de ferramentas que possui ampla aplicabilidade além do produto base.

Por fim, o Pentaho é um pacote de BI completo que cobre várias operações de relatórios para mineração de dados, tendo como principal componente o Pentaho Reporting, que é um rico conjunto de recursos.

4.4.3 Considerações sobre o SpagoBI

O SpagoBI é o único pacote de BI completo que elimina qualquer software e bloqueio de fornecedores: os direitos dos usuários de usar o software estão claramente separados da compra de serviços profissionais, sem a obrigação de compra do cliente.

Este modelo de negócio específico está impulsionando a crescente adoção da suíte em todo o mundo. Ultimamente, o SpagoBI foi o número um dos principais projetos de código aberto baixados do OW2 Consortium.

5 CONSIDERAÇÕES FINAIS

No início deste trabalho buscou-se responder a seguinte pergunta de pesquisa: Como a utilização de softwares de Inteligência de Negócios *Open Source* pode contribuir nos custos de licenciamento para as organizações?

Sendo assim, as ferramentas livres analisadas neste trabalho oferecem uma ampla variedade de relatórios e funcionalidades, sendo utilizadas em diversas organizações.

Algumas ferramentas possuem um maior destaque como a Pentaho Community e SpagoBI pois fornecem capacidades extras como um bom processamento OLAP, uma interface efetiva, suporte para diferentes áreas, entre outras funcionalidades.

As ferramentas *open source* citadas neste trabalho podem ser usadas, dentro de suas capacidades, para substituir poderosas ferramentas pagas de BI como por exemplo Oracle *Business Intelligence*, Microsoft *Business Intelligence*, SAP *Business Intelligence*, SAS *Business Intelligence* e IBM Cognos.

Como é possível observar, algumas ferramentas se destacaram mais em alguns critérios do que em outros e algumas observações importantes devem ser consideradas na hora da escolha de uma ferramenta de BI, como observar a data das informações, pois esta é uma tecnologia que vem avançando com muita rapidez e algumas fontes de pesquisa podem apresentar relatórios e informações com data desatualizada, o que não representa a situação atual.

É importante observar também se as informações se referem à última versão do produto e atentar-se aos pontos fortes e fracos de cada solução. Talvez alguns desses pontos sejam mais relevantes para o negócio de cada organização.

Alguns excelentes fabricantes podem não estar sendo considerados nas pesquisas por diversos motivos como, por exemplo, possuir uma tecnologia recente. Se possível é recomendável acompanhar fontes de pesquisas confiáveis como, por exemplo, os relatórios do Gartner, e observar a manutenção da posição da ferramenta pretendida no quadrante.

Muitas vezes uma ferramenta de BI *open source* consegue resolver diversos problemas e são em grande parte produtos que não têm grande visibilidade no mercado e acabam sendo dispensadas erroneamente. Cada necessidade deve ser

analisada do ponto de vista do cliente, e não do fabricante. A experiência de outrem é sempre boa referência, portanto deve-se procurar saber se outros projetos já implementaram a solução pretendida na ferramenta em questão e descobrir quais foram os benefícios e dificuldades ocorridos.

As três ferramentas analisadas neste trabalho fornecem um conjunto de recursos rico pronto para uso corporativo, sendo necessário o usuário final fazer uma comparação completa para selecionar qual satisfaz suas necessidades.

5.1 LIMITAÇÕES DA PESQUISA E TRABALHOS FUTUROS

Este trabalho teve como limitação a pesquisa a utilização de apenas três ferramentas para a exposição e análise comparativa, sendo todas elas com ao menos uma versão *open source*. Sugere-se para pesquisas futuras que sejam feitas mais análises comparativas utilizando-se os mesmos critérios, porém com ferramentas tanto *open source* quanto pagas.

REFERÊNCIAS

BOUMAN, Roland; DONGEN, Jos van (2009). **Business Intelligence and Data Warehousing with Pentaho and MySQL**. Wiley Publishing Inc, Indianapolis, Indiana.

BARBIERI, Carlos (2011). **BI2 - Business Intelligence Modelagem e Qualidade**. Ed Elsevier.

CECI, Flávio (2012). **Business Intelligence, livro digital**. Universidade do Sul de Santa Catarina. Unisul Virtual.

CHEN, Hsinchun; CHIANG, Roger HL (2012). **Business Intelligence and Analytics From Big Data to Big Impact**. MIS quarterly v36 n4.

CHEN, Hsinchun (2012). **Dark Web Exploring and Data Mining the Dark Side of the Web**. Integrated Series in Information Systems v30. Springer Science Business Media.

EXAME (2016). **Business Intelligence x Business Discovery: especialista explica a diferença**. Disponível em: <<http://exame.abril.com.br/negocios/dino/business-intelligence-x-business-discovery-especialista-explica-a-diferenca-dino890107269131/>>. Acesso em Setembro de 2017.

FELLER, Joseph *et al* (2005). **Perspectives on Free and Open Source Software**. MIT Press. The MIT Press Cambridge, Massachusetts London, England.

FORBES INSIGHTS (2016). **Breakthrough Business Intelligence - How Stronger Governance Becomes a Force For Enablement**.

INMON W. H.; LINSTEDT, Dan (2014). **Data Architecture - A Primer for the Data Scientist Big Data, Data Warehouse and Data Vault**. Morgan Kaufmann.

JASPERSOFT (2017). **Jaspersoft Business Intelligence Software**. Disponível em: <<https://www.jaspersoft.com/>> Acesso em Julho de 2017.

JASPER COMMUNITY (2017). **Jaspersoft Community**. Disponível em: <<https://community.jaspersoft.com/wiki/jasperreports-server-performance-reference>> Acesso em setembro de 2017.

KIMBALL, Ralph; ROSS, Margy (2013). **The Data Warehouse Toolkit**. Ed. John Wiley & Sons.

KOTLER, Philip (2015). **Confronting Capitalism, Real Solutions for a Troubled Economic System.**

LACY, Miguel Koren O'Brien de (2010). **SpagoBI - Plataforma de BI livre e aberta.** Revista Espírito Livre, Junho de 2010. Disponível em: <ftp://ftp.feis.unesp.br/softwarelivre/Revistas/Espirito%20Livre/Revista_EspiritoLivre_015_junho2010.pdf>. Acesso em Agosto de 2017.

LOGZ.IO (2017). **Business Intelligence Tools.** Disponível em: <<https://logz.io/blog/business-intelligence-tools/>> Acesso em Setembro de 2017.

MADSEN, Mark (2009). **BeyeNETWORK Research Report - Open Source Solutions: Managing, Analyzing and Delivering Business Information.** Disponível em: <http://www.b-eye-network.com/files/BeyeNETWORK_Open_Source_Research_Report.pdf>. Acesso em Setembro de 2017.

NAVEGA, Sérgio (2002). **Princípios essenciais do Data Mining.** In. Infoimagem. Anais Cenadem, 2002. Disponível em: <<http://www.intelliwise.com/reports/i2002.htm>>. Acesso em Agosto de 2017.

OPENIN (2017). **Pentaho Community.** Disponível em: <<https://openin.com.br/parceiros/pentaho-community/>>. Acesso em Setembro de 2017.

ORACLE (2017). **Business Intelligence Oracle.** Disponível em: <<https://www.oracle.com/br/solutions/business-analytics/business-intelligence/index.html>>. Acesso em julho de 2017.

PENTAHO (2017). **Pentaho: Data Integration, Business Analytics and Big Data.** Disponível em: <<http://www.pentaho.com/>> Acesso em julho de 2017.

PENTAHO PERFORMANCE (2017). **Performance and Scalability Overview.** Disp. em: <<http://www.pentaho.com/sites/default/files/uploads/resources/pentaho-performance-and-scalability-overview-august2014.pdf>>. Acesso em Setembro de 2017.

PORTER, Michael E. (2008). **On Competition.** Harvard Business Review Press.

POZZEBON, Marlei; FREITAS, Henrique M. R. de; PETRINI, Maira (1997). **Pela integração da inteligência competitiva nos Enterprise Information Systems (EIS).** Ci. Inf., Brasília, v. 26, n. 3, p., Set. 1997. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651997000300003&lng=en&nrm=iso>. Acesso em julho de 2017.

PRASAD, K. V. K. K. (2007). **Data Warehouse Development Tools. Covering Informatica, Cognos, Business Objects and Datastage with Case Studies.** Dreamtech Press.

PRIMAK, Fábio Vinicius (2008). **Decisões com B.I. Business Intelligence.** Ed. Ciência Moderna.

RATNER, Bruce (2011). **Statistical and Machine-Learning Data Mining Techniques for Better Predictive Modeling and Analysis of Big Data.** Second Edition.

SÁ, Thays *et al* (2012). **Uma Análise Comparativa entre as Ferramentas OLAP como Apoio a Soluções de BI nas Empresas.** V Congresso Tecnologia da Informação, Brasil.

SANTOS, Maribel Yasmina; RAMOS, Isabel (2006). **Business Intelligence: Tecnologias da Informação na gestão de conhecimento.** FCA - Editora de Informática.

SPAGOBI (2017). **Spago BI - 100% Open Source Business Intelligence.** Disponível em: <<http://www.spagobi.org/>> Acesso em julho de 2017.

SPAGOWORLD (2017). **SpagoBI cited as a Strong Performer by independent Research Firm.** Disponível em: <<https://www.spagoworld.org/xwiki/bin/view/SpagoBI/PressRoom?id=SpagoBI-ForresterWave-July2012&srid=XibdUp2E>> Acesso em Setembro de 2017.

STACKOWIAK, Robert *et al* (2007). **Oracle Data Warehousing and Business Intelligence Solutions.**

TOBACOWALA, Rishad; GUPTA, Sunil (2016). **Extracting Insights from Vast Stores of Data.** Harvard Business Review, Agosto de 2016.

TIBCO (2017). **Jaspersoft OLAP User Guide.** Disponível em: <<https://docs.tibco.com/pub/js-jrs/6.2.0/doc/pdf/Jaspersoft-OLAP-User-Guide.pdf>> Acesso em Setembro de 2017.

WHITE, Tom (2015). **Hadoop: The Definitive Guide.** 2015. 4th Edition. O'Reilly Media.

ZIKOPOULOS, Paul e EATON, Chris (2011). **Understanding Big Data Analytics for Enterprise Class Hadoop and Streaming Data.**