

Desenvolvimento de uma Ferramenta de Business Intelligence para o Ambiente de Ensino a Distância Amadeus: Um Estudo de Caso.

Marcelo Silva de Moraes Filho, Milton Guedes Duarte, Eduardo José Marcelino Vicente dos Santos.

Faculdade Estácio de Sá do Recife
Av. Eng. Abdias de Carvalho, 1678 – Recife – Pernambuco – Brazil.
marcelosilvafilho@gmail.com, milton.g.duarte@gmail.com,
ejmvs@cin.ufpe.br

***Abstract.** This paper presents the effective use of Business Intelligence (BI) in the environment of open source ODL Amadeus, in order to study and apply its concepts in addition to disclosing the open source community, for the implementation of the project was used the methodologies of research exploratory and qualitative, and as a result was the provision of an analytical tool for the environment of the Amadeus where was carried out some performance tests, using the suite Pentaho for generation of artifacts of this tool. After the end of the development of this project, it was concluded that this tool will bring benefits for educational institution that the use, such as the agility in decision-making or autonomy to carry out the consultations.*

***Resumo.** Este trabalho apresenta o uso efetivo do Business Intelligence (BI) no ambiente open source de Ensino a Distância Amadeus, com o intuito de estudar e aplicar seus conceitos além de divulgar a comunidade open source, para a execução do projeto foi usado as metodologias de pesquisas exploratória e qualitativa, e como resultado foi a disponibilização de uma ferramenta analítica para o ambiente do Amadeus onde foi realizado alguns testes de desempenho, utilizando-se a suíte Pentaho para geração dos artefatos dessa ferramenta. Após o término do desenvolvimento desse projeto, foi concluído que essa ferramenta trará benefícios para instituição de ensino que a usar, como a agilidade nas tomadas de decisões ou autonomia para realizar as consultas.*

1. Introdução.

O mercado atual e as empresas vem aumentando de forma expressiva, e consequentemente a geração de dados provenientes acompanha este desenvolvimento, de certa maneira as informações geradas nestes últimos anos devem ser usadas não apenas de forma inteligente, mas também de maneira ágil a obter uma visão estratégica e prospecção de investimentos futuros, caso contrário poderia ocasionar perdas e dificuldades de sobrevivência.

Segundo Costa e Santos (2012a), o mundo de negócios contemporâneo está cada vez mais competitivo e o seu ambiente organizacional se mostra em constante mudança, as organizações também estão num nível de pressão elevada, com tudo isso, existe a obrigação de obter respostas rápidas a todas essas mudanças, entender as necessidades

dos seus clientes de forma inovadora, e como elas atuam perante o mercado, torna-se elemento decisivo para o sucesso da organização.

Também segundo Costa e Santos (2012b), atuando num mundo global, as organizações ficam numa constante necessidade de informação e conhecimento, sendo a transformação da informação em conhecimento, um recurso essencial para a função central do negócio, para os gestores isso é bem aceito, principalmente quando a informação é disposta numa quantidade aceitável, com qualidade, confiável e no momento certo, alcançando assim uma vantagem competitiva. Contudo, a falta dessas informações leva ao caos e assim ao insucesso.

Existem muitas ferramentas EAD open source, como o Moodle, Amadeus, *WebMatrix*, *Bigbluebutton*, entre outros. O crescente uso desse tipo de ferramenta tem levantado à importância de preparar uma ferramenta de BI.

Dentro deste mercado enfatizamos o Ensino à Distância (EAD) que devido às inovações tecnológicas tem se expandido bastante pelo país e que não difere do mercado de uma maneira geral produzindo uma quantidade de dados significativa e complexa de ser analisada, sendo viabilizada em geral por ferramentas computacionais para o controle dessas informações, dentre essas ferramentas, existe uma desenvolvida pelo grupo de pesquisa em tecnologia educacional do Centro de Informática da UFPE de código *open source* denominada Amadeus.

O principal objetivo do Amadeus é ampliar as possibilidades que os professores podem fazer de trabalho e assim permitir formas criativas de relacionamento, e também promover a comunicação e colaboração entre os participantes (Queiroz, 2010). Considerando essas características do Amadeus, aliadas ao fato de ser um ambiente muito estudado pelo meio acadêmico, faz do Amadeus uma boa escolha a construção de uma ferramenta de *Business Intelligence*, viabilizando seus conceitos e resultados para futuros trabalhos acadêmicos.

2. Business Intelligence, componentes e ferramentas.

O conceito *Business Intelligence* surgiu em 1996 com a *Gartner Research Group*, em algumas organizações já se usava essa ideia que para concretizar esse conceito se usava diferentes sistemas. Com a evolução das Tecnologias e Sistemas de Informação e também com as mudanças das organizações surgiu o conceito *Executive Information Systems*. Este novo conceito implantou novas tecnologias e muitas técnicas de extração, transformação, processamento e apresentação de dados, que veio para melhorar a qualidade da gestão estratégica nas organizações e ajudar a tomada de decisão dos gestores com fornecimento de informação em tempo real. [Costa & Santos, 2012c]

O Business Intelligence é a área da tecnologia da informação com um estudo interdisciplinar, e tem como objeto de estudo a elaboração de sistemas de informação computacionais, como os *data warehouse* (responsáveis por organizar grande volume de dados), os *datas minings* (responsáveis por facilitar a descoberta de relações entre os dados), e oferecer interfaces que facilitem para o usuário o entendimento das relações entre os dados, com a intenção de ajudar, por exemplo, na tomada de decisão. [Turban, 2009 apud Matheus & Parreiras, 2004]

Dentre seus componentes podemos destacar como principais o *data warehouse* (DW), as ferramentas de processo de carga (ETL), as ferramentas de acesso aos dados (OLAP), metadados, relatórios e consultas e *dashboards*.

Inmon (1997) foi quem criou o termo *data warehouse* (DW) ou armazenagem de dados, como sendo "uma coleção de dados orientada por assunto, integrada, variante e não volátil, que tem por objetivo dar suporte aos processos de tomada de decisão".

O momento de extração dos dados e a inclusão no DW é conhecido como o processo de carga, sendo esta uma das fases mais críticas da construção de um DW, uma vez que, nesse processo é realizado procedimento de limpeza, transformação dos dados e integração. Esses procedimentos são realizados por ferramentas de ETL (*Extract, Transform and Load*). [Santos, 2004]

É através das ferramentas OLAP que é possível disponibilizar ao usuário os dados contidos no DW em forma de informação gerencial, essas ferramentas são as melhores de se trabalhar com o modelo dimensional, o que permite através de várias visões complexas elaborarem análises sofisticadas. [Santos, 2004]

Metadados segundo Berson e Smith (1997 apud Santos, 2004) são dados que são utilizados para dizer o que são descritos de dados contidos num DW/BI, e informa também quais são as técnicas necessárias para sua administração, isso tudo irá facilitar na elaboração de relatórios e consultas pelo usuário final.

Os relatórios são uma das ferramentas mais usada para análise num sistema de BI, elas geralmente são um documento que apresenta um conjunto específico de dados de forma estruturada, o que permite entregar a informação necessária aos seus utilizadores, num formato o qual eles escolheram, através de informação resumida, com a utilização de tabelas dinâmicas (*pivot tables*), gráficos ou outras formas [Janus & Misner, 2011 apud Cavaco, 2013]

Diante da necessidade de agregar toda a informação relevante de um BI num único painel de leitura rápida, imediata e acessível, surgiu os *dashboards*, ajudando o monitoramento o desempenho da organização, podendo em alguns casos, substituir algumas abordagens de relatórios. [Watson, 2006 apud Val, 2010]

2.1. PENTAHO

A plataforma *Pentaho* é um projeto *open source* para BI, composto de vários produtos e subprodutos. O projeto também é conhecido como *Open BI Suite*, onde fornece desde a relatórios abrangentes, a análises OLAP, mineração de dados, *dashboards* e integração de dados. [Pentaho, 2010 apud De Souza Fernandes & Kantorski, 2012]

Todos os programas no *Pentaho*, praticamente, foram programados em Java, alguns dos componentes no lado do servidor web foram implementados na tecnologia AJAX. O Java tem como propriedade ser uma linguagem portátil entre arquiteturas de hardware e sistemas, e como consequência o *Pentaho* fica disponível para muitos sistemas operacionais diferentes. [Bouman & Dongen, 2009]

3. EAD e o Ambiente Amadeus.

São ambientes de Educação a Distância, também conhecidos como EAD, eles são baseados em Sistemas de Gerenciamento de Aprendizado (LMS, abreviatura em inglês)

que utilizam importantes ferramentas de comunicação entre o Aluno e Tutor. Essas ferramentas que são acopladas ao LMS têm como objetivo oferecer um ambiente de armazenamento de dados e também a troca de informações durante o processo de ensino e aprendizagem a partir de mensagens do tipo assíncronas (armazenadas) e síncronas (em tempo real). [Borges Do Val et. al., 2011]

O crescimento do EAD no Brasil teve seu ápice a partir de 2004, com a grande inserção de universitários e no ano de 2008 com cerca de um milhão de Alunos estavam regularmente matriculados (cerca de um sexto do total matriculado dessa época). Esse sucesso tem como fator a “extraterritorialidade”, o que permite que Aluno de outros estados se matriculem. [Val, 2010]

O Amadeus é um LMS de segunda geração, baseado no conceito *Blended Learning*, esse conceito trata-se da combinação de diversas plataformas de forma harmônica. O objetivo desse LMS é estender ao máximo os vários tipos de interação entre os usuários. [Gomes, 2011]

Ainda segundo Gomes (2011), os principais diferenciais do Amadeus são:

- Código aberto, tendo como consequência a redução dos custos no momento de sua aquisição e implantação, e também a constante atualização da ferramenta.
- O uso de diferentes mídias, que inclui desde os *chats* até vídeos para assistir e discutir de forma colaborativa.
- Disponibilização de formas alternativas de interação, através de jogos, telefones celulares ou até de experimentos de laboratórios que podem ser realizados e analisados de forma remota.
- Integração com o Sistema Brasileiro de TV Digital.

4. Metodologia

Neste trabalho, foram adotadas metodologias de pesquisas exploratória e qualitativa. Dentro dos procedimentos técnicos de coleta de dados, esse trabalho se classifica como um estudo de caso, para realizar esse procedimento foi usado o modelo proposto por Kimball (2002) e também fora utilizado um questionário.

A seleção do modelo proposto por Kimball (2002) se deu pelo fato de ser o mais completo, detalhado, de implementação mais rápida e que tem como foco a criação de incrementos com os *data mart*, disponibilizando uma maior interação com o usuário e consequentemente um feedback mais rápido dos mesmos. O modelo é composto de doze etapas, conforme representado na figura 1 abaixo:

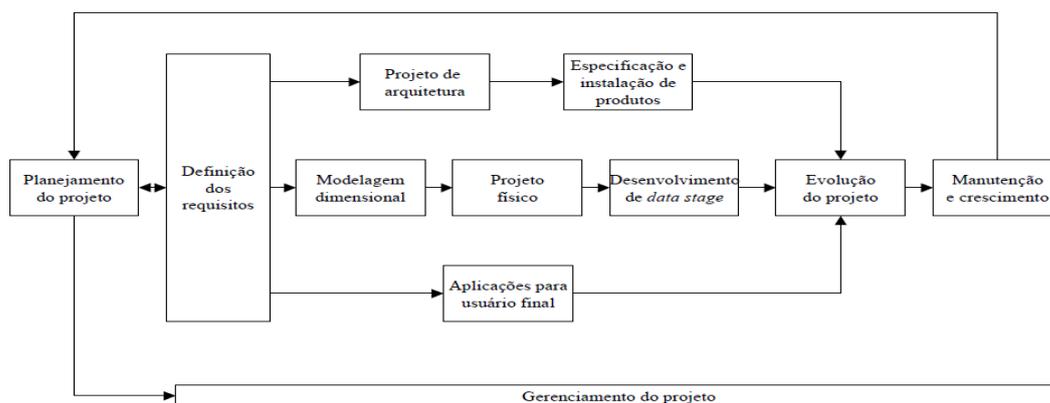


Figura 1. Modelo do processo de Construção de um DW
Fonte: Kimbal (2002)

Desses processos não foram utilizadas as etapas de gerenciamento de projeto, projeto de arquitetura e projeto físico pelo fato do projeto ter uma curta duração.

5. Projeto

O Projeto oferece ao usuário do Amadeus uma ferramenta completa. Ele disponibiliza para o usuário os seguintes artefatos: scripts SQL de criação das tabelas FATO e dimensão, ambiente pré-configurado para fazer o todo o processo ETL nos dados, XML para publicar os cubos e exemplo de relatórios e *dashboards*.

5.1. Planejamento

A primeira ação foi entrar em contato com os desenvolvedores do Amadeus, ambiente de estudo em questão, onde indicaram um usuário na UNIVASF. O mesmo foi informado sobre a ferramenta onde o ajudaria na gerência do ambiente, aprovando a iniciativa do projeto e provendo informações para sua elaboração.

5.2. Definição de requisitos de negócio

No portal SPB do governo federal, existe um fórum de discussão, onde foi divulgado o projeto, e solicitada a participação de usuários experientes na ferramenta, para que estes pudessem participar do levantamento dos FATOS. O instrumento usado para levantamento dos requisitos foi o questionário, o qual foi enviado para o corpo acadêmico da UNIVASF. Ainda neste questionário foram levantados os assuntos mais relevantes e o público-alvo.

5.3. Seleção e instalação de produtos

Neste processo foram definidas as ferramentas, todas *open source*, levando em consideração o foco acadêmico e a própria ferramenta Amadeus ser de código livre.

O SGBD do Amadeus é o *PostgreSQL* e para fazer a engenharia reversa e modelagem dimensional foi utilizado a ferramenta *SQL Power Architect*. Para atender todas as fases de construção do DW será usada a suíte Pentaho, que é uma ferramenta de arquitetura *ROLAP*. As ferramentas presentes nessa suíte são: PDI (extração, transformação e carga), *Schema Workbench* (publicação dos cubos OLAP), e *BI-Server* (consultas OLAP).

5.4. Modelagem Dimensional

Dentro deste processo foi realizada a engenharia reversa do banco de dados do Amadeus, que contém um total de quarenta e sete tabelas, foram utilizados os quatro passos da modelagem dimensional de Kimball (2002): a escolha do processo que será modelado, definição da granularidade, escolha das dimensões e a identificação dos FATOS.

Para realizar a modelagem dimensional foi utilizada a técnica definida por Machado (2008). Essa técnica consiste em dividir as dimensões em quatro tipos: Onde, O quê, Quando e Quem.

- Dimensão Onde: Onde foi realizada a venda?
- Dimensão O Quê: O que foi vendido?
- Dimensão Quem: Quem realizou a venda?
- Dimensão Quando: Quando foi realizada a venda?

Aplicando a metodologia, foi possível com os requisitos especificados, a geração de cinco cubos, com suas respectivas tabelas FATO e de DIMENSÃO são eles:

- Cubo de Pedidos para Docência: apresenta os registros dos pedidos realizados para ser um docente (professor) junto com seus respectivos currículos, e ele possui como métrica a quantidade dos pedidos realizados.
- Cubo de Histórico dos Materiais: apresenta os registros dos históricos dos materiais ao longo do tempo, e ele possui como métrica a sua quantidade.
- Cubo Avaliações Realizadas: apresenta os registros das avaliações realizadas ao longo do tempo, e ele possui como métrica a quantidade de avaliações.
- Cubo Histórico do Curso: apresenta os registros dos cursos realizados ao longo do tempo, e ele possui como métrica a quantidade e a média da quantidade de alunos.
- Cubo Mensagens do Professor: Esse cubo apresenta os registros das mensagens realizadas pelo professor ao longo do tempo, e ele possui como métrica a quantidade.

Dentro do processo, no momento da divisão, percebeu-se o compartilhamento de dimensões entre as Fatos. A figura abaixo apresenta um dos cubos que foram citados acima, ao centro observamos a tabela Fato e ao redor as tabelas de Dimensão.

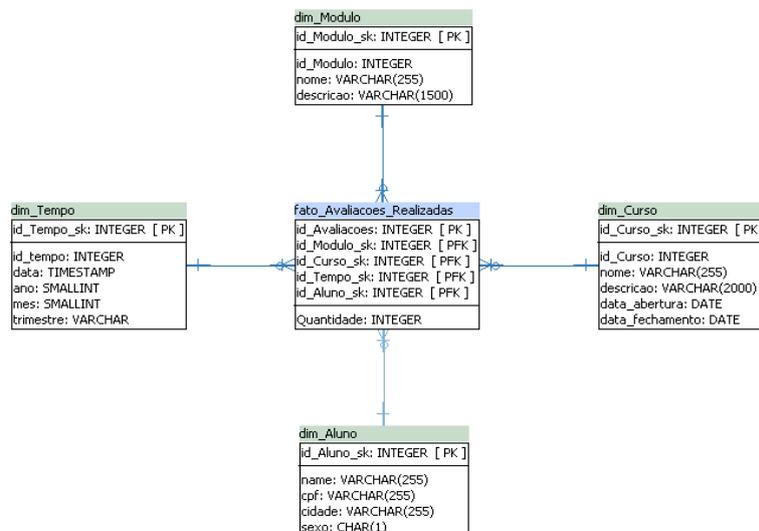


Figura 2. Cubo de Avaliações Realizadas
Fonte: Elaborada pelos autores

5.5. Projeto de Desenvolvimento da Organização dos Dados

O modelo dimensional final do projeto apresenta um total de sete tabelas dimensão e cinco tabelas FATO, sendo quatro dessas dimensões compartilhadas com outros FATOS. Para cada tabela foi elaborado um processo de ETL usando a ferramenta, da suíte *Pentaho*, chamado PDI.

Para criar a tabela FATO, é necessário que todas as transformações das dimensões no PDI já tenham sido executadas.

5.6. Especificação e Desenvolvimento de Aplicações do Usuário Final

Com os cubos gerados a partir do questionário o qual foi levantado algumas necessidades do usuário, o desenvolvimento se deu com a criação e publicação dos cubos OLAP. Esse cubos são arquivos XML que serão interpretados pelo *BI-SERVER*. Foi utilizado a ferramenta da *Schema Workbench* da suíte *Pentaho* e depois de publicados, serão apresentados no *BI-SERVER*

6. Resultados

Os resultados foram obtidos através de simulações para a geração de relatórios e *dashboards*. Foi usado dados fictícios obtidos no site *Mockaroo* (MOCKROO, 2015), e também alguns poucos dados gerados pela UNIVASF.

O *BI-SERVER* permite a geração de relatórios ad-hoc, com isso foi possível analisar cada cubo citados anteriormente, a figura a seguir mostra o relatório do cubo Avaliações Realizadas:

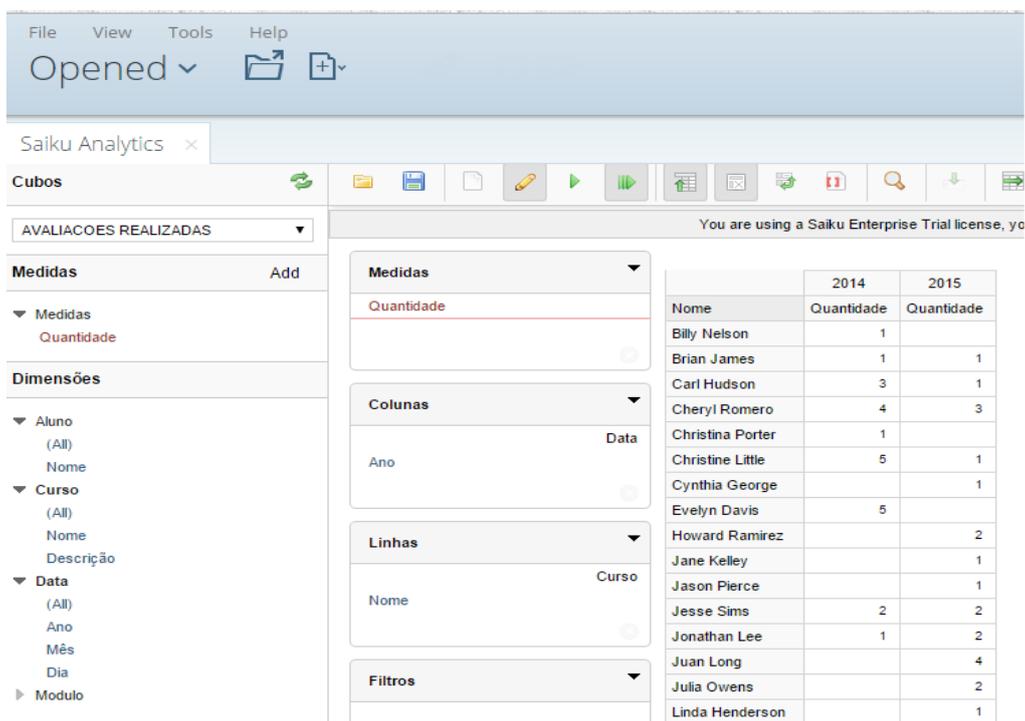


Figura 3. Relatório Ad-hoc no Bi-Server do cubo Avaliações Realizadas
Fonte: Elaborada pelos autores

Agora passando a dimensão curso que era linha para coluna e adicionando a dimensão módulo como linha geramos um novo relatório, mudando a visão do cubo como é mostrado na figura a seguir:

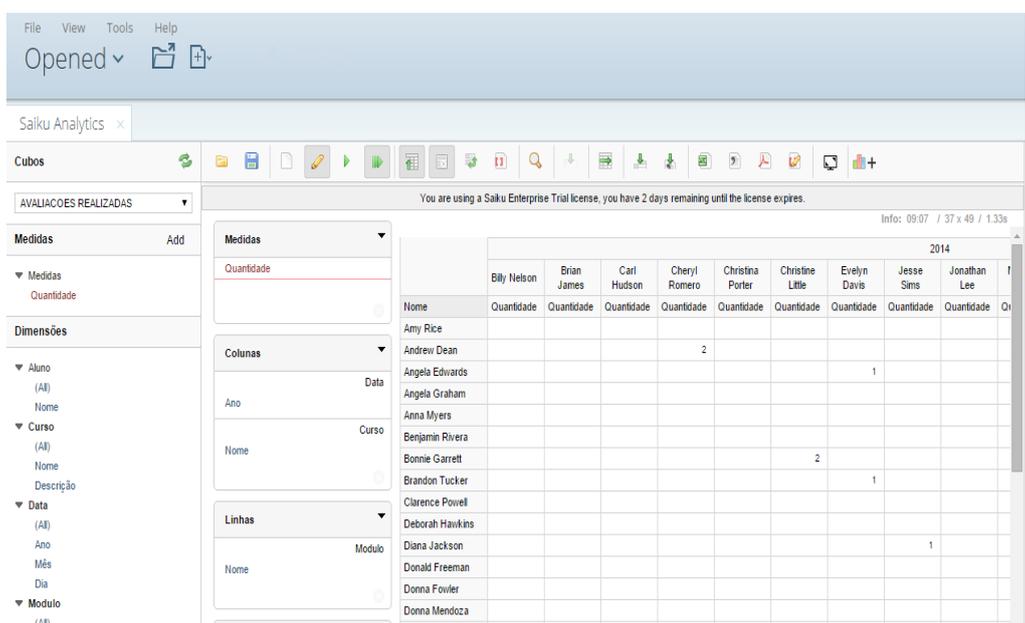


Figura 4. Novo Relatório Ad-hoc no Bi-Server do cubo Avaliações Realizadas
Fonte: Elaborada pelos autores

O procedimento visto acima aplica-se também aos outros cubos gerados pela ferramenta.

6.1. Análise Dashboard Avaliações Realizadas

O *Bi-Server* disponibiliza formas de gerar *dashboards* customizados com os cubos OLAP do projeto, tem também a opção de vários tipos de gráficos, e no caso dos mesmos não atenderem a necessidade, existe a possibilidade de adicionar novos da comunidade ou desenvolver novos.

Neste projeto criado uma *dashboard*, figura 5 abaixo, usando os gráficos pré-definidos. Para tanto, foi inserida em cada uma as consultas OLAP do exemplo a seguir. Se for desejado mais consultas, basta criar outro gráfico e inserir novas consultas.

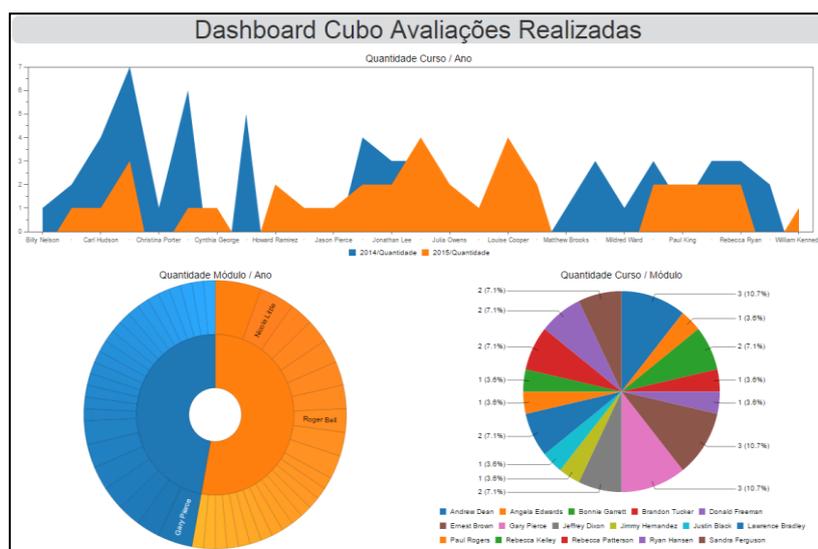


Figura 5. Dashboard no Bi-Server do Projeto
Fonte: Elaborado pelo autores

Foi realizado nesse projeto uma *dashboard* a partir do Cubo Avaliações Realizadas. Nele foram colocadas as seguintes consultas: quantidade curso por ano, quantidade de módulo por ano e quantidade de curso por módulo e ano. Na Figura 5, a são apresentados o *dashboard* desenvolvido.

Essas *dashboards* fornecerão uma visão abrangente e visual das medidas (indicadores-chave de desempenho), como também tendências e exceções de desempenho.

7. Conclusão

A implantação da tecnologia de *Business Intelligence* na ferramenta Amadeus irá apoiar nas decisões estratégicas para os alunos a partir da interpretação das informações armazenadas e sem uso, agregando valor a informação gerada pelas instituições.

Ao usar o sistema de BI o usuário do Amadeus terá mais uma opção de realizar consultas de forma independente, flexível e intuitiva, a informação será apresentada para ele de forma mais compreensível, o que pode trazer uma segurança na análise do negócio e também com os indicadores de aprendizagem dos alunos.

A abordagem em *data marts* proposto por Kimball (2002) irá permitir que o projeto cresça de forma exponencial, novas FATOS podem ser sugeridos por gestores

mais experientes das instituições, a consequência disso é o envolvimento de todos na instituição com o intuito de melhorar cada vez mais o BI.

O BI permite uma agilidade nas consultas, além de ajudar nas decisões. Também contribuirá para as outras áreas da instituição, o que dará subsídios para um melhor planejamento de toda a organização.

Dependendo com o crescimento do projeto, o BI pode ser integrado a outros sistemas, como banco de dados, sistemas legados, outros sistemas de BI, sistemas empresariais (como ERP e CRM), comércio eletrônico, e muitos outros.

O projeto foi compartilhado no serviço de *web hosting* de compartilhamento de projeto, conhecido como *GITHUB*, permitindo o acesso livre do mesmo através do endereço: <https://github.com/celofilho/amadeuspentaho>.

Conclui-se que o Amadeus ganhou uma ferramenta analítica que possibilita consultas mais rápidas, flexíveis, e seguras; os seus insumos darão mais subsídios para o apoio à decisão com uma agilidade e também autonomia para realizar as consultas, o que torna crucial o uso do Amadeus aliado a um BI.

7.1. Trabalhos Futuros

Como foi desenvolvida uma ferramenta de Business Intelligence e disponibilizada para a comunidade, esta ferramenta pode ser adaptada para outros ambientes EAD, uma vez que todos os artefatos do projeto e a forma de construção foram disponibilizados a comunidade.

Referências

- Berson, Alex; Smith, Stephen J. Data warehousing, data mining, and OLAP. McGraw-Hill, Inc., 1997.
- Borges do Val, Ronaldo et al. Visualização de dados aplicados em educação à distância no processo de avaliação ao aluno. 2011
- Cavaco, João Manuel de Leça Pereira. Estágio profissional em Business Intelligence na ANA aeroportos. 2013.
- Costa, Sérgio; Santos, M. Sistema de Business Intelligence no suporte à Gestão Estratégica. In: Conferência da Associação Portuguesa de Sistemas de Informação (CAPSI). 2012. P. 1-10.
- De Souza Fernandes, Mikael; Kantorski, Gustavo Zanini. Uma Ferramenta Web para BI focada no Gestor de Informação, 2012.
- Gomes, Alex Sandro et al. Instalação, Configuração e Uso da Plataforma de Gestão de Aprendizagem Amadeus. 2011.
- Inmon, William H. Como Construir o Data Warehouse. Rio de Janeiro: Campus, 1997.
- Inmon, William H. UNSTRUCTURED APPLICATIONS: UNLOCKING THE POTENTIAL. Disponível em:<<http://www.inmoncif.com/registration/whitepapers/unlockingunstructuredapps.pdf>> Acesso em 16 set. 2014.

- Janus, Philo; Misner, Stacia. Building integrated business intelligence solutions with SQL Server 2008 R2 & Office 2010. McGraw-Hill, 2011.
- Kimball, Ralph. The data warehouse toolkit: Guia complete para modelagem dimensional. Rio de Janeiro: Campus, 2002.
- Machodo, Felipe Nery Rodrigues. Tecnologia E Projeto De Data Warehouse. 4ª. Edição. São Paulo: Editora Érica, 2008.
- Matheus, Renato F.; Parreiras, Fernando S. Inteligência empresarial versus Business Intelligence: abordagens complementares para o apoio à tomada de decisão no Brasil. In: KMBRASIL 2004-Congresso Anual da Sociedade Brasileira de Gestão do Conhecimento. 2004.
- Mockaroo. Gerador de Dados. 2015. Disponível em: <www.mockaroo.com>. Acesso em: 21 maio 2015.
- Pentaho. Open Source Business Intelligence. 2010. Disponível em: <<http://www.Pentaho.com>>. Acesso em: 09 abr. 2015.
- SURVEY. Formulário Online. 2015. Disponível em: <www.survey.com>. Acesso em: 21 maio 2015.
- Queiroz, Douglas. Desenvolvimento de um componente de visualização de informação para a plataforma de ensino a distancia Amadeus. 2010. 43 f. TCC (Graduação) - Curso de Ciências da Computação, CIn UFPE, Recife, 2010.
- Santos, Ricardo S. et al. Projeto de data warehouse para a saúde pública. In: Anais do IX Congresso Brasileiro de Informática em Saúde. 2004. p. 131-136.
- Watson, H. J. Dashboards and Scorecards. Business Intelligence Journal, v. 11, n. 2, p. 4, 2006.