



**UNIVERSIDADE DO ESTADO DA BAHIA
DEPARTAMENTO DE EDUCAÇÃO - CAMPUS I
PROGRAMA DE PÓS-GRADUAÇÃO GESTÃO E TECNOLOGIAS
APLICADAS À EDUCAÇÃO – GESTEC**

BRUNO CÉSAR PEREIRA LEITE

**UNEB.I.: UMA PROPOSTA DE OTIMIZAÇÃO NO
LEVANTAMENTO DE DADOS PARA O ANUÁRIO
“UNEB EM DADOS” A PARTIR DA UTILIZAÇÃO DE
TÉCNICAS DE *BUSINESS INTELLIGENCE***

**SALVADOR
2016**

BRUNO CÉSAR PEREIRA LEITE

**UNEB.I.: UMA PROPOSTA DE OTIMIZAÇÃO NO
LEVANTAMENTO DE DADOS PARA O ANUÁRIO
“UNEB EM DADOS” A PARTIR DA UTILIZAÇÃO DE
TÉCNICAS DE *BUSINESS INTELLIGENCE***

Dissertação apresentada ao Programa De Pós-Graduação
Gestão e Tecnologias Aplicadas à Educação – GESTEC,
Departamento de Educação, Universidade do Estado da
Bahia, como requisito parcial para obtenção do grau de Mestre
em Gestão e Tecnologias Aplicadas à Educação.

ORIENTADOR: PROF DR MARCUS TÚLIO DE FREITAS PINHEIRO

**SALVADOR
2016**

L533

Leite, Bruno César Pereira

UNEB.I: uma proposta de otimização no levantamento de dados para o anuário “UNEB em dados” a partir da utilização de técnicas de Business Intelligence/ Bruno César Pereira Leite.

86f. il.

Orientador: Prof. Dr. Marcus Túlio de Freitas Pinheiro

Dissertação (mestrado) - Universidade do Estado da Bahia.

Departamento de Educação. Programa de Pós-Graduação Gestão e Tecnologias Aplicadas à Educação, Salvador, 2016.

1. Inteligência Competitiva (Administração). 2. Banco de Dados. 3. Data Warehouse. I. Pinheiro, Marcus Túlio de Freitas. II. Universidade do Estado da Bahia. III. Título.


CDD: 658.4038

FOLHA DE APROVAÇÃO

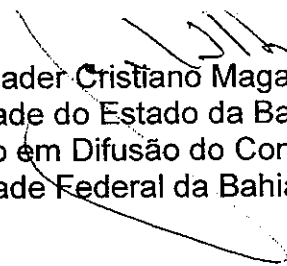
“UNEB. I: UMA PROPOSTA DE OTIMIZAÇÃO NO LEVANTAMENTO DE DADOS PARA O ANUÁRIO “UNEB EM DADOS” A PARTIR DA UTILIZAÇÃO DE TÉCNICAS DE BUSINESS INTELLIGENCE”

BRUNO CÉSAR PEREIRA LEITE


Trabalho de Conclusão de Curso apresentado ao Programa de Pós-Graduação (*Stricto Sensu*) Gestão e Tecnologias Aplicadas à Educação, Área de Concentração II – Processos Tecnológicos e Redes Sociais, em 16 de dezembro de 2016, como requisito parcial para obtenção do grau de Mestre em Gestão e Tecnologias Aplicadas à Educação, pela Universidade do Estado da Bahia, composta pela Banca Examinadora:



Prof. Dr. Marcus Túlio de Freitas Pinheiro
Universidade do Estado da Bahia - UNEB
Doutorado em Educação
Universidade Federal da Bahia - UFBA



Prof. Dr. Jader Cristiano Magalhães de Albuquerque
Universidade do Estado da Bahia - UNEB
Doutorado em Difusão do Conhecimento
Universidade Federal da Bahia - UFBA



Prof. Dr. Ramon Oliveira Vidal
German Center for Neurodegenerative Diseases - DZNE/Alemanha
Doutorado em Genética e Biologia Molecular
Universidade Estadual de Campinas - UNICAMP

Às mulheres da minha vida.

Minha mãe Lizete Brício, minha irmã
Vanessa Leite, minha esposa Joanne
Leite, minha avó Maria Celeste e minhas
tias Liana Brício e Liamara Brício. Amo
vocês!

AGRADECIMENTOS

A toda minha família por todo amor, carinho, formação ética e moral e pela importância que sempre deram à educação.

À minha esposa pelo amor e apoio tão importantes no decorrer da pesquisa.

Ao meu orientador Marcus Túlio de Freitas Pinheiro, pelas intervenções sempre assertivas. Por ter me dado liberdade quando possível e limites quando necessário.

Ao grupo DCETM pelas participações e contribuições nos encontros semanais das orientações coletivas que ajudaram a moldar o trabalho desde a fase de anteprojeto até o formato final.

Ao professor Ivan Novaes que à frente da Secretaria de Avaliação Institucional e através do Programa de Reorganização e Reorientação de Procedimentos Acadêmicos ajudou de forma vital nessa pesquisa, tanto nas orientações relacionadas à metodologia de pesquisa científica como dando norte às funcionalidades da ferramenta.

À equipe da SEAVI e SGC da UNEB que viabilizaram o presente trabalho através dos encontros onde explicavam os significados dos dados oriundos do Sistema SAGRES.

Aos professores do Programa do Mestrado em Gestão e Tecnologia Aplicadas à Educação da Universidade do Estado da Bahia.

Aos sempre solícitos funcionários do Programa do Mestrado em Gestão e Tecnologia Aplicadas à Educação da Universidade do Estado da Bahia.

À Universidade do Estado da Bahia pela oportunidade da realização do curso de mestrado.

Aos colegas de trabalho da equipe UDO/GERINF cujo empenho, compromisso, responsabilidade e companheirismo, permitiram que eu pudesse me ausentar nos momentos de aula e reuniões de grupo de pesquisa sem que os serviços de TIC

deixassem de ser prestados à comunidade Unebiana. Vocês todos foram fundamentais para que pudesse conciliar as atividades do trabalho à pesquisa. Apesar de agradecer a todos, preciso destacar alguns desses. Pedro Herrera que me influenciou diretamente na realização da seleção para o programa, deu a ideia inicial para anteprojeto e também auxiliou em várias etapas da pesquisa. Iury Barreto, que foi também colega de curso e não apenas participou de alguns trabalhos em equipe, como fez sugestões importantes para a melhoria do projeto e foi um grande revisor durante toda a jornada. Flávio Farias que atuou como DBA da GERINF no início do projeto e deu dicas importantes no pontapé inicial do trabalho. Izaak Candeias e Daniel Pelegrini que além de grandes amigos pessoais, foram peças fundamentais para que o trabalho fosse concluído. Daniel entrou no meio do projeto como o novo DBA da GERINF e me surpreendeu com a velocidade com a qual se empolgou e abraçou a causa, trazendo sempre novas sugestões e gerando discussões valiosas para o aprimoramento da ferramenta. Izaak, ninja do Excel, com sua ajuda na ferramenta adotada para visualização dos dados possibilitou que fizéssemos alguns ajustes vitais na reta final. E Benjamin Ramos, mais do que um chefe de trabalho, um líder e amigo que acreditou em meu potencial e viabilizou a construção deste projeto.

Aos amigos que conviveram com as angústias, alegrias e expectativas. Alguns próximos fisicamente e outros já morando em outro estado ou país, mas todos sempre acreditando e dando forças para que eu pudesse chegar até aqui. Mesmo com medo de esquecer alguém, não posso deixar de citar César Lawinsky, Daniel Pelegrini, Tâmara Sleyne, Antônio Júnior, Izaak Candeias, Valdemir Junior, Matheus Thomaz e Ramon Vidal, amigo que admiro demais e que me inspirou a seguir na carreira científica.

A todos que direta ou indiretamente contribuíram para a realização desta pesquisa.

Muito obrigado a todos vocês!

RESUMO

A presente pesquisa tem como objetivo aperfeiçoar o processo de levantamento dos dados para elaboração do Anuário UNEB em DADOS a partir do desenvolvimento do protótipo de uma ferramenta de Business Intelligence. Para um melhor entendimento dos conceitos utilizados no desenvolvimento do projeto, alguns autores foram consultados como Almeri Finger e Takeshy Tachizawa (Gestão de Instituições de ensino Superior), Marta Valentim, Nonaka e Takeuchi (Gestão da Informação e do conhecimento) e Ralph Kimball (Data Warehouse e Business Intelligence).

No intuito de verificar as dificuldades e oportunidades intrínsecas do referido processo, particularmente numa instituição de Ensino superior pública, e, possivelmente, auxiliar interessados no tema, descrevo todas as etapas, desde as entrevistas e reuniões com as partes interessadas no projeto, até o resultado final alcançado que são os relatórios propriamente ditos, passando pelo desenvolvimento da ferramenta em si (levantamento de requisitos, modelagem e implementação). Após implantação e testes, a ferramenta desenvolvida demonstrou possibilidades além das previstas, como geração de relatórios dinâmicos e a possibilidade de análise histórica, uma vez que pode acessar dados relativos, não apenas aos anos de publicação do Anuário, mas podendo retroagir até o início do uso do sistema de informações acadêmicas da instituição, o SAGRES. Futuramente, o projeto poderá ser expandido para obtenção de informações de outras categorias que compõem a comunidade acadêmica da universidade (docentes e técnicos) e possivelmente, outros métodos de acesso às informações não apenas ao grupo gestor, mas para toda a comunidade interessada, auxiliando a UNEB no conhecimento sobre seus dados institucionais e no atendimento pleno à lei de acesso a informação.

Palavras-chave: Inteligência de Negócios; Business Intelligence; Bancos de Dados; Data Warehouse; Data Mart; Gestão Universitária.

ABSTRACT

The present research aims to improve the process of data collection for the elaboration of the UNEB em Dados yearbook from the development of a Business Intelligence tool prototype. For a better understanding of the concepts used in the project development, some authors were consulted, such as Almeri Finger and Takeshy Tachizawa (Management of Higher Education Institutions), Marta Valentim, Nonaka and Takeuchi (Information and Knowledge Management) and Ralph Kimball (Data Warehouse and Business Intelligence).

In order to verify the difficulties and intrinsic opportunities of this process, particularly in a public higher education institution, and possibly help researchers, I describe all the steps, from the interviews and meetings with the stakeholders, until the achieved result, which are the reports themselves, through the development of the tool itself (requirements survey, modeling and implementation). After development and tests, the developed tool demonstrated possibilities beyond what was foreseen, such as dynamic reporting and the possibility of historical analysis, since it can access data related not only to the published yearbook timespan but being able to retroact until the adoption of SAGRES, the university academic information software. In the future, the project may be expanded to obtain information from other categories that make up the university's academic community (faculty and staff members) and possibly other methods of accessing information not only to the management group, but to the entire interested community, helping UNEB to understand its own data and in full compliance with the free information access law.

Keywords: Business Intelligence; Databases; Data Warehouse; Data Mart; University Management.

LISTA DE ABREVIATURAS E SIGLAS

- BI – Business Intelligence
- DBA – Administrador de Banco de Dados (Do inglês: Database Administrator)
- DM – Data Mart
- DW – Data Warehouse
- ETL – Extração, Transformação e Carga (Do inglês: Extract, Transform and Load)
- GERINF – Gerência de Informática da UNEB
- IES – Instituição de Ensino Superior
- INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
- KPI – Key Performance Indicators
- OLTP – Online Transaction Processing
- OLAP – Online Analytical Processing
- SEAVI – Secretaria de Avaliação Institucional
- SGC – Secretaria Geral de Cursos
- SIG – Sistemas de Informações Gerenciais
- SQL – Structured Query Language, ou Linguagem de Consulta Estruturada. É a linguagem de pesquisa declarativa padrão para banco de dados relacional.
- TI – Tecnologia da Informação
- TIC – Tecnologia da Informação e Comunicação
- UDO – Unidade de Desenvolvimento Organizacional da UNEB
- UNEB – Universidade do Estado da Bahia

LISTA DE TABELAS

Tabela 1: dicionário de dados do Cubo OLAP desenvolvido	71
Tabela 2: Quadro resumo das etapas de desenvolvimento do protótipo	80

LISTA DE FIGURAS

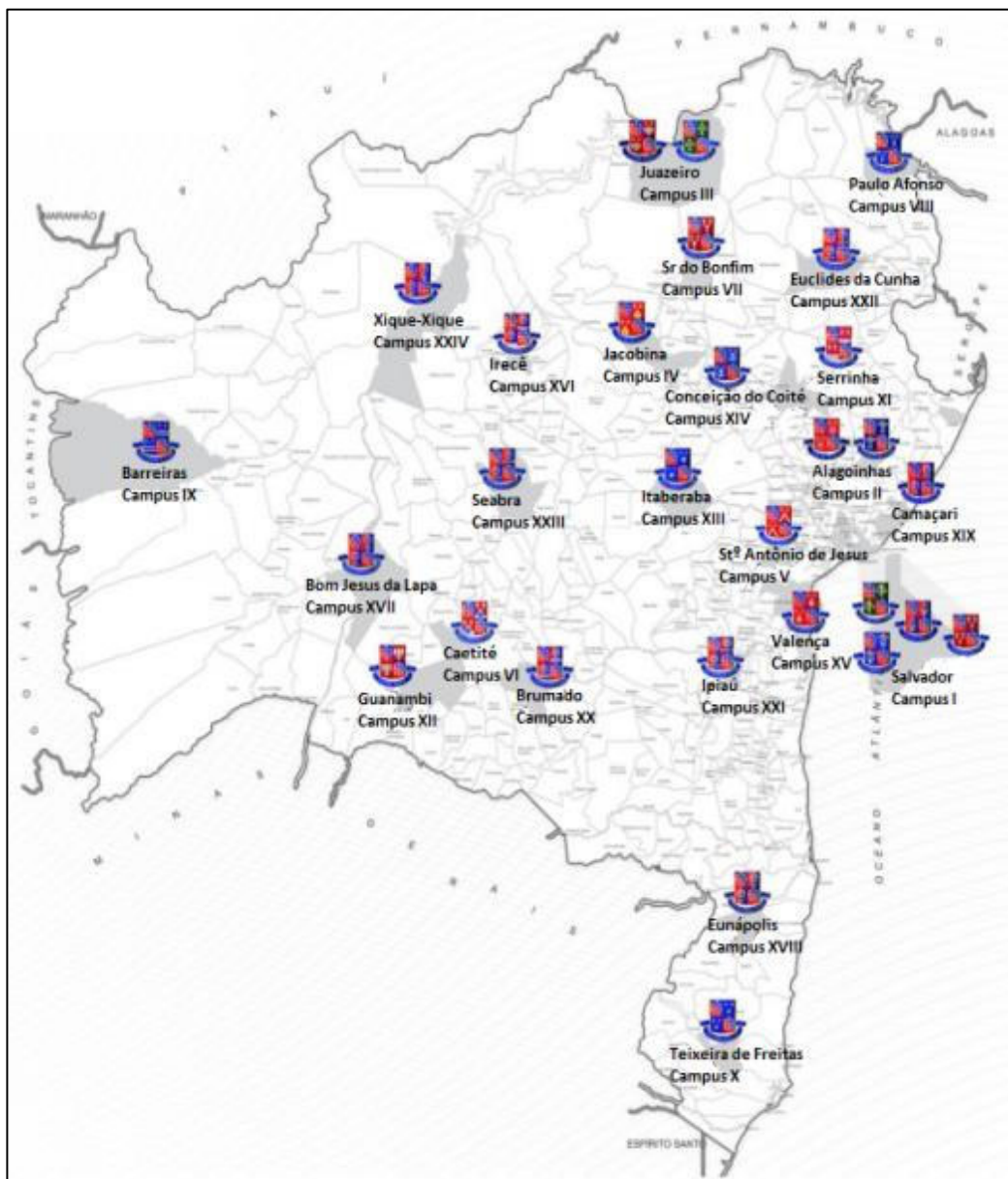
Figura 1: Localização dos Campi da UNEB	11
Figura 2: Exemplo de integração das aplicações e bases de dados na UNEB	15
Figura 3: Fluxo básico de uma IES	25
Figura 4: Fundamentos de um BI	33
Figura 5: abstração do conceito de BI	33
Figura 6: Modelo estrela	35
Figura 7: Modelo floco de neve	36
Figura 8: Ciclo de construção do DW	37
Figura 9: Fluxo de dados/informações na SEAVI	43
Figura 10: Capa do Anuário UNEB em Dados 2015 - Base 2014	45
Figura 11: Exemplo de dados Cursos Graduação do Anuário	46
Figura 12: arquitetura básica do BI (protótipo)	47
Figura 13: Diagrama do modelo de dados sobre os candidatos do vestibular da UNEB	50
Figura 14: exemplo de relatório dinâmico no Excel	52
Figura 15: parte da matriz necessidade preliminar	53
Figura 16: resumo do desenvolvimento do protótipo	55
Figura 17: primeiro rascunho da modelagem do DW	56
Figura 18: primeira versão do modelo do Data Mart	57
Figura 19: desenho do fluxo de dados (Data Flow) da Dimensão Curso no SSIS	58
Figura 20: caixa de ferramentas do SSIS	60
Figura 21: abstração da hierarquia definida no protótipo	62
Figura 22: exemplo de relatório dinâmico extraído na ferramenta	63
Figura 23: exemplo de relatório de escolaridade da primeira iteração do protótipo	64
Figura 24: exemplo de relatório de período letivo da primeira iteração do protótipo	65
Figura 25: fluxo de dados da tabela fato extraído do SSIS	67
Figura 26: modelo de dados final do DW extraído do Sql Server 2012	68
Figura 27: modelo de dados final do cubo extraído do SSAS	69
Figura 28: parte do anuário destacando divergência de ano de criação do curso de agronomia	72
Figura 29: parte da planilha elaborada através da ferramenta desenvolvida destacando divergência de ano de criação do curso de agronomia	72
Figura 31: Cursos de Gradação Regular Presencial, DTCS-III	74
Figura 30: Cursos de Gradação Regular Presencial, DTCS-III na ferramenta (visualização padrão)	74
Figura 32: Cursos de Gradação Regular Presencial, DTCS-III na ferramenta (visualização clássica)	75
Figura 33: gráfico de evolução das matrículas dos Cursos de Gradação Regular Presencial, DTCS-III na ferramenta	78

SUMÁRIO

1. INTRODUÇÃO	11
1.1. OBJETIVO GERAL	21
1.2. OBJETIVOS ESPECÍFICOS	21
2. REFERENCIAL TEÓRICO	22
2.1. GESTÃO DE INSTITUIÇÕES DE ENSINO SUPERIOR	22
2.2. GESTÃO DA INFORMAÇÃO E DO CONHECIMENTO	26
2.3. BUSINESS INTELLIGENCE: UMA METODOLOGIA DE APOIO A DECISÃO	30
3. METODOLOGIA E LOCUS DA PESQUISA	40
3.1. CARACTERIZAÇÃO DA PESQUISA	40
3.2. LOCUS DA PESQUISA	42
4. DESENVOLVIMENTO	48
4.1. LEVANTANDO REQUISITOS	48
4.2. DESENVOLVENDO A FERRAMENTA	55
4.2.1. DESENHANDO E POPULANDO O DATA WAREHOUSE	55
4.2.2. CRIANDO, VISUALISANDO E AJUSTANDO O CUBO	60
4.2.3. FINALIZANDO O PROTÓTIPO	65
4.3. ANÁLISE DOS RESULTADOS	71
5. CONSIDERAÇÕES FINAIS	81
6. REFERÊNCIAS BIBLIOGRÁFICAS	83

1. INTRODUÇÃO

A UNEB é a maior Instituição de Ensino Superior (IES) da Bahia e está geograficamente presente em praticamente todo o estado. Além de possuir 29 departamentos distribuídos em 24 campi, sendo que 25 desses departamentos estão espalhados em 23 municípios do interior do estado e outros 4 no campus da capital - conforme pode ser visto na Figura 1 - a UNEB ainda se espalha por todo o estado através de ações extensionistas que executa por meio de convênios com organizações públicas e privadas que alcançam a quase a totalidade dos 417 municípios baianos, beneficiando principalmente cidadãos pertencentes a segmentos social e economicamente desfavorecidos (www.uneb.br – 2016).



Fonte: Projeto de Reconhecimento do Curso de Administração - 2011

Figura 1: Localização dos Campi da UNEB

Atuando desde o ano de 2014 como Gerente de Tecnologia da Informação (TI) na Gerencia de Informática (GERINF), setor que está ligado à Unidade de Desenvolvimento Organizacional (UDO) da Universidade do Estado da Bahia (UNEB), acompanho e direciono as atividades de uma equipe responsável pelo desenvolvimento de sistemas administrativos para auxiliar os diversos setores da instituição no desenvolvimento e acompanhamento de suas atividades. Nessa atividade pude perceber um aumento na demanda por novos sistemas por parte das diversas áreas da Administração Central nos últimos anos. E em todas as reuniões iniciais, buscando entender a ideia central do sistema e o motivo pelo qual o mesmo se faz necessário, sempre ouvimos sobre a necessidade por sistematizar os dados e responder rapidamente às perguntas que são feitas por outros setores.

Ciente do papel social da UNEB e da capilaridade inerente à sua multicampia foi natural a percepção da importância que sistemas automatizados têm para aumentar a confiabilidade das informações passadas para os gestores no momento de uma tomada de decisão, seja para a criação de novos cursos de graduação e pós-graduação - uma das propostas do projeto de gestão do atual reitor é, por exemplo, *“Definir áreas prioritárias para proposição e implantação de Programas de Pós-Graduação em conformidade com as especificidades da multicampia da UNEB, das RGDs e das demandas locais/regionais”* - ou definir onde investir os escassos recursos disponíveis.

Com o aumento do número de sistemas desenvolvidos, é importante que seja feito um trabalho de rastreio das aplicações legadas para evitar a criação de bases de dados que gerem redundância de informação ou, pior, inconsistência das mesmas. Ciente desse cenário, a Gerência de Informática da UNEB – GERINF, que possuía em seu portfólio até o ano de 2015, 34 sistemas, sendo que destes, 8 ainda encontravam-se em fase de desenvolvimento, vem envidando esforços para a integração entre suas diversas aplicações e desenvolveu um total de nove

Webservices¹ para evitar a redundância e inconsistência de dados nas bases da instituição.

A UNEB vem se aperfeiçoando e buscando o constante aprimoramento no desenvolvimento de seus sistemas informatizados. Podemos citar como exemplo o SIGP² (Sistema Integrado de Gestão de Pessoas), um sistema informatizado que através de uma carga diária de dados funcionais dos servidores (docentes, técnicos e analistas, cargos temporários e contratados pelo Regime Especial de Direito Administrativo - REDA) e da folha de pagamento dos funcionários da UNEB enviada pela SAEB (Secretaria de Administração do Estado da Bahia), permite obter uma visão gerencial em relação à gestão de pessoas da Universidade, como por exemplo: quantidade de docentes, analistas, técnicos, custos de folha de pagamento, bem como outras informações. Além disso, todas as portarias são também importadas do Sistema de Publicações On-Line – SPO, outro sistema desenvolvido na Universidade, o que permite uma visão mais detalhada do corpo de servidores da instituição para a Pró-Reitoria de Gestão de Pessoas.

Todos os dados funcionais e de folha de pagamento fornecidos pela SAEB são oriundos do Sistema Integrado de Recursos Humanos – SIRH, sistema informatizado utilizado por todas as secretarias do governo do Estado da Bahia. O SIRH é um sistema informatizado desenvolvido pela Prodeb (Companhia de Processamento de Dados da Bahia) para³:

[...] suprir a carência de informações confiáveis referentes ao quadro de servidores e empregados, que era um entrave histórico para gestão da política de recursos humanos no estado. Bancos de dados isolados, sem integração, registros inconsistentes, sem padronização, impediam a Secretaria de Administração, realizar as atividades de planejamento, coordenação, supervisão, controle e desenvolvimento de recursos humanos da administração pública estadual do Poder Executivo.

Tendo a base de dados do SIRH relacionadas aos servidores da UNEB sob o domínio da Universidade, ela serve como consulta para outros sistemas, permitindo assim uma integração mais efetiva e evitando ilhas de informação,

¹ Solução utilizada na integração de sistemas e na comunicação entre aplicações diferentes. Com esta tecnologia é possível que novas aplicações possam interagir trocando dados com aquelas que já existem e que sistemas desenvolvidos em plataformas diferentes sejam compatíveis.

² Disponível em <<http://www.sigp.uneb.br>> Acesso em 21 de novembro de 2014

³ Disponível em <<http://www.portalprodeb.ba.gov.br/sirh.asp>> Acesso em 21 de novembro de 2014

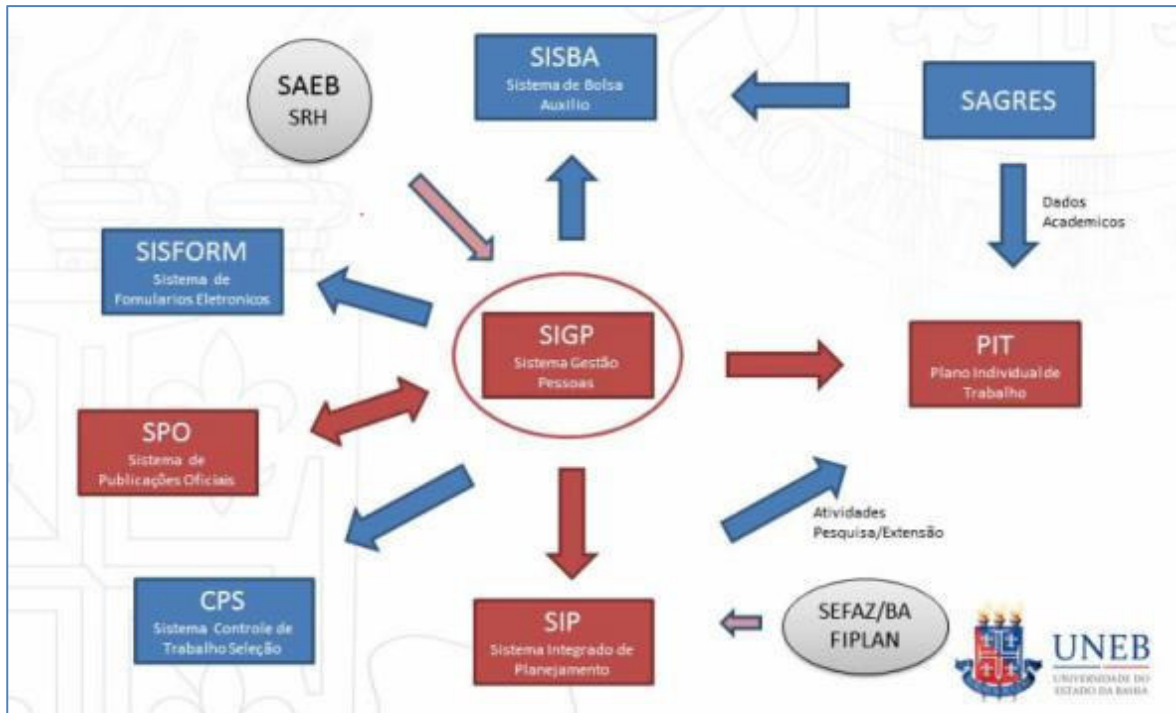
problema comum a várias organizações. Com isso, apenas os dados que são específicos da UNEB são inseridos como complemento, evitando redundância de dados e retrabalho da equipe operacional.

Atualmente, a UNEB utiliza o sistema informatizado Sagres como software para acompanhar a vida acadêmica dos estudantes na instituição. O Sagres, já disponível na sua versão 2.0, foi desenvolvido pela TecnoTrends e é descrito pela própria empresa como⁴:

[...] um produto resultante de mais de 15 anos de experiência e conhecimento na área de administração acadêmica para ensino fundamental, médio e superior. Simples e flexível, é capaz de adequar-se a instituições de qualquer porte, desde uma escola ou uma pequena faculdade, até uma universidade multi-campi.

O Sagres realiza pré-matrícula, matrícula por turma e matrícula via WEB. Além disso, os alunos acessam assuntos de seu interesse como emissão de atestados e históricos escolares, os professores inserem notas e frequência de suas classes, enquanto gestores podem obter alguns dados estatísticos sobre assuntos pertinentes ao alunado da instituição. Seguindo o princípio do uso de webservices para evitar inconsistências de informações entre os sistemas, a GERINF vem integrando as diversas bases de dados de suas aplicações, não apenas entre as que são desenvolvidas localmente, mas também com a de sistemas de terceiros, como é o caso do SAGRES, o SIRH e também o FIPLAN, sistema de gestão orçamentária e financeira da Secretaria da Fazenda do Estado da BAHIA – SEFAZ. Essa integração pode ser compreendida mais facilmente através da Figura 2.

⁴ Disponível em < <http://www.tecotrends.com.br/novoportal/Modules/Extranet/instituicao.aspx> > Acesso em 21 de novembro de 2014



Fonte: RAMIREZ, 2016, p. 20.

Figura 2: Exemplo de integração das aplicações e bases de dados na UNEB

Desde 2011 a Universidade do Estado da Bahia – UNEB vem sendo a responsável pelo processo de inscrição e publicação dos resultados do seu vestibular. A cada seleção realizada, o volume de dados armazenados dos candidatos vem aumentando significativamente e os relatórios solicitados pela Comissão Permanente de Seleção da UNEB – CPS (o setor responsável pela seleção) dizem respeito apenas ao processo seletivo em si. Para apoiar a CPS, a GERINF desenvolveu o SIV⁵ (Sistema Integrado de Vestibular) e o Inscrição Vestibular, dois sistemas que, em conjunto, gerenciam todo o processo de inscrição do vestibular, envolvendo desde o formulário eletrônico para inscrição dos candidatos, passando pelo cadastro dos cursos e suas respectivas vagas ofertadas, até a homologação de todos os candidatos que tiveram suas inscrições deferidas. A emissão do boleto e confirmação do pagamento também é feita pelo próprio sistema, assim como o pré-cadastro dos candidatos à isenção do pagamento da taxa de inscrição.

Com isso, informações como “Quantitativo de inscritos”, “Inscrições por curso”, “Inscritos por local de prova” são alguns exemplos dos relatórios que o SIV

⁵ Disponível em <<http://www.siv.uneb.br/>> Acesso em 21 de novembro de 2014

fornece para os gestores da CPS. Esses relatórios são úteis para que a Universidade saiba qual a demanda que existe para os cursos que oferta nas diversas regiões do estado, mas funcionam apenas como um “retrato” do processo seletivo de um ano específico. Apenas alguns relatórios vão um pouco além e fornecem um comparativo do total de inscritos por curso entre o ano corrente e o anterior.

A partir de algumas reuniões com o coordenador da equipe de desenvolvimento de sistemas da GERINF, surgiu a ideia de utilizar essa noção de integração que vem sendo adotada no setor para cruzar diferentes bases com informações do corpo discente da UNEB, de modo a obter informações úteis para a alta gestão universitária. Logo essa conversa se estendeu e alcançou o administrador das bases de dados que passou a incorporar conceitos de Data Warehouse para o corpo do projeto: um depósito de dados digitais para armazenar informações detalhadas, criando e organizando relatórios através de históricos que são depois usados para ajudar a tomar decisões importantes.

Nasce então a ideia de se desenvolver uma ferramenta de Business Intelligence (ou *BI*, termo que refere-se ao processo de coleta, organização, análise, compartilhamento e monitoramento de informações que oferecem suporte a gestão de negócios de uma organização) para a UNEB que possa cruzar todas as bases existentes da instituição e fornecer informações preciosas sobre as diversas áreas e públicos atendidos pela universidade, desde o corpo docente e discente, passando pela equipe de técnicos e analistas da área administrativa até a infraestrutura da própria UNEB.

Ciente da impossibilidade de se realizar uma pesquisa completa para a implantação de uma ferramenta desse porte no espaço de tempo disponível do mestrado, um recorte foi feito para demonstrar o potencial que esse tipo de ferramenta possui. O recorte em questão contempla apenas o cruzamento de informações relativas ao alunado, coletando, sumarizando e consolidando dados do sistema de informações acadêmicas da UNEB (SAGRES). A escolha desse recorte se dá por entender que o corpo discente é a razão de existir da Universidade e, portanto, deve ser o primeiro segmento a ser analisado.

Em 2014, ano de criação deste projeto, a UNEB possuía 104 cursos de graduação presenciais espalhados em vinte e nove departamentos, que por sua vez eram organizados em 24 cidades da Bahia. É um desafio constante para a UNEB manter a demanda pelos cursos existentes e propor novos cursos que se adaptem e contribuam para o desenvolvimento das diversas regiões do estado. Para isso, é preciso que a instituição conheça melhor seu público e é nesse ponto que uma ferramenta de BI poderá fornecer informações valiosas para ajudar na tomada de decisões dos gestores da universidade.

Apesar do esforço já em andamento na UNEB para a migração de informações entre seus sistemas, não existe hoje um cruzamento entre as informações existentes no Sistema de Inscrição do Vestibular e os dados do Sagres. Apenas esse cruzamento já permitiria extrair algumas informações do questionário socioeconômico com os acadêmicos e gerar relatórios relevantes para os gestores da universidade. Apenas para citar alguns exemplos, é possível responder perguntas como: qual a cidade que mais “envia” candidatos para cursos oferecidos apenas na capital? Qual a cidade com maior número de candidatos cotistas? Sem falar que é possível categorizar por curso de modo a entender melhor a demanda dos candidatos.

Se pensarmos que esse cruzamento de informações é o início para um futuro relacionamento com a base de dados dos resultados das provas realizadas pela empresa Consultec⁶, isso possibilita uma análise mais precisa e aprofundada do perfil dos alunos ingressantes na Universidade desde o ano de 2011, ano em que a UNEB iniciou o gerenciamento das inscrições para seu processo seletivo.

A UNEB possui diversas bases com potencial para cruzamento de informações para gerar indicadores sobre candidatos, técnicos, discentes e docentes. No entanto, ainda há muito para coletar, organizar e relacionar os dados das diferentes aplicações.

Não bastasse o trabalho de ordem técnica - que diz respeito aos analistas e desenvolvedores da GERINF, há uma necessidade premente de se (re)pensar a

⁶ Empresa responsável pela elaboração e processamento do resultado das provas do vestibular da UNEB de 2011 até o momento de elaboração deste projeto.

gestão do conhecimento da UNEB sobre a UNEB. Conforme dito anteriormente, as universidades vivem um momento de alta competitividade e as instituições públicas, pelo seu papel social, precisam se manter atraentes (competitivas) de modo a atenderem melhor a sociedade. Para que essa gestão ocorra, uma etapa anterior é imprescindível: a criação do conhecimento sobre a organização. E aqui faço referência ao conceito dado por Nonaka e Takeuchi (1998), que tratam o conhecimento como a unidade básica de análise para explicar o comportamento da organização.

Conforme aponta Weber (2014), o Censo 2013 demonstrou que, apesar de o número de matrículas no ensino superior continuar aumentando, o número de concluintes caiu cerca de 5,7%. Além disso, o mesmo censo revelou que 73,5% dos estudantes de graduação frequentam instituições particulares. É neste cenário que as universidades públicas como a UNEB precisam trabalhar para identificar oportunidades para não apenas aumentar a oferta de vagas, mas garantir que essas vagas sejam preenchidas e seu corpo discente conclua seus respectivos cursos.

Segundo Pinheiro (2002), “O conhecimento em sua variedade de formas vem a cada dia assumindo um papel determinante nos processos competitivos das empresas”. Mais uma vez, no atual cenário econômico-social, uma universidade, ainda que pública (e talvez, principalmente as públicas), precisa ser cada vez mais competitiva. Ou seja, é preciso gerar e gerir melhor seu conhecimento.

No ano de 2014, a Secretaria de Avaliação Institucional – SEAVI, coordenada pelo Prof. Ivan Luiz Novaes⁷, iniciou o planejamento para a publicação do Anuário UNEB em DADOS 2015, ano base 2014, como parte do Programa de Reorganização e Reorientação dos Procedimentos Acadêmicos, cuja origem decorre da articulação entre a Secretaria Especial de Avaliação Institucional (SEAVI), a Unidade de Desenvolvimento Organizacional (UDO)/Gerência de

⁷ Doutorado em Educação (PhD) - Université de Sherbrooke - Canadá, mestrado em Educação pela Universidade Federal da Bahia - UFBA e licenciatura em Educação Física pela Universidade Católica do Salvador. Atua como professor do Programa de Mestrado Profissional Gestão e Tecnologias Aplicadas à Educação. Professor do Programa de pós-graduação Educação e Contemporaneidade, ambos vinculados ao Departamento de Educação Campus I da Universidade do Estado da Bahia - Uneb. Desenvolve estudos, atividades técnicas e pesquisas nos campos da gestão escolar e universitária, planejamento educacional e avaliação institucional.

Informática (GERINF), setor do qual faço parte como gerente de TI, e a Secretaria Geral de Cursos (SGC) da UNEB.

A SEAVI foi criada em 2014, instituída pela Resolução CONSU/UNEB 1.026/2014, conforme relata Novaes e Carneiro (2014), como “...órgão vinculado ao Gabinete do Reitor e com a finalidade de apresentar diretrizes além de oferecer apoio acadêmico e administrativo aos processos avaliativos no âmbito da Universidade.” A Assessoria de Comunicação da UNEB – ASCOM, destaca no portal da Universidade⁸ que a SEAVI objetiva integrar os processos avaliativos internos e externos a partir da articulação entre os setores acadêmicos e administrativos e ressalta que essa articulação se viabiliza através da sistematização e disponibilização de informações, dados e instrumentos necessários à melhoria dos processos de gestão e desempenho acadêmico e organizacional.

Novaes e Carneiro (2014) explicam ainda que a SAEVI “se apresenta no atual contexto da gestão universitária da Uneb, como um setor que visa atender às necessidades de integrar e melhor articular órgãos, departamentos e setores em torno dos processos de planejamento e avaliação institucional”. Falando em Avaliação Institucional cabe recorrer ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP, órgão federal com missão de promover estudos, pesquisas e avaliações sobre o Sistema Educacional Brasileiro com o objetivo de subsidiar a formulação e implementação de políticas públicas para a área educacional a partir de parâmetros de qualidade e equidade, bem como produzir informações claras e confiáveis aos gestores, pesquisadores, educadores e público em geral⁹.

Para gerar seus dados e estudos educacionais o INEP realiza levantamentos estatísticos e avaliativos em todos os níveis e modalidades de ensino, sendo que aqui nos interessa apenas os que dizem respeito ao nível superior de ensino. Particularmente o Censo Superior, coleta, anual, de uma série de dados do ensino superior no País, incluindo cursos de graduação, presenciais

⁸ Disponível em <<http://www.uneb.br/2016/02/29/a-seavi-uneb/>> Acesso em 19 de setembro de 2016

⁹ Disponível em <<http://portal.inep.gov.br/conheca-o-inep>> Acesso em 24 de setembro de 2016

e à distância. O Anuário UNEB em DADOS visa justamente fornecer dados gerais sobre os cursos da UNEB, a partir de sua vinculação aos Campi e Departamentos Universitários.

O Anuário UNEB em DADOS auxilia a gestão da Universidade a ter uma visão macro da situação da UNEB frente a indicadores que são levados em conta para o Censo INEP da Educação Superior. Carneiro e Novaes (2008) alertam da importância de considerar os processos adotados em torno da avaliação para compreender as transformações da educação superior no decorrer dos anos. E, para essa compreensão, é necessário realizar uma coleta e posterior análise dos dados coletados. É nesse ponto que o ANUÁRIO entra como ferramenta de suma importância.

Durante o processo de levantamento de informações para o referido Anuário, em conjunto com o Administrador de Banco de Dados e o coordenador da equipe de desenvolvimento de sistemas da GERINF, percebeu-se que um mapeamento dos indicadores e a aplicação das técnicas de *BI* não apenas automatizaria o processo de obtenção das informações necessárias para os próximos Anuários, como auxiliaria na identificação de possíveis pontos de falhas nas bases de dados da Universidade.

Esse projeto se justifica por auxiliar a gestão da UNEB a, não apenas obter informações relevantes sobre a Universidade para melhorar seus indicadores junto ao INEP, mas a (re)pensar as diversas bases de dados que vem sendo utilizadas, visto que, certamente, no momento da homologação do Data Warehouse - base de dados consolidada que é utilizada pela ferramenta de *BI* - serão encontradas incongruências nos dados. Isso se dá em função da falta de conscientização do corpo operacional sobre a importância dos dados inseridos nos sistemas para a geração de relatórios estratégicos. Em outras palavras, o não cadastro de algumas informações, ou um cadastro incorreto, pode acarretar em relatórios inconsistentes que não reflitam a realidade da universidade para os gestores.

Diante de todo o cenário descrito e dos benefícios intrínsecos da integração de sistemas e da utilização de ferramentas automatizadas de análise de dados, este projeto de pesquisa apresenta a seguinte pergunta que irá nortear seu desenvolvimento: como uma ferramenta de Business Intelligence (*BI*) através

de seus processos de coleta, sumarização e consolidação pode auxiliar na composição do Anuário UNEB em Dados?

1.1. OBJETIVO GERAL

Otimização do processo de levantamento dos dados para elaboração do Anuário UNEB em DADOS a partir do desenvolvimento do protótipo de uma ferramenta de *Business Intelligence*.

1.2. OBJETIVOS ESPECÍFICOS

- Conhecer as bases de dados que são fontes para a composição do Anuário UNEB em Dados;
- Estruturar um Data Mart (sub-conjunto de um Warehouse) que consolide dados relativos aos discentes Matriculados desde 2012;
- Realizar levantamento das perguntas que precisam ser respondidas pelo Anuário UNEB em Dados;
- Desenvolver protótipo de ferramenta de Business Intelligence (BI) que auxilie a SEAVI no contexto do Anuário UNEB em Dados;
- Apresentar através de planilhas e gráficos os dados do Anuário UNEB em Dados na ferramenta;

2. REFERENCIAL TEÓRICO

Nesse capítulo busco embasar a pesquisa abordando os principais conceitos utilizados durante o projeto. Ainda que aqui e ali tenha um pouco de minhas observações nesse capítulo, a ideia principal é explorar nesse ponto do texto o estado da arte das três áreas de conhecimento que norteiam o trabalho ora desenvolvido, a saber, Gestão de Instituições de Ensino Superior, Gestão da Informação e Conhecimento e por fim o conceito de Business Intelligence (*BI*), para no próximo capítulo entrar finalmente na descrição do desenvolvimento da pesquisa com minhas contribuições de maneira mais contundente.

Sendo a Universidade lócus deste projeto, o subcapítulo 4.1 traz um diálogo com autores sobre Gestão Universitária para entender melhor como uma Universidade funciona e o que é realmente importante para instituições desse tipo, pontuando a importância que deve ser dada ao público discente, razão de ser das Universidades. Em seguida, no subcapítulo 4.2, a Gestão da Informação e Conhecimento é abordada trazendo conceitos e teorias que embasem o direcionamento dado no trabalho, que se debruça justamente no levantamento e gestão de informações para a Universidade. Por fim, no subcapítulo 4.3 conceituo e apresento metodologias de Business Intelligence que serão utilizadas para o desenvolvimento do protótipo que é o produto deste projeto.

2.1. GESTÃO DE INSTITUIÇÕES DE ENSINO SUPERIOR

O mercado vem mudando e tornando-se cada vez mais competitivo nas diversas áreas. Em verdade, as próprias barreiras entre as áreas de negócios vêm sendo derrubadas. Como afirma Tachizawa et al (1999, p.22), “As organizações não se limitam mais às suas tradicionais bases de clientes.”. Hoje é lugar comum encontrar, por exemplo, serviços financeiros dentro de uma agência dos correios. A concorrência é feroz e pode surgir de onde menos se espera.

Neste contexto, as próprias Instituições de Ensino Superior – IES precisam estar preparadas para enfrentar o mercado que a circunda. Não há mais como se sentir confortável e sossegado. Se já não houvesse essa

inexorável desintegração de barreiras que antes separavam a área educacional das demais áreas, a própria competição entre as IES públicas e privadas da atualidade já seria suficiente para que os gestores e executivos dessas tradicionais instituições se preocupassem com um modelo de gestão que garantisse melhor qualidade de ensino e maior eficiência administrativa.

Finger (1997, p.13) cita dois problemas básicos que servem para as organizações acadêmicas buscarem uma eficiência maior e processos mais elaborados. São eles:

- a) Escassos recursos financeiros e necessidade de fazer com que o produto organizacional tenha um preço, monetário e social, que o torne atraente; e
- b) Visão a mais longo prazo da instituição, onde a mesma encontre caminhos que possam fundamentar ações permanentes.

O próprio Finger (1997) alerta que as universidades brasileiras nunca atentaram para esses dois princípios. Sendo nas instituições públicas o maior descaso com programas de médio e longo prazo, visto que é característico uma gestão descontinuar o que a anterior iniciou, apenas para não “apoiar gestões anteriores” ao passo que tenta “deixar a sua marca”.

De acordo com Torres (2012), as Universidades são organizações com peculiaridades em relação a outros tipos de Instituições. Seja uma IES Privada ou Pública, o modelo de funcionamento das instituições demandam características flexíveis, com forte presença do corpo Docente.

Para prosseguirmos com as afirmações de Finger e de Torres, devemos então pensar as IES em termos de organização. Para tal, utilizemos o conceito dado por Robbins (2001:31 *apud* SOUZA, 2007:54), cuja descrição diz que “a organização é um arranjo sistemático de duas ou mais pessoas que cumprem papéis formais e compartilham um propósito comum”.

O autor ainda complementa que, tal qual indústrias e outras empresas, as universidades ou faculdades, enquanto organizações, compartilham de algumas características comuns. São elas:

- Têm um propósito normalmente expresso em termos de metas ou conjuntos de metas;
- São compostas de pessoas;
- Desenvolvem uma estrutura sistemática, que define papéis formais e limita o comportamento de seus membros.

Na tentativa de compreender como as universidades funcionam, Hardy e Fachin (2000), apresentam os seguintes modelos básicos de gestão universitária: o Burocrático, o Político, o da Anarquia Organizada e o Colegiado.

O modelo burocrático possui forte presença do formalismo, da impessoalidade e do profissionalismo. No caso das IES, há uma considerável padronização de habilidades e de procedimentos que possibilitam a execução do trabalho nas instituições que tiveram tamanho e complexidade aumentados. Além disso, observa-se a existência de unidades autônomas na estrutura ligadas entre si e com a Reitoria. Essa descentralização e democracia exigem quantidade considerável de hierarquia. Por fim, há a burocracia hierárquica do corpo administrativo que difere da autonomia e liberdade acadêmica do profissional acadêmico.

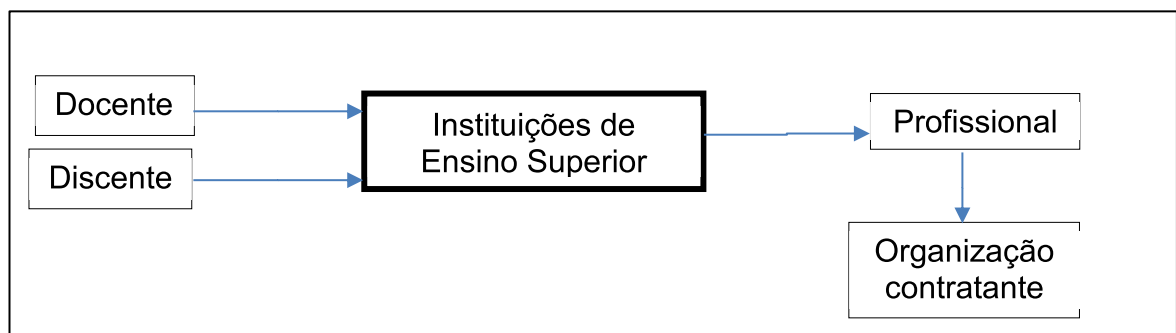
Já no modelo político, os departamentos com maior capacidade de captar recursos externos tem maior alocação de recursos internos. Nota-se também que a rotatividade dos chefes de departamentos nesse modelo é analisada como um processo político.

Por outro lado, o modelo colegiado, também chamado de “comunidade de homens cultos”, baseia-se na decisão obtida pelo consenso. Observa-se a ideia de autoridade profissional, baseada na competência e não na posição, o que gera uma hierarquia mais achatada. No modelo colegiado, há uma estrutura bastante igualitária, com relações predominantemente informais.

A Anarquia Organizada, ou modelo “lata de lixo”, caracteriza organizações como situações de decisão com objetivos indefinidos, tecnologia indistinta, falta de conexão entre problemas e soluções. É o modelo adotado quando as decisões não são tomadas intencionalmente, sendo determinadas pela ausência

de ação ou por acidente. Parece próprio para assuntos menos importantes e outro modelo de gestão pode ser adotado quando a situação exigir.

Qualquer que seja o modelo de gestão universitária a ser implantado numa determinada IES, o mesmo deverá levar em consideração os objetivos e as funções conflitantes da mesma. Também não se pode esquecer que o gestor deverá ter acesso a informações sobre a instituição para tomar suas decisões de forma embasada e não apenas utilizando a intuição. Além disso, é preciso ter em mente que, em qualquer que seja o curso escolhido, o aluno formado por toda IES, normalmente, irá exercer uma profissão correlata a esse curso, em uma determinada organização empregadora desse profissional, presumivelmente preparado para exercer uma determinada função (TACHIZAWA, ANDRADE 1999). O contexto descrito acima pode ser descrito visualmente na Figura 3.



Fonte: autor

Figura 3: Fluxo básico de uma IES

Conforme Mezomo (1994 *apud* TACHIZAWA e ANDRADE, 1999:24) “os alunos compram e utilizam os serviços prestados pela instituição de ensino, sendo, portanto seus *clientes*”. O mesmo autor afirma que não apenas os alunos, mas todas as pessoas envolvidas no processo educacional (estejam elas dentro ou fora das IES) são consideradas clientes das IES.

Observando o diagrama acima e levando em consideração a afirmação de Mezomo, e a feroz concorrência citada no início do presente capítulo, parece sensato sugerir que uma IES deve adotar o modelo de gestão universitária que melhor atenda seus clientes. Por outro lado, também parece lógico inferir que uma IES para ser mais competitiva precisa conhecer muito bem seu cliente. O presente trabalho propõe uma ferramenta que auxilia a Universidade a entender melhor o cliente discente.

Pimenta (2007) afirma que, dentre outras coisas, a missão da instituição universitária engloba a responsabilidade na qualificação e capacitação permanente das pessoas, bem como impulsionar a pesquisa e o desenvolvimento científico. Para tal, importa saber em primeira instância, quais resultados vem sendo obtidos nesse processo de qualificação e capacitação. Quem são os atores que fazem uma IES funcionar? Temos os técnicos/analistas administrativos, temos os docentes e os discentes. Como esses atores se relacionam e qual a eficiência dessas relações? É preciso ter indicadores sobre os três segmentos, mas dada a importância que os discentes têm para uma Universidade, é possível inferir que este seja o primeiro grupo a ser analisado. Isto porque eles, os discentes, são atores que, num primeiro momento são os clientes das Universidades, mas logo em seguida (Figura 3), após a conclusão de seus cursos, “transformam-se” em profissionais, produto para o segundo cliente de uma IES, a sociedade. Conhecer o público discente ajuda não apenas a compreender uma IES, mas também a melhorar eficiência das mesmas, de modo a atender plenamente o seu público. Para tal, é preciso otimizar a criação e gestão das informações organizacionais.

A UNEB possui diversos sistemas e bases de dados com informações sobre seus discentes, mas como se dá a gestão dessa informação? Para entender como se dá essa gestão precisamos compreender o que é Gestão da Informação. E mais do que isso, verificar se essa informação é transformada em conhecimento útil para a Gestão Universitária. A ferramenta aqui proposta busca auxiliar a UNEB na coleta, organização e análise da informação organizacional para facilitar a apropriação dessa informação e, conseqüentemente, propiciar a geração de “novo” conhecimento.

2.2. GESTÃO DA INFORMAÇÃO E DO CONHECIMENTO

De acordo com Terra (2000), “vivemos um momento de importante transição do ambiente econômico, onde a gestão proativa do conhecimento adquire um papel central para a competitividade tanto das empresas, como dos países”. Tomando essa afirmativa como verdadeira, se tanto empresas como países passam a dar destaque à gestão do conhecimento, o que dizer de uma IES, local cuja razão de ser

é justamente a criação e difusão do conhecimento? E se a competição entre as universidades tem se tornado cada vez mais acirrada, é necessária uma gestão efetiva do conhecimento das universidades. Tanto o conhecimento científico (produto das IES) quanto o conhecimento organizacional (vital para o funcionamento das instituições e foco do presente trabalho).

Para dar prosseguimento ao estudo do conhecimento organizacional, é deveras importante dar uns passos para trás e entender um pouco sobre o conhecimento em si. Na literatura é comum encontrarmos afirmações por parte de diversos teóricos que estamos vivendo na era do conhecimento. Mas o que é o conhecimento? Para Davenport (1998), o “conhecimento é a informação mais valiosa (...) valiosa precisamente porque alguém deu à informação um contexto, um significado, uma interpretação (...)”. Partindo desse conceito, podemos dizer que o conhecimento diz respeito a uma informação que foi previamente processada por um indivíduo. Mas se isso nos diz o que é o conhecimento, nos deixa ainda sem um entendimento pleno, visto que ainda não sabemos o que é a informação.

Rumo à luz da questão, recorremos ao estudioso Peter Drucker que define informação como sendo “dados dotados de relevância e propósito” (DRUCKER 1988, apud Davenport, 1998:19). E nesse ponto sou obrigado a concordar com Davenport que afirma serem os humanos quem dotam os dados de tais atributos. Afinal de contas, mesmo quando uma máquina transforma (processa) automaticamente uma folha de custos num gráfico de pizza ou outro qualquer, em última instância, é necessário que uma pessoa (o usuário) escolha qual tipo de gráfico utilizar. Fadel et al (2010, p. 14) traz outra definição, dizendo que “...a informação é insumo para qualquer fazer, seja no âmbito acadêmico, seja no âmbito empresarial. A geração de “novo” conhecimento somente é possível quando a informação é apropriada pelo indivíduo por meio do estabelecimento de relações cognitivas”.

Chegamos mais perto de um entendimento completo do que seja conhecimento, mas ainda nos resta entender o que seriam os dados. Afinal, partindo dos conceitos dados por Drucker (1988) e Davenport (1998), os dados parecem ser a menor unidade utilizada para gerar o conhecimento. E o próprio Davenport define enfaticamente dados como “observações sobre o estado do mundo”. Essa definição

é interessante, mas ainda considero um tanto vaga. Recorro então a Miranda (1999) que conceitua de maneira mais objetiva dizendo: “dados é o conjunto de registros qualitativos ou quantitativos conhecido que organizado, agrupado, categorizado e padronizado adequadamente transforma-se em informação;”.

Agora com as definições postas, temos que, de acordo com Miranda (1999), Drucker (1988) e Davenport (1998), existe o dado, que organizado, categorizado e padronizado, em outras palavras, processado (como já vimos, por um ser humano), transforma-se em informação. Já a informação, ao ser dotada de contexto, significado e interpretação, vira o que chamamos de conhecimento.

Sobre a diferença entre dado, informação e conhecimento, há de se deixar registrado que essa diferenciação ainda não é uma unanimidade entre os estudiosos, conforme Russo (2010), que relata o seguinte:

(...) seus significados não são tão distintos e, por vezes, se confundem pela proximidade de sua aplicação em um determinado contexto. Eles formam, ainda, um sistema hierárquico de difícil delimitação, pois algumas vezes, o que é um dado para um indivíduo pode ser informação ou conhecimento para outro.

Também é necessário conceituar os termos 'gestão da informação' e 'gestão do conhecimento' que, apesar de possuírem relação estreita, não são termos sinônimos e, não raramente, são confundidos. A gestão da informação tem como foco o negócio da organização e sua ação é restrita aos fluxos formais; já a gestão do conhecimento tem como foco o capital intelectual e sua ação é restrita aos fluxos informais; (VALENTIM, 2002).

O presente projeto possui foco na gestão da informação, mas, invariavelmente, esbarra nos fluxos informacionais informais da Universidade para levantamento das informações necessárias ao desenvolvimento do protótipo e, portanto, um conhecimento básico dos conceitos de gestão do conhecimento é necessário para entender melhor a dinâmica desses fluxos informacionais.

Nonaka (1997) classifica o conhecimento em dois tipos: o explícito, que pode ser expresso em afirmações gramaticais ou até mesmo em expressões

matemáticas ou manuais, e aquele considerado pelo autor como o mais importante, o tácito, que é o conhecimento pessoal incorporado à experiência do indivíduo e envolve fatores intangíveis como crenças, emoções e valores.

A diferença entre conhecimento tácito e explícito fica mais clara nas palavras de Pinheiro (2002):

O processo de identificar, compartilhar o conhecimento explícito é uma tarefa, de certa forma, simples, pois o conhecimento dessa natureza se apresenta em meio formal, claro e estruturado. Existe outra forma de conhecimento, denominado conhecimento tácito, ou seja, o conhecimento subjetivo ligado às experiências, habilidades adquiridas pelo indivíduo ou pelo setor. Este tipo de conhecimento apresenta uma dificuldade maior na sua identificação e compartilhamento.

Apesar de dar maior importância ao conhecimento tácito, Nonaka (1997) trata os dois tipos de conhecimento como unidades estruturais básicas que se complementam. Tendo foco maior na gestão da informação, iremos nos ater ao conhecimento explícito, uma vez que trata do conhecimento objetivo e “mais baseado” na informação estruturada.

As organizações que pretendem sobreviver na chamada “era do conhecimento”, precisam estar atentas à gestão do conhecimento organizacional, ou seja, devem analisar e fomentar normas e práticas que estimulem a criação, o armazenamento e a difusão desse conhecimento. Cabe à alta direção pensar e gerar políticas para a aquisição de conhecimento interno ou externo à organização. Terra (2002), diz que:

A Gestão do Conhecimento centra-se em três aspectos principais: foco nos ativos intangíveis (principalmente o fator humano), tornar a gestão do conhecimento algo explícito, incentivar e criar mecanismos que facilitem aos empregados o compartilhamento de seus conhecimentos.

Podemos perceber a importância de tornar o conhecimento explícito dentro da organização, pois, conforme visto anteriormente, esse é o tipo de conhecimento mais facilmente “transferido”, ou seja, é a forma como o conhecimento pode ser disseminado dentro da organização. Por outro lado, os

gestores precisam também fomentar a internalização, em outras palavras, a transformação desse conhecimento (explícito) disseminado na organização em conhecimento tácito. Pois é nesse ponto que o indivíduo apreende de fato o conhecimento.

Terra (2002) defende que o desafio da adoção das práticas e modelos associados à Gestão do Conhecimento precisa necessariamente ser apoiado por mudanças de processos, estruturas e sistemas de informação. É nesse ponto que os gestores devem entender a importância estratégica da área de Tecnologia da Informação e Comunicação (TIC).

2.3. BUSINESS INTELLIGENCE: UMA METODOLOGIA DE APOIO A DECISÃO

Nos dias atuais, é cada vez mais essencial a utilização dos sistemas informatizados e a sua progressiva atualização por parte das organizações. A quantidade de dados existente nas organizações vem crescendo vertiginosamente e encontrar informações relevantes nesse emaranhado torna-se uma tarefa extremamente exaustiva. Várias são as técnicas e tecnologias utilizadas para essa “extração de conhecimento”, tendo o conceito de Bancos de Dados possivelmente como o principal elemento envolvido em todo o processo.

O Banco de Dados é parte fundamental de todo e qualquer sistema informacional utilizado nas diversas organizações. É nele que os registros (dados) são armazenados para posterior manipulação por parte dos usuários. A evolução dos sistemas computacionais passa pela evolução dos próprios Sistemas de Bancos de Dados.

Conforme narra Elmasri e Navathe (2005), a maioria dos sistemas pioneiros de banco de dados foi implementada em computadores grandes (mainframes) e caros, começando em meados de 1960. É importante ressaltar que esses sistemas eram baseados em três paradigmas principais: os sistemas hierárquicos, aqueles baseados em modelo de rede, e os de arquivos invertidos.

Porém, provendo uma fundamentação matemática e separando o armazenamento físico dos dados da modelagem conceitual dos mesmos, entre o final da década de 70 e no início dos anos 80, os primeiros sistemas baseados no modelo relacional começavam a surgir no mercado (DATE, 2000). O modelo viria a ser praticamente o padrão do mercado para o desenvolvimento de sistemas transacionais.

Os sistemas transacionais, também conhecidos como Sistemas de Processamento de Transações em Tempo Real, ou *Online Transaction Processing* – OLTP são os sistemas utilizados pela parte operacional das organizações para registrar os dados de todas as transações das mesmas. Seja o registro das vendas de um supermercado, ou a matrícula dos alunos de uma universidade. Esses sistemas são utilizados basicamente para tornar mais eficiente as tarefas de registro de informações sobre o negócio.

Com o passar do tempo, o volume de informações inseridos nos sistemas tornou-se tão alto que apenas informatizar essa inserção de dados não era mais suficiente. Murakami (2003) afirma que atualmente, a estratégia de TIC transcende ao seu tradicional papel de *back office* e evoluiu para o nível estratégico. Na visão de Luftman (2000), o papel estratégico de TI não deve apenas apoiar estratégias empresariais, mas também amoldar estratégias empresariais novas.

Dito isso, os sistemas transacionais continuaram tendo sua utilidade, mas, por atuarem basicamente no nível operacional, não mais eram suficientes para manter uma organização competitiva na chamada “era da informação”. Logo, sistemas com novos enfoques precisaram ser criados para auxiliar os gestores. Nessa leva surgiram sistemas que a partir dos dados operacionais forneciam relatórios contendo uma visão macro das operações da empresa aos administradores. Em outras palavras, surgiam os chamados Sistemas de Informações Gerenciais – SIG.

Carmo e Pontes (1999) afirmam que é necessário que um sistema de informações possa atender às principais necessidades de informações gerenciais, considerando aspectos importantes utilizados para o gerenciamento das

empresas. O que só confirma que um sistema de informações não apenas atue no nível operacional das empresas. Importante ressaltar que os SIG são imprescindíveis em organizações de qualquer porte. Se por um lado uma grande empresa depende dos sistemas em função do alto volume de dados gerados pelos seus sistemas transacionais, por outro, os autores alertam para o fato das pequenas organizações possuírem um quantitativo escasso de recursos materiais e humanos, o que aumenta a concentração de decisões no gestor.

Importante aqui explicitar a distinção entre dado e informação do ponto de vista da TIC. Conforme elucidada Date (2000), vários escritores descrevem “dados” como sendo o que está realmente armazenado no banco de dados e conceituam “informação” como o *significado* desses dados para o usuário. Por exemplo, podemos ter o dado “Fábio” armazenado no banco de dados e esse dado pode dizer respeito a informação “nome” de um determinado aluno.

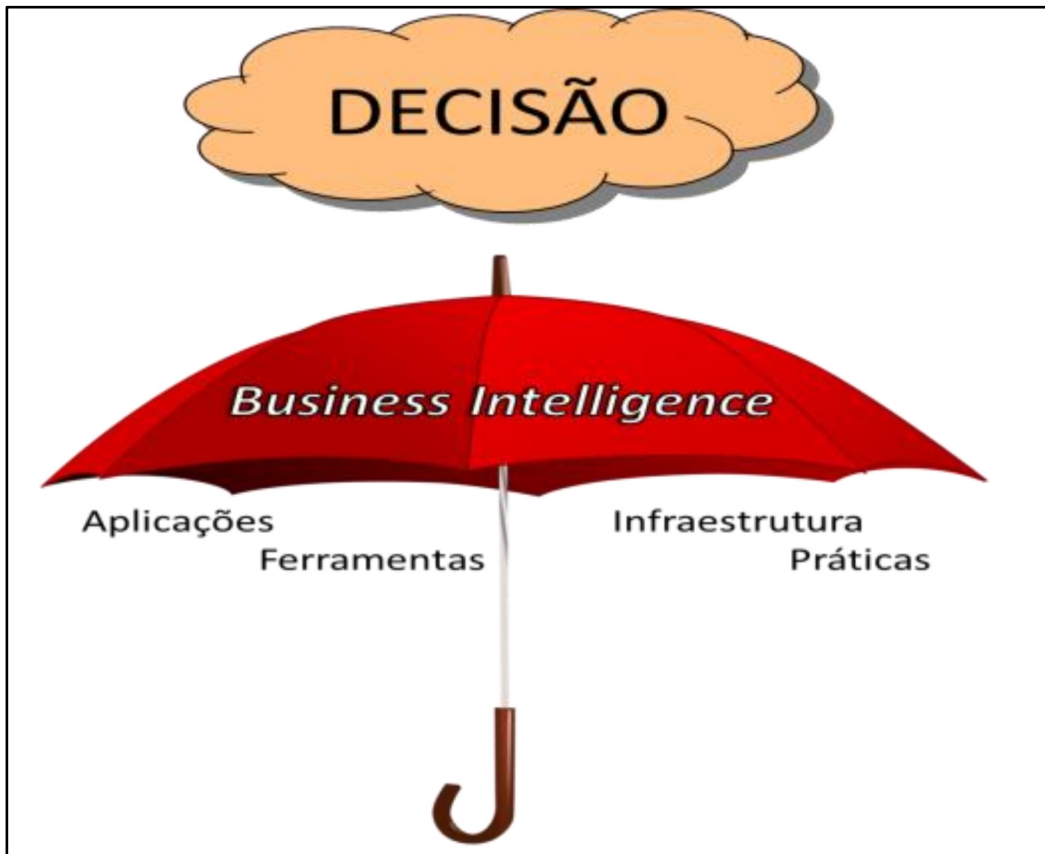
Buscando auxiliar na consulta e, principalmente, na análise desses dados, incorpora-se o conceito de *Business Intelligence* ou *BI*, que, segundo Sezões et al (2006, p.10), consiste no “conjunto de aplicações de apoio à tomada de decisão que possibilitam um acesso rápido, compartilhado e interativo das informações, bem com a sua análise e manipulação”.

Gartner Inc. traz uma definição mais completa de Business Intelligence dizendo¹⁰ com a qual concordo e compartilho a seguir:

Business intelligence (BI) é um termo guarda-chuva (Figura 5) que inclui as aplicações, infraestrutura e ferramentas, e melhores práticas que permitem o acesso e análise de informação para aumentar e otimizar performance e decisão.

Precisamos então ter em mente que BI não envolve apenas a parte ferramental, mas um conjunto de técnicas, ferramentas e procedimentos que buscam prover o corpo estratégico de informações importantes para a tomada de decisão. Tratarei, todavia, no presente trabalho, basicamente da parte ferramental do processo.

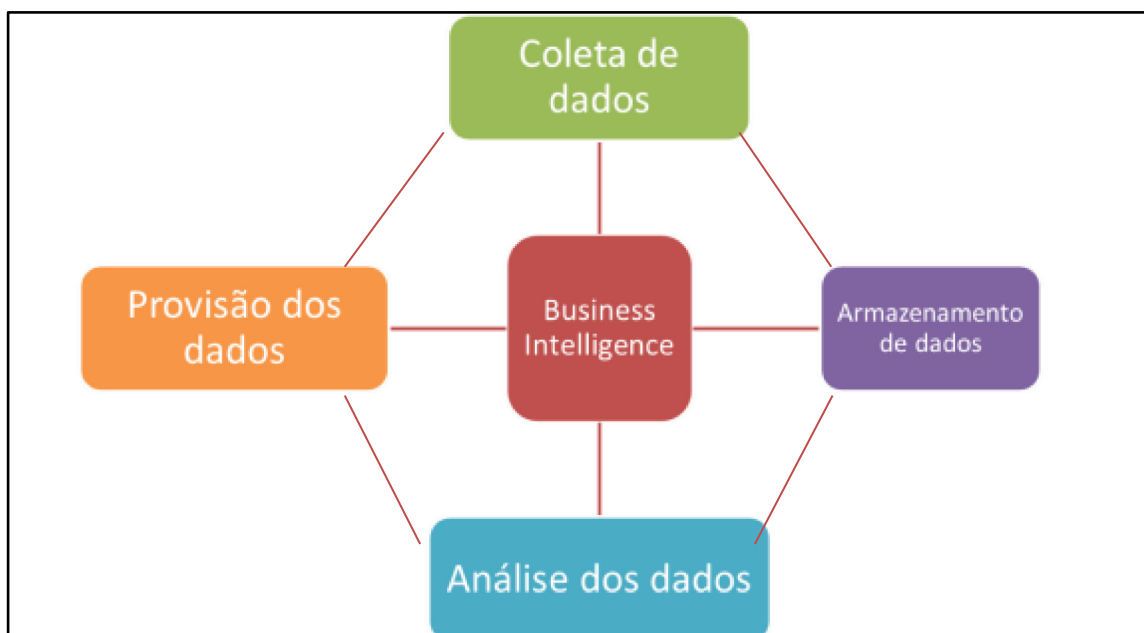
¹⁰ Disponível em < <http://www.gartner.com/it-glossary/business-intelligence-bi> > Acesso em 1 de junho de 2015



Fonte: autor

Figura 5: abstração do conceito de BI

Uma ferramenta de *BI*, ou Inteligência de Negócios, se apoia primordialmente em quatro fundamentos para se concretizar como tal (Figura 4). São eles: Coleta de dados, armazenamento de dados, Análise dos dados e Provisão dos dados. Não existe uma ferramenta de *BI* se qualquer desses fundamentos inexistir no processo.



Fonte: autor

Figura 4: Fundamentos de um BI

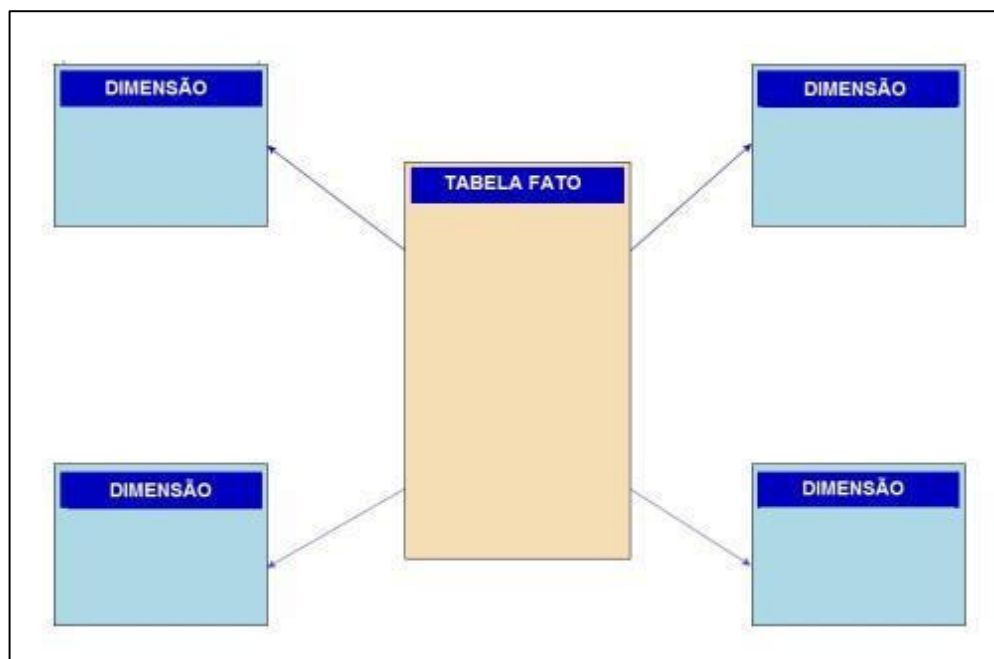
Na coleta de dados, técnicas e processos são realizados para buscar nas diversas bases transacionais os dados que são importantes para gerar informações relevantes aos gestores no processo decisório. Essa “fase” do *BI* exige análise minuciosa das bases de dados da organização e constante conversa com os envolvidos na operação dos sistemas transacionais para compreender e relacionar os dados que migrarão para o *BI*.

Podemos dizer que a fase de Armazenamento de dados é uma das fases mais importantes, pois nela reside uma mudança importante de paradigma tecnológico que possibilita ou, pelo menos, facilita e muito as próximas etapas do processo. Isso porque no armazenamento de dados de um *BI* nós consolidamos todos os dados coletados nas bases transacionais num banco com modelo dimensional chamado de Data Warehouse (DW). O DW é uma das principais e mais utilizadas ferramentas de *BI* do mercado. Conforme veremos mais à frente, algumas soluções de *BI* afirmam não prescindir do DW, mas explicarei também que a verdade não é bem essa. Kimball (2013) diz que o modelo dimensional é mais amplamente aceito para soluções de *BI* por duas razões: entrega dados “entendíveis” para o usuário do negócio e possui uma alta performance de consulta.

A definição clássica e mais difundida do que é um DW é dada por Inmon (1992), que diz ser “[...] uma coleção de dados integrados, orientados por assunto, não voláteis, e variáveis com relação ao tempo, de apoio às tomadas de decisões gerenciais”.

Um modelo dimensional conta basicamente com uma tabela fato contendo o assunto principal e tabelas dimensionais (também conhecidas como dimensões) vinculadas a essa tabela central. Existem dois tipos básicos de modelos dimensionais, o modelo estrela (Figura 6) e o floco de neve (Figura 7).

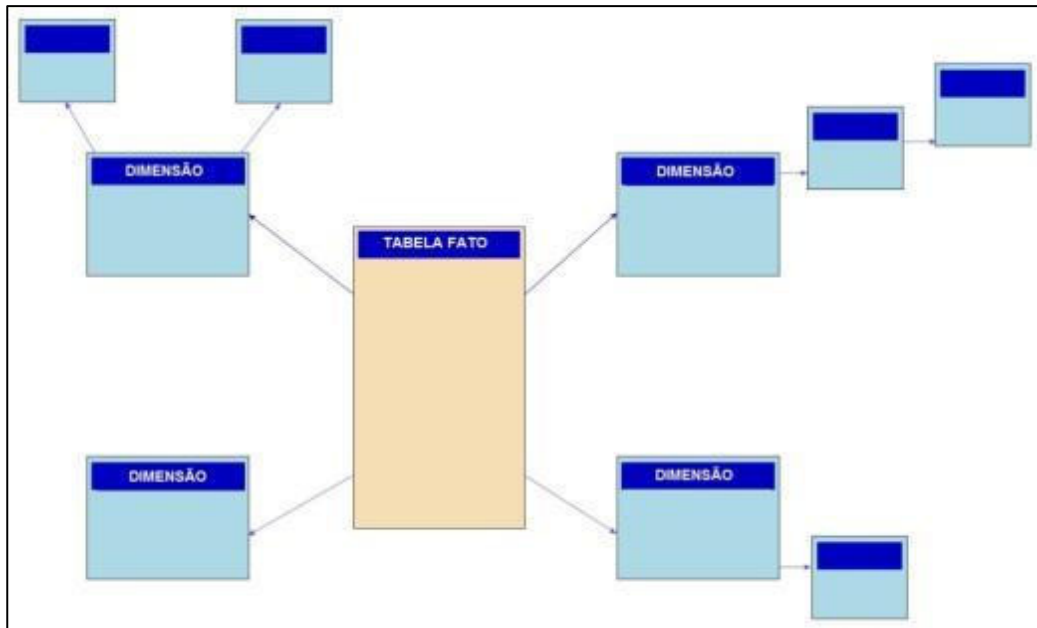
No modelo estrela, todas as dimensões ligam-se diretamente à tabela fato (Figura 6), enquanto que no modelo floco de neve algumas dimensões não estão ligadas à fato, mas sim a alguma dimensão que esteja ligada a tabela fato (Figura 7). A ideia é diminuir o tamanho de algumas tabelas e conseqüentemente sua ocupação na base de dados. Porém, Kimball (2013) defende que o ganho de espaço em disco é insignificante e, além de possuir uma perda de desempenho em função do aumento de junções entre as tabelas, o que aumenta a quantidade de JOINS nos códigos das consultas em SQL¹¹, isso aumentaria a complexidade do modelo. Sendo assim, o autor deixa claro a sua preferência pelo modelo estrela e esse é o modelo adotado no presente projeto.



Fonte: autor

Figura 6: Modelo estrela

¹¹ Structured Query Language, ou Linguagem de Consulta Estruturada, é a linguagem de pesquisa declarativa padrão para banco de dados relacional. Muitas das características originais do SQL foram inspiradas na álgebra relacional.



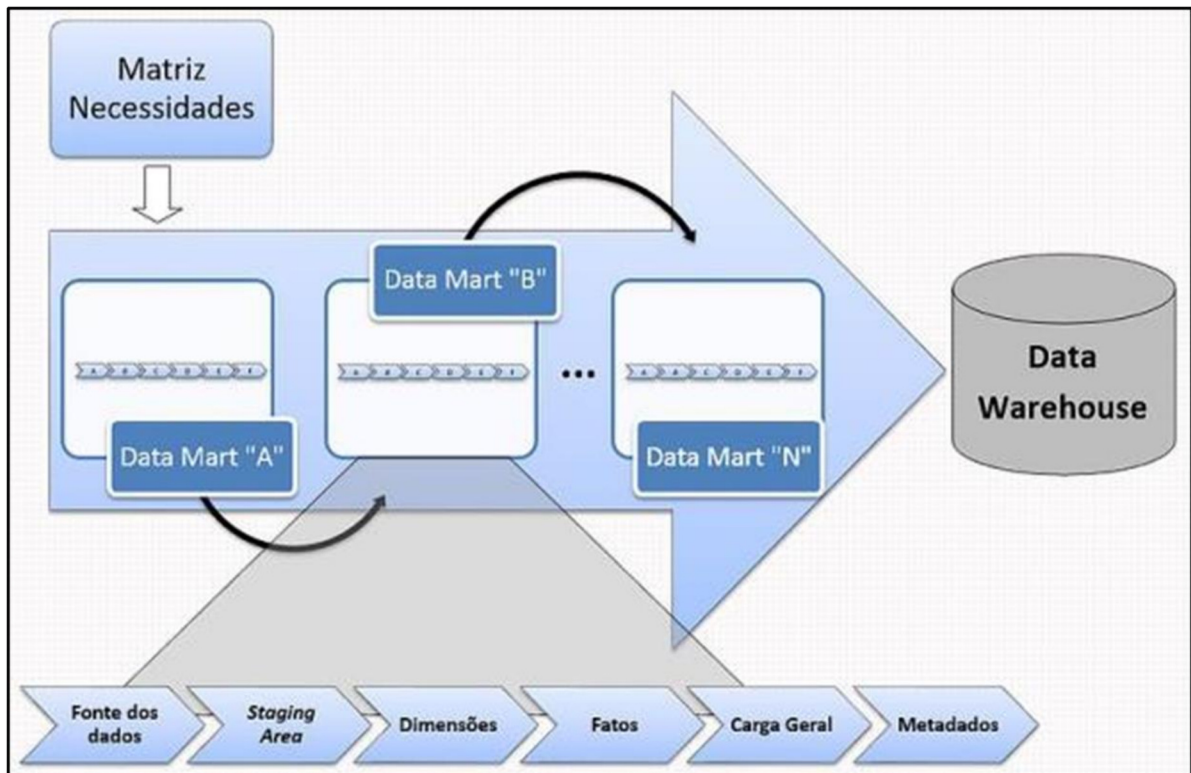
Fonte: autor

Figura 7: Modelo floco de neve

Conforme abordagem proposta por Kimball (2013), o Data Warehouse é composto por diversos assuntos ou departamentos como se fossem “mini” Data Warehouses. Esses pequenos DWs são chamados de Data Marts (DM). Logo, ainda que não seja a única maneira, o processo de implementação de um DW pode ser composto de diversas iterações de um ciclo de desenvolvimento de um DM. Para uma melhor visualização, segue uma ilustração (Figura 8). Inmon prefere que se utilize uma abordagem inversa, onde seja desenvolvido um DW primeiro, para que os DM possam consumir os dados a partir desse grande armazém de dados consolidados. Essa pode ser uma metodologia mais eficiente a longo prazo, porém, alguns teóricos, como o próprio Kimball, acreditam ser essa abordagem demasiadamente lenta para apresentar algum resultado. Ralph Kimball defende a implementação de Data Marts “normalizados”. Isto é, com tabelas fato compartilhando dimensões comuns. Dessa forma, pode-se ter um DM com resultados mais rápidos, porém já pensando num DW consolidado.

Como o recorte aqui proposto visa uma ferramenta de *BI* para auxiliar no levantamento dos dados relativos aos discentes para a elaboração do Anuário UNEB em Dados, definimos então que nossa ferramenta abordará apenas o “assunto” matrícula. Assim, optamos pelo desenvolvimento de um Data Mart como ponto de partida para um futuro Data Warehouse completo da Universidade. A

partir deste ponto do texto, para fins de padronização e no intuito de facilitar a compreensão do trabalho, irei me referir à ferramenta como um Data Warehouse (DW), ao invés de Data Mart (DM), ainda que tratemos apenas de um assunto específico da instituição.



Fonte: <http://www.binapratca.com.br/>

Figura 8: Ciclo de construção do DW

Os termos Data Warehouse e *Business Intelligence* são frequentemente relacionados e confundidos como sendo a mesma coisa. Porém, é importante ter em mente que o DW é apenas uma ferramenta, ainda que amplamente utilizada, de BI. É um depósito de dados projetado especificamente para facilitar o acesso às informações estratégicas. Já o termo de *Business Intelligence*, conforme exemplificado na Figura 5, refere-se a um conceito abrangente que vai além das ferramentas, incluindo metodologias, infraestrutura e práticas que auxiliem no processo decisório das organizações. Com o avanço das ferramentas de BI, nem sempre é necessário o desenvolvimento de um DW, pois é possível que os dados transacionais sejam diretamente acessados para a análise e geração de relatórios/gráficos. Porém, o que na verdade acontece é que essas ferramentas usam metodologias próprias para coletar e organizar os dados internamente. Em

outras palavras, essas ferramentas automatizam o processo de criação do DW. Ou seja, de uma maneira ou de outra, o Data Warehouse será necessário para que o BI funcione.

Durante o desenvolvimento de um DW, os dados operacionais, por estarem vindo de diversas bases ditas operacionais (dos sistemas transacionais da organização), precisam passar por um pré-processamento para só então serem carregados no DW. Essa é a etapa de extração transformação e carga, comumente chamada entre os profissionais da área de TI de ETL, do inglês Extract, Transform and Load.

Algumas etapas estão envolvidas no processo de desenvolvimento do DW. Algumas delas estão embutidas no próprio ETL, outras são agregadas ao mesmo. São elas:

- Avaliação da Fonte de Dados
- Construção da Staging Area (área temporária)
- Construção das Dimensões (tabelas com informações sobre o fato a ser analisado)
 - Construção dos Fatos (integrando as dimensões e acrescentando as métricas que dizem respeito ao fato)
 - Definição do Processo Geral de Carga
 - Criação dos Metadados

O terceiro fundamento de um BI, a Análise dos dados, aparece no momento em que o usuário (ou uma ferramenta de automatizada, através de processo de data mining, por exemplo) passa a estudar os dados já armazenados. Aqui os dados já armazenados, são organizados no que é chamado de cubo OLAP (Online Analytical Processing) para descoberta de padrões e tendências, facilitando a percepção (*insights*) de novas informações/conhecimento para a tomada de decisão.

A biblioteca MSDN (Microsoft Developer Network) diz o seguinte sobre o cubo OLAP¹²:

Um cubo OLAP é uma estrutura de dados que supera as limitações dos bancos de dados relacionais, proporcionando rápida análise de dados. Os cubos podem exibir e somar grandes quantidades de dados enquanto fornecem aos usuários acesso pesquisável a quaisquer pontos de dados. Dessa forma, os dados podem ser acumulados, segmentados e separados, conforme necessário, para controlar a mais ampla variedade de perguntas relevantes para a área de interesse de um usuário.

Por fim, temos o fundamento de Provisão dos dados. Aqui temos o que vem sendo chamado de “empoderamento do usuário”, onde os profissionais de TI buscam dar independência aos usuários finais através de interfaces mais simples e intuitivas, permitindo que eles acessem e analisem os dados diretamente sem a necessidade de inserção de código por parte dos programadores. Kimball (2013) frisa que o DW deve tornar a informação facilmente acessível, apresentá-la consistentemente, se adaptar às mudanças, apresentar a informação em um tempo razoável e servir como uma fonte confiável e autorizada para uma melhor tomada de decisões. Mas o mais importante: não importa o quão elegante, segura e veloz a solução seja, ela precisa ser “abraçada” pela área do negócio. Isso porque, diferente de um sistema operacional (transacional), em que os usuários não têm escolha quanto a usar ou não, algumas vezes o Data Warehouse é de uso opcional. Se a gestão não se interessar pelo sistema, o projeto está fadado ao fracasso. Os dados não devem estar mais em estruturas onde apenas desenvolvedores e Administradores de Bancos de Dados entendam, mas armazenados de forma que o usuário final possa acessar e analisar diretamente as informações que precisa.

E é em função deste último fundamento que aqui utilizaremos o Microsoft Office Excel¹³ para análise dos dados organizados no cubo que será gerado previamente. Escolhemos essa ferramenta em razão do amplo uso da mesma por grande parte dos usuários finais da UNEB, o que, em tese, facilita a assimilação da nova forma de acesso e visualização dos dados que serão disponibilizados.

¹² Disponível em < [https://msdn.microsoft.com/pt-br/library/hh916543\(v=sc.12\).aspx](https://msdn.microsoft.com/pt-br/library/hh916543(v=sc.12).aspx) > Acesso em 29 de dezembro de 2016

¹³ Editor de planilhas produzido pela Microsoft para computadores que utilizam o sistema operacional Microsoft Windows.

3. METODOLOGIA E LOCUS DA PESQUISA

Nesse capítulo é apresentada a metodologia de pesquisa adotada no desenvolvimento do trabalho, bem como explicado porque utilizar a metodologia de desenvolvimento da ferramenta. No subcapítulo 5.1 a pesquisa é caracterizada a partir da portaria da CAPES que dispõe sobre o reconhecimento dos Mestrados Profissionais e de um diálogo com alguns teóricos de metodologia de pesquisa, em particular Thiollent, que explana sobre a pesquisa Ação. Já o subcapítulo 5.2, trata de explicar o porquê da escolha da SEAVI como Locus da Pesquisa.

3.1. CARACTERIZAÇÃO DA PESQUISA

Como o curso de Mestrado em Gestão e Tecnologias Aplicadas à Educação – GESTEC é um curso *Stricto Sensu* na modalidade Profissional, o mesmo, segundo a Capes¹⁴ “ênfatisa estudos e técnicas diretamente voltadas ao desempenho de um alto nível de qualificação profissional”. Essa modalidade de Mestrado tem, de acordo com a PORTARIA Nº 080, de 16 de dezembro de 1998, dentre suas características:

d) exigência de apresentação de trabalho final que demonstre domínio do objeto de estudo, (sob a forma de dissertação, projeto, análise de casos, performance, produção artística, desenvolvimento de instrumentos, equipamentos, protótipos, entre outras, de acordo com a natureza da área e os fins do curso) e capacidade de expressar-se lucidamente sobre ele.

Com essas informações em mãos e buscando o desenvolvimento de um protótipo de ferramenta tecnológica (o *BI*) como produto, recorreremos à bibliografia disponível sobre metodologia da pesquisa para a escolha daquela que será utilizada. Segundo Paranhos (2014), quanto à natureza, a pesquisa pode ser classificada em básica ou aplicada, sendo que a aplicada, além de gerar

¹⁴ Disponível em < <http://www.capes.gov.br/acessoainformacao/perguntas-frequentes/avaliacao-da-pos-graduacao/7419-mestrado-profissional> > Acesso em 24 de setembro de 2016

conhecimentos novos úteis para o avanço da ciência, parte da aplicação prática, com enfoque numa realidade que possui interesses locais e pontuais. Thiollent (1986) complementa dizendo que “a Pesquisa Aplicada concentra-se em torno dos problemas presentes nas atividades de instituições, organizações, grupos ou atores sociais. Ela está empenhada na elaboração de diagnósticos, identificação de problemas e busca de soluções.”.

A pesquisa também pode ser classificada quanto aos procedimentos adotados durante a mesma. De acordo com Tozoni-Reis (2009):

“...o envolvimento do estudioso com o campo não impede o processo de investigação, ao contrário, cria condições concretas para que se possa captar os significados dos fenômenos estudados. Assim como o pesquisador é um elemento importante no processo de pesquisa, também o campo se destaca como determinante do conhecimento a ser produzido. O campo é o lugar no qual o pesquisador coleta os dados que, interpretados, discutidos e analisados, constroem os significados buscados.”

Dentre os tipos de pesquisa classificados quanto aos procedimentos, Thiollent (1986) descreve a Pesquisa Ação como:

“...um tipo de pesquisa social com base empírica que é concebida e realizada em estreita associação com uma ação ou com a resolução de um problema coletivo e no qual os pesquisadores e os participantes representativos da situação ou do problema estão envolvidos de modo cooperativo ou participativo”.

Thiollent (1986) explica ainda que a pesquisa-ação possui início (fase exploratória) e fim (divulgação dos resultados) bem definidos, porém, em função das circunstâncias e da dinâmica interna do grupo de pesquisadores no relacionamento com a situação investigada, o planejamento da mesma é muito flexível e não segue uma ordem rígida de fases.

A partir dessas definições e ciente do ambiente complexo que uma Universidade Multicampi possui, a presente pesquisa adotou a pesquisa Aplicada com uma abordagem de pesquisa-ação, justamente por propor uma ferramenta que auxilie na solução prática (Pesquisa Aplicada) de um problema investigado e pelos múltiplos caminhos que podem ser percorridos entre a fase exploratória e a divulgação dos resultados a depender das circunstâncias (pesquisa-ação).

O projeto obedecerá as seguintes etapas metodológicas em sua abordagem: a) Identificação do Problema (fase exploratória); b) Definição do tema; c) coleta de dados; d) modelagem do protótipo da ferramenta a partir do recorte desta pesquisa: dados da SEAVI para o Anuário UNEB em Dados; e) apresentação dos dados levantados aos sujeitos envolvidos (SEAVI), para avaliação e contribuições; f) aprendizagem e, g) apresentação dos resultados desta pesquisa.

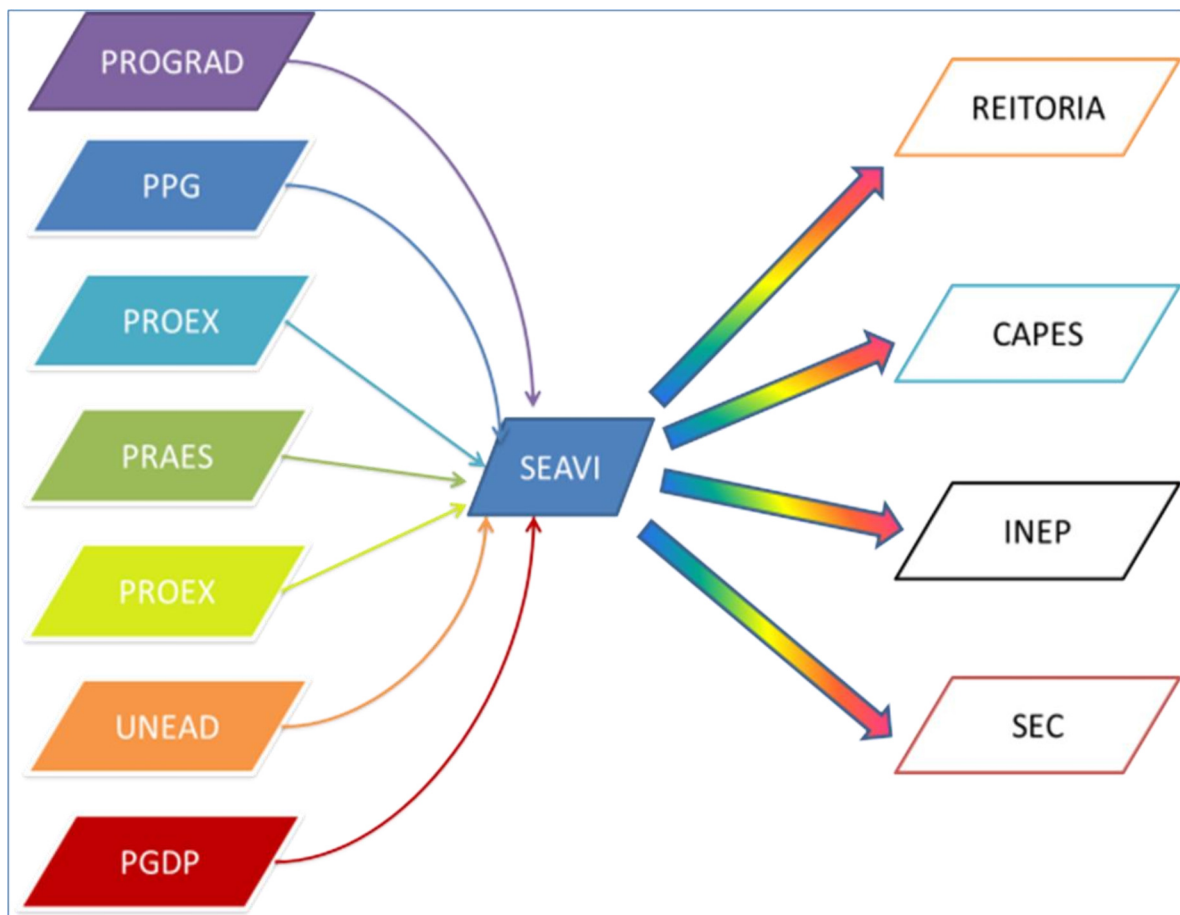
Para viabilizar o desenvolvimento deste projeto, elegeu-se como lócus de pesquisa a SEAVI pelos motivos apresentados no próximo subcapítulo.

3.2. LOCUS DA PESQUISA

Porque utilizar a SEAVI como locus da pesquisa e não as Pró-Reitorias ou os Departamentos Acadêmicos que são as verdadeiras fontes (e consumidores) dos dados da instituição? Ocorre que a SEAVI, sendo a secretaria responsável pela avaliação institucional, em tese, deve ter acesso aos diversos dados institucionais e, justamente por esse acesso aos dados, é constantemente solicitada para fornecer diversas informações estatísticas sobre a UNEB seja para setores internos, pesquisadores ou órgãos reguladores das Instituições de Ensino Superior.

Para responder as solicitações, a Secretaria busca informações nos setores responsáveis pelos dados necessários, que são às próprias Pró-Reitorias, a saber: Pró-Reitoria de Ensino de Graduação - PROGRAD (dados relativos aos alunos dos cursos de graduação), Pró-Reitoria de Pesquisa e Ensino de Pós Graduação - PPG (dados relativos aos alunos dos cursos de Pós-Graduação, bolsistas de

Iniciação Científica e Grupos de Pesquisa), Pró-Reitoria de Assistência Estudantil - PRAES (dados relativos às bolsas de auxílio permanência e afins), Pró-Reitoria de Extensão - PROEX (dados relativos às ações de extensão), Unidade Acadêmica de Educação à Distância – UNEAD (dados relativos aos alunos de cursos de Educação à Distância) e Pró-Reitoria de Gestão e Desenvolvimento de Pessoal – PGDP (dados relativos aos servidores da Universidade, Docentes, Técnicos e Analistas). Somente após o retorno dos dados, a SEAVI faz uma análise dos mesmos, criticando e qualificando para transformá-lo em informação em responder a solicitação num fluxo que se assemelha a Figura 9. Essa “centralidade” otimiza o acesso aos dados, ao passo que ajuda a entender quais são as diversas necessidades por dados institucionais.



Fonte: autor

Figura 9: Fluxo de dados/informações na SEAVI

É exatamente nessa problemática que a SEAVI vem atuando através de seu programa de reorganização das bases de dados e reorientação dos processos acadêmicos de modo a qualificar a informação gerada pela Universidade sobre a

Universidade. Uma parte desse programa fundamental é a elaboração e publicação do Anuário UNEB em Dados, documento que cataloga e sistematiza informações acadêmicas a partir de diversas bases informacionais da Universidade.

Idealizado pelo Secretário Especial de Avaliação Institucional da UNEB (SEAVI), professor Ivan Novaes, o Anuário começou a ser construído ainda no início de 2015. Planejado, tal qual o programa que integra, para ter natureza plurianual, teve sua primeira edição publicada em 2015, trazendo os dados relativos ao ano de 2014. Fruto de articulação com as diversas Pró-Reitorias e Departamentos Acadêmicos, o documento só foi finalizado após diversas reuniões para coleta, análise, apresentação e revisão dos dados. O processo, mesmo que utilize consultas em ferramentas informatizadas, ainda é um processo *ad hoc*, pois para cada “assunto” (graduação, extensão, vestibular, etc...) é necessário escrever uma série de consultas em linguagem SQL que, além de não ser algo inteligível para o usuário final, invariavelmente requer ajuste na codificação.

Durante o levantamento, alguns problemas foram detectados, como informações incompletas, campos não preenchidos ou informações incorretas, e pontos de melhorias foram identificados para o processo de inserção de dados nos sistemas que impedissem novas inconsistências e/ou redundância de informação nas bases dos sistemas. Além disso, um trabalho de reorganização das bases teve início para tratar as informações incorretas cadastradas.

Em dezembro de 2015, o Anuário UNEB em Dados 2015 - Base 2014¹⁵ (Figura 10) foi oficialmente lançado numa reunião do Conselho Universitário (CONSU) da UNEB representando, segundo o Núcleo de Jornalismo da própria Assessoria de Comunicação da UNEB (ASCOM), “uma conquista para os diversos segmentos da Instituição, pois viabiliza uma visão geral sobre os Departamentos da Universidade e os cursos ofertados¹⁶”.

¹⁵ Disponível em: http://www.uneb.br/files/2015/09/anuario_uneb-web.pdf

¹⁶ Disponível em: <http://www.uneb.br/2015/12/11/uneb-investe-em-transparencia-e-divulga-documento-com-dados-academicos/>



Fonte: Anuário Uneb em Dados 2015 – Base 2014

Figura 10: Capa do Anuário UNEB em Dados 2015 - Base 2014

A edição inicial já trazia informações gerais sobre os diversos departamentos como Contato, breve caracterização e site para maiores informações, bem como dados sobre seus respectivos cursos de graduação (Figura 11), tais quais Código INEP, Nome, ano de início, Grau Acadêmico, Turno, Vagas (Vestibular e SiSU), Concorrência (Vestibular e SiSU), Discentes Matriculados (Não Cotistas, Cotista Negro, Cotista Indígena e Total), percentual de Evasão e Conceito (ENADE e CPC).

Além dos dados relativos aos cursos de graduação, a primeira edição do Anuário também já trazia algumas informações sobre Processo Seletivo, Apoio Social a Estudantes (Bolsa Auxílio), Pesquisa e Pós-graduação, Corpo Docente e

DEPARTAMENTO DE EDUCAÇÃO (DEDC - XI)

Contato: (75) 3261-2062

Caracterização: o Departamento de Educação (DEDC – XI) do Campus XI da UNEB, em Serrinha, originou-se da Faculdade Estadual de Serrinha (FES), criada no ano de 1988, e, posteriormente, integrada à Universidade do Estado da Bahia. A partir da reestruturação das Universidades Estaduais da Bahia, em 1997, a FES passou a ser denominada Departamento de Educação (DEDC – XI).

Informações: <http://www.uneb.br/serrinha/dedc/>

Cursos de Graduação Regular Ofertados: Bacharelado em Administração, Licenciatura em Geografia, Licenciatura em Pedagogia.

Cursos do Programa Nacional de Formação de Professores (PARFOR): Licenciatura em Pedagogia, Licenciatura em Ciências Biológicas, Licenciatura de Educação Física, Licenciatura em Geografia, Licenciatura em História, Licenciatura em Letras – Língua Portuguesa, Licenciatura em Matemática.

Tabela 45 - Cursos de Graduação Regular Presencial, DEDC - XI

Código INEP	Nome	Ano Início	Grau Acadêmico	Turno	Vagas		Concorrência		Discentes Matriculados				Evasão %	Conceito	
					Vestibular	SiSU	Vestibular	SiSU	Não Colistas	Colista Negro	Colista Indígena	Total		ENADE	CPC
68567	Administração	2003	Bacharelado	N	45	5	7,18	434,40	107	78	1	186	5,91	3	3
91797	Geografia	2006	Licenciatura	M	32	8	3,53	140,00	94	82	2	178	10,67	4	4
83394	Pedagogia	2004	Licenciatura	M	30	10	5,97	179,90	222	179	-	401	2,23	2	3
				V	30	10	3,20	206,10							
TOTAL					137	33	-	-	423	339	3	765	-	-	-

Fonte: Sistema SAGRES

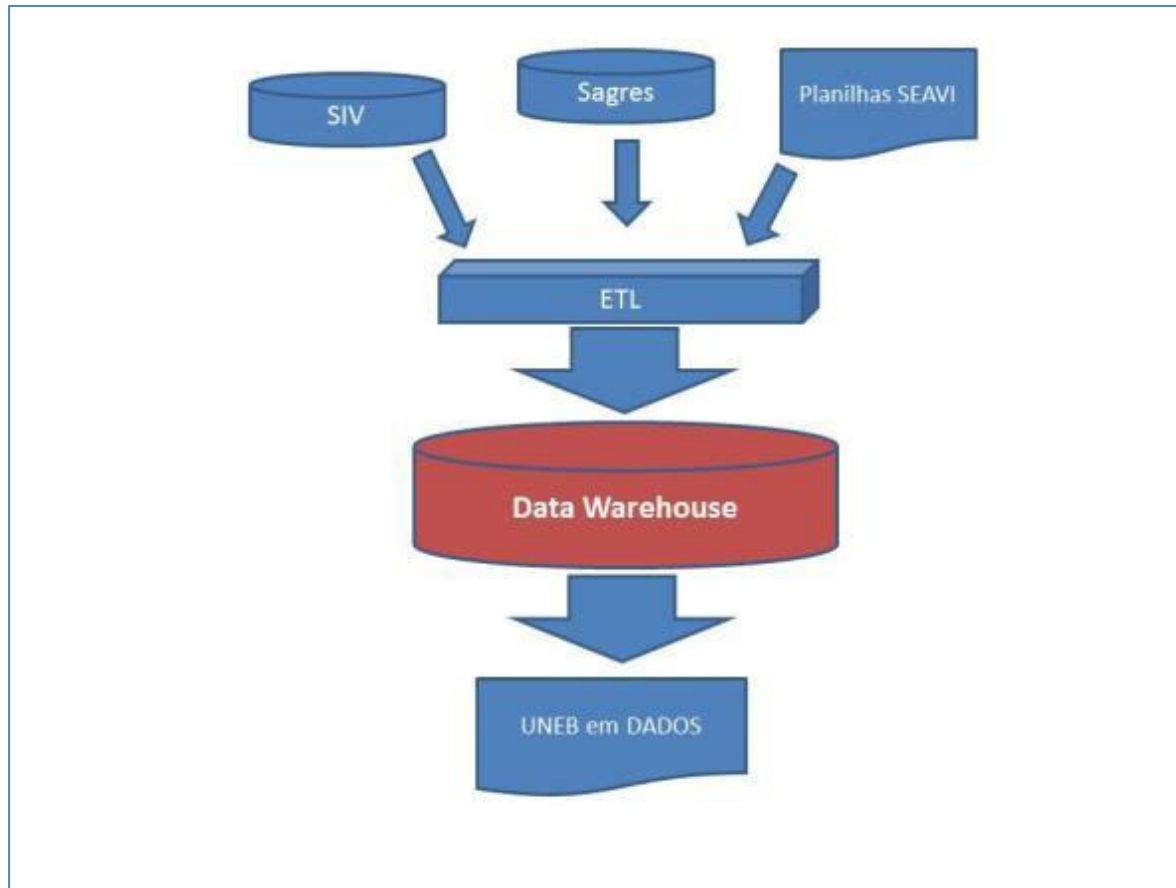
Fonte: Anuário Uneb em Dados 2015 – Base 2014

Figura 11: Exemplo de dados Cursos Graduação do Anuário

Corpo Técnico. A segunda edição do Anuário (Anuário UNEB em Dados 2016 – Base 2015) estava em fase de conclusão quando do término do presente trabalho e já acrescentava algumas novas informações como: dados sobre a Extensão universitária, o Sistema de Bibliotecas (SISB), a Editora da UNEB (EDUNEB) e o Serviço Médico, Odontológico e Social (SMOS). Com sua consolidação e a ampliação do escopo apresentada já na segunda edição, o Anuário UNEB em Dados apresenta-se como uma ferramenta valiosa, não apenas para o grupo gestor da Universidade, mas para toda a comunidade científica que passa a ter uma fonte única e consolidada de dados sobre a instituição.

A ferramenta de B.I. aqui proposta (um Data Warehouse) buscará atuar na coleta tratamento e disponibilização de dados a partir da base de dois sistemas atualmente isolados, a saber: Sistema de Inscrição do Vestibular e o Sagres, além de planilhas que a SEAVI disponha para elaboração do Anuário UNEB em DADOS

(Figura 12) de modo que a coleta de informações para o referido Anuário passe a ser mais automatizada e não mais um processo *ad hoc* como hoje é feito.



Fonte: autor

Figura 12: arquitetura básica do BI (protótipo)

O desenvolvimento do DW poderá auxiliar a gestão da UNEB na verificação de possíveis inconsistências que dificultam o acesso às informações necessárias para tomada de decisão, como a distribuição das matrículas nos diversos cursos/departamentos/campi e a evolução desse quantitativo ao longo dos anos, segmentados por sexo, , ao passo que possibilitará a identificação de pontos de melhorias nos indicadores da instituição frente aos processos avaliativos aos quais é submetida e facilitará o acesso de pesquisadores aos dados sobre a Universidade.

4. DESENVOLVIMENTO

Nesse capítulo é descrito o processo de desenvolvimento do projeto. Desde a fase exploratória até a apresentação dos resultados da pesquisa e do produto. No subcapítulo 6.1 (Levantando Requisitos) narro a fase exploratória inicial junto a SEAVI e SGC, a definição do tema da pesquisa e a colocação do problema, bem como a especificação do escopo do protótipo. Em seguida, no subcapítulo 6.2 (Desenvolvendo a Ferramenta), descrevo todo o processo de desenvolvimento da ferramenta em si, desde a modelagem do protótipo até os ajustes finais. Já no capítulo 6.3 (Análise dos Resultados), é relatada a análise e apresentação dos resultados desta pesquisa.

4.1. LEVANTANDO REQUISITOS

O primeiro contato com a SEAVI se deu em 01 de abril de 2015 para, ciente da necessidade premente de acesso ágil a indicadores relevantes e confiáveis, conhecer mais sobre a Secretaria. Na oportunidade, a reunião se deu informalmente e além do então Secretário, Prof. Ivan Novaes, participaram também os três membros da equipe da Secretaria, bem como o Administrador de Banco de Dados (DBA) da GERINF, que à época, desenvolvia um trabalho de conclusão de curso para Graduação com foco num Data Warehouse baseado nos dados do Vestibular da UNEB.

Apresentados o projeto do DBA e a presente proposta, o secretário explanou sobre o programa de reorganização das bases de dados e reorientação dos processos acadêmicos, que possuía os seguintes propósitos:

- Identificar informações de âmbito acadêmico;
- Identificar inconsistências nas bases de dados;
- Buscar confiabilidade;
- Reorganizar os dados, para justamente garantir a confiabilidade das informações;

- Reorientar os processos de modo a evitar o surgimento de novas inconsistências nas bases.

Na ocasião, o Prof Ivan Novaes discorreu sobre a importância do apoio da UDO/GERINF para o projeto e fez uma breve apresentação de sua ideia para o Anuário UNEB em Dados, à época ainda chamado de “Revista UNEB em Dados”, como proposta de divulgação de dados gerais sobre a UNEB que até então não estavam disponíveis numa única fonte e com foco na confiabilidade dos dados, uma vez que o trabalho envolveria o levantamento, crítica e tratamento dos dados antes da publicação (praticamente o desenvolvimento de um DW).

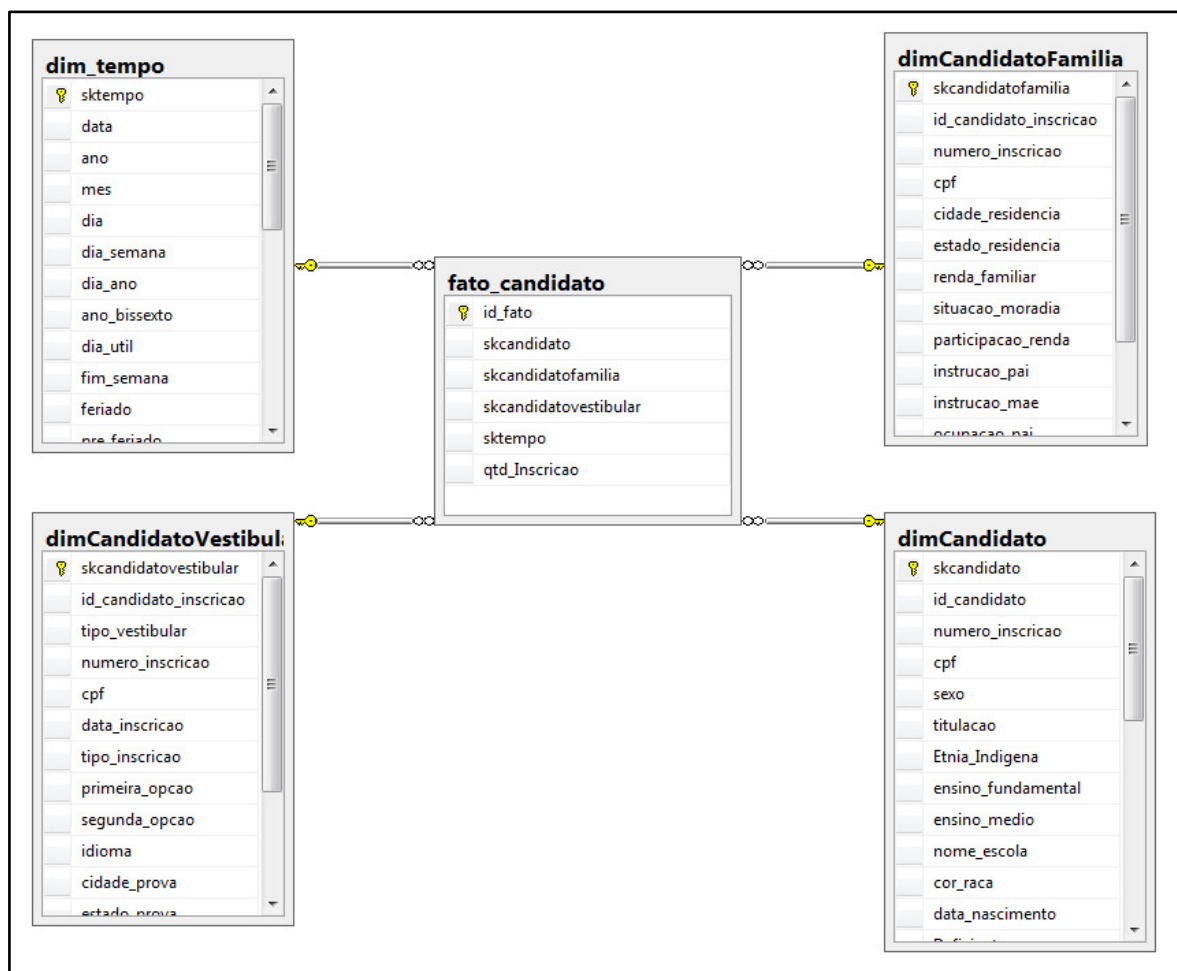
Finalizada a reunião, ficou claro que os interesses do presente trabalho e do programa da SEAVI eram bastante comuns. Até então, a ideia do projeto era auxiliar de forma pontual num problema genérico de dados que sempre eram solicitados à SEAVI, mas com o programa em andamento, estava estabelecido o tema da pesquisa: o levantamento de dados para a composição do Anuário UNEB em Dados. E um grupo de trabalho, ainda que de forma incipiente, começava a se formar entre membros da SEAVI e GERINF/UDO.

No segundo encontro do grupo, já com a participação do Secretário Geral de Cursos, iniciaram-se as discussões sobre a melhor forma de apresentação dos dados na Revista. Essa definição de estrutura seria importante para entender quais bases seriam acessadas e quais dados seriam levantados numa primeira consulta. Enquanto era definido o formato da revista, todos concordaram com uma questão: os dados consultados iriam popular uma base centralizada exclusiva para o projeto como modelo de referência.

Ao término da reunião, ficou decidido que os esforços seriam envidados para coletar os dados que fossem possíveis de levantar, criticar e tratar ainda no horizonte do ano corrente (2015), levando em consideração que haveria o tempo de diagramação, revisão e impressão da revista. Com isso em mente e tendo o professor definido que a revista seria anual, então optou-se inicialmente por levantar os dados sobre os departamentos, como histórico e endereço, além dos dados de 2014 sobre o vestibular e matrículas da Graduação, uma vez que, na época, nem todos os programas de Pós-Graduação tinham todos os dados

cadastrados no sistema acadêmico. Um ponto de melhoria para o sistema acadêmico foi encontrado na reunião, uma vez que foi levantado que havia um campo para o endereço do campus que não vinha sendo preenchido e o Secretário encarregou-se de providenciar a atualização dos referidos campos.

Na terceira reunião, Aldo Barbosa, analista de sistemas da GERINF que ficava alocado na SEAVI justamente para apoiar a secretaria nas questões mais técnicas do sistema acadêmico, juntamente com nosso DBA, informou que haviam feito algumas consultas nas bases do Sistema Acadêmico para iniciar o desenvolvimento do Data Warehouse e encontraram alguns problemas como, campos não preenchidos ou preenchidos de forma não padronizada, o que dificultava a sumarização dos dados. Foi inclusive informado de que o processo de ETL poderia tentar sanitizar alguns dados inconsistentes, mas além de ser um trabalho mais demorado, inconsistências não previstas poderiam gerar novas inconsistências no futuro.



Fonte: obtido da documentação interna da UDO/GERINF

Figura 13: Diagrama do modelo de dados sobre os candidatos do vestibular da UNEB

Em tempo, o DBA apresentou a todos o os presentes o modelo de dados do DW de seu projeto (Figura 13), com foco no candidato de Vestibular, a partir da base de dados do sistema de vestibular da UNEB. Na oportunidade, também demonstrou como utilizar funções de Business Intelligence para relatórios dinâmicos do Microsoft Office Excel (Figura 14).

No diagrama apresentado (Figura 13) temos a tabela fato_candidato que armazena o total de inscrições (qtd_inscricoes), única métrica utilizada no protótipo em questão, e o código de cada uma das tabelas de dimensões do modelo, a saber: tempo (onde ficavam armazenadas as informações sobre o tempo das inscrições como ano, mês, dia, se era um ano bissexto, etc...), candidato (com os dados pessoais dos candidatos), candidato_familia (com os dados relativos ao grupo familiar dos candidatos) e candidato_vestibular (com os dados da inscrição do candidato, como a primeira e segunda opção de curso, cidade de prova escolhida, idioma da prova estrangeira, etc...).

O DBA então demonstrou como, utilizando o recurso de Tabela Dinâmica do Excel (Pivot Tables), ficava fácil e intuitivo analisar os dados que já haviam sido coletados previamente da base transacional. No exemplo da Figura 14 é possível verificar o total de inscrições por grupo étnico nos vestibulares entre os anos de 2012 e 2014, sendo que nenhuma linha de código é necessária, dando total liberdade ao usuário final para elaborar relatórios dinâmicos. O processo de extração, transformação e carga dos dados será melhor explicado no próximo subcapítulo, quando descrevo o desenvolvimento da ferramenta que foi utilizada nesta pesquisa.

Rótulos de Linha	Total de inscrições	Percentual Candidato
2012	41301	100,00%
AMARELO(A) (DE ORIGEM ASIÁTICA)	283	0,69%
BRANCO(A)	7424	17,98%
INDÍGENA	316	0,77%
NÃO INFORMADO	1	0,00%
PARDO(A)	21689	52,51%
PRETO(A)	11588	28,00%
2013	39664	100,00%
AMARELO(A) (DE ORIGEM ASIÁTICA)	315	0,79%
BRANCO(A)	7165	18,06%
INDÍGENA	297	0,75%
NÃO INFORMADO	6	0,02%
PARDO(A)	20476	51,62%
PRETO(A)	11405	28,75%
2014	53942	100,00%
AMARELO(A) (DE ORIGEM ASIÁTICA)	393	0,73%
BRANCO(A)	9708	18,00%
INDÍGENA	440	0,82%
NÃO INFORMADO	18	0,03%
PARDO(A)	28168	52,22%
PRETO(A)	15215	28,21%
Total Geral	134907	100,00%

Fonte: Microsoft Excel 2010

Figura 14: exemplo de relatório dinâmico no Excel

Após várias reuniões com diversas consultas às bases de dados de alguns sistemas, seguidas de análises dos resultados e consequentes ajustes, ao final de 2015 nasce o Anuário UNEB em Dados 2015 – Ano base 2014¹⁷ com dados consolidados sobre o vestibular, cursos de graduação regular presencial, cursos ofertados pelo Programa de Formação de Professores (PARFOR), bolsa auxílio, alguns dados sobre a pesquisa e pós-graduação, além de informações sobre o corpo docente e técnico da universidade.

Com as exaustivas consultas nas bases de dados, seguidas de análises das informações coletadas e melhorias dos códigos das consultas, fruto de conversas com os setores fontes dos dados, emerge então o problema a ser pesquisado: a otimização do processo de coleta para o Anuário UNEB em Dados a partir do uso de técnicas e ferramentas de Business Intelligence.

A primeira aproximação para o levantamento dos requisitos do protótipo, por sugestão do DBA, numa tentativa de acelerar o desenvolvimento da ferramenta, foi feita com a utilização de uma matriz de necessidades, técnica desenvolvida por

¹⁷ Disponível em < <http://www.uneb.br/anuario-uneb-em-dados> Acesso em 30 de Julho de 2016.

Oliveira (2013) para apresentar o mapeamento dos dados que estão armazenados nos sistemas transacionais das organizações (banco de dados, planilhas eletrônicas, arquivos texto, dentre outros) e que servirão de base para a construção de um Data Warehouse.

Para tal, uma reunião inicial com a equipe da SEAVI foi realizada para apresentar a matriz e solicitar que preenchessem a mesma e nos enviassem por email para análise e implementação. No entanto, a abordagem se mostrou ineficaz, visto que houve divergência no entendimento dos campos e forma de preenchimento da matriz. Ainda assim, após uma nova reunião para sanar algumas dúvidas e novo prazo para entrega, uma versão preliminar da planilha (Figura 15) foi elaborada.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	
1	Matriz de necessidades / UNEB																																	
2	Fato	Técnico										departamento	Curso										Docentes											
3	Dimensão	Total_Técnicos_Campus	Total_Técnicos_Departamento	Total_Técnicos_Gênero_Departamento	Total_Técnicos_Formação_Departamento	Total_Técnicos_Gênero_Formação_Departamento	Total_Técnicos_Deficiência_Departamento	Total_Técnicos_Cor/Raça_Departamento	Total_Técnicos_DiscentePG_Departamento	Total_Técnicos_DiscenteGRAD_Departamento	Total_Técnicos_Idade_Departamento	Total_Departamentos_Campus	Total_vagas_departamento	Total_Cursos_Departamento	Total_Cursos_Nivel_Grau_Acadêmico_Modalidade	Total_Cursos_Nivel_Turno_Departamento	Total_Cursos_Programas_Especiais_Departamento	Total_Docentes_Departamento	Total_Docentes_Curso	Total_Docentes_Titulação_Departamento	Total_Docentes_Titulação_Curso	Total_Docentes_Titulação_Nivel_Curso_Departamento	Total_Docentes_Titulação_PIBID_Departamento	Total_Docentes_Titulação_PIBIC_Departamento	Total_Docentes_DiscentePG_Departamento	Total_Docentes_Titulação_CH_Departamento	Total_Docentes_Titulação_Área de Conhecimento	Total_Docentes_Titulação_Área de Conhecimento	Total_Docentes_Área de Conhecimento (Inicial)	Total_Docentes_Área de Conhecimento (Inicial)	Total_Docentes_Área de Conhecimento (Inicial)	Total_Docentes_Gênero_Departamento	Total_Docentes_Gênero_Curso	Total_Docentes_Gênero_Titulação_Departamento
4	Campi																																	
5	Descrição dos Campi																																	
6	Departamentos																																	
7	Descrição dos Departamentos																																	
8	Cursos																																	
9	Descrição do Curso																																	
10	Vagas																																	
11	Docentes																																	
12	Descrição dos docentes																																	

Fonte: Microsoft Excel 2010

Figura 15: parte da matriz necessidade preliminar

Apesar do potencial da matriz necessidade, adequar-se a uma nova metodologia iria requerer tempo extra e ficou então decidido que uma abordagem mais direta e objetiva seria necessária. Como o objetivo maior era otimizar o levantamento de dados do Anuário UNEB em Dados, surgiu então a idéia de levantar os dados a partir das planilhas que compuseram o Anuário 2015 – Base 2014, em particular as que diziam respeito às matrículas dos departamentos, visto que o foco da pesquisa seria justamente o corpo discente. Em outras palavras, os requisitos para o DW que utilizados no protótipo de *BI* desse projeto foram

levantados tomando como base os dados sobre o corpo discente, previamente formatados para o Anuário UNEB em Dados 2015. Mais especificamente, as seguintes perguntas foram utilizadas como ponto de partida para o desenvolvimento da ferramenta:

- Qual o total de discentes matriculados nos diversos cursos da UNEB entre os anos de 2012 e 2016?
- Quantos desses alunos estão matriculados em cursos de graduação?
- Quantos desses alunos estão matriculados em cursos de pós-graduação?
- Quantos dos alunos matriculados em cursos de graduação estão em cursos de bacharelado?
- Quantos dos alunos matriculados em cursos de graduação estão em cursos de licenciatura?
- Quantos desses alunos são de ampla concorrência?
- Quantos desses alunos são cotistas indígenas?
- Quantos desses alunos são cotistas étnicos?
- Qual o código Inep dos cursos de graduação da UNEB?
- Qual o ano de criação dos cursos de graduação da UNEB?

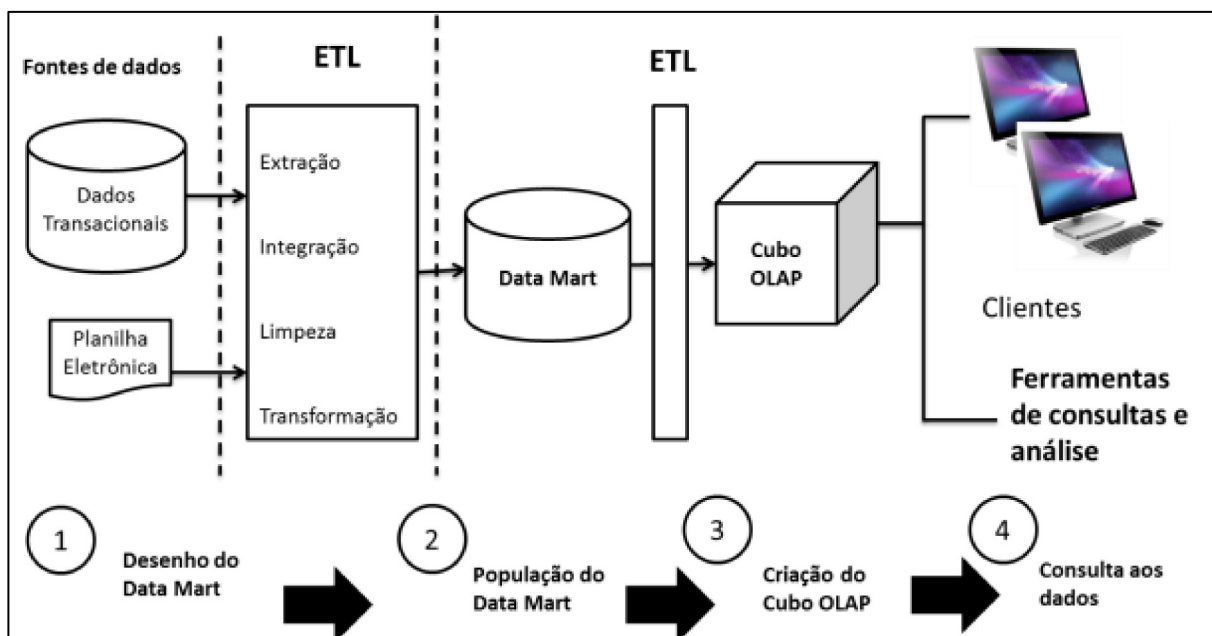
Essas foram as perguntas que nortearam o desenvolvimento da ferramenta que poderá auxiliar a gestão da UNEB aprimorando a visualização e análise de seus dados para um melhor atendimento das demandas dos serviços da Universidade que estão, conforme Plano de Ações Prioritárias 2014 (PAP)¹⁸, em permanente ampliação. Essa mesma ferramenta propicia o empoderamento de seus gestores que terão acesso direto e ágil aos dados detalhados sobre os discentes matriculados da UNEB, tanto para tomada de decisão quanto para atendimento às instituições reguladoras externas.

¹⁸ Disponível em < http://www.uneb.br/files/2015/03/relatorio_pap.pdf>. Acesso em 30 de Novembro de 2016

4.2. DESENVOLVENDO A FERRAMENTA

Para o desenvolvimento do protótipo, foram utilizadas ferramentas da Microsoft¹⁹ em todas as etapas (Figura 16). Mais especificamente, foi utilizado o SQL Server Integration Services (SSIS) para o processo de ETL, desde a coleta dos dados armazenados na base transacional até a integração no Data Warehouse. Para a montagem do cubo OLAP foi utilizado o SQL Server Analysis Services (SSAS), para acessar os dados armazenados no cubo e gerar relatórios e gráficos, foram utilizados os recursos de Tabela Dinâmica e Gráfico Dinâmico do Excel 2010. Por fim, foi utilizado o SQL Server para modelagem e armazenamento dos dados do DW. Salientando que os dados originalmente estavam armazenados numa base transacional do SQL Server 2008 e em planilhas do Microsoft Excel.

4.2.1. DESENHANDO E POPULANDO O DATA WAREHOUSE



Fonte: autor

Figura 16: resumo do desenvolvimento do protótipo

¹⁹ Empresa transnacional norte americana com sede em Redmond, Washington, que desenvolve, fabrica, licencia, apoia e vende softwares de computador, produtos eletrônicos, computadores e serviços pessoais.

A primeira etapa para o desenvolvimento do Data Warehouse (que, conforme já explicado, aqui será um Data Mart, dado o escopo de informações que serão elencados) é o desenho do modelo que será implementado. Seguindo as orientações de Kimball (2013), foi adotado o modelo estrela para a modelagem. Algumas iterações foram feitas até a modelagem final. Partindo da planilha dos cursos de graduação que consta no Anuário UNEB em Dados 2015- Base 2014 (Figura 11), alguns rascunhos foram feitos apenas para fins de teste da plataforma (Figura 17).

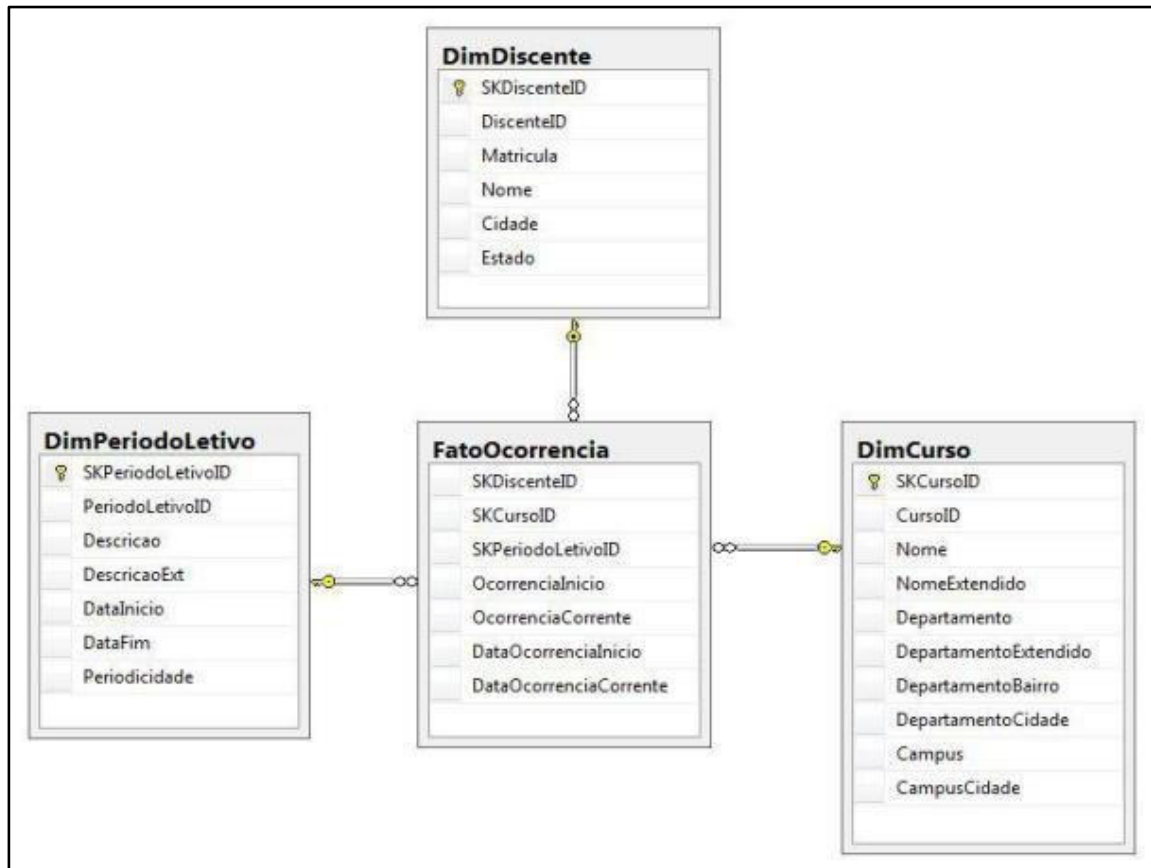


Fonte: autor

Figura 17: primeiro rascunho da modelagem do DW

Nesse primeiro rascunho, ainda havia certa dúvida entre Matrícula e Ocorrência para o nome da tabela fato. No fim, foi escolhido o nome matrícula e desenhadas as dimensões Discente (dados relativos aos discentes), Período Letivo, Curso, Campus, Departamento e Estado, que descreve o status da matrícula no período letivo (matriculada, trancada, abandonada, etc...). Ao desenhar o modelo no SQL Server, optou-se por agregar as informações de departamento e campus dentro da dimensão curso para simplificar o modelo. No

final da primeira iteração, tínhamos um modelo simplificado que adotamos para verificar a viabilidade da ferramenta (Figura 18).



Fonte: obtido da documentação interna da UDO/GERINF

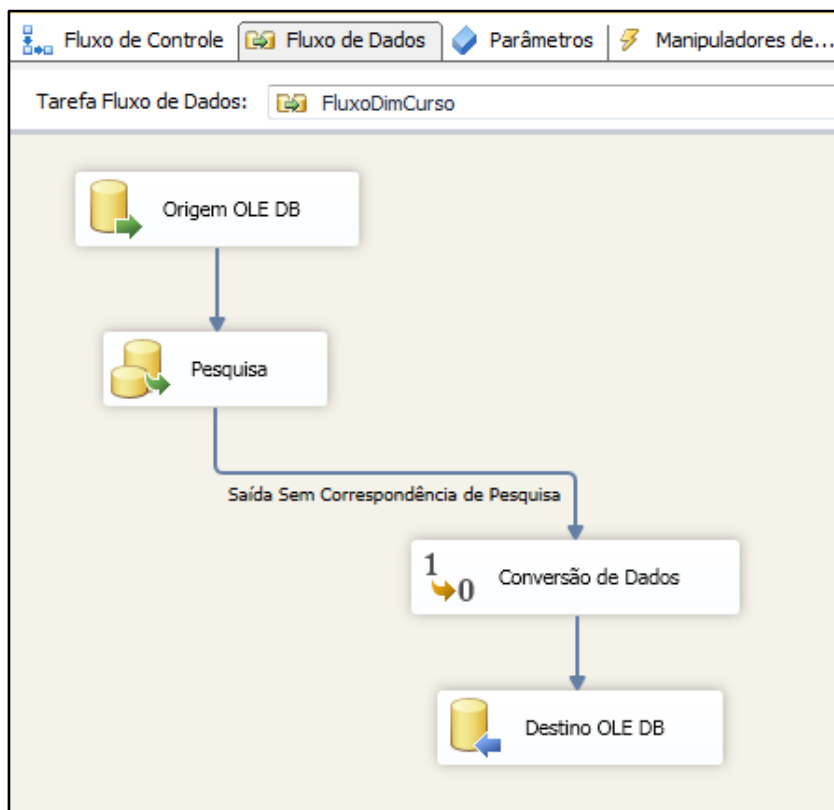
Figura 18: primeira versão do modelo do Data Mart

Uma vez desenhado o modelo, seguimos para a extração dos dados das bases de dados transacionais de origem (aqui temos a base de dados relacional do SAGRES e algumas planilhas eletrônicas do MEC) para popular a base de dados analítica de destino que será nosso DW. A idéia é escrever as consultas em SQL apenas uma vez, para que os usuários finais possam abstrair desse processo de codificação que até então era realizado para cada Anuário. Com as consultas já elaboradas e devidamente testadas e validadas juntamente com a equipe da área de negócio (no caso, a Secretaria Geral de Cursos), passamos para a etapa de ETL propriamente dita, utilizando o SSIS. A Microsoft dá as seguintes informações sobre o SSIS²⁰:

²⁰ Disponível em: <https://msdn.microsoft.com/pt-br/library/ms169917.aspx>

[...]é uma plataforma para a criação de soluções para integração de dados de alto desempenho, incluindo também pacotes ETL (extração, transformação e carregamento) para armazenamento de dados. O SSIS inclui ferramentas gráficas e assistentes para criação e depuração de pacotes; tarefas para execução de funções de fluxo de trabalho como, por exemplo, operações de FTP, execução de instruções SQL e envio de mensagens de email; fontes de dados e destinos para extração e carregamento de dados; transformações para limpeza, agregação, junção e cópia de dados;

O SSIS basicamente trabalha preparando pacotes e fluxos de trabalho. Nele nós definimos pacotes com fluxos de controle (chamado de control flow no SSIS) e fluxos de dados (chamado de data flow no SSIS) para extração, transformação e carga dos dados. Um fluxo de dados básico tem uma conexão com a base de origem (Extração), uma transformação intermediária (Transformação) que pode ou não existir e uma conexão com a base de destino (Carga). A Figura 19 exemplifica visualmente um fluxo de dados básico com as três etapas.

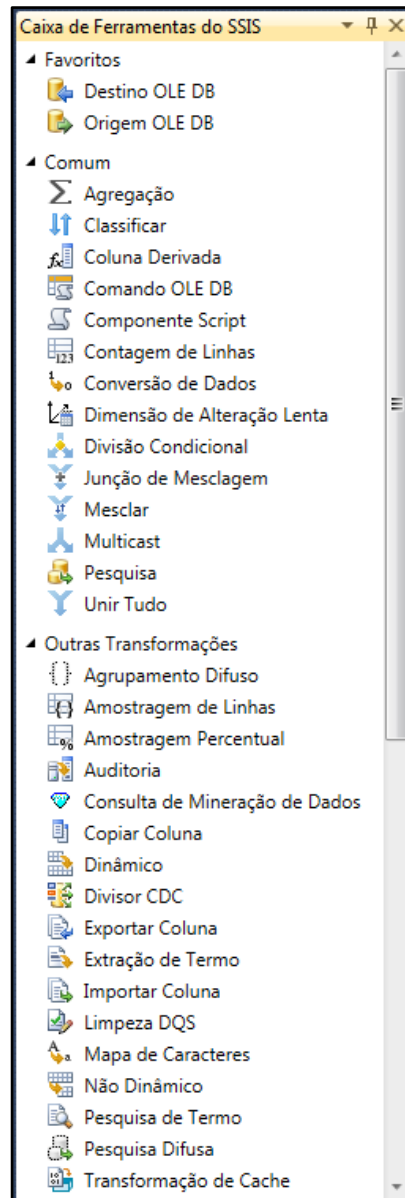


Fonte: SQL Server Integration Services

Figura 19: desenho do fluxo de dados (Data Flow) da Dimensão Curso no SSIS

Aqui a ferramenta inicialmente se conecta ao banco de origem, que pode ser um banco relacional (que foi o nosso caso), um arquivo de texto ou uma planilha eletrônica, e utiliza as consultas elaboradas para a extração dos dados. Na segunda etapa, são definidas as conversões e padronizações de dados, que se fizerem necessárias para que os mesmos sejam devidamente sumarizados e preparados para a terceira e última etapa que é a carga na base consolidada, o Data Warehouse em si previamente modelado no SQK Server. Na Figura 20, temos o menu de caixa ferramentas do SSIS onde podemos identificar a ferramenta “Origem OLE DB” para a Extração, “Destino OLE DB” para a carga e uma série de outras destinadas a transformações dos dados agrupadas no sub-menu “Outras Transformações”.

Na primeira iteração, optamos por realizar a carga dos dados sem qualquer transformação para ter uma visão de como os dados estavam gravados na base operacional. Decidimos que apenas após uma primeira análise de como os dados estavam na base, faríamos as devidas transformações. Sendo assim, podemos dizer que na primeira iteração do protótipo, tivemos apenas um processo de EL (Extração e Carga, do inglês Extract and Load).



caixa de ferramentas do SSIS

Figura 20: caixa de ferramentas do SSIS

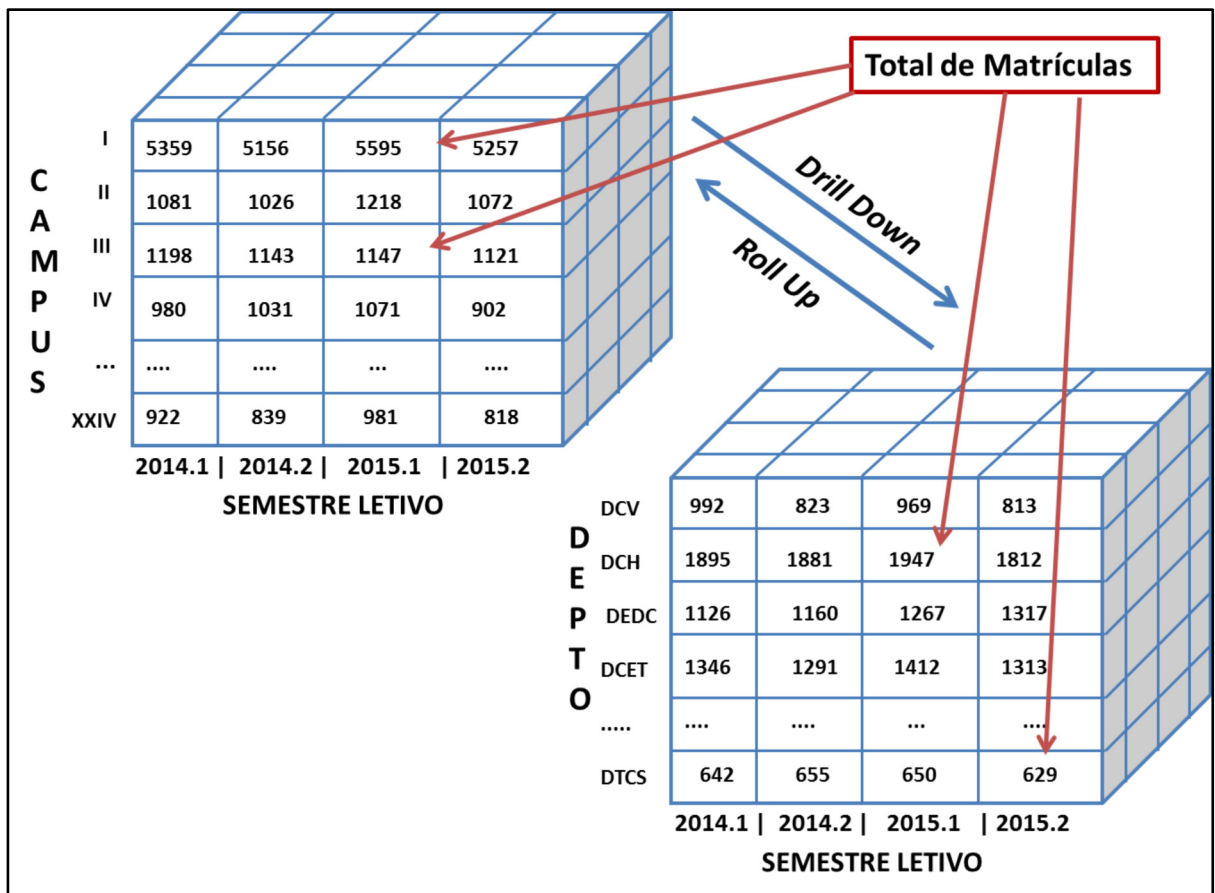
4.2.2. CRIANDO, VISUALISANDO E AJUSTANDO O CUBO

Após povoar o DW com os dados, partimos para a fase de criação do cubo OLAP. Nessa etapa, utilizamos o SQL Server Analysis Services (SSAS) para preparar os dados para visualização analítica. A Microsoft define o SSAS como:

“[...]um mecanismo de dados analíticos online usado em soluções de apoio à decisão e análise de negócios, que fornece os dados analíticos para relatórios de negócios e aplicativos cliente, tais como Excel, relatórios do Reporting Services e outras ferramentas de visualização de dados de terceiros.”.

É no SSAS que criamos o modelo de dados multidimensional (o Cubo OLAP), implantamos o modelo como um banco de dados em uma instância do próprio Analysis Service, processamos o banco de dados para carregá-lo com os dados armazenados previamente no DW, configuramos a atualização dos dados e, em seguida, atribuímos permissões para permitir o acesso a dados pelos usuários finais. É também no SSAS que definimos os cálculos e métricas que encapsulam a lógica de negócios e as interações, como comportamentos de detalhamento e navegação. Abstraindo um pouco da parte técnica, é aqui que relacionamos a tabela fato com as tabelas de dimensões e conseguimos abstrair o modelo relacional de dados para um modelo dimensional, cruzando as diversas dimensões como numa espécie de “cubo mágico” onde podemos alterar as dimensões (atributos) através dos quais analisamos os dados.

Na criação do cubo definimos as medidas que serão analisadas pelo usuário e podemos criar hierarquias entre os atributos para que possamos agrupar os dados e analisá-los num maior ou menor nível de detalhe (técnica chamada de *roll up* e *drill down*). Na primeira iteração do protótipo, criamos como medida principal na tabela fato, o total de matrículas e uma hierarquia entre campus, departamento e curso. Uma abstração da hierarquia definida nessa implementação pode ser vista nos cubos da Figura 21. No primeiro cubo, numa visão mais ampla, temos o total de matrículas por Campus ao longo dos semestres. Fazendo um aprofundamento (*drill down*) nos dados, podemos detalhar essas matrículas por departamento e se quisermos detalhar ainda mais, podemos ter a informação separada por curso.



Fonte: autor

Figura 21: abstração da hierarquia definida no protótipo

Para visualização dos dados, escolhemos o Microsoft Excel 2010, que consegue se conectar com o cubo gerado pelo SSAS e permite criar relatórios dinâmicos através do uso de pivot tables (tabelas dinâmicas). A escolha da ferramenta se dá em função da mesma, sendo amplamente utilizada no âmbito da UNEB, possuir uma interface já reconhecível para os usuários, eliminando a necessidade de treinamentos extensos para o início do acesso e análise dos dados.

Para iniciar o acesso, basta acessar a aba “Dados”, clicar no botão “De Outras Fontes”, escolher a opção “Dos Serviços de Análises” e informar o servidor no qual o Cubo está hospedado. Feito isso, basta escolher o banco de dados (que neste caso é um DW) e o cubo. No momento de definir a forma de exibir os dados na pasta de trabalho, é possível já escolher, além do Relatório de Tabela Dinâmica, o uso de Gráfico Dinâmico. Com isso, ao navegar e alterar as

dimensões de análise, um gráfico vai sendo alterado dinamicamente para facilitar a visualização.

Após escolher a forma como deseja exibir os dados, o usuário terá um menu do lado direito com os campos do cubo (Fato e dimensões) e 4 áreas para onde poderá arrastar os campos para analisar os dados: Filtro, Rótulos de Linha, Rótulos de Colunas e Valores. A área de valores é onde os usuário informa qual a medida da tabela fato será analisada. A partir daí, os dados poderão ser agrupados e cruzados a partir dos atributos definidos nas dimensões, arrastando-os para os Rótulos de Linha e/ou Coluna. É também possível filtrar e ordenar os dados a partir de algum atributo, arrastando o respectivo campo para a área de Filtro de Relatório e escolhendo o intervalo requerido. Na Figura 22 temos o total de matriculas (Valores) nos Departamentos (Rótulo de Coluna), por Cidade de Origem dos Discentes (Rótulos de Linha), Filtrando apenas os registros de 2015.1 e 2015.2 no Campus de Salvador (Filtro de Relatório).

Rótulos de Linha	DCET - CAMPUS I	DCH - CAMPUS I	DCV - CAMPUS I	DEDC - CAMPUS I	Total Geral
Salvador	729	1748	994	1146	4617
Lauro de Freitas	34	61	48	74	217
Camaçari	20	20	24	42	106
	17	75	3	8	103
Feira de Santana	15	25	56	3	99
Símoes Filho	13	16	7	13	49
Planaltino				44	44
Maracas				42	42
Salinas da Margarida	8	24	1	1	34
Lajedo do Tabocal				31	31
Dias d'Ávila	8	6	9	7	30
Vera Cruz	5	8	9	7	29
Conceição do Coité	8	5	8	3	24
Santo Antonio de Jesus	3	2	15		20
Riachão do Jacuípe	3	7	9	1	20
Alagoinhas	5	5	8	1	19
Jequié	4	1	11	3	19
Itaparica	2	8	5	2	17
Valença	1	5	8	1	15
Itaberaba	6	3	5		14

Fonte: Microsoft Excel 2010

Figura 22: exemplo de relatório dinâmico extraído na ferramenta

Como na primeira iteração do protótipo, propositalmente não foi realizada, praticamente nenhuma transformação nos dados, foram então detectadas algumas redundâncias, como o campo escolaridade que apresentava tanto a descrição “nível superior” quanto “graduado”, assim como as descrições “especialista” e “pós-graduado”. Além disso, identificamos uma quantidade significativa de registros relativos ao nível de Pós-Graduação sem a identificação da escolaridade. Na Figura 23 podemos verificar que havia na base 30 registros de Pós-Graduação Stricto Sensu e 238 de Pós-Graduação Lato Sensu sem identificação de escolaridade.

	A	B
1	Rótulos de Linha	Contagem Distinta SK Discente ID
2	Graduação	44792
3	Nível Superior	44792
4	Graduação - EAD	8710
5	Nível Superior	8710
6	Graduação-Programas Especiais	8473
7	Nível Superior	8473
8	Pós-Graduação EAD (Lato Sensu)	1254
9	Especialista	171
10	Pós-Graduado	1083
11	Pós-Graduação Lato Sensu	4404
12		238
13	Graduado	285
14	Pós-Graduado	3881
15	Pós-Graduação Stricto Sensu	1645
16		30
17	Doutor	126
18	Graduado	10
19	Mestre	1479
20	Total Geral	69267

Fonte: Microsoft Excel 2010

Figura 23: exemplo de relatório de escolaridade da primeira iteração do protótipo

A fase de transformação é de fundamental importância para se ter um DW inteligível para o usuário final. É nessa etapa que, além de estruturar os tipos de dados, definimos a padronização das informações a serem apresentadas. No caso do protótipo aqui desenvolvido, nessa etapa, foi padronizado o período letivo, que no sistema de informações acadêmicas trazia descrições diferentes para cursos de níveis diferentes. Na primeira iteração tínhamos os semestres dos cursos de Graduação contendo apenas números e os Cursos de Pós-Graduação com as

letras “PG” para identificação. Dessa forma, para sumarizar os dados de matriculados do semestre 2015.1 de todos os níveis de cursos, seria necessário selecionar os semestres 20151, relativo aos cursos de Graduação e PG151 para os cursos de Pós-Graduação (Figura 24).

	A	B	C	D
1	Contagem Distinta SK Discente ID	Rótulos de Coluna		
2	Rótulos de Linha	20151	PG151	Total Geral
3	Graduação	21635		21635
4	Graduação - EAD	1	1	2
5	Graduação-Programas Especiais	1344		1344
6	Pós-Graduação EAD (Lato Sensu)		640	640
7	Pós-Graduação Lato Sensu	1	604	604
8	Pós-Graduação Stricto Sensu		641	641
9	Total Geral	22981	1886	24866

Fonte: Microsoft Excel 2010

Figura 24: exemplo de relatório de período letivo da primeira iteração do protótipo

Diversos pontos de melhorias foram identificados e passados para a Secretaria Geral de Cursos, uma vez que alguns deles precisam ser alterados na base de origem. Após o uso e verificação da primeira versão do protótipo, novas alterações foram feitas e uma versão mais próxima da idealizada foi desenvolvida.

4.2.3. FINALIZANDO O PROTÓTIPO

Apesar de a primeira versão do protótipo apresentar um avanço no processo de coleta de dados para a elaboração do Anuário UNEB em Dados, várias melhorias ainda eram necessárias para disponibilizar a ferramenta para o usuário final da forma que a mesma havia sido idealizada. No entanto, dada a exiguidade de tempo, algumas melhorias foram implementadas e outras ficaram reservadas para trabalhos futuros. Apresento nesse subcapítulo o modelo de dados da versão final do protótipo desenvolvido durante a pesquisa e um dicionário que facilite sua compreensão.

Importante ressaltar que no desenvolvimento da versão final, novos dados foram inseridos no processo de ETL ainda no SSIS e, diferente da primeira

aproximação, algumas transformações foram realizadas. Essas mudanças impactaram diretamente no processamento do novo Cubo, gerando algumas mensagens de erro.

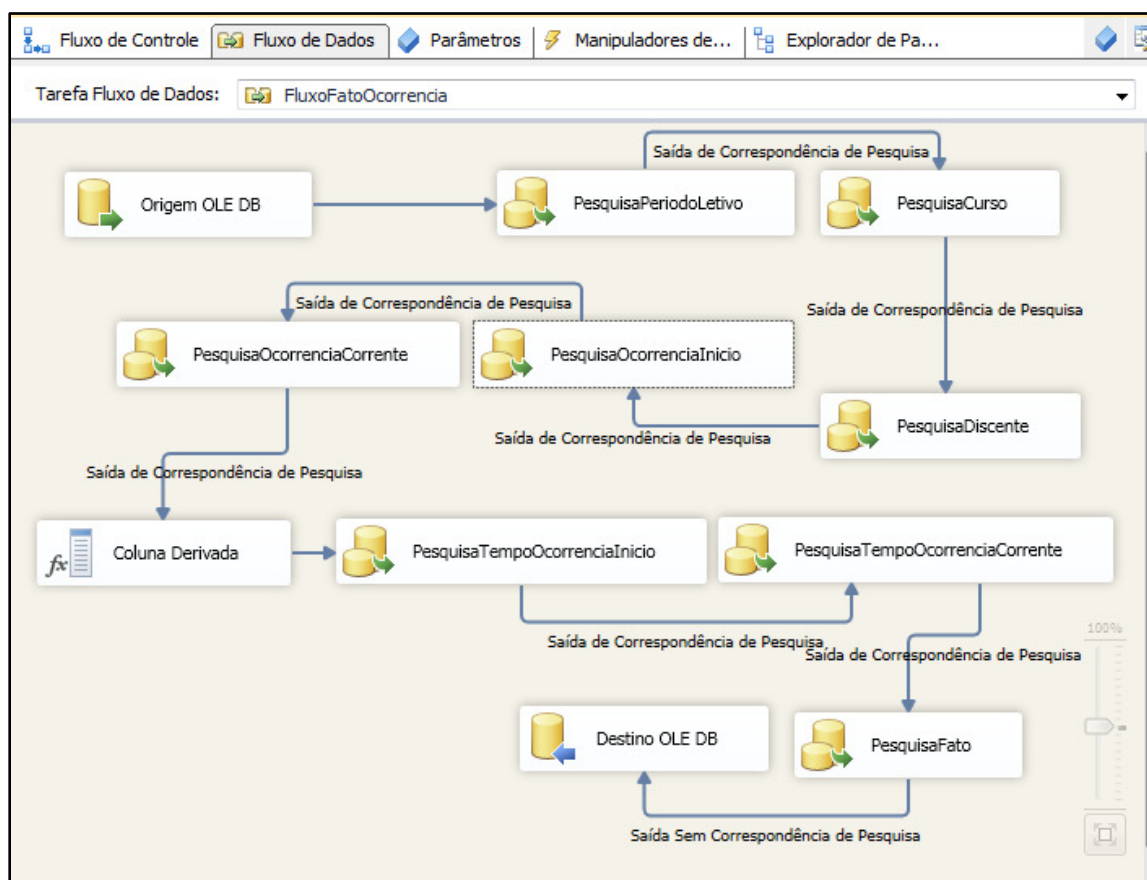
Analisando as mensagens, percebemos que várias delas eram bastante genéricas e apontavam para uma parada no processamento em função de um erro anterior. Precisamos então encontrar qual mensagem referia-se ao erro que iniciou a parada. Após ler todas, encontramos uma que informava sobre o campo “Codigo Inep” que apresentava registros duplicados. Isso ocorre porque na compilação padrão do cubo, o SSAS efetua uma série de verificações e validações da estrutura do DW, sendo que entre estas verificações está uma verificação de duplicidades nos atributos de uma dimensão. Como não poderíamos alterar os dados na base, buscamos uma solução de contorno para finalizar o processamento.

Para solucionar o problema, precisamos alterar as configurações de processamento antes de executá-lo. No caso de um DW, diferentemente de um banco relacional transacional, em algumas circunstâncias é necessário admitirmos duplicidades nos atributos. E quando falamos de atributos estamos falando de qualquer campo da dimensão, ainda que não seja a chave primária. As alterações nas configurações do processamento são feitas na tela de execução do Cubo do SSAS. Para realizá-las, basta clicar em “Alterar Configurações”.

Ao entrar na tela de configurações do processamento, verificamos que há uma aba identificada como “Erros de chave de dimensão” onde está marcada a opção “Usar configuração de erro padrão”. Essa configuração faz com que o SSAS pare o processamento em caso de erro encontrado. Para contornar o problema, alteramos a opção “Ação se houver erro” de “Parar processamento” para “Parar log” e alteramos a opção “Chave duplicada” de “Ignorar erro” para “Relatar e continuar”.

Desta forma, ao encontrar os atributos duplicados, o SSAS reportou o erro encontrado, mas continuou o processamento do Cubo sem gerar novos erros.

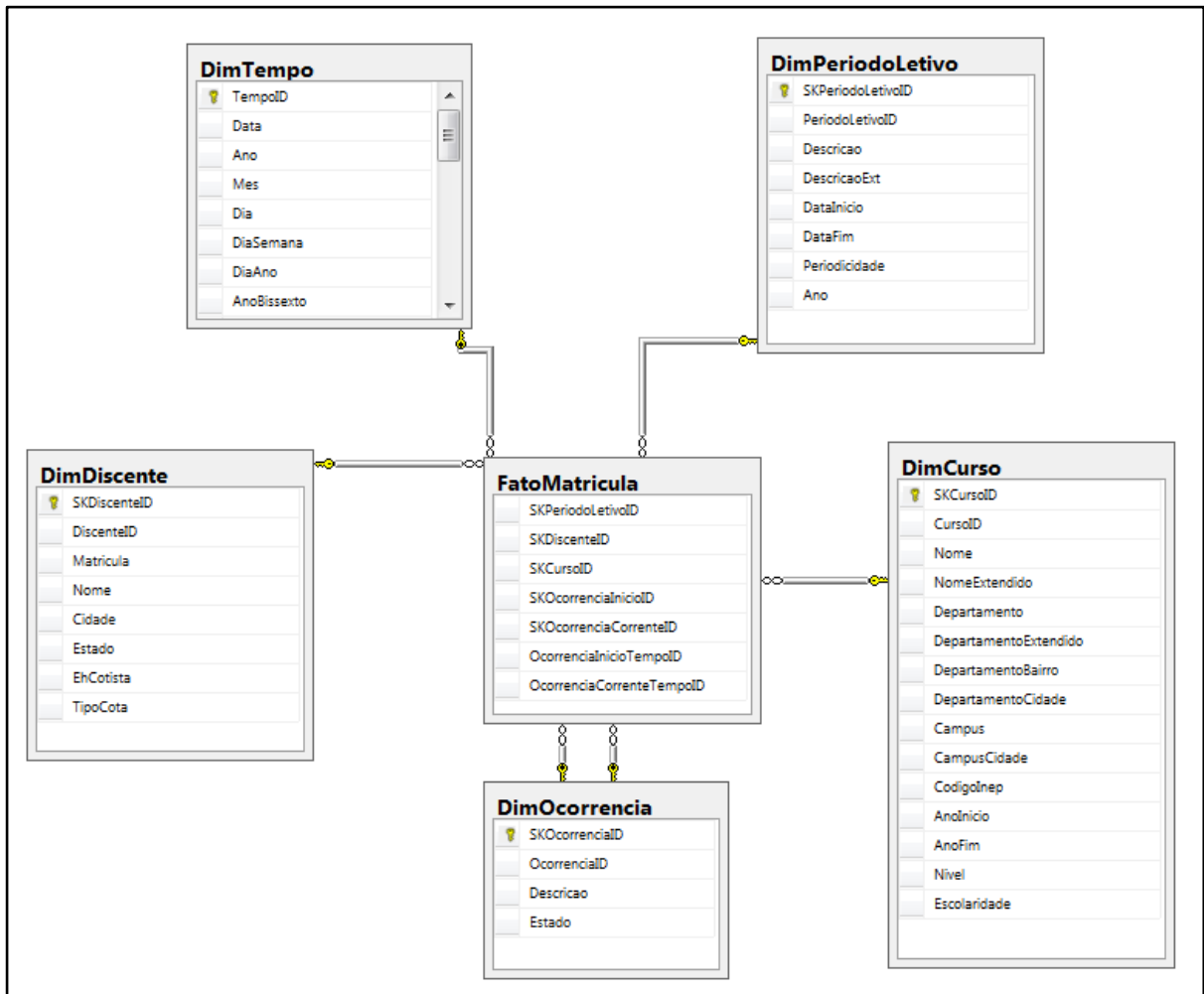
Dentre as mudanças realizadas no processo de ETL, a mais significativa foi a criação de uma coluna derivada na dimensão Ocorrência. A coluna gerada serviu para armazenar o Estado da ocorrência, podendo ser inativo, ativo matriculado ou ativo não matriculado. Dessa forma, o usuário pode rapidamente analisar as ocorrências correspondentes a qualquer desses três estados. O desenho geral do fluxo de dados da tabela Fato no SSIS pode ser visto na Figura 25. Nessa imagem é possível identificar o momento de criação da Coluna derivada logo após o fluxo de controle PesquisaOcorrenciaCorrente.



Fonte: SQL Server Integration Services

Figura 25: fluxo de dados da tabela fato extraído do SSIS

Em relação à versão inicial do modelo de dados, poucas alterações significativas foram realizadas em termos de dimensões. As dimensões discentes, período letivo e curso foram mantidas. Da mesma forma, a tabela fato continuou sendo chamada de matrículas. Porém, conforme pode ser visto na Figura 26 as dimensões Tempo e Ocorrência foram incluídas no novo modelo.



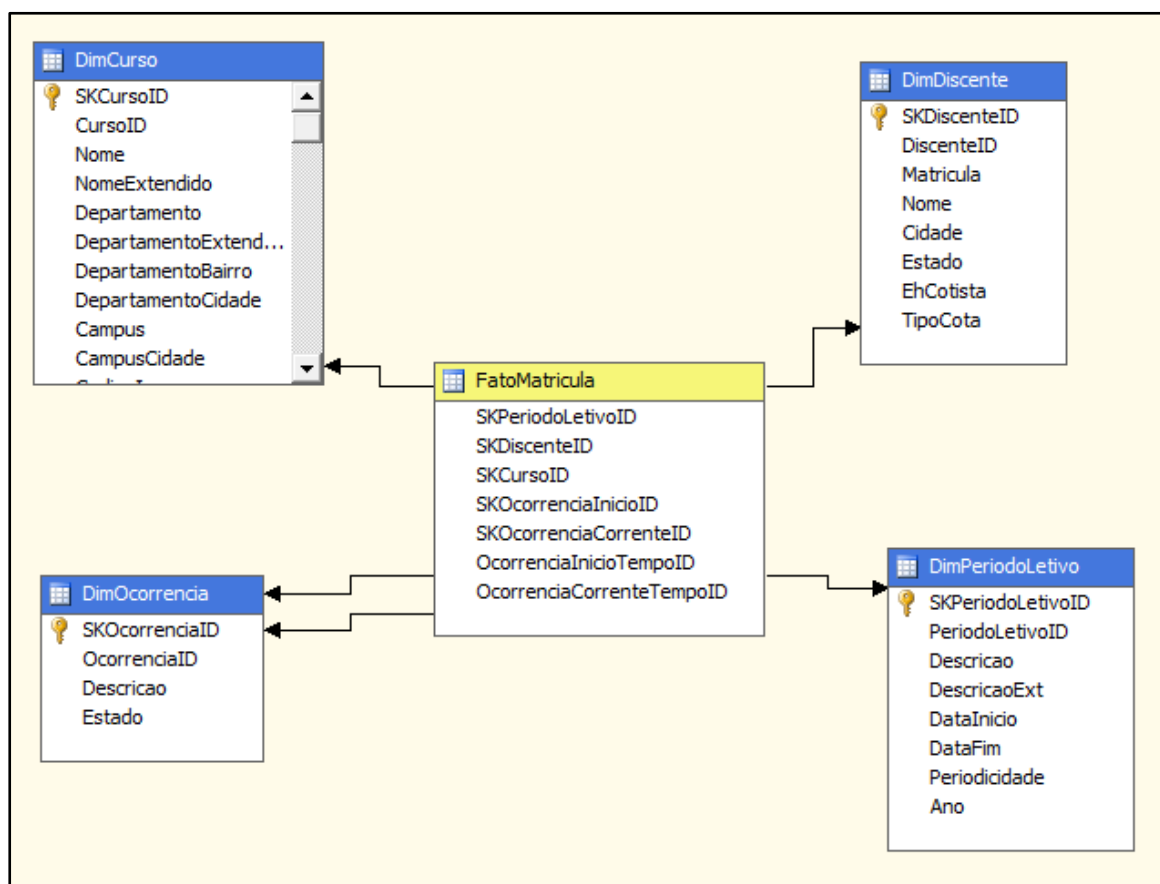
Fonte: Microsoft Excel 2010

Figura 26: modelo de dados final do DW extraído do Sql Server 2012

A inclusão da dimensão Ocorrência se deu em função da dúvida existente no primeiro rascunho elaborado para o desenvolvimento do protótipo. Normalmente pensamos nas matrículas numa unidade temporal de semestres. Porém, o Anuário UNEB em Dados apresenta em seus relatórios as matrículas anualmente. Para atender a essa necessidade, tivemos que diferenciar uma matrícula, que é única para cada aluno, de sua ocorrência, que é o ato do aluno se matricular (ou não) semestralmente. Assim, na tabela dimensão ocorrência armazenamos o estado da matrícula em cada semestre.

Já a dimensão Tempo, apesar de ser incluída no modelo de dados do Datawarehouse, ficou de fora do modelo do cubo, uma vez que não teríamos uma utilidade prática para ela nesse momento. Importa saber que nessa dimensão ficarão armazenadas as informações de data, dia, mês, ano, dia da semana e se o ano é bissexto.

Na modelagem do Cubo (Figura 27), podemos observar que a tabela fato apenas armazena as referências das tabelas das dimensões, as chamadas *Surrogate Keys*, ou chaves substitutas. Essas chaves “substituem” as chaves que existem (ou deveriam existir) no modelo de dados da base transacional e servem para garantir a integridade do modelo de dados dimensional e independem da base de dados de origem. Ao processar o cubo, além do modelo, também são geradas as medidas que serão analisadas. No caso de nossa ferramenta, temos as seguintes medidas: total de discentes, total de ocorrências, percentual de discentes por curso e percentual de discentes por cota.



Fonte: SQL Server Analysis Services

Figura 27: modelo de dados final do cubo extraído do SSAS

Após todos os ajustes da ferramenta e o processamento do Cubo, seguimos para a fase de análise dos resultados. Para um melhor entendimento sobre as informações que podem ser acessadas e analisadas na ferramenta segue um dicionário de dados resumido com descrição e exemplo dos dados disponíveis nas dimensões do cubo gerado e das métricas disponibilizadas.

Dimensão	CAMPO	DESCRIÇÃO DO CAMPO	EXEMPLO DE CONTEUDO DO CAMPO
Curso	Campus	Identificação do Campus	Campus XIX
Curso	Campus Cidade	Nome da cidade	Camaçari
Curso	Código Inep	Identificador do curso junto ao Inep	1160430
Curso	Curso	Nome do Curso	Administração
Curso	Curso Ano Criação	Ano de criação do curso	1998
Curso	Curso Ano Extinção	Ano de extinção do curso	2013
Curso	Curso Código Interno	Junção do nome do curso com códigos identificadores criados pelas secretarias	C-I-038-19971Administração
Curso	Curso Escolaridade	Título conferido ao discente que conclui curso de certo nível	Doutor
Curso	Cursos ID	Código identificador do curso na base de origem	1000000377
Curso	Curso Nível	Identificação do nível dos cursos	Graduação
Curso	Departamento Cidade	Cidade do departamento	Alagoinhas
Curso	Departamento Nome	Nome completo do departamento, seguido de sua sigla	Departamento de Ciências Humanas e Tecnologias – DCHT
Curso	Departamento Sigla	Sigla do departamento seguida do nome do campus ao qual pertence	DCET – Campus I
Discente	Discente Cidade	Cidade do discente	
Discente	Discente Estado	Sigla da Unidade Federativa do discente	BA
Discente	Discente Matrícula	Nº da matrícula	231510203
Discente	Discente Tipo Cota	Forma de ingresso do discente	Ampla concorrência / Cota étnica / etc...
Ocorrência	Ocorrência Descrição	Descrição da ocorrência da matrícula no semestre	Matriculado / Abandono / Cancelado / etc...
Ocorrência	Ocorrência Estado	Tipo de ocorrência da matrícula	Ativo Matriculado / Ativo não Matriculado / Inativo
Ocorrência	Ocorrência ID	Código identificador da	2000000005

Dimensão	CAMPO	DESCRIÇÃO DO CAMPO	EXEMPLO DE CONTEUDO DO CAMPO
		ocorrência na base de origem	
Período Letivo	Período Letivo	Identificação do período	20151
Período Letivo	Período Letivo Ano	Ano do período letivo	2015
Período Letivo	Período Letivo Data Fim	Data final do período letivo	2012-07-31 00:00:00.000
Período Letivo	Período Letivo Data Início	Data inicial do período letivo	2012-07-31 00:00:00.000
Período Letivo	Período Letivo Descrição	Uma melhor descrição do período letivo	20121 FÉRIAS
Período Letivo	Período Letivo ID	Código identificador do período letivo na base de origem	2000000210
Período Letivo	Período Letivo Periodicidade	Identificação do tipo de periodicidade do período letivo	Anual / Semestral
Tabela Fato	Total de ocorrências	Métrica que informa quantas ocorrências de matrículas existem na base de origem	23215
Tabela Fato	Total de Discentes	Métrica que informa quantos alunos existem na base de origem	69436

Fonte: autor

Tabela 1: dicionário de dados do Cubo OLAP desenvolvido

4.3. ANÁLISE DOS RESULTADOS

Após a finalização do protótipo, uma comparação entre os números de matriculados expostos no Anuário UNEB em Dados 2015 – Base 2014 e os apresentados pela ferramenta foi necessária para validar a consistência das informações disponibilizadas no DW. Para a comparação, foram escolhidos como amostra os departamentos dos campi I (Salvador), II (Alagoinhas), III (Juazeiro) e IX (Barreiras) por serem os que possuíam maior número de alunos ativos no ano de 2014.

Para a amostra observada, a ferramenta apresentou os mesmos 46 cursos disponíveis no Anuário, igualmente organizados em seus respectivos departamentos, sendo que desses, 100% possuíam o mesmo código Inep e apenas o curso de Agronomia do DTCS – Campus III (Juazeiro) apresentou uma diferença em relação ao ano de criação do curso (Figuras 28 e 29), o que representa uma diferença de 2,17% entre as informações referente ao ano de criação de cursos apresentadas pela ferramenta e as disponíveis no Anuário. Porém, uma consulta na base do sistema de origem foi realizada e ficou confirmado que a base apresenta a informação levantada pela ferramenta. O dado foi passado para a SGC para que fosse feita a devida correção.

Código INEP	Nome	Ano Início	Grau Acadêmico	Turno	Vagas		Concorrência		Discentes Matriculados				Evasão %	Conceito	
					Vestibular	SISU	Vestibular	SISU	Não Cotistas	Cotista Negro	Cotista Indígena	Total		ENADE	CPC
2442	Agronomia	1962	Bacharelado	I	58	12	7,74	283,42	197	158	17	372	7,52	3	4
19498	Direito	1998	Bacharelado	V	45	5	30,51	324,00	159	99	15	273	6,23	2	1
1268537	Engenharia de Bioprocessos e Biotecnologia	2014	Bacharelado	I	29	6	7,76	405,67	15	12	1	28	32,14	-	-
TOTAL					132	23	-	-	371	269	33	673	-	-	-

Fonte: Anuário Uneb em Dados 2015 – Base 2014

Figura 28: parte do anuário destacando divergência de ano de criação do curso de agronomia

Total de discentes		Discente		Cota		Total Geral				
Departamento	Sigla	Código Inep	Curso	Ano Criação	Grau	Turno	Ampla concorrência			
DCH - CAMPUS III		83353	Docência e Gestão de Processos Educativos	2007	Licenciatura	Integral	1			
			Pedagogia	2004	Licenciatura	Vespertino	184			
		68566	Comunicação Social	2003	Bacharelado	Vespertino	100			
DTCS - CAMPUS III		2442	Agronomia	1960	Bacharelado	Integral	128			
			Engenharia Agrônômica	2013	Bacharelado	Integral	69			
		19498	Direito	1998	Bacharelado	Integral	159			
		1268537	Engenharia de Bioprocessos e Biotecnologia	2014	Bacharelado	Integral	15			
Total Geral							655	480	43	1178

Fonte: Anuário Uneb em Dados 2015 – Base 2014

Figura 29: parte da planilha elaborada através da ferramenta desenvolvida destacando divergência de ano de criação do curso de agronomia

Em relação aos números de matrículas, que foi a métrica definida para o escopo da ferramenta nesse protótipo, 11 dos 46 cursos apresentaram uma pequena diferença entre os dados. A ferramenta apresentou um total de 8.838 matrículas registradas, contra 8.854 disponíveis no anuário, uma diferença de

0,18%. Dentre essas matrículas divergentes, 8 dizem respeito a categoria “ampla concorrência” que possui um total de 5.022 matrículas registradas, uma diferença de 0,16%, 7 dizem respeito a categoria “cota étnica” que possui um total de 3.679 matrículas registradas, uma diferença de 0,19% e apenas 1 matrícula diz respeito a categoria “cota indígena” que possui um total de 135 matrículas registradas, uma diferença de 0,74% entre os dados da ferramenta e os expostos no Anuário. Os dados relativos aos cursos dos departamentos dos campi III e IX não apresentaram qualquer índice de divergência entre os levantados pela ferramenta e os apresentados no Anuário.

Um dado bastante discrepante surgiu ao analisar, o curso de Pedagogia do Campus I, código Inep 83344, que apresentou uma diferença de 19 matrículas na categoria “cota étnica” entre a ferramenta e o Anuário. Porém, observando o total de matrículas, a diferença era de apenas 1 registro, o que nos leva a crer que tratava-se de um erro de digitação no momento da elaboração do anuário.

Para entender a diferença entre os dados comparados, foi solicitada a consulta que gerou os dados do Anuário. Executada a consulta, encontramos alguns casos onde os números passaram a ser iguais. Na busca pela compreensão dessas situações, chegamos à conclusão de que pode ter existido algum erro de digitação no momento da preparação do Anuário, ou pode ter existido alguma matrícula retroativa no sistema. Adaptamos então a para levantar, não o quantitativo, mas os números das matrículas dos discentes que compunham o quantitativo. Percebemos então que isso pode ter ocorrido porque até o ano de 2014, o sistema permitia lançar algumas disciplinas diretamente no histórico, sem realizar a matrícula do aluno no componente. Desta forma, o campo utilizado para a categorização da matrícula que é recuperado pelo DW não era preenchido e a ferramenta não teria como identificar. Atualmente a SGC corrigiu o problema bloqueando o lançamento de disciplinas sem a devida efetivação da matrícula. Quanto aos casos em que ainda existiam divergências, a explicação da matrícula retroativa também se aplica, visto que a SGC ainda estava realizando as correções na base.

Superada a questão de confiabilidade nos quantitativos, verificamos o formato de disponibilização dos dados na ferramenta. Comparando com a Tabela apresentada no Anuário percebemos que o protótipo permite uma formatação bastante próxima com pequenas diferenças de Design. Apresento a seguir uma amostra das tabelas apresentadas no Anuário (Figura 31) e dois formatos de relatório elaborados na Ferramenta (Figura 30 e Figura 32).

Total de discentes		Rótulos de Coluna ▾			
Rótulos de Linha		Ampla concorrência ▾	Cota étnica	Cota indígena	Total Geral
1268537		15	12	1	28
Engenharia de Bioprocessos e Biotecnologia		15	12	1	28
2014		15	12	1	28
Bacharelado		15	12	1	28
Integral		15	12	1	28
19498		159	99	15	273
Direito		159	99	15	273
1998		159	99	15	273
Bacharelado		159	99	15	273
Integral		159	99	15	273
2442		197	153	17	367
Agronomia		128	102	10	240
1960		128	102	10	240
Bacharelado		128	102	10	240
Integral		128	102	10	240
Engenharia Agrônômica		69	51	7	127
2013		69	51	7	127
Bacharelado		69	51	7	127
Integral		69	51	7	127
Total Geral		371	264	33	668

Fonte: Microsoft Excel 2010

Figura 31: Cursos de Graduação Regular Presencial, DTCS-III na ferramenta (visualização padrão)

Código INEP	Nome	Ano Início	Grau Acadêmico	Turno	Vagas		Concorrência		Discentes Matriculados				Evasão %	Conceito	
					Vestibular	SISU	Vestibular	SISU	Não Cotistas	Cotista Negro	Cotista Indígena	Total		ENADE	CPC
2442	Agronomia	1962	Bacharelado	I	58	12	7,74	283,42	197	158	17	372	7,52	3	4
19498	Direito	1998	Bacharelado	V	45	5	30,51	324,00	159	99	15	273	6,23	2	1
1268537	Engenharia de Bioprocessos e Biotecnologia	2014	Bacharelado	I	29	6	7,76	405,67	15	12	1	28	32,14	-	-
TOTAL					132	23	-	-	371	269	33	673	-	-	-

Fonte: Anuário UNEB em Dados 2015 – Base 2014

Figura 30: Cursos de Graduação Regular Presencial, DTCS-III

Total de discentes						Discente Tipo Cota			Total
Codigo Inep	Curso	Curso Ano Criação	Grau	Turno	Ampla concorrência	Cota étnica	Cota indígena	Total Geral	
1268537	Engenharia de Bioprocessos e Biotecnologia	2014	Bacharelado	Integral		15	12	1 28	
19498	Direito	1998	Bacharelado	Integral		159	99	15 273	
2442	Agronomia	1960	Bacharelado	Integral		128	102	10 240	
	Engenharia Agrônômica	2013	Bacharelado	Integral		69	51	7 127	
Total Geral						371	264	33 668	

Fonte: Microsoft Excel 2010

Figura 32: Cursos de Graduação Regular Presencial, DTCS-III na ferramenta (visualização clássica)

Importante salientar que a ferramenta não pretende substituir o Anuário em questões de design, mas facilitar o levantamento dos dados para sua publicação. Dito isto, ressalto que os dados apresentados na ferramenta foram levantados apenas com alguns cliques selecionando os campos necessários na lista de campos da tabela dinâmica do excel. Para fins de comparação, a tabela acima foi criada em poucos minutos apenas selecionando a métrica “Total de discentes”, os campos “Código Inep”, “Curso” e “Curso Ano Criação” como rótulos de linha e o campo “Discente Tipo Cota” como rótulo de coluna. Em seguida foram aplicados filtros de ano letivo, nível de curso, departamento e estado da ocorrência. Para se obter os mesmos dados sem a ferramenta, seria necessário escrever uma consulta em linguagem SQL semelhante a que se segue:

```

Select DISTINCT Count(DISTINCT alu_nu_matricula) as Qtd_Matriculado,
Forma_Ingresso, crs_id_curso_inep , tpc_ds_tp_curso, fac_nm_faculdade
, crr_ch_turno_matutino, crr_ch_turno_vespertino, crr_ch_turno_noturno
, crs_nm_resumido
from
(
select DISTINCT alu.alu_nu_matricula,
case when (fmi.fmi_ds_forma_ingresso = 'Concurso Vestibular (2)' or
fmi.fmi_ds_forma_ingresso = 'SISU (2)') then 'COTISTA NEGRO'
when (fmi.fmi_ds_forma_ingresso = 'Concurso Vestibular (3)' or
fmi.fmi_ds_forma_ingresso = 'SISU (3)') then 'COTISTA INDIO'
when (fmi.fmi_ds_forma_ingresso = 'Concurso Vestibular (5)' or
fmi.fmi_ds_forma_ingresso = 'SISU (5)') then 'COTISTA NEGRO'
when (fmi.fmi_ds_forma_ingresso = 'Concurso Vestibular (6)' or
fmi.fmi_ds_forma_ingresso = 'SISU (6)') then 'COTISTA INDIO'
ELSE
'AMPLA CONCORRENCIA'
END Forma_Ingresso
, crr.crr_ch_turno_matutino, crr.crr_ch_turno_vespertino,
crr.crr_ch_turno_noturno,
crs.crs_nm_resumido, oca.oca_ds_ocorrencia, crs.crs_id_curso_inep,
crs.crs_ds_codigo_capes, tpc.tpc_ds_tp_curso, fac.fac_nm_faculdade

from CRS_curso crs
inner join TPC_tipo_curso tpc on tpc.tpc_id_tp_curso =
crs.crs_id_tp_curso
inner join RCA_registro_curso_aluno rca on rca.rca_id_curso =
crs.crs_id_curso
inner join CRR_curriculo crr on crr.crr_id_curso = crs.crs_id_curso

```

```

inner join ALU_aluno alu on alu.alu_id_pessoa = rca.rca_id_aluno
inner join FMI_forma_ingresso fmi on fmi.fmi_id_forma_ingresso =
rca.rca_id_forma_ingresso
left join FMS_forma_saida fms on fms.fms_id_forma_saida =
rca.rca_id_forma_saida
inner join OCP_ocorrendia_por_periodo ocp on ocp.ocp_id_registro_curso
= rca.rca_id_registro_curso
inner join HIS_historico_ingresso_saida his on
his.his_id_registro_curso = ocp.ocp_id_registro_curso and
his.his_id_registro_curso = rca.rca_id_registro_curso
inner join PEL_periodo_letivo pele on pele.pel_id_periodo_letivo =
ocp.ocp_id_periodo_letivo
inner join OCA_ocorrendia_academica oca on oca.oca_id_ocorrendia =
ocp.ocp_id_ocorrendia_corrente
inner join FAC_faculdade fac on fac.fac_id_faculdade =
crs.crs_id_faculdade
where --his.his_id_periodo_fim is null
      pele.pel_nu_ano = 2015
      and oca.oca_id_ocorrendia not in (2000000002, 2000000003, 2000000004,
2000000009, 2000000021)
      --and crs.crs_id_curso_inep = 2473

UNION ALL

select DISTINCT alu.alu_nu_matricula,
      case when (fmi.fmi_ds_forma_ingresso = 'Concurso Vestibular (2)' or
fmi.fmi_ds_forma_ingresso = 'SISU (2)') then 'COTISTA NEGRO'
      when (fmi.fmi_ds_forma_ingresso = 'Concurso Vestibular (3)' or
fmi.fmi_ds_forma_ingresso = 'SISU (3)') then 'COTISTA INDIO'
      when (fmi.fmi_ds_forma_ingresso = 'Concurso Vestibular (5)' or
fmi.fmi_ds_forma_ingresso = 'SISU (5)') then 'COTISTA NEGRO'
      when (fmi.fmi_ds_forma_ingresso = 'Concurso Vestibular (6)' or
fmi.fmi_ds_forma_ingresso = 'SISU (6)') then 'COTISTA INDIO'
      ELSE
      'AMPLA CONCORRENCIA'
      END Forma_Ingresso
, crr.crr_ch_turno_matutino, crr.crr_ch_turno_vespertino,
crr.crr_ch_turno_noturno,
  crs.crs_nm_resumido, oca.oca_ds_ocorrendia, crs.crs_id_curso_inep,
crs.crs_ds_codigo_capes, tpc.tpc_ds_tp_curso, fac.fac_nm_faculdade

from CRS_curso crs
inner join TPC_tipo_curso tpc on tpc.tpc_id_tp_curso =
crs.crs_id_tp_curso
inner join RCA_registro_curso_aluno rca on rca.rca_id_curso =
crs.crs_id_curso
inner join CRR_curriculo crr on crr.crr_id_curso = crs.crs_id_curso
inner join ALU_aluno alu on alu.alu_id_pessoa = rca.rca_id_aluno
inner join FMI_forma_ingresso fmi on fmi.fmi_id_forma_ingresso =
rca.rca_id_forma_ingresso
inner join OCP_ocorrendia_por_periodo ocp on ocp.ocp_id_registro_curso
= rca.rca_id_registro_curso
inner join HIS_historico_ingresso_saida his on
his.his_id_registro_curso = ocp.ocp_id_registro_curso and
his.his_id_registro_curso = rca.rca_id_registro_curso
inner join PEL_periodo_letivo pele on pele.pel_id_periodo_letivo =
ocp.ocp_id_periodo_letivo
inner join OCA_ocorrendia_academica oca on oca.oca_id_ocorrendia =
ocp.ocp_id_ocorrendia_corrente
inner join FAC_faculdade fac on fac.fac_id_faculdade =
crs.crs_id_faculdade

```

```

where pele.pel_nu_ano = 2015
      --and oca_id_ocorrencia not in (2000000002, 2000000003, 2000000004,
2000000009, 2000000021)
      and exists (select 1 from MTR_matricula mtr
                  where mtr.mtr_id_registro_curso =
rca.rca_id_registro_curso
                  and mtr.mtr_id_periodo_letivo = pele.pel_id_periodo_letivo
                  and mtr.mtr_id_situacao_matricula not in (2000000006,
2000000009, 2000000011))
      --and crs.crs_id_curso_inep = 2473
      ) as tabela
Where crs_id_curso_inep = 83233
--Where crs_ds_codigo_capes = '28005015003P3'

Group by tabela.Forma_Ingresso,
crr_ch_turno_matutino, crr_ch_turno_vespertino, crr_ch_turno_noturno,
crs_id_curso_inep, crs_id_curso_inep, tpc_ds_tp_curso, fac_nm_faculdade,
crs_nm_resumido
Order by tabela.Forma_Ingresso

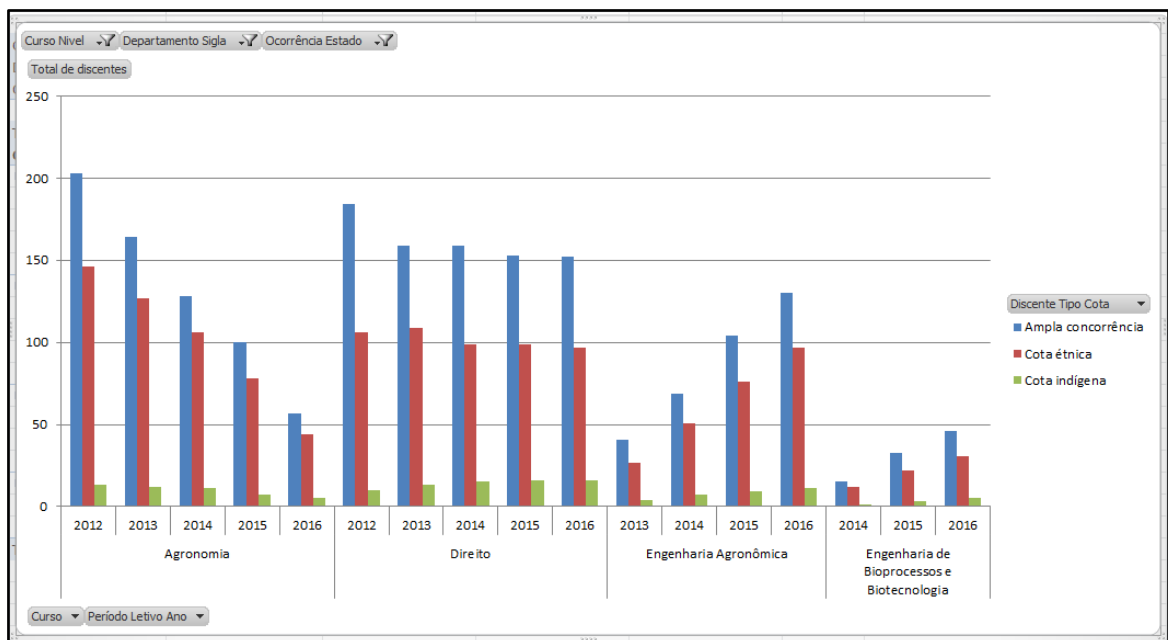
```

Além de trabalhosa, a consulta não é inteligível para o usuário final e necessita que se saiba os códigos dos cursos, dos tipos de ocorrências e dos tipos de cotas como cadastrados no sistema. Sem falar que a consulta é executada na mesma base de dados em que o sistema está rodando, concorrendo com o processamento transacional que ocorre a todo instante, ou numa instância de testes ou homologação onde os dados nem sempre estão consolidados.

Além de mais prática, a ferramenta possui um caráter dinâmico e permite elaborar relatórios a partir de diversos atributos das dimensões disponibilizadas, enquanto que o relatório do Anuário, apesar de possuir um design diferenciado, é estático e sem a possibilidade de alterações. Na ferramenta é possível, após percepção de uma situação desconhecida a partir de um dado relatório, realizar novos filtros e cruzamentos dinâmicos que permitam nova análise das informações.

Outra possibilidade permitida pela ferramenta é uma análise de série histórica dos dados em tempo real. Como não possui as amarras do papel, o DW pode disponibilizar os dados da Universidade desde o início do uso de seus sistemas informatizados. Para o desenvolvimento deste projeto, no entanto, foram coletados dados amostrais apenas dos últimos 5 anos. Com isso, é possível analisar os mesmos dados disponibilizados pelo Anuário desde o ano de 2012. Na Fonte: Microsoft Excel 2010

Figura 33 utilizamos o recurso de gráfico dinâmico do Excel para analisar visualmente os mesmos dados apresentados nas figuras Figura 30 e Figura 32,



Fonte: Microsoft Excel 2010

Figura 33: gráfico de evolução das matrículas dos Cursos de Graduação Regular Presencial, DTCS-III na ferramenta

porém acompanhando a evolução das matrículas desde 2012 até o semestre atual.

Desta forma, facilmente o usuário consegue visualizar o aumento do número de matrículas nos cursos de Engenharia Agrônoma e Engenharia de Bioprocessos e Biotecnologia nos últimos anos, enquanto o curso de Agronomia apresentou uma queda acentuada no mesmo período. No caso da queda das matrículas do curso de Agronomia, entendemos que há uma migração para o curso de Engenharia Agrônoma que o substituiu, mas é importante analisar o gráfico do curso de Direito que teve uma queda significativa em seu número de matrículas no ano de 2013 e, apesar de uma aparente estabilidade, ainda continua reduzindo o

número de matrículas ativas entre os anos de 2013 e 2016. Tudo isso sem a necessidade de uma linha de código, já que esse processo é realizado previamente pela equipe de TI durante o desenvolvimento do DW.

Após os testes e comparações realizados, o protótipo demonstrou ser uma ferramenta poderosa para análise de dados, não apenas podendo auxiliar no levantamento de informações para as próximas edições do Anuário UNEB em Dados, mas abrindo possibilidades para criação de relatórios dinâmicos e diversos dos constantes no mesmo. No entanto, algumas observações se fazem necessárias. Enquanto construto social, a ferramenta de *BI* necessita do envolvimento dos diversos setores da instituição. Desde o nível estratégico, que precisa demonstrar interesse e apoiar o desenvolvimento da ferramenta, quer seja investindo em pessoal, hardware e software, quer seja definindo as métricas importantes para a Universidade, até o nível operacional que precisa ter conhecimento da importância da forma como os dados são inseridos, modificados ou apagados dos sistemas institucionais.

Para se ter uma idéia do cuidado que é necessário com os dados nos sistemas de origem, após desenvolver as primeiras consultas nas bases de dados do Sistema Acadêmico SAGRES, algumas inconsistências foram detectadas. Abaixo listo algumas delas:

1. Existiam registros de Alunos sem correspondentes na tabela Pessoas. Possivelmente, um script foi executado diretamente na base de dados para excluir alguns alunos, mas por alguma falha de integridade referencial, os “registros pais” foram excluídos sem que os “registros filhos” fossem excluídos em cascata. Aparentemente esse tipo de inconsistência não gera problemas para a área operacional do sistema, mas dificulta enormemente o trabalho do analista que precisa gerar relatórios estratégicos a partir das mesmas bases.

2. Existiam 3818 registros da tabela de “alunos nos cursos da instituição” que estão com o campo referente a curso vazio. Sendo que é inerente a este tipo de tabela (uma tabela de relacionamento) que os campos referentes aos alunos e aos cursos estejam preenchidos. Provavelmente, o mesmo problema do exemplo anterior se aplica aqui.

3. Existiam registros na tabela de “ocorrência por período do aluno no curso” que indicavam saída do aluno (ocorrência “abandono”, por exemplo), mas o campo “forma de saída” não estava preenchido. Esse problema gerou inconsistências na primeira abordagem realizada, uma vez que tentamos buscar os desligamentos dos alunos no referido campo.

Vale ressaltar que, por se tratar de um protótipo, a ferramenta não possui todas as informações disponibilizadas no Anuário. E ainda que venha a possuir tais informações, a mesma não possui pretensão de substituir a publicação, mas auxiliar no processo de levantamento dos dados que a compõe e ampliar o leque de possibilidades para análise dos dados institucionais da Universidade, tanto por parte dos gestores, quanto de pesquisadores interessados.

Retomando e ampliando um pouco o que está disposto na Figura 16, segue um quadro resumo das etapas utilizadas para o desenvolvimento do protótipo desde sua concepção até a implementação.

Planejamento	Levantamento de Indicadores	Modelagem Lógica	Modelagem Física	Desenvolvimento
✓ Análise da situação atual	✓ Análise de Indicadores existentes	✓ Identificação das dimensões	✓ Mapeamento de Fontes de Dados	✓ Desenvolvimento da ferramenta
✓ Coleta de materiais	✓ Revisão de indicadores	✓ Mapeamento do sistema de origem	✓ Avaliação de plataforma de BI	✓ Apresentação da ferramenta
✓ Definição de Escopo	✓ Detalhamento de indicadores		✓ Definição de estratégia de atualização	
✓ Entregáveis	✓ Entregáveis	✓ Entregáveis	✓ Entregáveis	✓ Entregáveis
✓ Plano de desenvolvimento	✓ Indicadores	✓ Modelo lógico	✓ Modelo físico de dados	✓ Ferramenta implementada
✓ Macro cronograma	✓ Conceitos dos indicadores		✓ Dicionário de dados	✓ Treinamento

Fonte: autor

Tabela 2: Quadro resumo das etapas de desenvolvimento do protótipo

5. CONSIDERAÇÕES FINAIS

Como foi demonstrado, uma ferramenta de Business Intelligence (BI) através de seus processos de coleta, sumarização e consolidação pode não apenas auxiliar na composição do Anuário UNEB em Dados, como ampliar o escopo de visão dos dados da instituição através de dados históricos e relatórios dinâmicos. O desenvolvimento do protótipo apresentado permitiu o alcance dos objetivos propostos. Como narrado, conhecer as bases de dados origem para a composição do Anuário UNEB em Dados e levantar as perguntas que precisam ser respondidas pelo Anuário UNEB em Dados foram condições *sine qua non* para estruturar o Data Mart relativo aos discentes Matriculados desde 2012. Além disso, somente com o Data Mart estruturado e modelado, foi possível desenvolver o protótipo que era, ele próprio, um dos objetivos específicos do projeto. Por fim, foram apresentados através de planilhas e gráficos alguns dados do Anuário UNEB em Dados na ferramenta com uma breve análise dos mesmos.

Apesar de todas as vantagens e possibilidades apresentadas pela ferramenta, ainda há muito para se melhorar. Como foi dito no início do projeto, um Data Warehouse apresenta informações de diversos assuntos e/ou departamentos, sendo que esses assuntos específicos são armazenados em partes menores chamadas de Data Marts. Assim posto, cabe deixar claro que o presente projeto desenvolveu um Data Mart com informações relativas apenas às matrículas dos discentes. Para trabalhos futuros, é possível desenvolver novos Data Marts para se aprofundar nas informações referentes às diversas áreas da instituição, como o desempenho acadêmico dos alunos, o vestibular, seu corpo docente e técnico, além de Data Marts sobre bolsas de iniciação científica e de auxílio estudantil.

Além de novos Data Marts, é importante desenvolver painéis de bordo (*dashboards*) com indicadores e gráficos previamente montados para os usuários de nível estratégico que precisam de informações mais rápidas e que responda a questões pré-definidas. Se por um lado, o fato de utilizar o Excel como ferramenta para visualização dos dados disponibilizados no Cubo OLAP elimina a necessidade de aquisição ou desenvolvimento de ferramenta específica, por outro, exige do usuário um maior empenho e envolvimento para minerar informações relevantes.

Por fim, fica como sugestão para trabalhos futuros, levantar os indicadores importantes da universidade que possam servir de base para o desenvolvimento dos *Key Performance Indicators – KPIs* da Universidade. De acordo com a Endeavor²¹, *KPIs* são as métricas eleitas como essenciais para avaliar um processo de gestão da instituição. Os *KPIs* podem ser categorizados em indicadores de produtividade, de qualidade, de capacidade ou estratégicos. Algumas universidades internacionais como a Queen’s University²² do Canadá utilizam alguns *KPIs* como Intensidade de Pesquisa, Engajamento de Estudantes de Pós-Graduação e Engajamento Internacional em Pesquisa. Faz-se necessária uma pesquisa junto ao grupo gestor da UNEB para levantar quais *KPIs* são relevantes para a instituição.

Sem dar um caráter definitivo ao trabalho, paro por aqui, vislumbrando esses três caminhos para o desenvolvimento de novas pesquisas que não apenas possibilitem um conhecimento mais amplo e assertivo acerca da Universidade do Estado da Bahia, mas auxiliem a fortalecer e aumentar o alcance desta instituição que possui como função social possibilitar o acesso a uma educação superior de qualidade nas diversas microrregiões da Bahia.

²¹ Disponível em < <https://endeavor.org.br/kpi/>>. Acesso em: 27 de Novembro de 2016

²² Disponível em < <http://www.queensu.ca/strategicplanning/framework/kpis/>>. Acesso em 28 de Novembro de 2016.

6. REFERÊNCIAS BIBLIOGRÁFICAS

CARMO, Vadson Bastos do; PONTES, Cecília Carmen Cunha. **Sistemas de informações gerenciais para programa de qualidade total em pequenas empresas da região de Campinas**. Ci. Inf, v. 28, n. 1, p. 49-58, 1999.

CARNEIRO, Breno Pádua Brandão; NOVAES, Ivan Luiz. As comissões próprias de avaliação frente ao processo de regulação do ensino superior privado. **Avaliação, Campinas, Sorocaba**, v. 13, n. 3, 2008.

DATE, C. J. **Introdução a sistemas de bancos de dados**. São Paulo: Elsevier Brasil, 2000.

DAVENPORT, T. H. **Ecologia da informação: por que só a tecnologia não basta para o sucesso na era da informação**. São Paulo: Futura, 1998.

DAVENPORT, T. H.; PRUSAK, L. **Conhecimento Empresarial: Como as organizações gerenciam o seu capital intelectual**. São Paulo: Campus, 1998.

ELMASRI, Ramez; NAVATHE, Shamkant B.; DE OLIVEIRA MORAIS, Rinaldo. **Sistemas de banco de dados**. 2005.

FADEL, Bárbara et al. **Gestão, mediação e uso da informação**. In: VALENTIM, Marta (Org.). **Gestão, mediação e uso da informação**. São Paulo: Cultura Acadêmica, 2010. p. 13-31.

FINGER, Almeri Paulo. **A Gestão Universitária no Brasil: A Busca de uma Identidade**. In: FINGER, Almeri Paulo (Org.) **Gestão de Universidades – Novas Abordagens**. Curitiba: Champagnat, 1997. 296 p.

GARTNER INC. Disponível em: <<http://www.gartner.com/it-glossary/business-intelligence-bi>> Acesso em 1 de junho de 2015

HARDY, Cyntia; FACHIN, Roberto. **Gestão Estratégica na Universidade Brasileira: teoria e casos**. Porto Alegre: Ed. Universidade/UFRGS, 2 ed., 2000. 233 p.

INMON, W. H. **Building the Data Warehouse**. 3ª ed. New York: John Wiley & Sons, 1992. 272p..

KIMBALL, Ralph. **Data Warehouse Toolkit**. 3ª ed. Indianapolis: John Wiley & Sons, Inc, 2013. 564 p. ISBN 978-1-118-53080-1.

LUFTMAN, Jerry. **Assessing business-IT alignment maturity. Strategies for information technology governance**, v. 4, p. 99, 2004.

MARQUES, Wagner Luiz. **Sistema De Informações Gerenciais**. Clube de Autores, 2007.

MIRANDA, R. C. da R. **O uso da informação na formulação de ações estratégicas pelas empresas**. Ciência da Informação, Brasília, DF, v. 28, n. 3, p. 284-290, set./dez. 1999.

MURAKAMI, Milton. **Decisão estratégica em TI: estudo de caso. 2003**. Tese de Doutorado. Universidade de São Paulo.

NONAKA, Ikujiro & TAKEUCHI, Hirotaka. **Criação de Conhecimento na Empresa**: como as empresas japonesas geram a dinâmica da inovação. Tradução de Ana Beatriz Rodrigues, Priscila Martins Celeste. Rio de Janeiro: Campus, 1997.

NOVAES, Ivan Luiz; CARNEIRO, Breno Pádua Brandão. **DILEMAS DA GESTÃO EM UMA UNIVERSIDADE ESTADUAL MULTICAMPI: A IMPLANTAÇÃO DA SECRETARIA ESPECIAL DE AVALIAÇÃO INSTITUCIONAL (SEAVI/UNEB)**. 2014.

OLIVEIRA, Grimaldo Lopes de. **GESMOODLE – FERRAMENTA DE MEDIAÇÃO TECNOLÓGICA NO ACOMPANHAMENTO DO PROCESSO DE APRENDIZAGEM DOS ALUNOS EM CURSOS DE GRADUAÇÃO A DISTÂNCIA DA UNIVERSIDADE DO ESTADO DA BAHIA (UNEB)**. 2013. 115 f. Dissertação (mestrado) – Universidade do Estado da BAHIA, Departamento de Educação. Programa de Pós-Graduação em Gestão e Tecnologias Aplicadas À Educação. Salvador. 2013.

PARANHOS, L. R. L.; RODOLPHO, P. J. . **Metodologia da pesquisa aplicada à tecnologia**. 1. ed. São Paulo: SENAI-SP, 2014. v. 1. 160p.

PIMENTA, Lídia Boaventura. **Processo decisório na universidade multicampi: dinâmica dos conselhos superiores e órgãos de execução**. 2007. 211 f. Dissertação (mestrado) - Universidade Federal da Bahia, Faculdade de Educação. Programa de Pós-Graduação em Educação. Salvador. 2007.

PINHEIRO, Marcus Tulio F. **Análise do conhecimento gerado na organização e sua influência na eficácia da gestão: estudo de caso – Consultec**. 2002. 112 f. Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia de Produção. Florianópolis. 2002.

PROGRAD UNEB. Disponível em: <<http://www.uneb.br/prograd/graduacao/>>. Acesso em: 09 de nov. 2014.

RAMIREZ, Pedro Alejandro Herrera. **Plataforma Pandora: Desenvolvimento do Sistema Integrado de Gestão da Informação da Pós-Graduação Stricto Sensu/UNEB**. 2016. 65 f. Dissertação (mestrado) – Universidade do Estado da BAHIA, Departamento de Educação. Programa de Pós-Graduação em Gestão e Tecnologias Aplicadas À Educação. Salvador. 2016.

TOZONI-REIS, Marília Freitas de Campos. **Metodologia da pesquisa**. Curitiba: IESDE Brasil SA, p. 15-36, 2009.

ROBBINS, Stephen P. **Administração: Mudanças e perspectivas**. São Paulo: Saraiva, 2001.

RUSSO, Mariza. **Fundamentos de Biblioteconomia e Ciência da Informação**, Coleção Biblioteconomia & Gestão, Série Didáticos, v. 1 (FACC/UFRJ). Rio de Janeiro: E-Papers, 2010. 177p.

SEZÕES, Carlos; OLIVEIRA, José e BAPTISTA, Miguel; **BUSINESS INTELLIGENCE**. SPI-Sociedade Portuguesa de Inovação, 2006.

TACHIZAWA, T e ANDRADE, R. O. B, **Gestão de instituições de ensino**. 1.ed. Rio de Janeiro, Editora Fundação Getúlio Vargas, 1999

TERRA, José Cláudio. **Gestão do conhecimento: o grande desafio empresarial**. São Paulo: Negócio Editora, 2000.

THIOLLENT, Michel. **Metodologia da Pesquisa-Ação**. 2.ed. Rio de Janeiro, RJ: Editora Cortez, 1986.

TORRES, Luiz Eduardo Santos. **Proposta de Modelo de Gestão Universitária Baseado em Sustentabilidade: Aplicação ao Caso PUC-Rio**. 2012. 70 f. Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial. Programa de Pós-Graduação em Engenharia de Produção. Florianópolis. 2012.

VALENTIM, M. L. P. **Inteligência competitiva em organizações: dado, informação e conhecimento**. DataGramZero, Rio de Janeiro, v.3., n.4, ago. 2002.

WEBER, Demétrio, Vieira, Leonardo. **Censo do ensino superior mostra queda no número de formandos em faculdades brasileiras**. *O Globo*. Disponível em <<http://oglobo.globo.com/sociedade/educacao/censo-do-ensino-superior-mostra-queda-no-numero-de-formandos-em-faculdades-brasileiras-13879540>>. Acesso em: 09 de set. 2014.