

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE MATEMÁTICA
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

LUCIANO ARANTES PINTO

AVALIAÇÃO DE UM AMBIENTE ANALÍTICO COM APOIO DE UM MODELO DE
CAPACIDADE E MATURIDADE

RIO DE JANEIRO

2018

LUCIANO ARANTES PINTO

AVALIAÇÃO DE UM AMBIENTE ANALÍTICO COM APOIO DE UM MODELO DE
CAPACIDADE E MATURIDADE

Trabalho de conclusão de curso de graduação apresentado ao Departamento de Ciência da Computação da Universidade Federal do Rio de Janeiro como parte dos requisitos para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Profa. Maria Luiza Machado Campos

RIO DE JANEIRO

2018

CIP - Catalogação na Publicação

P659a Pinto, Luciano Arantes
Avaliação de um Ambiente Analítico com apoio de um
Modelo de Capacidade e Maturidade / Luciano Arantes
Pinto. -- Rio de Janeiro, 2018.
119 f.

Orientadora: Maria Luiza Machado Campos.
Trabalho de conclusão de curso (graduação) -
Universidade Federal do Rio de Janeiro, Instituto
de Matemática, Bacharel em Ciência da Computação,
2018.

1. Inteligência de Negócios. 2. Ambiente
Analítico. 3. Modelo de Capacidade e Maturidade. I.
Campos, Maria Luiza Machado, orient. II. Título.

LUCIANO ARANTES PINTO

AVALIAÇÃO DE UM AMBIENTE ANALÍTICO COM APOIO DE UM MODELO DE
CAPACIDADE E MATURIDADE

Trabalho de conclusão de curso de graduação
apresentado ao Departamento de Ciência da
Computação da Universidade Federal do Rio de
Janeiro como parte dos requisitos para obtenção do
grau de Bacharel em Ciência da Computação.

Aprovado em 14 de dezembro de 2018.

BANCA EXAMINADORA:

Maria Luiza Machado Campos, Ph.D. (UFRJ)

Pedro Manoel da Silveira, Ph.D. (UFRJ)

Daniel Schneider, D.Sc. (UFRJ)

Dedicatória: Aos meus pais Heraldo e Dalva e minha Tia Néia, que são o alicerce da minha vida.

AGRADECIMENTOS

À minha orientadora, professora Maria Luiza Machado Campos, por ter me concedido a oportunidade e ter acreditado na conclusão deste trabalho.

“Nunca largue mãos de seus sonhos, pois se eles se acabam ou deixam de existir, a sua vida fica como a de um pássaro que não tem asas e não consegue voar.”

RESUMO

De uma forma geral, nos dias de hoje, a estrutura tradicional de Inteligência de Negócios, do inglês de Business Intelligence (BI), não consegue atender aos anseios e demandas dos usuários de negócios e acaba sendo insuficiente para alcançar os objetivos das áreas e da própria corporação. A estrutura atual, que é centralizada pela equipe da área de Tecnologia da Informação (TI) da empresa, oferece como capacidades analíticas: relatórios, dashboards, processamento analítico online (OLAP), ferramenta para consultas e análises ad hoc.

Para superar as limitações existentes, as unidades de negócios começam a implantar soluções unilaterais, como ferramentas de reconhecimento de dados, aplicações analíticas verticais, bancos de dados analíticos de código aberto ou ainda, plataformas alternativas de BI na nuvem sem o suporte ou a aprovação da TI. Neste cenário, é frequente termos redundância de trabalho com geração de respostas conflitantes para perguntas semelhantes.

Este panorama indica uma necessidade premente de reestruturar a plataforma de BI e análise e ainda revisar as habilidades e responsabilidades dos usuários e reorganizar o processo de criação e implantação de conteúdo. Neste contexto, ainda surge o conceito de Analíticos de Negócios, do inglês Business Analytics (BA), como elemento capaz de fornecer flexibilidade e agilidade que os usuários esperam para seus processos de análise e suporte a decisão. Porém o uso desgovernado de analíticos pode aumentar significativamente a geração de informações redundantes e conflitantes.

Uma forma eficiente para lidar com uma transformação como esta é através da implantação de um modelo de capacidade e maturidade, que pode trazer como benefícios uma maior confiabilidade no que refere ao cumprimento de prazos e custos, uma melhor qualidade nos produtos gerados além de uma busca por melhoria contínua.

O produto deste trabalho é mostrar como um modelo de capacidade e maturidade pode ser importante para uma organização aprimorar o seu processo de exploração de dados e informações, através da modernização de suas práticas em BI e ainda com a introdução dos novos conceitos de BA.

Palavras Chave: Business Intelligence. Business Analytics. Arquitetura. Análises. Suporte a Decisão. Modelo de Capacidade e Maturidade. Exploração de Dados e Informações.

ABSTRACT

In general, nowadays, the traditional Business Intelligence (BI) framework can not answer the expectations and demands of business users and is insufficient to reach the objectives of the areas and the corporation itself. The current structure, which is centralized by the company's Information Technology (IT) team, offers analytical capabilities: reporting, dashboards, online analytical processing (OLAP), query tool and ad hoc analysis.

To overcome existing limitations, business units begin deploying unilateral solutions such as data recognition tools, vertical analytic applications, open source analytic databases, or alternative BI platforms in the cloud without the support or approval of the IT staff. In this scenario, it is often redundant to generate conflicting answers to similar questions.

This scenario indicates a pressing need to restructure the BI platform and analysis and further review user skills and responsibilities and reorganize the content creation and deployment process. In this context, the concept of Business Analytics (BA) still emerges as an element capable of providing flexibility and agility that users expect for their analysis and decision support processes. But the ungoverned use of analytics can significantly increase the generation of redundant and conflicting information.

An efficient way for a transformation like this is through the implementation of a capacity and maturity model, which can bring as benefits a greater reliability regarding the fulfillment of deadlines and costs, a better quality in the products generated besides a search for improvement to be continued.

The product of this work is to present how a model of capacity and maturity may be important for an organization to improve its process of data and information exploration, by modernizing its practices in BI and with the introduction of new BA concepts.

Key words: Business Intelligence. Business Analytics. Architecture. Analysis. Decision Support. Capability and Maturity Model. Data Exploration and Information.

LISTA DE FIGURAS

Figura 1 – Representação contínua do CMMI (NUNES, 2013).....	21
Figura 2 – Representação por estágios do CMMI (NUNES, 2013).....	22
Figura 3 - Modelo estrela (RESENDE, 2017).....	33
Figura 4 - Etapas do processo KDD (Adaptação de FAYYAD 1996).....	36
Figura 5 - Business Analytics: integração entre múltiplos campos (NEOINFINITO, 2015)...	38
Figura 6 - Espectro de capacidades analíticas (TAPADINHAS, 2016).....	40
Figura 7 - Fluxo de questões analítica (LEEK, 2015).....	41
Figura 8 - Diagrama conceitual de LDW (LEWIS, 2017).....	50
Figura 9 - Estilos de virtualização de dados (LEWIS, 2017).....	54
Figura 10 - Potencial de análise para impulsionar a mudança organizacional (LEWIS, 2017)	55
Figura 11 - Comparação de Citizen Data Science com Self-Service Analytics e Data Science (IDIONE e TAPADINHAS, 2016).....	57
Figura 12 - Extremos de BI/BA (WAYNE, 2016).....	60
Figura 13- Modelo de Maturidade Gartner para BI e BA (HOWSON e DUNCAN, 2016)....	64
Figura 14 - Dimensões do modelo de maturidade do TDWI (HALPER e STODDER, 2014)	68
Figura 15 - Estágios de maturidade em BI e BA do TDWI (HALPER e STODDER, 2014)..	70
Figura 16 – Modelo de maturidade para BIA da HP (HP, 2009).....	77
Figura 17 - Estágio de maturidade na dimensão organização.....	86
Figura 18 - Estágio de maturidade na dimensão infraestrutura.....	88
Figura 19 - Estágio de maturidade na dimensão gestão de dados.....	91
Figura 20 - Estágio de maturidade na dimensão analíticos.....	93
Figura 21- Estágio de maturidade na dimensão governança.....	96

LISTA DE QUADROS

Quadro 1 - Definição de maturidade (ALMEIDA, 2005)	24
Quadro 2 - Diferenças entre sistema OLTP e OLAP (ELIAS, 2014)	31
Quadro 3 - Linhas de desenvolvimento em BI/BA (WAYNE, 2016).....	58
Quadro 4 - Características das linhas de BI/BA (WAYNE, 2016)	58
Quadro 5 - Vantagens e desvantagens top-down x botton-up (WAYNE, 2016)	61
Quadro 6 - Modelo de maturidade em capacidade (Adaptação de WELLS, 2016)	62
Quadro 7- Interpretação da pontuação da avaliação de maturidade do TDWI	83
Quadro 8 - Avaliação da dimensão organização	86
Quadro 9 – Referências para a dimensão organização	86
Quadro 10- Avaliação da dimensão infraestrutura	88
Quadro 11- Referências para a dimensão infraestrutura	89
Quadro 12- Avaliação da dimensão gestão de dados	91
Quadro 13- Pontuação obtida na dimensão gestão de dados.....	91
Quadro 14- Avaliação da dimensão analíticos	93
Quadro 15 - Pontuação obtida na dimensão analíticos.....	94
Quadro 16- Avaliação da dimensão governança	96
Quadro 17- Pontuação obtida na dimensão governança.....	96

LISTA DE ABREVIATURAS E SIGLAS

API	Application Programming Interface
BA	Business Analytics
BI	Business Intelligence
BICC	Business Intelligence Competency Center
CAO	Chief Analytics Officer
CEO	Chief Executive Officer
CFO	Chief Financial Officer
COO	Chief Operating Officer
CTO	Chief Technical Officer
CMM	Capability Maturity Model
CMMI	Capability Maturity Model Integration
COE	Center of Excellence
DM	Data Mart
DoD	United States Department of Defense
DV	Data Virtualization
DW	Data Warehouse
EDW	Enterprise Data Warehouse
ETL	Extract, Transform and Load
IA	Inteligência Artificial
KDD	Knowledge Discovery in Databases
KPI	Key Performance Indicator
LDW	Logical Data Warehouse
ML	Maturity Levels
RFID	Radio-Frequency Identification
OLAP	On Line Analytical Processing
OLTP	On Line Transaction Processing
PA	Process Area
PMO	Project Management Office
POC	Proof of Concept
SA	Staging Area
TDWI	The Data Warehousing Institute
TI	Tecnologia da Informação

SUMÁRIO

1 INTRODUÇÃO	13
1.1 MOTIVAÇÃO.....	13
1.2 OBJETIVOS.....	14
1.3 JUSTIFICATIVA.....	15
1.4 ESTRUTURA DO TRABALHO	17
2 USO DE MODELOS EM DESENVOLVIMENTO DE SOFTWARE	18
2.1 VISÃO GERAL.....	18
2.2 O MODELO CMMI.....	20
2.2.1 Apresentação.....	20
2.2.2 Representações.....	21
2.2.3 Níveis de Capacidade	22
2.2.4 Níveis de Maturidade	24
2.2.5 Áreas de Processos.....	26
3 CONCEITOS TRADICIONAIS DE AMBIENTE ANALÍTICO.....	27
3.1 DEFINIÇÕES NA LITERATURA	27
3.2 CONCEITOS CORRELATOS	28
3.2.1 Data Warehouse (DW).....	28
3.2.2 Data Mart (DM).....	28
3.2.3 Extract, Transform and Load (ETL).....	29
3.2.4 Operational Data Store (ODS)	29
3.2.5 Staging Area (SA).....	30
3.2.6 Sistemas OLTP e OLAP	30
3.2.7 Modelo Dimensional.....	31
3.2.8 Cubo de Dados	34
3.2.9 Dashboards	35
3.2.10 Mineração de Dados	35
4 CONCEITOS COMPLEMENTARES EM AMBIENTE ANALÍTICO.....	38
4.1 INTRODUÇÃO.....	38
4.2 ESPECTRO DE CAPACIDADES ANALÍTICAS.....	39

4.3 SELF-SERVICE BI.....	42
4.4 USUÁRIOS	43
4.5 VISUALIZAÇÃO DE DADOS	44
4.6 NARRATIVA DE DADOS.....	45
4.7 PREPARAÇÃO DE DADOS NO FORMATO SELF-SERVICE	46
4.8 BIG DATA	47
4.9 DATA WAREHOUSE LÓGICO (LDW).....	48
4.10 DATA LAKE	51
4.11 VIRTUALIZAÇÃO DE DADOS (DV).....	52
4.12 CIÊNCIA DE DADOS.....	54
4.13 BI TRADICIONAL X BA MODERNO	57
5 MODELOS DE CAPACIDADE E MATURIDADE EM PRÁTICAS ANALÍTICAS.	62
5.1 INTRODUÇÃO.....	62
5.2 MODELO GARTNER	63
5.3 MODELO TDWI.....	68
5.4 MODELO HP	77
6 ESTUDO DE CASO	83
6.1 DESCRIÇÃO DO CONTEXTO	84
6.2 DIMENSÕES APLICADAS SEGUNDO O MODELO DO TDWI	85
6.2.1 Dimensão Organização.....	85
6.2.2 Dimensão Infraestrutura	87
6.2.3 Dimensão Gestão de Dados.....	89
6.2.4 Dimensão Analíticos	92
6.2.5 Dimensão Governança	94
7 CONCLUSÕES.....	98
7.1 TENDÊNCIAS ANALÍTICAS	98
7.2 IMPORTÂNCIA DA AVALIAÇÃO DA MATURIDADE ANALÍTICA	98
7.3 PROPOSTAS DE TRABALHOS FUTUROS	99
<u>REFERÊNCIAS</u>	100
<u>ANEXO A - AVALIAÇÃO DO MODELO DE MATURIDADE DO TDWI.....</u>	108

1 INTRODUÇÃO

1.1 MOTIVAÇÃO

Ao longo dos anos, a competição ficou cada vez mais presente nos negócios, e com isso a disputa por mercados passou a exigir das corporações cada vez mais tomadas de decisões melhores e precisas. A informação atualizada e exata passou a ser fundamental para uma gestão ágil, veloz e com a menor margem possível de erro. Para tal, surgiram os ambientes de Inteligência de Negócios, do inglês Business Intelligence (BI)¹, e os grandes bancos de dados organizacionais começaram a produzir variantes, como os depósitos de dados, chamados de Data Warehouse (DW), exatamente com a finalidade de entregar aos tomadores de decisão a informação na forma mais precisa e palatável possível (BARBIERE, 2011).

Recentemente, uma nova era do BI foi inaugurada, com o aprofundamento em campos nos quais já atuava e inserindo-se em outros. Dados de projetos, de blogs e de redes sociais, e de comportamento das pessoas nos seus mais variados papéis, vêm para se juntar ao mundo do varejo, onde tudo começou. O mundo do BI começa a também falar sobre BA, que passa a misturar o BI tradicional voltado para análises descritivas e de diagnóstico com novas técnicas e disciplinas e começam a explorar um novo mundo de análises como preditivas e prescritivas através de um linguajar cada vez mais visual. Este novo cenário expõe a necessidade de criação de mecanismos inferenciais de negócios com latência quase zero, e ainda exigirá maior conteúdo qualitativo das informações.

Constatamos que dados estão presentes em todos os lugares, em todas as interações de negócios e que o uso de analíticos cada vez mais faz parte de nosso dia a dia e como consequência não é mais possível armazená-los apenas em uma estrutura tradicional como o DW. Desta forma, torna-se premente uma nova forma de desenvolver e entregar BI com o intuito de atender o dinamismo necessário para melhorar o impacto na organização e aumentar a satisfação dos usuários.

Um ponto importante a acrescentar é que em muitas organizações a plataforma de BI tradicional não está atendendo de forma eficaz as demandas dos usuários de negócios e acaba sendo insuficiente para atingir os objetivos do próprio negócio (TAPADINHAS, 2016).

¹ Existem termos em inglês, como Business Intelligence, que são amplamente utilizados no mercado e na literatura sem a necessidade de serem traduzidos para a língua portuguesa. Por esta razão, utilizo estes termos, neste trabalho, em sua forma nativa na língua inglesa.

Hoje, muitas organizações percebem que a utilização de técnicas de BA pode fornecer uma vantagem competitiva importante. Essas organizações querem evoluir suas estratégias de análise para além das planilhas ou dos painéis simples, muitas empresas procuram construir uma ampla cultura analítica, na qual a análise de dados desempenha um papel essencial em todas as decisões e é fundamental para a colaboração comercial. O uso de analíticos está rapidamente se tornando mais self-service e democratizado. Mais frequentemente, métodos de desenvolvimento ágil estão sendo usados em implantações de BA para que os usuários possam obter o valor comercial dos projetos mais cedo (HALPER e STODDER, 2014).

1.2 OBJETIVOS

Este trabalho tem como objetivo apresentar um modelo de avaliação de capacidade e maturidade para ambiente analítico em uma empresa. Este modelo deverá ser capaz de identificar o nível de maturidade que a iniciativa analítica precisa alcançar para suportar os objetivos da empresa. O modelo apresentado também deverá ser capaz de ajudar na elaboração de um roteiro de melhorias para o processo de evolução de todo o ambiente analítico. Ainda como parte do objetivo principal deste trabalho, será apresentado um questionário para realização de uma pesquisa cujas respostas permitam determinar o nível de maturidade da empresa. Vale destacar que os resultados e recomendações desta pesquisa são baseados em respostas e problemas comuns descobertos nas próprias respostas, e que resultados e recomendações, embora tendam a ser precisos, variam de acordo com as condições únicas de cada organização.

Como objetivo secundário, este trabalho fará uma revisão bibliográfica nos conceitos de Modelo de Maturidade e Capacidade, BI e BA. Este trabalho irá explorar os motivos pelos quais a abordagem tradicional de BI com a criação de um grande DW corporativo não mais atende as expectativas das áreas de negócio de uma empresa. Também serão discutidos os principais conceitos envolvidos em uma arquitetura moderna de BI e BA, como self-service, visualização de dados, descoberta de dados, virtualização de dados e ciência de dados que trazem a agilidade e o dinamismo necessários para atender à diferentes capacidades analíticas (descritiva, diagnóstica, preditiva e prescritiva).

1.3 JUSTIFICATIVA

Cada vez mais, os usuários de negócios, aqueles usuários que trabalham fora do departamento de TI, necessitam realizar análises rápidas que levem à descoberta de informações e a geração de conhecimento. Para atender a este tipo de necessidade, é imprescindível uma arquitetura de BI que permita, de uma forma fácil e rápida, acesso a dados de uma ampla gama de fontes e ainda possibilite uma exploração ágil de capacidades analíticas, principalmente, no formato self-service.

Os usuários de negócios querem que uma plataforma moderna de BI e BA seja implementada e operada de forma que lhes permita se beneficiar de uma maior confiança nas análises que eles mesmo criam, ganhar maior autonomia e aumentar a profundidade de percepção de seus processos de análise (TAPADINHAS, 2016).

O uso de analíticos tornou-se, rapidamente, o tópico mais quente no mundo de BI, o seu uso difundiu-se, mas com graus de adoção e sucesso amplamente variados. O trabalho de criação de analíticos está se deslocando rapidamente da área de TI para as unidades de negócios. Nas áreas de negócios, percebe-se uma troca, por parte dos usuários, no uso do Excel com análises baseadas em fórmulas para o uso de ferramentas de análise no formato self-service como Qlik, Spotfire e Tableau (WELLS, 2016).

Estamos presenciando a expansão das práticas de self-service BI, onde a criação de analíticos deixa de ser centralizada pela área de TI e passa a ser distribuída, departamental e conduzida por usuários de negócios. Neste cenário, é comum encontramos organizações cuja demanda pelo uso do self-service seja superior à capacidade de usuários de negócios para desenvolverem analíticos e realizarem as análises. A maioria das organizações se configuram com bolsões ou grupos de indivíduos que executam trabalhos de análise e estes acabam por desconhecerem outros que fazem trabalhos similares. Neste contexto, é muito frequente a existência de trabalho redundante, com frequentes ações de retrabalho, que geram respostas conflitantes e com baixa qualidade de dados a perguntas semelhantes. Adiciona-se a este cenário uma total ausência de governança, o que traz à tona problemas relacionados à segurança dos dados.

Certamente, a idéia não é inibir o crescimento do uso de analíticos e práticas de self-service BI e sim estimulá-los através de uma mudança do crescimento orgânico para o crescimento gerenciado. O gerenciamento anda de mãos dadas com a medição, de modo que

um modelo de capacidade e maturidade analítico é um bom lugar para começar. (WELLS, 2016).

Os modelos de maturidade de BI identificam pontos fortes e fracos na organização, além de fornecer informações comparativas. Os modelos são usados para descrever, explicar e avaliar ciclos de vida de crescimento. O conceito básico de todos os modelos baseia-se no fato de que as coisas mudam ao longo do tempo e que a maioria dessas mudanças pode ser prevista e regulada. O objetivo, é claro, é aumentar o nível de maturidade ao longo do tempo.

Segundo o TDWI (The Data Warehousing Institute), a utilização de um modelo de maturidade para ambiente analítico, através de uma ferramenta de validação, pode ser vista como uma resposta às necessidades das organizações para compreender como elas estão posicionadas em relação à empresas concorrentes ou de mesmo segmento no âmbito de práticas analíticas. O uso de uma ferramenta de avaliação mede a maturidade das ações com analíticos, de uma forma objetiva, através de dimensões que são fundamentais para identificar o seu verdadeiro valor para corporação (HALPER e STODDER, 2014).

Para descrever ou quantificar a capacidade analítica de uma forma significativa, um modelo de capacidade e maturidade deve ser utilizado. O modelo de maturidade em capacitação da Universidade Carnegie Mellon fornece uma boa base e por este motivo, este trabalho explorará a metodologia do CMMI (Capability Maturity Model Integration) como guia destinado a melhorar os processos e a organizar a prática de BI com a adoção dos recursos modernos trazidos por BA. Desta forma, o uso do CMMI não tem como objetivo limitar o crescimento das práticas analíticas, muito pelo contrário, o pensamento é incentivá-las para que se expanda de uma forma gerenciada evitando que cresça de uma forma caótica (WELLS, 2016).

1.4 ESTRUTURA DO TRABALHO

Este trabalho está organizado e estruturado para que o leitor possa entender os conceitos em estudo, o problema existente e os resultados alcançados. Sendo assim, esta monografia está dividida nos seguintes capítulos:

- **Capítulo 1:** Apresenta a contextualização do problema, bem como os objetivos que se pretende alcançar bem como as justificativas por trás do projeto.
- **Capítulo 2:** Discute a importância do uso de modelos para trazer qualidade em processos de desenvolvimento de software e descreve o uso do CMMI (Capability Maturity Model Integration).
- **Capítulo 3:** Contém uma revisão bibliográfica sobre Business Intelligence na ótica dos seus conceitos, definições e componentes tradicionais.
- **Capítulo 4:** Aborda os conceitos que estão envolvidos em uma arquitetura de Business Intelligence e Business Analytics.
- **Capítulo 5:** Traz um roteiro para construção de um modelo de capacidade e maturidade com foco em ambiente analítico, além de apresentar alguns modelos de mercado.
- **Capítulo 6:** Apresenta um estudo de caso de aplicação de um modelo de capacidade e maturidade para ambiente analítico em uma organização.
- **Capítulo 7:** Apresenta as conclusões extraídas deste trabalho, bem como possíveis evoluções que podem ser adicionadas no futuro.

2 USO DE MODELOS EM DESENVOLVIMENTO DE SOFTWARE

2.1 VISÃO GERAL

Com o passar dos anos, a tecnologia cada vez mais influencia as mais diversas áreas da sociedade, em um processo aparentemente irreversível. Vários setores da sociedade apresentam um nível de dependência cada vez maior com relação aos sistemas de informação e em algumas situações chega a ser uma necessidade. Desta forma, uma estruturação do setor de TI com o intuito de aperfeiçoar a qualidade e confiabilidade de seus produtos não só se tornou um diferencial competitivo, mas principalmente um pré-requisito para sua existência. É notório que um grande conjunto de novas metodologias, ferramentas de produtividade e soluções vêm surgindo ao longo dos últimos anos com o objetivo de auxiliar os profissionais da área de TI na consolidação de um processo eficiente que garanta a qualidade de seus produtos (BARTIÉ, 2002).

No atual cenário competitivo de mercado, uma empresa deve prezar cada vez mais pela qualidade de software, pois não se trata apenas de um diferencial de mercado, mas de um pré-requisito que a mesma deve ter para poder obter colocação do seu produto no mercado global.

Segundo Presman (2002), qualidade de software pode ser caracterizada como um conjunto de conformidades contendo requisitos funcionais e de desempenho, apresentando padrões de desenvolvimento previamente documentados e que são utilizados em todo desenvolvimento de software.

Segundo Bartié (2002), qualquer decisão tomada durante o desenvolvimento do software pode afetar a sua qualidade final. Neste sentido, todas as decisões e ações geradas durante todo o ciclo de desenvolvimento caracterizam o produto final. Para produzir com qualidade, é necessário investir em qualidade em todos os pontos do processo de desenvolvimento. A qualidade de software pode ser definida como um processo sistemático que trata em cada etapa os artefatos gerados com o objetivo de garantir a conformidade de processos e produtos, prevenindo e eliminando possíveis problemas e erros.

Para Paulk (2004), a utilização de modelos pode ajudar as organizações a evoluírem de forma sistemática em sua competência para efetuar os compromissos e desenvolver softwares de uma maneira eficaz e eficiente.

A utilização de modelos de maturidade de processos municia as organizações com informações que direcionam a definição de seu plano de melhoria da qualidade e produtividade.

A maturidade de uma empresa afere a capacidade técnica, gerencial e a qualificação que essa corporação tem para desenvolver software. Algumas vantagens que são facilmente percebidas com o uso de modelos de maturidade de processos são o fornecimento de um vocabulário comum de comunicação, a definição de critérios objetivos para avaliação do produto, além de uma maior segurança de que práticas de garantia da qualidade foram aplicadas no desenvolvimento dos produtos e serviços.

Nos dias de hoje, o ambiente analítico tem sido considerado de importância estratégica para as organizações e como elemento facilitador na obtenção de vantagem competitiva. Por esta razão, a utilização de modelos de maturidade poder exercer um papel decisivo, permitindo a avaliação e orientação para um melhor desenvolvimento de aplicações analíticas (DE BRUIN, 2005).

Para Halper e Stodder (2014), a orientação e o controle que os modelos de maturidade proporcionam são totalmente aplicáveis ao desenvolvimento analítico. Estes modelos são de grande importância para organizações interessadas em iniciar a sua implantação ou orientar o seu curso caso já possua alguma iniciativa analítica. A maturidade em ambiente analítico pode ser descrita como a evolução de uma organização para integrar, gerenciar e alavancar todas as fontes de dados internas e externas relevantes, o que significa criar um ecossistema que permita visões e ações. Em outras palavras, a maturidade em práticas analíticas não é simplesmente sobre ter alguma tecnologia instalada, envolve tecnologias, gerenciamento de dados, análise, governança e componentes organizacionais e pode levar anos para criar e introduzir uma cultura analítica na organização. O uso de um modelo de maturidade é útil para qualquer empresa considerando o processo de implementação de um projeto de ambiente analítico, pois ajuda a criar estrutura em torno de um programa e determina por onde começar, também ajuda a identificar e definir os objetivos deste programa, além de criar um processo para comunicar essa visão por toda a organização. Um modelo de maturidade fornecerá uma metodologia para medir e monitorar o estado do programa além do esforço necessário para completar o estágio atual, bem como as etapas necessárias para atingir o próximo estágio de maturidade. O modelo serve como um odômetro para medir e gerenciar o seu progresso e adoção dentro da empresa.

2.2 O MODELO CMMI

2.2.1 Apresentação

O Modelo de Maturidade e Capacidade conhecido como CMM (Capability Maturity Model) descreve uma sequência de características, baseadas em situações atuais de organizações, que obedecem a processos comuns e repetitivos na execução de trabalhos. A extremidade mais baixa da escala descreve as empresas que não usam processos repetitivos, e onde a maioria dos trabalhos são caóticos. A extremidade mais elevada descreve as empresas que usam processos definidos e repetitivos, coletam métricas para auxiliá-los, continuamente, a aprimorar seus processos, e procuram modos criativos para executar suas tarefas com maior eficiência e efetividade (ALMEIDA, 2005).

O CMM foi desenvolvido nos anos de 1984 a 1987 pelo Senhor Watt Humphrey no Instituto de Engenharia de Software (Software Engineering Institute – SEI), que é uma parte da Universidade de Carnegie Mellon nos Estados Unidos. O trabalho foi e continua sendo financiado pelo Departamento de Defesa dos Estados Unidos (DoD), que procurava maneiras para comparar e medir os vários fornecedores que desenvolviam seus softwares (SAMARANI, 2005).

Antigamente, havia vários modelos de CMM, como por exemplo:

- SW-CMM – Capability Maturity Model for Software
- SE-CMM – System Engineering CMM
- EIA 731 SECM – System Engineering Capability Model
- IPD-CMM – Integrated Product Development CMM

Estes modelos eram úteis apesar de utilizados de forma isolada, porém a implementação de vários modelos, dentro da mesma organização, criava uma série de dificuldades, pois as diferenças entre eles limitavam significativamente a possibilidade das empresas focarem nas melhorias destes modelos. Por este motivo, a adoção de um modelo integrado que contivesse as disciplinas necessárias, com suporte para treinamento, avaliação e medição de resultados, poderia ter um efeito melhor na qualidade dos softwares desenvolvidos nas organizações. Como resultado surgiu o CMMI, aonde a letra “I” significa integrado. Este foi criado com dois objetivos distintos. O primeiro, de curto prazo, focava na integração de três modelos

específicos: SW-CMM, SE-CMM e IPD-CMM, já o segundo objetivo, de mais longo prazo, focava na criação de uma base onde fosse possível agregar novas disciplinas ao CMMI.

2.2.2 Representações

Segundo Nunes (2013), com a intenção de medir a qualidade e avaliar as organizações ou partes dela, o modelo CMMI foi elaborado por componentes que estão organizados segundo duas formas de representação: representação contínua e representação por estágios.

A representação contínua possui níveis de capacidade por área de processo (PA), ou seja, a avaliação é feita de maneira independente entre as áreas de processos, conforme mostra o exemplo da figura 1.

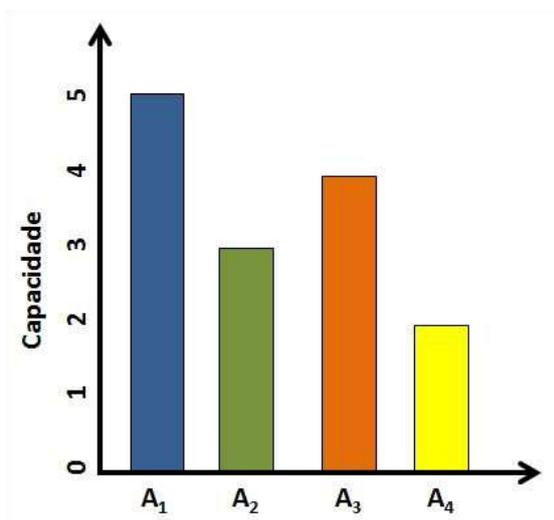


Figura 1 – Representação contínua do CMMI (NUNES, 2013)

Com a representação contínua, a organização consegue ter maior flexibilidade, pois a escolha da área de processo que será avaliada e melhorada é definida com suporte dos objetivos de negócio da organização. Vale destacar, que é possível fazer comparações intra e interorganizações para uma determinada área de processo e ainda realizar uma comparação de resultados por equivalência.

Na representação por estágios a organização é medida como um todo através de níveis de maturidade (ML). Cada estágio de maturidade é composto por uma série de áreas de processos que são avaliadas em conjunto, levando com que a organização progrida como um todo, aonde cada nível se torne a base para o próximo, formando assim uma pirâmide, como mostra o exemplo da figura 2.

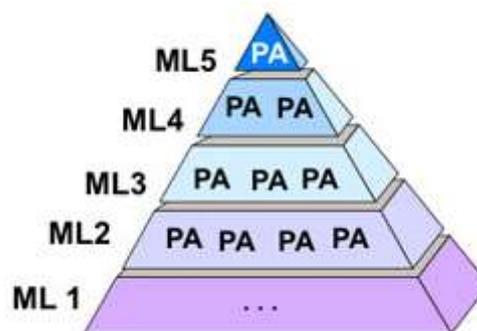


Figura 2 – Representação por estágios do CMMI (NUNES, 2013)

Com a representação por estágios, o próprio modelo já define uma sucessão de melhorias que vão sendo somadas a cada nível. Da mesma forma que na representação contínua, é possível fazer comparações intra e interorganizações, mas desta vez, através de níveis de maturidade, ou seja, em relação a um conjunto de áreas de processos que estão inter-relacionadas.

2.2.3 Níveis de Capacidade

Para Samarani (2005), todos os modelos do CMMI com a representação contínua refletem os níveis de capacidade no seu projeto e conteúdo. Um nível de capacidade consiste em práticas genéricas e específicas, para uma área de processo, que podem melhorar processos organizacionais associados a esta área. Os benefícios de melhoria do processo são alcançados assim que forem atingidos os objetivos genéricos e específicos de uma área de processo em um nível particular de capacidade. Os níveis de capacidade definidos no CMMI são:

Nível 0 – Incompleto: Um processo incompleto é um processo que não é executado ou é executado parcialmente. Objetivos específicos da área de processo não são satisfeitos e não existe nenhum objetivo genérico para este nível, visto que não existe razão para institucionalizar um processo executado parcialmente.

Nível 1 – Executado: Um processo executado é um processo que satisfaz os objetivos específicos para a área de processo, suportando e habilitando as tarefas necessárias para gerar produtos de trabalho. Apesar de uma melhoria em relação ao Nível 0, os processos executados estão em risco devido ao funcionamento sem uma forte ligação com as metas corporativas.

Nível 2 – Gerenciado: Um processo gerenciado é aquele executado que possui a infraestrutura básica para suportar o processo. É executado e planejado de acordo com as políticas existentes, emprega especialistas que possuem recursos adequados para produzir saídas controladas, envolve os usuários mais relevantes, é monitorado, controlado e revisado, possui metas corporativas bem como metas de área de processo e ainda é avaliado pela aderência a descrição do processo.

Nível 3 – Definido: Um processo definido é um gerenciado que tem o rigor necessário em padrões, descrições de processos e procedimentos para serem aprendidos, repetidos, facilmente auditados, consistentes em resultados e capazes de produzir resultados idênticos em circunstâncias idênticas.

Nível 4 - Gerenciado Quantitativamente: Um processo gerenciado quantitativamente é um definido que é controlado usando técnicas estatísticas e outras técnicas quantitativas. Metas mensuráveis de qualidade e desempenho são estabelecidas e usadas para gerenciar o processo. Qualidade e desempenho são medidos e gerenciados ao longo da vida do processo. Este nível foi removido na versão 1.3 do CMMI.

Nível 5 – Otimização: Um processo em otimização é aquele gerenciado quantitativamente que é constantemente aprimorado através da análise e compreensão das causas de variação no processo. O foco de um processo em otimização é a melhoria contínua do desempenho através de melhorias incrementais e inovadoras. Este nível foi removido na versão 1.3 do CMMI.

2.2.4 Níveis de Maturidade

Um nível de maturidade é um patamar evolutivo bem definido, visando alcançar um processo de software maduro. Os níveis são uma forma de priorizar as ações de melhoria de tal forma que se aumente a maturidade do processo de software. Cada nível de maturidade abrange um conjunto de objetivos de processo que quando satisfeitos, consolidam um importante componente do processo de software. Para se estar em um determinado nível, o processo deverá atender a todas as áreas deste nível e de seus níveis inferiores.

O CMM é um modelo para medição da maturidade no que diz respeito a desenvolvimento e administração de software de uma corporação. Seu objetivo é estabelecer o nível de maturidade e projetar os planos de ação para o desenvolvimento da empresa. Uma definição a respeito de maturidade pode ser melhor compreendida através da análise do quadro 1 (ALMEIDA, 2005).

Quadro 1- Definição de maturidade (ALMEIDA, 2005)

Organizações Maduras	Organizações Imaturas
Papéis e responsabilidades são bem definidos	Processos são improvisados
Existe base histórica	Não existe base histórica
É possível julgar a qualidade do processo	Não há maneira objetiva de julgar a qualidade do processo
A qualidade dos produtos e processos é monitorada	Qualidade e funcionalidade do produto são sacrificadas
O processo pode ser atualizado	Não há rigor no processo a ser seguido
Existe comunicação entre o gerente e seu grupo	Resolução de crises imediatas

Segundo Samarani (2005), os níveis de maturidade definidos no CMMI são:

Nível 1 – Inicial: Os processos são geralmente ad hoc e caóticos, ou seja, a corporação não oferece um ambiente estável que suporte os processos. O sucesso nessas corporações depende da competência e heroísmo das pessoas e não de processos experimentados. Aqui, as organizações até desenvolvem produtos que funcionam, entretanto, eles frequentemente excedem os prazos e orçamentos. Esse nível ainda se notabiliza pela vocação de excesso de

comprometimentos, abandono do processo durante as crises e uma incapacidade de repetir os sucessos.

Nível 2 – Gerenciado: Os projetos da corporação têm assegurado que as demandas são gerenciadas e os processos são planejados, executados, medidos e controlados. A disciplina de processo ajuda a assegurar que rotinas atuais são mantidas durante possíveis crises. Com estas práticas, os projetos são executados e gerenciados de acordo com a documentação estabelecida. No nível gerenciado, acordos são estabelecidos entre os usuários chave e são revisados quando necessário e os produtos de trabalho e serviços atendem as suas respectivas especificações.

Nível 3 – Definido: Os processos são bem definidos e entendidos e são descritos em padrões, procedimentos, ferramentas e métodos. O quadro de processos padrão da empresa são a base para a maturidade definida, além de serem aprimorados periodicamente, esses processos padrão fundam uma consistência através da empresa de acordo com as diretrizes organizacionais. Uma distinção clara entre os níveis de maturidade 2 e 3 é com relação ao escopo dos padrões, descrições dos processos e procedimentos. No nível 2, os padrões, as descrições dos processos e procedimentos podem ser diferentes em cada instância específica do processo, já no nível 3, são adaptados a partir do conjunto de processos padrão da organização sendo de uma forma geral, mais consistentes, além de mais rigorosos que no nível 2.

Nível 4 - Gerenciado Quantitativamente: A organização e projetos definem objetivos quantitativos para a qualidade e desempenho do processo e os utiliza como critério para gerenciamento de processos. Estes objetivos quantitativos são fundamentados em necessidades dos usuários de negócios da empresa e dos implementadores do processo. A qualidade e desempenho do processo são entendidos em termos estatísticos e são gerenciados ao longo da vida dos processos. As medidas de qualidade e desempenho do processo são integradas ao repositório de medidas da corporação com a intenção de suportar a tomada de decisão com base em fatos. Uma diferença nítida entre a maturidade nos níveis 3 e 4 é a previsibilidade de desempenho do processo, aonde no nível 4, o desempenho é controlado utilizando técnicas estatísticas ou quantitativas e é qualitativamente previsível, já no nível 3, os processos são previsíveis apenas qualitativamente.

Nível 5 - Em Otimização: Uma organização melhora continuamente seus processos com base na compreensão quantitativa de causas comuns de variação ao processo. A maturidade

nível 5 trabalha na melhoria contínua de desempenho do processo através de um processo incremental e inovador e de melhorias tecnológicas. Com a intenção de refletir as mudanças dos objetivos de negócio, o escopo de melhoria de processo quantitativo para a corporação é definido e continuamente revisado. Uma distinção entre os níveis de maturidade 4 e 5 é com relação ao tratamento do tipo de variação do processo. No nível 4, a organização está preocupada com o tratamento de causas especiais da variação e em providenciar previsibilidade estatística dos resultados, já no nível 5, a preocupação é com o tratamento das causas comuns da variação e com a mudança no processo visando a melhora no seu desempenho para assim atingir os objetivos estabelecidos.

2.2.5 Áreas de Processos

É um conjunto de práticas relacionadas em uma área de negócios da empresa que satisfaz um conjunto de objetivos considerados importantes para se ter melhoria na respectiva área quando implementada coletivamente. Para atingir um determinado nível de capacidade ou maturidade, a corporação deve satisfazer todos os objetivos da área ou conjunto de áreas de processos que se quer evoluir (SEI, 2010).

Pensando em melhoria do processo de desenvolvimento de software, as áreas de processos são exatamente aquilo no que a organização deve focar. Para que uma organização possa se graduar em um determinado nível de maturidade CMMI, a empresa deve executar os processos relacionados às áreas chaves daquele determinado nível (ALMEIDA, 2005).

Segundo o SEI (2010), as áreas de processos são agrupadas em quatro categorias:

- ✓ Áreas de processos de gerenciamento de processos
- ✓ Áreas de processos de gerenciamento de projetos
- ✓ Áreas de processos de engenharia
- ✓ Áreas de processos de suporte

3 CONCEITOS TRADICIONAIS DE AMBIENTE ANALÍTICO

3.1 DEFINIÇÕES NA LITERATURA

Segundo o Gartner Group, BI pode se definido como o processo de transformar dados em informação para no momento seguinte transformá-la em conhecimento. Essa estratégia pode ser tratada como um conjunto de técnicas utilizadas para coletar dados de diversas fontes, organizá-los, analisá-los e compartilhá-los com os responsáveis pela tomada de decisão dentro das organizações.

O termo BI foi utilizado pela primeira vez na década de 50 por Hans Peter Luhn, um pesquisador da IBM, no artigo intitulado “A Business Intelligence System” (ELENA, 2011).

Segundo Hans Peter Luhn (1958), BI é um sistema automático para propagar informação para diversas áreas de qualquer corporação, através do uso de máquinas de processamento de dados, com auto abstração e auto codificação de documentos e ainda criando perfis para cada ponto de ação da empresa por palavra padrão.

Para Santos (2009), BI é a aplicação de um conjunto de técnicas e ferramentas que são ofertadas para apoiar a gestão de um negócio através de uma melhor tomada de decisão.

De acordo com Vercellis (2009), o conceito de BI pode ser definido como técnicas que utilizam modelos matemáticos e metodologias de análise para explorar os dados disponíveis e assim produzir informação e conhecimento para processos de tomada de decisões complexos.

Para o TDWI, é uma estratégia que combina dados, tecnologia, análises e conhecimento humano para potencializar melhores decisões nos negócios e assim guiar o sucesso das empresas. Para isso, uma estratégia é combinar um Enterprise Data Warehouse (EDW) com ferramentas de BI com o objetivo de transformar dados em informações úteis para o negócio (TDWI, 2013).

Para o Gartner, é um termo guarda-chuva que inclui aplicações, infraestrutura e ferramentas, que aliadas com melhores práticas, permitem acesso e análise de informações para promover e otimizar decisões e performance (GARTNER, 2013).

De acordo com Duan e Xu (2012), é o processo de transformação de dados brutos em informações que podem trazer uma nova realidade estratégica para a organização, através de novas descobertas operacionais que produzem benefícios reais para o processo de tomada de decisão nos negócios.

3.2 CONCEITOS CORRELATOS

3.2.1 Data Warehouse (DW)

O termo DW foi criado por Willian H. Inmon no início da década de 1990 e ele o definiu como uma coleção de dados, orientado a um assunto, integrado, variante no tempo e não volátil, para suporte ao gerenciamento dos processos de tomada de decisão (INMON, 1995). Para Inmon estes termos significam:

- ✓ Orientado a Assunto: Os dados trazem informações sobre um assunto particular em vez de sobre operações contínuas da empresa.
- ✓ Integrado: Os dados que são reunidos no DW a partir de uma variedade de origens e integrados de forma coerente.
- ✓ Variante no tempo: Todos os dados no DW são identificados em um período de tempo próprio.
- ✓ Não volátil: Os dados são estáveis em um DW, o que quer dizer que novos dados são adicionados, mas nunca removidos, o que proporciona uma visão consistente dos negócios.

3.2.2 Data Mart (DM)

Para Inmon (2005), um DM é um subconjunto de um DW no que tange a um assunto em especial, evidenciando uma ou mais áreas específicas para assim fornecer suporte à decisão para um pequeno grupo de usuários. A diferença entre DM e um DW pode ser medida no que diz respeito ao tamanho e ao escopo do problema em questão, ou seja, as definições dos problemas e os requisitos de dados os mesmos para ambos.

Segundo Primak (2008) quando comparados a um DW, os DMs possuem algumas vantagens como custo mais baixo, escopo mais simplificado e um menor tempo de implementação, pois o DM atende apenas uma área ou assunto específico de uma organização, já o DW requer um tempo de desenvolvimento maior pois envolve toda a organização e assim os custos são maiores.

3.2.3 Extract, Transform and Load (ETL)

O processo de ETL tem como objetivo extrair, transformar e carregar dados de uma ou mais bases de origem para uma ou mais bases de destino, sendo o processo mais demorado e crítico na construção de um DW. A maior parte do esforço exigido na implementação de um DW é gasto no desenvolvimento de soluções de ETL. Estima-se que oitenta por cento de todos os esforços sejam empregados em processos de ETL (INMON, 2005).

Segundo Barbieri (2001), o processo de ETL se resume basicamente em 5 passos:

1. Mapear a origem dos dados a serem coletados, considerando que as fontes podem estar espalhadas em diversos sistemas transacionais e banco de dados da empresa.
2. Limpar os dados para facilitar a posterior transformação, realizando os ajustes necessários para corrigir imperfeições e assim oferecer um melhor resultado para o usuário final.
3. Transformar os dados com o objetivo de padronizá-los em um único formato.
4. Executar a carga dos dados para o DW.
5. Atualizar os dados no DW a partir das alterações sofridas nos sistemas operacionais da empresa.

Segundo Primak (2008), é importante ter em mente que a etapa de ETL é uma das mais críticas na construção de um DW, pois envolve a fase de movimentação dos dados.

3.2.4 Operational Data Store (ODS)

Segundo Barbieri (2001) o ODS está diretamente ligado ao armazenamento e tratamento de dados transacionais, de forma consolidada, e sem as características dimensionais, ou seja, é um repositório integrado e consolidado com dados de diversas origens, que contem os detalhes do dia a dia de trabalho da empresa. O ODS pode ser consultado por outros sistemas ou explorado na realização de análises e ainda encarado como a visão integrada do mundo operacional. Normalmente, a sua implementação adota bases de dados relacionais e o seu armazenamento se dá através de ferramentas de ETL, que extraem dados de banco de dados de

diversas origens e os coloca, de forma integrada. Além disso, há estratégias que utilizam o ODS como base origem para a construção de um DW, pois nele é feita a uniformização e conformação de todos os dados da organização.

3.2.5 Staging Area (SA)

Também chamada por Inmon (1995) como camada de integração e transformação, a SA é uma área de tratamento, padronização e transformação dos dados oriundos de aplicações transacionais para carga na arquitetura de dados do BI (ODS, DM, DW).

Um bom entendimento para SA é considerá-la como sendo uma área de armazenamento temporário, onde os dados de origem são copiados e preparados para a carga final no ambiente analítico. Nesta área, os dados são previamente validados, limpos ou rejeitados de acordo com regras genéricas como formatos, campos obrigatórios e regras de integridade em geral.

Normalmente, a SA fica separada do DW para não haver concorrência em acessos e por possuírem estratégias de infraestrutura, backup e log diferentes.

3.2.6 Sistemas OLTP e OLAP

As siglas OLTP e OLAP são bastante utilizadas no contexto analítico, mas possuem conceitos divergentes e são aplicadas em contextos diferentes.

O OLTP, do inglês "On-line Transaction Processing", é a sigla utilizada em alusão aos sistemas operacionais das corporações, ou seja, se referem ao processamento dos dados das rotinas diárias dos sistemas informacionais da empresa dando suporte à atividade finalística da organização.

Já o OLAP, do inglês "On-line Analytical Processing", é a sigla utilizada em referência a capacidade de se analisar grandes volumes de dados por diferentes óticas dentro de um DW.

Desta forma, podemos fazer uma classificação de sistemas de acordo com o seu tipo de utilização, e assim dividimos em dois grandes grupos: Sistemas Transacionais e Sistemas Analíticos. O quadro 2 traz as principais diferenças entre sistema OLTP e OLAP.

Quadro 2 - Diferenças entre sistema OLTP e OLAP (ELIAS, 2014)

	OLAP	OLTP
Foco	Foco no nível estratégico da organização. Visa a análise empresarial e tomada de decisão.	Foco no nível operacional da organização. Visa a execução operacional do negócio.
Performance	Otimização para a leitura e geração de análises e relatórios gerenciais.	Alta velocidade na manipulação de dados operacionais, porém ineficiente para geração de análises gerenciais.
Estrutura dos dados	Os dados estão estruturados na modelagem dimensional. Os dados normalmente possuem alto nível de sumarização.	Os dados são normalmente estruturados em um modelo relacional normalizado, otimizado para a utilização transacional. Os dados possuem alto nível de detalhes.
Armazenamento	O armazenamento é feito em estruturas de <i>Data Warehouse</i> com otimização no desempenho em grandes volumes de dados.	O armazenamento é feito em sistemas convencionais de banco de dados através dos sistemas de informações da organização.
Abrangência	É utilizado pelos gestores e analistas para a tomada de decisão.	É utilizado por técnicos e analistas e engloba vários usuários da organização.
Frequência de atualização	A atualização das informações é feita no processo de carga dos dados. Frequência baixa, podendo ser diária, semanal, mensal ou anual (ou critério específico).	A atualização dos dados é feita no momento da transação. Frequência muito alta de atualizações.
Volatilidade	Dados históricos e não voláteis. Os dados não sofrem alterações, salvo necessidades específicas (por motivos de erros ou inconsistências de informações).	Dados voláteis, passíveis de modificação e exclusão.
Tipos de permissões nos dados	É permitido apenas a inserção e leitura. Sendo que para o usuário está apenas disponível a leitura.	Podem ser feito leitura, inserção, modificação e exclusão dos dados.

3.2.7 Modelo Dimensional

Segundo Kimball (2002), modelagem dimensional é uma técnica de estruturação de base de dados cuja finalidade é suportar consultas executadas por usuários finais em um DW.

O modelo multidimensional é uma técnica de modelagem conceitual de negócios, que simplifica a investigação, o extrato e a organização de dados para a análise de negócios. Este tipo de modelagem permite visualizar dados abstratos de uma maneira simplificada ao

relacionar informações de diferentes segmentos da empresa de forma eficaz. Este modelo relaciona tabelas de fatos com tabelas de dimensões em um banco de dados do tipo OLAP, possibilitando combinações ilimitadas e um desempenho melhor na execução de consultas.

É pertinente lembrar que apesar de se tratar de modelagem dimensional os dados que serão refletidos nessa modelagem continuam sendo relacionais.

Segundo Kimball (2002), os seguintes elementos fazem parte do modelo dimensional:

- ✓ Tabela Fato: É a principal tabela de um modelo dimensional, é o local aonde as medidas de interesse da empresa estão armazenadas. O fato representa uma medida de algum assunto modelado, como quantidades, valores e indicadores. A tabela de fatos é composta por uma chave primária, formada por uma combinação única de valores de chaves de dimensão e pelas métricas relativas ao assunto. Vale destacar que a tabela de fatos não deve misturar assuntos diferentes na mesma estrutura.
- ✓ Tabela Dimensão: É composta de atributos e contém as perspectivas de negócio pelas quais se deseja analisar um assunto ou fato. Seus atributos são utilizados nas restrições de consultas, agrupamento dos resultados e cabeçalhos de relatórios. Em termos de volume de dados, as tabelas de dimensão costumam ser bem menores que as tabelas de fato.
- ✓ Métricas: São as informações armazenadas nas tabelas fato e que são usadas para medir o desempenho dos processos de negócio. As métricas são informações numéricas, que podem ou não estarem agregadas e que na maioria das vezes são aditivas, permitindo a realização de operações como adição, subtração e médias. Podem existir outros dois tipos de métricas, as não aditivas e as semi-aditivas. As métricas não aditivas não podem ser sumarizadas ou agregadas através das dimensões. Já as métricas semi-aditivas só podem ser sumarizadas ou agregadas por um conjunto restrito de dimensões.

De acordo com Kimball (2002), é possível implementar um modelo dimensional através de dois esquemas:

- ✓ Esquema Estrela, do inglês Star Schema é uma estrutura simples, com poucas tabelas e com relacionamentos bem definidos e com grande proximidade ao modelo de negócio, o que facilita a leitura e entendimento, não só pelos analistas, como por usuários finais não familiarizados com estruturas de banco de dados. Uma de suas características é a presença de dados altamente redundantes, o que proporciona um melhor desempenho. O nome estrela vem da disposição das tabelas no modelo, que apresenta uma tabela central, a tabela de fatos, que se relaciona com diversas outras tabelas, as tabelas de dimensão. A figura 3 apresenta a estrutura geral de um esquema estrela.

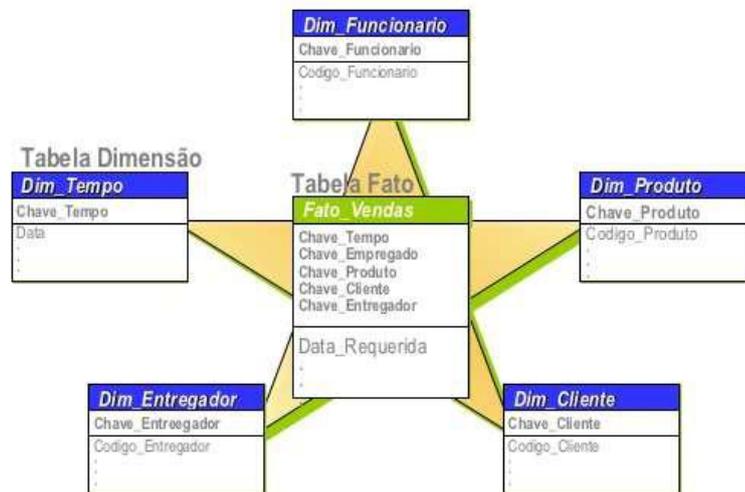


Figura 3 - Modelo estrela (RESENDE, 2017)

- ✓ Esquema Floco de Neve, do inglês Snow Flake, é uma variação do esquema estrela, aonde tabelas de dimensão são normalizadas na terceira forma normal. O esquema floco de neve deve ser utilizado apenas em situações em que a linha da dimensão estiver muito longa, o que pode ser relevante do ponto de vista de armazenamento. Um ponto relevante no uso deste esquema é que a normalização das tabelas de dimensão traz uma maior complexidade ao modelo de dados, diminuindo assim a compreensão desse modelo por parte dos usuários. Uma outra questão é que a normalização traz um aumento no número de relacionamentos o que pode vir a diminuir a performance das consultas realizadas pelos usuários. Esta consequência pode ser crucial para o sucesso do

ambiente analítico, pois os usuários esperam respostas cada vez mais rápidas em suas análises.

3.2.8 Cubo de Dados

O princípio de uso para a modelagem dimensional é que a maioria dos tipos de dados de negócio possam ser representados em um cubo de dados, onde as células deste cubo apresentem valores medidos enquanto que os lados do cubo apresentem as dimensões dos dados. Um cubo de dados pode ter mais que três dimensões, neste caso, podendo ser chamado de hipercubo. De uma forma em geral, os termos cubo e cubo de dados são usados como sinônimos de hipercubo (KIMBALL, 2002).

Cubo é a estrutura multidimensional de dados que exterioriza o jeito pelo qual os tipos de informações se relacionam entre si. Um cubo é constituído por tabela de fatos e por tabelas de dimensão que a contornam e representam possíveis maneiras de visualizar e consultar os dados. O cubo armazena todas as informações relacionadas a um assunto qualquer, de maneira que diversas combinações possam ser estabelecidas entre elas, resultando na geração de várias visões sobre o mesmo tema.

Podemos considerar que um cubo de dados é um tipo de matriz, no formato multidimensional, que permite a exploração e análise de uma coleção de dados sob diversas perspectivas, considerando algumas dimensões ou fatores de cada vez.

Podemos identificar alguns tipos de operações básicas que são utilizados em cubos de dados para realizar análises:

- ✓ Drill up/down – Utilizado para aumentar ou reduzir o nível de detalhe da informação acessada.
- ✓ Slicing – Utilizado para selecionar as dimensões a serem consideradas na consulta.
- ✓ Dicing – Utilizado para limitar o conjunto de valores a ser mostrado, fixando-se algumas dimensões.
- ✓ Pivoting – Utilizado para inverter as dimensões entre linhas e colunas.

3.2.9 Dashboards

O termo dashboard ou painel de bordo é utilizado para apontar um painel de indicadores, como por exemplo em um automóvel, aonde encontramos um visor gráfico com informações de velocidade, de rotações do motor, de temperatura do motor, do nível do óleo, da quantidade de combustível e que em determinada situação pode emitir algum tipo de alerta.

De acordo com Nascimento (2017), os dashboards são painéis que apresentam métricas e indicadores importantes com o objetivo de atingir metas traçadas em um formato visual e que facilite a compreensão das informações geradas. Neste aspecto, apesar de existirem diversos tipos de apresentação de informação, o formato visual se destaca quando se trata de entender como estão indo resultados, possibilitando o consumo da informação por qualquer usuário. O uso de dashboards, possibilita o monitoramento dos resultados de uma empresa através de diversos indicadores. Neste contexto surge o contexto de indicador chave de performance ou KPI, do inglês “Key Performance Indicator”.

Em uma organização, podemos considerar a existência de três diferentes tipos de dashboards que atendem a diferentes escalões da empresa, cada qual com suas responsabilidades e necessidades de monitoramento:

- ✓ Os operacionais, que mostram as métricas que devem ser acompanhadas por analistas para otimizarem seus trabalhos no seu dia a dia e assim obterem uma maior agilidade em correções necessárias.
- ✓ Os táticos, que permitem ter uma visão de como estão as operações em relação as estratégias criadas
- ✓ Os estratégicos, que normalmente são KPIs criados de acordo com metas estabelecidas.

3.2.10 Mineração de Dados

Segundo Addrians e Zantinge (1996), mineração de dados, do inglês data mining é uma tecnologia que nasceu da confluência de três áreas: estatística clássica, inteligência artificial e aprendizado de máquina. Pode-se afirmar que a mineração de dados pertence a um processo maior conhecido como KDD (Knowledge Discovery in Databases), em português, Descoberta

de Conhecimento em Bases de Dados. O KDD permite a extração de conhecimento previamente desconhecido e potencialmente útil de um banco de dados. Esse conceito é enfatizado por Fayaad (1996), ao afirmar que mineração de dados é um processo complexo de identificação de padrões desconhecidos, válidos, e potencialmente úteis que são compreensíveis em dados. Um processo de KDD tem por objetivo descobrir informações valiosas que estão escondidas em um grande conjunto de dados e que podem ser de grande importância para apoiar tomadas de decisões estratégicas. Este processo pode ser dividido em cinco etapas conforme a figura 4.

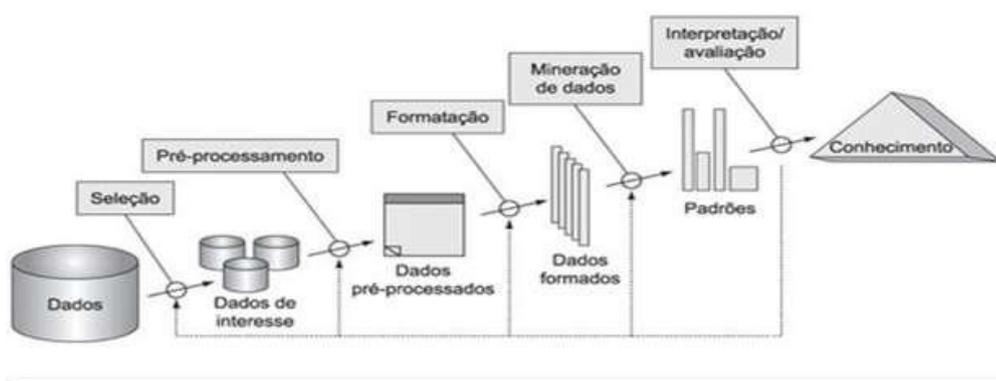


Figura 4 - Etapas do processo KDD (Adaptação de FAYYAD 1996)

- ✓ Seleção: A etapa de seleção é a primeira etapa do KDD, e é nesta que serão avaliados os conjuntos de dados relevantes que serão utilizados na tentativa de gerar resultados com informações úteis.
- ✓ Pré-processamento: Nesta etapa, ocorre um trabalho de limpeza dos dados, seleção de atributos e correção de informações ausentes, errôneas ou inconsistentes nas bases de dados, para que não haja comprometimento com a qualidade dos modelos de conhecimento a serem extraídos ao final do processo de KDD.
- ✓ Formatação: Nesta etapa, os dados obtidos na etapa anterior são reorganizados em uma forma específica para que possam ser interpretados na etapa posterior.
- ✓ Mineração de dados: Nesta etapa é onde meros dados são transformados em informações através de algoritmos específicos e com uma metodologia que procura uma descrição lógica ou matemática, eventualmente de natureza complexa, de padrões e regularidades.

- ✓ **Interpretação:** A última etapa do KDD é a de interpretação de resultados, aqui as regras indicadas pelo processo anterior serão interpretadas e avaliadas. Após a interpretação, poderão aparecer novos padrões e descoberta e relacionamentos de novos fatos que poderão vir a ser utilizados para pesquisas, otimização e outros.

4 CONCEITOS COMPLEMENTARES EM AMBIENTE ANALÍTICO

4.1 INTRODUÇÃO

Business Analytics (BA) pode ser definido como um conjunto de tecnologias utilizadas para modelar e resolver problemas reais de decisão. Sua natureza é multidisciplinar pois utiliza conhecimentos da matemática, estatística, engenharia de produção, computação, ciências da administração que são combinados com regras de negócio específicas de cada empresa, como pode ser observado na figura 5. A utilização de BA permite um desempenho em tomada de decisão nunca antes alcançado por organizações dos mais diversos setores no mundo todo (NEOINFINITO, 2015).

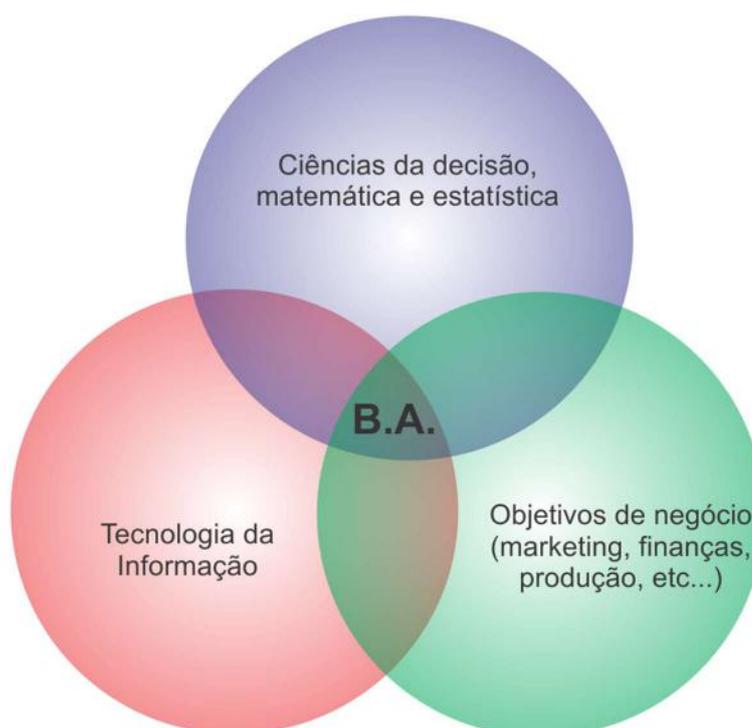


Figura 5 - Business Analytics: integração entre múltiplos campos (NEOINFINITO, 2015)

De acordo com Guerra (2017), BA pode ser definido como uma solução que permite obter conhecimento do negócio em tempo real, por meio da análise dos dados. Apresenta capacidade de analisar os dados em tempo real, permitindo e otimizando as análises preditivas. Facilita o cruzamento de informações históricas que não necessariamente estejam na base da

ferramenta. Normalmente é utilizado por companhias que necessitam antecipar cenários e prever o que está por vir. No varejo, por exemplo, auxilia na previsão de tendências sazonais.

Com a tecnologia atual disponível, as empresas são capazes de coletar quantidades enormes de dados com relativa facilidade, com isso, muitas empresas têm agora mais dados do que podem lidar, no entanto, estes dados são geralmente sem sentido até que sejam analisados para descoberta de tendências, padrões, relações e outras informações úteis (ALBRIGHT e WINSTON, 2013).

Para Kosani e Kadre (2015), BA é tudo que envolve dados, metodologias, TI, aplicações, técnicas matemáticas, métodos estatísticos e habilidades necessárias para obter novos conhecimentos de negócios e assim compreender o seu desempenho. O BA é um processo de transformação de dados em ações através de análises e descobertas no contexto da tomada de decisão organizacional e da resolução de problemas que utiliza a exploração iterativa e metódica de dados passados para suportar decisões.

O uso de analíticos é atualmente o termo do momento, não porque seu uso é exagerado na imprensa especializada, mas porque dezenas de gestores de BI estão buscando maneiras de entregar mais valor a partir de seus investimentos em dados. Esses gestores criaram DW's e entregaram relatórios e painéis, mas agora querem modernizar seus programas de BI. Infelizmente, o caminho de um BI tradicional para o BA não é uma linha direta, ao invés disso exigem que os gestores repensem suas arquiteturas organizacionais, de dados e analíticas. Neste contexto, pode ser necessário que os gestores descartem seus programas de BI atuais e evangelizem novas maneiras de capturar e usar dados para alavancar negócios (ECKERSON, 2015).

Para o Gartner, BA é um conjunto de soluções usadas para construir modelos de análise e simulações, para criar cenários, compreender realidades e prever estados futuros. O BA inclui a mineração de dados, análises preditivas, análises prescritivas, análises diagnósticas e estatísticas, e é entregue como uma aplicação ágil e flexível para usuários de negócios.

4.2 ESPECTRO DE CAPACIDADES ANALÍTICAS

Segundo o Gartner, uma plataforma analítica para ser bem-sucedida deve oferecer soluções diferentes para as diversas necessidades dos usuários, apresentar um conjunto diversificado de níveis de integração e proporcionar um equilíbrio entre confiança e agilidade.

De acordo com Tapadinhas (2016), o objetivo consiste em ajudar os usuários a alcançar seus objetivos através do uso da tecnologia adequada, não erradicando as soluções de BI tradicionais que resolvem parcialmente seus problemas atuais. Neste cenário, podemos identificar quatro estilos analíticos para compor um espectro de capacidades analíticas: Descritivo, Diagnóstico, Preditivo e Prescritivo, conforme a figura 6.



Figura 6 - Espectro de capacidades analíticas (TAPADINHAS, 2016)

- ✓ **Análise Descritiva:** Este tipo de análise responde à pergunta: “O que aconteceu?”. A análise descritiva é baseada em dados históricos e dados atuais da empresa, aonde um analista de negócios cruza informações com o objetivo de gerar um panorama claro e preciso dos temas relevantes no presente momento.
- ✓ **Análise Diagnóstica:** Usa-se esse tipo de análise para encontrar respostas para a pergunta: “Por que algo específico aconteceu?” Ou “O que deu errado?”. Nesta capacidade, o foco está na relação de causas e consequências percebidas ao longo do tempo, dentro de um determinado tema. Assim, a análise diagnóstica funciona com base na coleta de dados relacionados a um determinado assunto, cruzando informações com o objetivo de entender quais fatores influenciaram o resultado atual.
- ✓ **Análise Preditiva:** Apesar deste tipo de análise ser baseado em dados históricos e atuais, ele é o mais indicado para quem precisa prever algum tipo de comportamento ou resultado. Esta técnica busca analisar dados relevantes ao

longo do tempo, buscando padrões comportamentais e suas variações de acordo com cada contexto, a fim de prever como será o comportamento de um determinado processo no futuro, dadas as condições atuais. Este tipo de análise envolve a construção de modelos complexos de análise, para assim prever um evento futuro ou tendência.

- ✓ **Análise Prescritiva:** Esta capacidade possui um modelo similar à preditiva, porém com objetivos ligeiramente diferentes, pois ao invés de tentar prever um determinado acontecimento, esta análise busca trazer informações de consequências deste acontecimento. Para efeitos de comparação, enquanto o modelo preditivo foca em tendências, a análise prescritiva busca trazer o impacto que essas tendências terão no futuro. Ou seja, este tipo de análise tem como objetivo otimizar processos, estruturas e sistemas através de ações baseadas no que você deve fazer com base em uma estimativa do que vai acontecer.

Antes de realizar uma análise de dados, a chave é definir o tipo da questão a ser perguntada. Com relação a dados, algumas perguntas são mais fáceis de responder e outras são mais difíceis. A figura 7 traz um gráfico com um fluxo de perguntas analíticas classificadas por sua complexidade. Este gráfico pode auxiliar a escolha do tipo de estilo do espectro de capacidade analíticas a ser utilizado (LEEK, 2015).

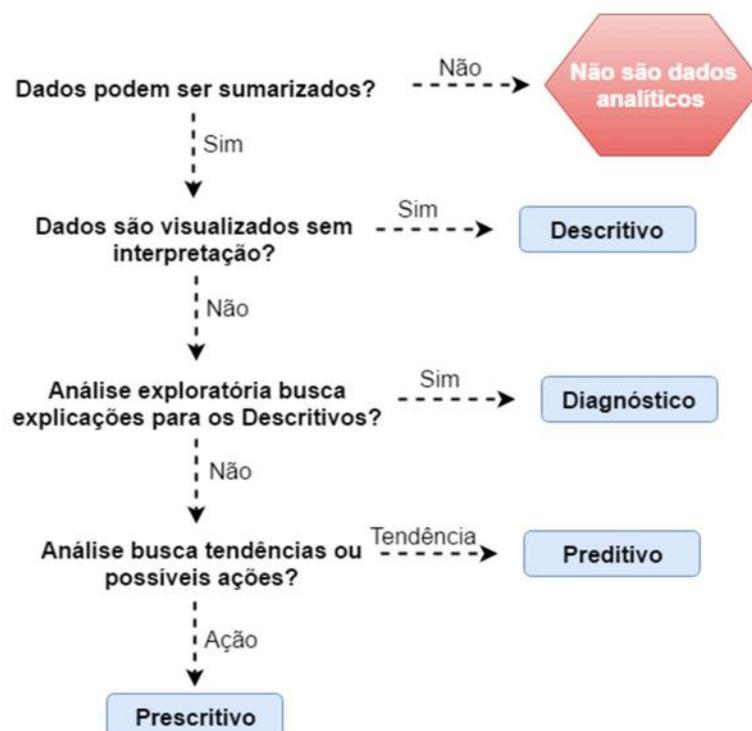


Figura 7 - Fluxo de questões analítica (LEEK, 2015)

4.3 SELF-SERVICE BI

Um dos termos mais badalados em BA, self-service pode ser traduzido do inglês como autosserviço ou autoatendimento e significa a possibilidade de conseguir um benefício a partir de uma atividade sem a necessidade de um intermediário (EXAME, 2016).

Segundo o Gartner, Self-Service BI é um conceito em que um usuário não técnico possa desenvolver e implantar suas próprias análises, relatórios e dashboards através de ferramentas de uma arquitetura que lhe entreguem a autonomia necessária para conceber e modelar a implementação de analíticos que lhes convém.

De acordo com Scheibler (2017), a essência do self-service BI deve garantir um modelo de autoatendimento para usuários de ferramentas de BI de modo a amplificar a obtenção de informações, realizar consultas, análises e gerar relatórios personalizados a qualquer tempo, a partir de qualquer lugar, agregando inteligência, autonomia e agilidade à gestão de diversos setores da corporação. O self-service BI concede aos usuários a liberdade necessária para que o foco do trabalho esteja apenas nas informações e análises que realmente interessam às suas atividades e sem desperdiçar seu tempo com foco na tecnologia, a qual fica exclusivamente com a TI. A área de TI tem como responsabilidade estabelecer a conectividade com os dados dos sistemas transacionais e habilitar os usuários sobre quais e onde estão as informações que lhes interessam, ou seja, enquanto a TI cuida do elenco e organização dos dados, o usuário foca apenas no significado deles. Dentro deste contexto, é imprescindível não perder de vista a segurança e a governança, sendo necessário abranger políticas de controle de qualidade dos dados, controle de acesso, recursos de privacidade, monitoramento e capacidade de gestão de todos estes itens.

Para Eckerson (2014), o self-service tem sido o Santo Graal do BI nos últimos 15 anos, pois promete libertar os usuários de negócio da sua dependência para com a equipe de TI, permitindo a criação de relatórios e análises que precisam, quando precisam e como querem. Porém, o problema com self-service BI é mais complexo do que parece pois não há uma ferramenta de self-service que atenda às necessidades de todos os usuários de negócios. Na verdade, existem tantas ferramentas de self-service quanto existem tipos de usuários. Para um executivo, self-service significa que ele pode clicar em um gráfico e analisar detalhes, enquanto um cientista de dados vê o self-service como uma ferramenta para suportar consultas avançadas com utilização de métodos estatísticos e visualização gráfica. Assim, a chave para o self-service

BI envolve o mapeamento dos tipos de usuários e seus requisitos de informação para os tipos de ferramentas no mercado.

4.4 USUÁRIOS

Para Eckerson (2016), uma plataforma analítica deve fornecer capacidades diferentes para perfis de usuários diferentes, cada qual com conhecimentos e necessidades também diferentes. Em geral, existem quatro classes de usuários analíticos, cada qual com uma noção de self-service BI:

- ✓ Administradores: Instalam, configuram e realizam tuning nas ferramentas, gerenciam usuários e permissões, monitoram performance e disponibilidade, solucionam problemas e gerenciam os contratos de licenciamento.
- ✓ Usuários casuais: Visualizam e interagem com relatórios, dashboards e demais analíticos pré-construídos. Podem ser divididos em dois grupos:
 - Visualizadores: Consomem relatórios, porém não interagem com estes, geralmente recebem estes relatórios estáticos por e-mail e não realizam questionamentos ou análises.
 - Exploradores: Interagem com relatórios e dashboards através de filtros, drills, agrupamentos, pivoteamento, ordenações, cálculos entre outras operações com os dados em questão para aumentar o entendimento sobre estes dados. Frequentemente salvam fotografias de relatórios e análises modificadas, porém nunca editam conexões e modelo de dados.
- ✓ Usuários avançados: São analistas de negócios e cientistas de dados com conhecimento em modelos de dados e ou estatística. Podem ser categorizados em:

- Iniciante: Analistas que customizam relatórios e dashboards usando o modelo de dados existente para criar novas métricas, dimensões, hierarquias e visualizações. Os analistas arrastam os objetos do modelo de dados para criar consultas ad-hoc e montar novas visualizações na forma de relatórios ou dashboards.
 - Intermediário: Os analistas vão um passo além dos iniciantes pois se conectam a novas fontes de dados que vão desde arquivos locais, aplicações corporativas e até fontes externas. Utilizam estas fontes para criar novos relatórios e dashboards ou alterar existentes, mesclando os dados originais com os novos em um novo modelo de dados.
 - Avançado: Os analistas são cientistas de dados que sabem realizar codificações em SQL, Java, Python ou outras linguagens de programação e ainda possuem conhecimento das fontes de dados, podendo criar relatórios e análises customizados.
- ✓ Desenvolvedores: Usam ferramentas de BI e suas API's para criar aplicações customizadas, construir relatórios e dashboards e embarcá-los dentro de outras aplicações.

4.5 VISUALIZAÇÃO DE DADOS

Segundo Eckerson (2015), os seres humanos são conectados para processar informações visualmente, nossos olhos contêm 70% de todos os receptores sensoriais em nosso corpo e metade dos recursos do nosso cérebro são dedicados ao processamento visual. No campo dos negócios, uma visualização bem projetada pode rapidamente comunicar uma quantidade grande de informações aos usuários que precisam tomar decisões e agir. Ou seja, a visualização de dados tem um enorme potencial tanto para acelerar o ritmo do negócio quanto melhorar a qualidade das decisões. Infelizmente, a análise visual de dados ainda está em sua infância e poucas pessoas estão capacitadas para utilizá-la em análises de informações quantitativas. A maioria das empresas ainda produz um número incontável de relatórios tradicionais que contêm

longas tabelas de informações numéricas, que são lidas por poucas pessoas e que agregam pouco valor ao negócio. O objetivo da visualização gráfica não é esconder a verdade, mas sim ampliá-la, ou seja, o que é importante é visto rapidamente. Um relatório visual bem projetado consegue destacar a mensagem por trás dos dados e instantaneamente impulsiona os usuários a discutir, decidir e agir. Esse é o poder da visualização de dados, mudando de ver para fazer em um nanosegundo.

Para Wells (2017), a visualização de dados está no coração do BA, é o meio pelo qual transformamos quantidades de dados complexos em comunicações empresariais compreensíveis, perspicazes e muitas vezes convincentes. A eficácia e o impacto no uso de painéis analíticos dependem da habilidade de criar e de compreender os recursos visuais. O valor real da visualização não vem da criação de imagens, mas da compreensão que elas podem dizer. Tornar painéis analíticos compreensíveis, depende da habilidade de quem os faz, mas o valor e o impacto derivam daqueles que os leem, interpretam e usam para conduzir conversação, colaboração e inovação. É correto afirmar que a análise visual de dados se tornou uma linguagem, a linguagem das imagens. Uma comparação com a linguagem das palavras que possui componentes como verbos, substantivos, adjetivos e regras de sintaxe e gramática, a linguagem visual também possui componentes e regras como sistemas de coordenadas, escalas, dicas visuais que são utilizados na elaboração dos painéis gráficos.

4.6 NARRATIVA DE DADOS

Segundo Wells (2017), a narrativa de dados tem se tornado um termo popular e importante no campo da análise de dados, pois é uma forma de olhar profundamente nos dados para encontrar as mensagens importantes e comunicá-las da maneira mais eficaz. É um conceito poderoso que às vezes é diminuído por aqueles que o posicionam como sinônimo de visualização de dados, como vendedores que o adotam como uma palavra-chave de marketing, ou especialistas que o equiparam às melhores práticas de visualização.

A visualização de dados é uma parte importante em um ambiente de BA, porém a narrativa de dados vai muito além da apresentação visual destes, pois ocorre uma mescla da linguagem das imagens com a linguagem das palavras através da complementação do visual gráfico com a sua narrativa, gerando uma história para os dados.

Histórias são coisas poderosas e nós as usamos ao longo dos tempos para capturar a atenção, transmitir idéias, instigar a imaginação e agitar a alma. Os dados podem ser

persuasivos, mas as histórias são muito mais. Uma história bem contada é uma narrativa inspiradora que é elaborada para envolver o público. Histórias produzem um impacto que não é possível de se obter com os dados isolados, pois elas mudam a maneira como interagimos com os dados, transformando-o de uma coleção seca de fatos para algo que pode ser divertido, atraente, provocador e inspirador. Existem quatro boas razões para perseguir a arte de contar histórias:

- ✓ As histórias são memoráveis, a maioria das pessoas retém a mensagem de uma história por mais tempo e com uma lembrança melhor do que a retenção de números e estatísticas.
- ✓ As histórias são passadas e recontadas, as pessoas gostam de compartilhar uma história boa, então o alcance e o impacto social das histórias são tipicamente maiores do que para os números.
- ✓ Histórias são convidativas para um envolvimento pessoal e participativo. Uma história bem contada faz com que o público se veja em algum lugar dela, trazendo um nível mais alto de engajamento, comunicação, conversa e envolvimento geral.
- ✓ As histórias motivam as pessoas a agir. Quando as histórias se tornam pessoais, elas também se tornam motivadoras e mais propensas a conduzir a ação do que são estatísticas e números. A narrativa faz muito mais do que apenas por um rosto bonito em fatos.

Criar uma narrativa pode ser a mola mestra para a alavancagem de BA em uma empresa, pois por si só os dados são ambíguos e contextualmente deficientes, além do que apenas as visualizações estão sujeitas a interpretações, e em algumas vezes, interpretações erradas. A narrativa reduz ambiguidade, conecta dados com contexto e descreve uma interpretação específica.

4.7 PREPARAÇÃO DE DADOS NO FORMATO SELF-SERVICE

A preparação de dados é um tópico importante nas áreas de TI e de negócios das organizações e atualmente, é também o foco nas indústrias de tecnologia de software que procuram inovar métodos para acelerar e automatizar os processos necessários para suportar análises de dados. Preparar, mesclar, integrar, limpar, transformar, governar e definir os

metadados de múltiplas fontes de dados, incluindo big data, tem sido um trabalho da área de TI. No entanto, o interesse crescente na ciência dos dados e na análise atraiu pessoas de outras áreas para a execução dessas tarefas. Usuários que não sejam de TI, como analistas de negócios e dados estão procurando ferramentas de self-service mais inteligentes que reduzam as dificuldades e tornem os processos de preparação de dados mais rápidos (STODDER, 2016).

Segundo o Gartner, a preparação de dados, em um modelo self-service, é um processo iterativo para explorar e transformar dados brutos em formatos adequados para experiências de ciência de dados, descoberta de dados e criação de analíticos em geral. As ferramentas utilizadas para esta finalidade devem ser orientadas a usuários de negócios e com capacidade para executar uma combinação de funções de integração e limpeza de dados. Como a integração de dados é uma das atividades mais desafiadoras e demoradas em um ambiente analítico, essas ferramentas visam reduzir o tempo e a complexidade nos processos de preparação de dados, melhorando assim a qualidade dos dados em um tempo mais adequado às necessidades de negócio. (SELVAGE e THOO, 2017).

4.8 BIG DATA

Estima-se que o mundo gere cerca de 2,5 quintilhões de bytes em diferentes ações no dia a dia na sociedade. Estas ações que vão desde as manifestações de usuários em redes sócias, movimentações financeiras a registros em sistemas corporativos tornaram-se ativos valiosos para corporações, que podem utilizá-los para conhecerem melhor seus clientes, entenderem seu comportamento de compra e até previrem uma crise no setor ou a perda de clientes para a concorrência (HEKIMA).

De acordo com o Gartner, big data pode ser definido como recursos de informação de alto volume, alta velocidade e ou de alta variedade que demandam formas de processamento de informações inovadoras e que habilitam a criação de novas percepções de negócio, e assim proporcionam uma melhor tomada de decisão e automação de processos.

Segundo Alecrim (2015), o conceito de big data pode ser explicado como conjuntos de dados excessivamente grandes que necessitam de ferramentas adequadas para lidar com esta volumetria, de forma que toda e qualquer informação neste contexto possa ser encontrada, analisada e utilizada em tempo hábil. Simplificando, este conceito pode ser explicado como a análise de grandes volumes de dados para a geração de resultados que em volumes menores dificilmente seriam alcançados.

A base do conceito de big data pode ser resumida em cinco aspectos: volume, velocidade, variedade, veracidade e valor e segundo Barony (2017), estes cinco ‘Vs’ podem ser descritos como:

- ✓ Volume: Atualmente, as empresas capturam dados de diferentes fontes, que incluem transações comerciais, redes sociais, sensores ou dados transmitidos de máquina a máquina. Armazenar esta volumetria de informações teria sido um problema no passado, mas as tecnologias novas, como o Hadoop, têm solucionado esta questão.
- ✓ Velocidade: Nos dias de hoje, os dados navegam em uma velocidade sem precedentes e é preciso lidar com grandes volumes em tempo real ou quase real. Informações de tags de radiofrequência (RFID), sensores, celulares e contadores inteligentes necessitam ser tratadas em tempo hábil.
- ✓ Variedade: Os dados para serem consumidos e analisados surgem em uma diferente gama de formatos como dados estruturados, dados numéricos em bancos de dados tradicionais, documentos de texto não estruturados, e-mail, vídeo, áudio, dados de cotações da bolsa, transações financeiras, entre outros.
- ✓ Veracidade: É necessário que os dados sejam confiáveis, façam sentido e sejam autênticos. Os dados devem ser classificados seguindo critérios para que possam ser estudados e aprimorados com o uso de inteligência artificial.
- ✓ Valor: Para investir em tecnologias de big data, a empresa deverá ter em mente o retorno sobre o investimento que as informações geradas pela solução trarão.

4.9 DATA WAREHOUSE LÓGICO (LDW)

Segundo o Gartner, os profissionais técnicos que implementam o LDW podem maximizar a eficácia do BA através do benefício comercial oferecido pela integração explícita de três fluxos de desenvolvimento analítico (COOK, 2017):

- ✓ DW clássico: Este fluxo cria o componente tradicional de DW que tipicamente é um sistema de gerenciamento de banco de dados relacional.
- ✓ Desenvolvimento ágil: Este fluxo usa a virtualização de dados combinada com DM's físicos e virtuais para adicionar novas formas de uso de dados.
- ✓ Data Lake: Este fluxo permite que as organizações trabalhem com dados em grande escala e com dados não estruturados.

Os três fluxos não são alternativos, na verdade, são complementares, um sistema analítico moderno necessita dos três fluxos. Planejar a implementação destes fluxos, desde o início, economizará tempo, dinheiro e esforço, além de reduzir o risco. Para Cook (2017), cada um destes fluxos, tem um uso diferente para seu modelo correspondente:

- ✓ Modelo de compromisso: Na maioria das vezes, os consumidores de informação precisam de um modelo de dados em que todos possam concordar. Ou seja, eles precisam de um modelo que represente uma visão coletiva dos principais itens de dados dentro da organização (o que está lá, seu formato e para que ele pode ser usado). Na linguagem do Gartner, isso é conhecido como um modelo de dados de compromisso porque as partes interessadas chegaram a um consenso sobre o modelo, e isso pode ter envolvido um compromisso entre essas partes interessadas.
- ✓ Modelo contendor: Às vezes, as pessoas querem usar dados existentes de novas maneiras, combinando ou aumentando o modelo atual para atender a novos requisitos. Isto é particularmente verdadeiro no desenvolvimento ágil, onde os dados podem ser combinados de novas formas ou aumentados por novas fontes de dados. Estas novas formas de analisar os dados podem trazer benefícios e não devem ser impedidas. A virtualização de dados é uma ótima maneira de habilitar esta mistura e combinação de visualizações de dados. Se esses modelos de contedores forem desenvolvidos e depois comprovados, é provável que sejam abraçados na visão de compromisso ou consenso.

- ✓ Modelo candidato: Aqui é onde novos dados estão disponíveis e são úteis para possíveis novas análises. Na verdade, pode ser inerentemente desestruturado e desta forma sendo impossível de ser descrito em um único formato. Este é o domínio de dados não estruturados.

De acordo com Lewis (2017), o LDW é um conceito baseado na fusão de componentes arquitetônicos, ou seja, não é uma solução comercial de prateleira. Os componentes que formam o LDW podem ser implementados por uma variedade de softwares de fornecedores diferentes e ainda através de componentes customizados. Desta forma, a estratégia LDW oferece uma separação de preocupações, com partes diferentes da arquitetura, atendendo à necessidades diferentes com base nas capacidades necessárias. A Figura 8 mostra o LDW como uma camada conceitual que une uma coleção de componentes arquitetônicos em uma visão lógica conectada. Esta camada fornece as definições lógicas, processos e repositórios que integram a arquitetura de armazenamento e persistência embaixo de onde reside o dado fonte.

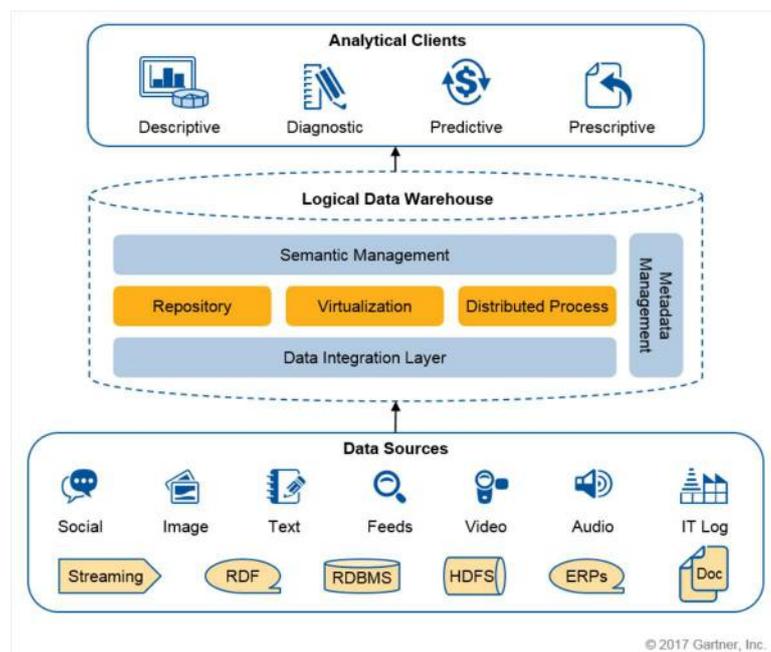


Figura 8 - Diagrama conceitual de LDW (LEWIS, 2017)

Segundo Cook (2017), as seguintes recomendações devem ser seguidas:

- ✓ Planeje e crie seus sistemas analíticos como componentes de um LDW desde o início, mesmo que a arquitetura não esteja totalmente definida neste momento, isso tornará a expansão mais suave e fácil.

- ✓ Para maximizar os resultados, utilize ciclos contínuos de desenvolvimento. Para cada ciclo, escolha um estilo de desenvolvimento ou tenha dois ou três estilos em paralelo.
- ✓ Gerencie e considere explicitamente cada estilo e fluxo. Use a transição dos entregáveis entre os fluxos para acrescentar ou remover controles de forma previsível e metódica.

4.10 DATA LAKE

Criado por James Dixon, CTO (Chief Technical Officer) do Pentaho, data lake é um vocábulo recente que descreve um componente importante no universo da análise de dados e de big data. Esta idéia parte do princípio de ser implantar um repositório único em que todos os dados brutos estejam disponíveis a qualquer pessoa da empresa para a realização de consultas e análises. A idéia central é que o data lake armazene todos os dados em seu formato bruto, sem qualquer processamento e sem governança. Data lake não se trata de uma tecnologia específica e sim de um conceito no qual para ser implementado pode necessitar de várias tecnologias. Em sua essência, é apenas um método de armazenamento de dados e o seu valor de negócio não depende das tecnologias escolhidas e sim das habilidades em ciência de dados que podem ser aplicadas sobre o data lake. Os data lakes não devem ser encarados como substitutos de plataformas analíticas ou de infraestruturas existentes, na verdade, complementam os esforços existentes e apoiam a descoberta de novas perguntas sobre os dados (MATOS, 2015).

Segundo Lewis (2017), um data lake pode ser utilizado para:

- ✓ Consolidar dados em seu estado bruto, não transformados, a partir de uma variedade de fontes de dados diferentes.
- ✓ Ter um propósito genérico de staging area.
- ✓ Coletar dados com a intenção de que novas idéias possam ser derivadas da análise de grandes volumes de dados ou novos tipos de dados.
- ✓ Ser um arquivo ativo de dados históricos.

Ainda segundo Lewis (2017), um data lake pode ser considerado útil quando utilizado para gerenciar dados em diferentes zonas ou camadas, com cada camada otimizada para diferentes estilos de consumo, como:

- ✓ Camada transitória: Atua como uma área para receber a ingestão de muitos dados de fontes diferentes. Esta camada contém todos os dados brutos, sem restrições, incluindo informações sensíveis em sua forma original e inalterada. A camada é considerada transitória porque serve apenas como uma Staging Area para preencher outras camadas.
- ✓ Camada de descoberta: Contém todos os dados brutos para promover a exploração e a descoberta. No entanto, os dados sensíveis são mascarados para evitar o consumo não autorizado, ficando em conformidade com os requisitos ou regulamentações para proteger informações confidenciais. Esta camada pode ser pensada como uma caixa de areia de dados brutos para exploração e análises avançadas.
- ✓ Camada refinada: Nesta camada, enriquecimentos e transformações são processados para criar conjuntos de dados novos, que ficam disponíveis para o consumo de outras aplicações e processos. Os dados são integrados em um formato comum através de técnicas de validação e limpeza aplicadas.
- ✓ Camada confiável: é um passo além da camada anterior, com os dados de referência reconciliados para garantir a consistência com as políticas de governança.

4.11 VIRTUALIZAÇÃO DE DADOS (DV)

Segundo o blog Vert, a virtualização é uma das mais fortes tendências do cenário atual da TI e muitas empresas estão descobrindo que ao utilizá-la da maneira correta, podem conseguir benefícios para o seu negócio. A virtualização de dados é capaz de fornecer à uma organização uma camada de abstração que permite acesso rápido e direto aos dados de várias fontes diferentes, sem a necessidade de movimento físico destes dados. Este acesso rápido se

torna fundamental nos dias de hoje, pois a informação pode ser vista como o bem mais valioso de uma empresa, e o tempo que se gasta para conseguir processar dados ou para dar uma resposta para os seus clientes podem significar o sucesso ou o fracasso. Vale destacar que quando os dados da empresa se encontram virtualizados, eles já se encontram preparados para trabalhos analíticos e disponíveis para os usuários de forma mais rápida.

Para Lewis (2017), quando combinado o DW com o Data Lake, o DV torna-se um componente integral do LDW. O DV pode fornecer uma interface uniforme para vários repositórios de dados, permitindo aos usuários acesso fácil a todos os dados da organização, sujeitos apenas a controles de segurança. O DV também pode ser usado por ferramentas analíticas para permitir a análise de dados que residam em repositórios distintos. Existem três estilos de virtualização de dados (Figura 9):

- ✓ Virtualização embutida: A tecnologia de virtualização está funcionalmente incorporada em uma ferramenta analítica, permitindo que o software de BI faça várias chamadas para os bancos de dados em questão fornecendo assim uma visão consolidada para análise ou relatórios.
- ✓ Virtualização física: Os dados são recuperados de fontes de dados diferentes e consolidados em novas estruturas físicas de dados para consumo de uma fonte unificada.
- ✓ Virtualização dinâmica: Neste modelo, um mecanismo de virtualização atua como um orquestrador de consulta, aceitando-as e decompondo-as em subconsultas para serem executadas contra as múltiplas fontes de dados. O mecanismo de virtualização é capaz de delegar estas subconsultas ao sistema de origem para processamento independente ou armazenamento em cache de dados de múltiplas fontes, para que ele possa executar tarefas de processamento por conta própria.

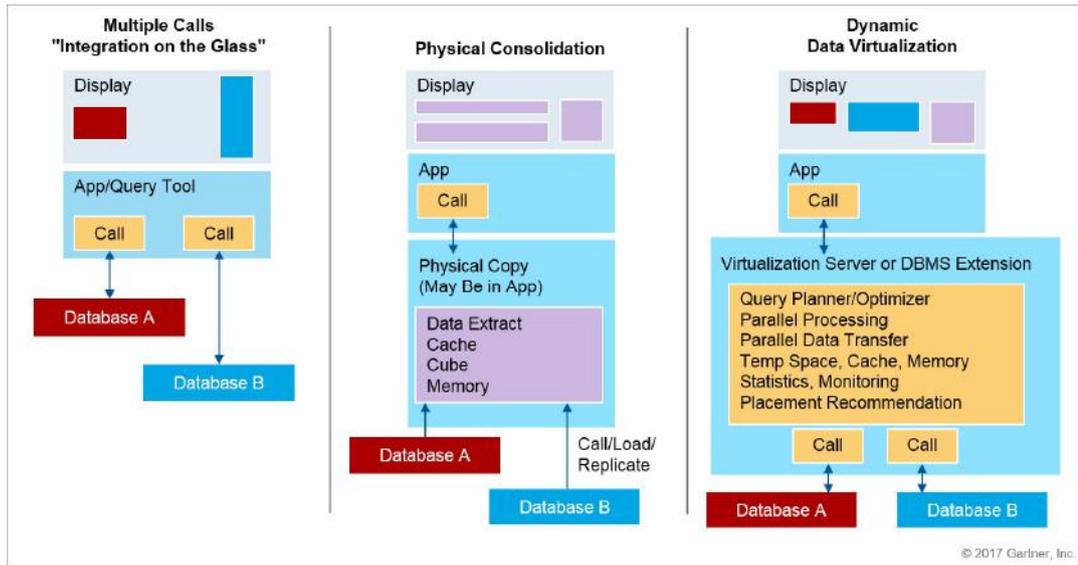


Figura 9 - Estilos de virtualização de dados (LEWIS, 2017)

4.12 CIÊNCIA DE DADOS

A cada dia, o volume, a variedade e a velocidade dos dados são cada vez maiores e para capitalizar as oportunidades que esta informação pode revelar, dados e analíticos estão assumindo um papel mais ativo e dinâmico no poder das atividades das organizações. Mais e mais organizações estão se tornando verdadeiramente orientadas a dados. Em uma recente pesquisa do Gartner com profissionais de TI, analíticos avançados foi a categoria de tecnologia identificada como tendo o maior potencial para mudar a organização nos próximos cinco anos, vide figura 10. Categorias relacionadas como algoritmos de negócios, aprendizado de máquina e inteligência artificial (IA) também obtiveram atenção significativa. Em conjunto, essas tecnologias representam uma tendência que não pode ser ignorada, as análises gerarão inovações significativas e trarão uma disrupção nos modelos de negócios estabelecidos ao longo dos próximos anos (LEWIS, 2017).

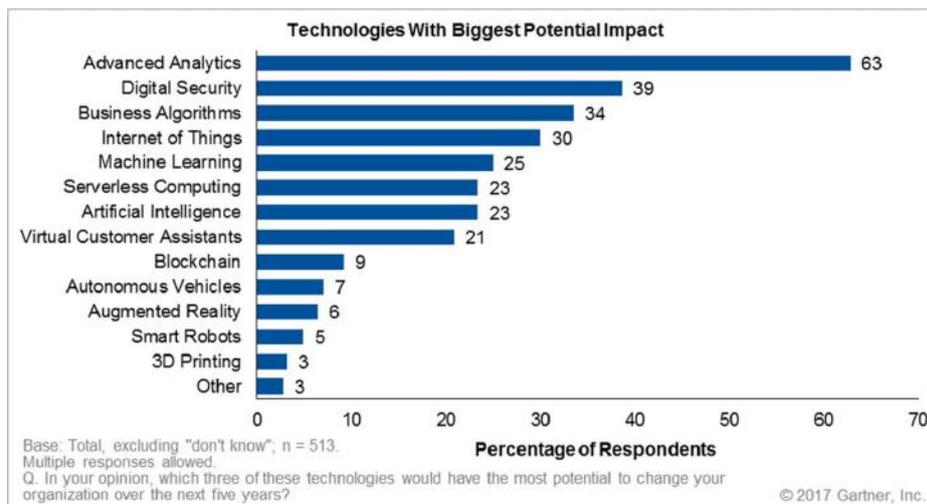


Figura 10 - Potencial de análise para impulsionar a mudança organizacional (LEWIS, 2017)

Segundo o Gartner, analíticos avançados podem ser definidos como um processo autônomo ou semi-autônomo de investigação de dados ou conteúdos usando técnicas e ferramentas sofisticadas, diferentes das utilizadas em um BI tradicional, com o intuito de descobrir percepções mais complexas, fazer previsões ou gerar recomendações. Estas técnicas analíticas avançadas incluem: mineração de dados e texto, aprendizado de máquina, correspondência de padrões, previsão, visualização, análise semântica, análise de sentimentos, análise de redes e cluster, estatística, análise gráfica, simulação, processamento de eventos complexos e redes neurais.

Este conjunto de técnicas avançadas para exploração com o objetivo de transformar dados em ações está diretamente associado ao tema Ciência de Dados, que apesar de ser uma área que já existe há mais de 30 anos, vem ganhando destaque nos últimos anos devido ao big data. No sentido literal, Ciência de Dados é um domínio interdisciplinar sobre processos e sistemas que visa descobrir novos conhecimentos ou percepções de dados através de modelos estruturadas ou não estruturadas, que funciona como o propulsor da análise de dados, como estatística, mineração de dados e análise preditiva (SILVEIRA, 2016).

De acordo com o Idoine e Tapadinhas (2016), existe uma diferença significativa entre as capacidades analíticas de self-service BI que são suportadas por ferramentas de descoberta de dados e as capacidades de analíticos avançados que são suportadas por Ciência de Dados. De forma semelhante, existe uma diferença entre analistas de informação e cientistas de dados tanto no que tange a sua qualificação profissional, nos casos de uso com os quais se envolvem quanto no impacto de suas realizações analíticas. Com intuito de reduzir esta lacuna, o Gartner define a figura do cidadão cientista de dados, do inglês citizen data scientist (CDS), que é uma

pessoa que cria ou gera modelos, que faz uso de análises diagnósticas avançadas ou ainda explora capacidades preditivas e prescritivas, porém sua principal função de trabalho não está no campo da estatística. Os CDS's são usuários avançados que podem executar tarefas analíticas das mais simples às mais sofisticadas.

Geralmente, o CDS é um membro de uma linha de negócios, de um departamento de fora da TI corporativa. A diferença entre o analista que pratica descoberta de dados e o CDS não é ampla, pois algumas semanas de treinamento sobre as ferramentas adequadas e melhores práticas pode ser tudo o que é necessário para transformar o primeiro no segundo. Em contrapartida, podem levar anos de treinamento intensivo e experiência do mundo real para transformar um analista de descoberta de dados em um cientista de dados altamente qualificado, tempo que poucas organizações podem pagar.

Segundo Idoine e Tapadinhas (2016), os gestores de dados e analíticos podem tirar partido dos CDS's para expandir e aprimorar suas iniciativas em BA. Para explorar o seu potencial, eles devem:

- ✓ Habilitar o acesso a novas fontes e tipos de dados, o que inclui tipos mais complexos e fontes de dados como big data. Além disso, eles devem treinar os CDS's sobre como entender esses dados e suas relações além de sua conexão com processos de negócios.
- ✓ Explorar recursos analíticos adicionais que incluem recursos avançados e alternativos para a exploração de informações e que possam trazer novas descobertas, representando um avanço em direção a processos de análise mais complexos, como análises preditivas e prescritivas.
- ✓ Habilitar e apoiar analistas avançados ao simplificar o acesso a recursos analíticos sofisticados para um número maior de usuários de negócio. Estes usuários podem vir a se tornarem novos CDS's através de um aprimoramento de seus conhecimentos e com a utilização de ferramentas e mecanismos fáceis de usar.

A figura 11 mostra que o CDS tem a capacidade de combinar uma gama de fontes de dados ampla, explorar novas capacidades analíticas e suportar grupos de analistas. A combinação de mais fontes de dados, mais capacidades e mais analistas traz um efeito

multiplicador, expandindo e aprimorando os casos de uso de self-service existentes, simplificando o acesso a análises avançadas e levando a novas descobertas e aplicativos capazes de gerar valor adicional para as organizações.

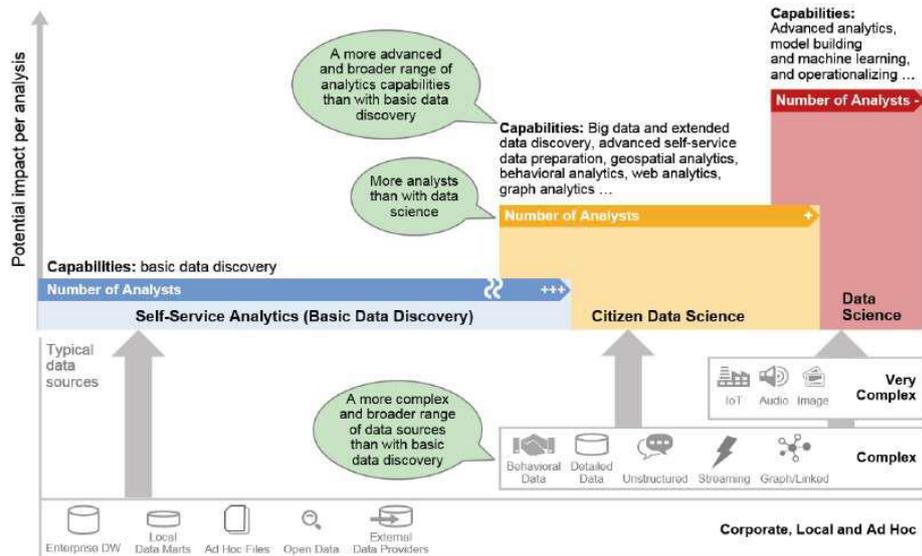


Figura 11 - Comparação de Citizen Data Science com Self-Service Analytics e Data Science (IDIONE e TAPADINHAS, 2016)

4.13 BI TRADICIONAL X BA MODERNO

Segundo Wayne (2016), o BI é alimentado por duas forças opostas: top-down, onde a TI corporativa impõe padrões na entrega de dados e relatórios, garantindo assim a verdade única da informação e botton-up, na qual analistas das próprias áreas de negócio criam seu próprios relatórios e análises a partir de dados locais, vide quadro 3. Na maioria das empresas, estas duas vertentes vivem em conflito pois a TI bloqueia o acesso aos dados o que frustra as áreas de negócio que contratam seus próprios analistas para criarem os relatórios e as análises que necessitam no momento em que necessitam. A interação ou ausência entre essas duas linhas determina a eficácia com que uma organização usa dados e análises para a tomada de decisões.

Quadro 3 - Linhas de desenvolvimento em BI/BA (WAYNE, 2016)

Top-Down	Botton-Up
Guiada pela TI	Guiada por usuários
Para usuários casuais	Para usuários avançados
Para monitorar processos	Para descobrir novas percepções
Reponde questões conhecidas	Responde novas questões
Usa dados certificados	Usa dados certificados e não certificados
Dados de um DW	Dados de uma fonte de dados qualquer

A linha de trabalho top-down é geralmente impulsionada pela TI, que fornece dados conformados e relatórios para usuários casuais como executivos, gerentes e funcionários de áreas operacionais que usam informações para monitorar os processos principais de negócios. A linha de trabalho botton-up é guiada por usuários avançados dos mais diversos departamentos que consultam, analisam, exploram, mineram e extraem todos os tipos de dados para responder às perguntas urgentes de gerentes e líderes de negócios.

Nas corporações, estas linhas de desenvolvimento, normalmente, se encontram em total desacordo, pois representam abordagens muito diferentes no que diz respeito à entrega de dados e à descoberta de novas percepções. A abordagem top-down enfatiza a centralização, economias de escala, governança, padrões e controle para garantir o fornecimento de dados precisos e consistentes em toda a organização. Em contrapartida, a abordagem botton-up concentra-se na velocidade, flexibilidade e liberdade para que as unidades de negócios possam atender às necessidades locais rapidamente, vide quadro 4.

Quadro 4 - Características das linhas de BI/BA (WAYNE, 2016)

Top-Down	Botton-Up
Centralizada	Descentralizada
Governança	Livre
Padrão	Velocidade
Controle	Liberdade
Estabilidade	Criatividade
TI	Negócios

A maioria das empresas não encontra o equilíbrio necessário entre esses dois modelos organizacionais. Para ter sucesso na esfera analítica, as organizações precisam, primeiramente, reconhecer que esses dois modelos existem, e em segundo lugar, precisam desenvolver

estruturas organizacionais, arquitetônicas e tecnológicas que permitam que esses dois mundos atuem de forma complementares um ao outro, o que é chamado pelo Gartner de BI multimodal.

Desde os anos 2000, as implementações na esfera analítica têm sido dominadas por abordagens e ferramentas que seguem a linha top-down. A partir do caos gerado pela disseminação de planilhas ao longo da década de 1990, muitos CEOs ou CFOs patrocinaram iniciativas corporativas de BI com intenção de entregar dados consistentes e métricas para toda a organização. O objetivo era gerar uma única versão da verdade e alinhar o negócio em torno de um conjunto comum de métricas que permitisse avaliar os processos de determinados segmentos da empresa. Implementações top-down, costumam ser acompanhadas por críticas negativas, pois a consolidação de dados corporativos, o consenso sobre métricas e definições de dados importantes demandam muito tempo, esforço e dinheiro. A implantação do DW corporativo pode levar anos além de consumir milhões de dólares. Uma vez construído, este ambiente é difícil e demorado para mudar e caro de manter.

A maioria das ferramentas de BI compradas pela TI corporativa reflete prioridades top-down, onde existe a necessidade de se ter administradores profissionais para instalar e configurar as ferramentas e de desenvolvedores profissionais para criar modelos de dados, processos de extração, transformação e carga de dados, além de relatórios e painéis. Estas ferramentas possuem limitadas funcionalidades de self-service, o que reduz a capacidade das unidades de negócios em obter dados para executar suas necessidades. Neste contexto, temos a maioria das equipes de BI corporativas sobrecarregadas com uma grande quantidade de solicitações de relatórios departamentais, painéis e conjuntos de dados.

Neste cenário, é comum encontrar usuários que não conseguem obter as informações necessárias para fazer seus trabalhos em tempo hábil. Estes usuários, geralmente, ficam frustrados quando o DW ou o DM não possui os dados que precisam para responder a perguntas urgentes e também ficam aborrecidos quando a ferramenta de BI não é flexível o suficiente para dar forma e comparar dados da maneira que eles querem. Para completar, os usuários casuais acham as ferramentas de BI pouco amigáveis e complexas de usar e assim acabam desistindo de usá-la, e passam a solicitar relatórios personalizados a uma equipe alternativa de BI, normalmente experts em planilhas lotados na própria área usuária. Esta equipe alternativa pode vir até a subutilizar a ferramenta de BI corporativa, gerando, a partir desta, extrações de dados para uso apenas em planilhas.

Neste contexto, as áreas de negócios começam a comprar as suas próprias ferramentas de self-service e a construir suas análises e dashboards, ao invés de esperar pela TI corporativa desenvolver DM's específicos no DW. Ou seja, o objetivo é realmente ficar independente da

TI corporativa, e assim conseguir a agilidade necessária na execução de seus processos de negócio.

Reconhecendo esta situação, novos e consolidados fornecedores de ferramentas de BI, começam a entregar ferramentas de self-service orientadas a descoberta com foco voltado para usuários de negócios. Essas ferramentas permitem que os usuários se conectem a uma ou mais fontes de dados, mesquem estes dados em um modelo lógico para que em seguida possam criar suas análises, relatórios e dashboards e ainda os publiquem para demais membros da equipe. Desta forma, podem acessar os dados que querem, quando querem e como querem, sem ter que lidar com a TI corporativa.

A grande difusão de ferramentas de self-service BI entre as áreas da corporação acaba por criar novos problemas de cunho executivo. Cada unidade de negócios define seu domínio de dados e métricas sem nenhum alinhamento estratégico. Este cenário traz uma proliferação de silos de dados entre as áreas da empresa, o que pode vir a gerar análises contraditórias e erradas. Neste contexto, a TI corporativa fica cada vez mais às margens das áreas de negócio, pois estas desejam maior autonomia e flexibilidade de dados enquanto a TI deseja controles e governança mais rigorosos.

Com esta dinâmica, as organizações podem flutuar entre os dois extremos à medida que experimentam os benefícios e, em seguida, as desvantagens de cada abordagem. Podemos verificar na figura 12 que muita liberdade pode levar ao caos, conflito e desordem, enquanto muito controle leva à estagnação, repressão e opressão.



Figura 12 - Extremos de BI/BA (WAYNE, 2016)

Em última análise, cada organização tem que descobrir como lidar com estas vertentes opostas em um ambiente harmônico e coerente, pois no final das contas, acabam por representar os diferentes lados da mesma moeda. Ao invés de realizar a opção por um dos extremos, seria bom para a corporação ganhar as vantagens de ambas as extremidades do espectro, sem qualquer das desvantagens. Seria ótimo ter liberdade e controle, velocidade e padrões, flexibilidade e estabilidade, criatividade e estrutura. As organizações que conseguem combinar

os impulsos top-down e botton-up alcançam o melhor de ambos os mundos enquanto minimizam suas desvantagens, vide o quadro 5.

Quadro 5 - Vantagens e desvantagens top-down x botton-up (WAYNE, 2016)

Top-Down	Botton-Up
Vantagens	Desvantagens
Consistência de dados	Falta de consistência de dados
Padronização de métricas	Falta de padronização de métricas
Alinhamento de negócios	Falta de alinhamento de negócios
Desvantagens	Vantagens
Construção complexa	Rápida implantação
Manutenção complexa	Manutenção simples
Alto custo	Baixo custo

Para criar uma harmonização, as empresas precisam desenvolver uma estratégia abrangente que aborde as dimensões organizacionais, arquitetônicas e tecnológicas do BI:

- ✓ Arquitetura organizacional que permita a criação de uma equipe de BI híbrida, composta por analistas de TI e funcionários das unidades de negócios.
- ✓ Arquitetura de dados que contenha acesso a dados corporativos onde especialistas de unidades de negócios possam acessar, mesclar e analisar dados sem criar silos que prejudiquem a consistência dos dados.
- ✓ Arquitetura de tecnologia que contemple uma plataforma analítica para suportar a exploração self-service com a criação de relatórios, análises e dashboards dentro de um ambiente governado que garanta a consistência dos dados e evite a proliferação de silos de dados.

Com esta visão, muitos fornecedores de ferramentas de BI começaram a vender ferramentas de descoberta de dados governadas que combinando o melhor do BI top-down e do botton-up conseguem amenizar seus efeitos colaterais. Alguns destes fornecedores de BI eram tradicionais fornecedores top-down que adicionaram recursos de descoberta de dados enquanto que outros são novos fornecedores de BI botton-up que adicionaram recursos de governança.

5 MODELOS DE CAPACIDADE E MATURIDADE EM PRÁTICAS ANALÍTICAS

5.1 INTRODUÇÃO

De acordo com Biguru (2012), para tentar entender a evolução da adoção e maturidade analítica nas organizações, devemos considerar que a maioria das empresas vem colocando o BI dentro de seu planejamento estratégico, que os custos de sua implantação são elevados e que existe uma dificuldade em provar o valor do próprio BI, o que torna difícil de justificar seus investimentos. Ter o conhecimento prévio sobre o que é possível conseguir com a implantação de um BI além de conhecer os desafios e as armadilhas envolvidos, permite que as organizações planejem com maior segurança a sua estratégia de implementação.

Ainda segundo Biguru (2012), A principal barreira à adoção de BI e a demonstrar seu valor como ferramenta estratégica, reside na sua complexidade. O BI é uma área ampla que compreende aspectos técnicos e não-técnicos, como pessoas e processos. Por estes aspectos, os modelos só podem fornecer um quadro prescritivo que precisa ser adaptado por cada organização. É importante entender que vários departamentos de uma organização podem estar em diferentes níveis de maturidade e que nem todas as organizações seguem a mesma trajetória de evolução ou têm que passar por cada etapa.

De acordo com Wells (2016), O CMMI faz uma distinção entre níveis de capacidade e níveis de maturidade como mostrado no quadro 6. E embora os nomes para níveis de capacidade e maturidade sejam semelhantes em alguns níveis, existem diferenças sutis, mas significativas. Os níveis de maturidade são definidos e medidos como um agregado para todos os processos em uma empresa. Os níveis de capacidade são definidos e medidos para processos específicos e direcionados.

Quadro 6 - Modelo de maturidade em capacidade (Adaptação de WELLS, 2016)

	Capacidade	Maturidade
Nível 5	Otimizado	Otimizado
Nível 4	Gerenciado Quantativamente	Gerenciado Quantativamente
Nível 3	Definido	Definido
Nível 2	Gerenciado	Gerenciado
Nível 1	Executado	Iniciado
Nível 0	Incompleto	Não se aplica
	Processo Específico	Em toda a empresa

Este tópico pretende dar uma visão global sobre a realidade dos modelos de capacidade e maturidade para ambientes analíticos mais referenciados no mercado. Os modelos são normalmente desenvolvidos por consultorias ou empresas de software que criam os seus próprios modelos com a intenção de fazer avaliações de capacidade e maturidade para fins comerciais. Ou seja, analisam e aferem o estado atual de uma determinada organização ou processo, propondo ações e serviços para que esta possa avançar para o próximo nível de capacidade ou maturidade.

5.2 MODELO GARTNER

Segundo Howson e Duncan (2016), o modelo de maturidade do Gartner fornece aos gestores de dados e analíticos uma maneira de identificar o nível de desenvolvimento que sua iniciativa analítica deve alcançar para apoiar os objetivos da empresa, além de ajudar a traçar um roteiro de melhorias.

Um programa de BI ou BA inclui resultados de negócios, pessoas, processos, dados e tecnologias desejados. À medida em que este programa vai ganhando maturidade, a arquitetura vai evoluindo junto com os processos e conhecimentos necessários para suportá-la. As organizações que possuem níveis de maturidade mais baixos, possuem processos analíticos concentrados em grupos isolados que acessam sistemas e fontes de dados não integrados e ainda são fortemente dependentes de planilhas. As organizações que estão em níveis de maturidade mais elevados têm processos integrados, utilizam desenvolvimento ágil, estabelecem padrões empresariais para dados e sistemas e combinam dados internos com dados externos (HOWSON e DUNCAN, 2016).

As organizações não devem adotar uma abordagem estratégica de um único passo para modernizar seus programas de dados e ambiente analítico, pois é necessário tempo para desenvolver todos os conhecimentos requeridos. Novos métodos e conceitos, como desenvolvimento ágil e TI bimodal, bem como inovações técnicas tais como nuvem, mobile, descoberta de dados, big data, etc., aumentam a necessidade das organizações desenvolverem sua maturidade analítica. Os gestores de dados e analíticos devem consultar o modelo de maturidade do Gartner para:

- ✓ Compreender os cinco níveis de maturidade analítica.
- ✓ Identificar o nível atual de maturidade da empresa.

- ✓ Determinar quais passos a empresa deve seguir para passar para o próximo nível de maturidade.

A ferramenta ITScore diagnostic, do Gartner, permite aos clientes diagnosticar a maturidade de vários aspectos do seu programa analítico, vide figura 13.

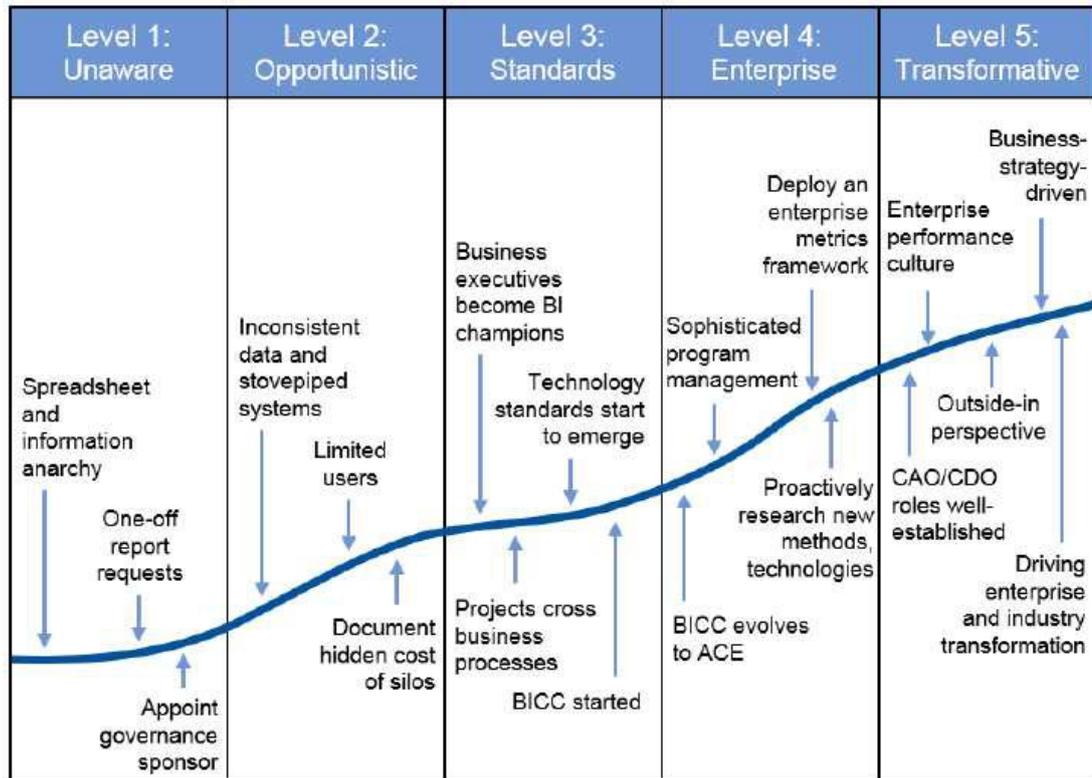


Figura 13- Modelo de Maturidade Gartner para BI e BA (HOWSON e DUNCAN, 2016)

Os níveis de maturidade propostos pela metodologia do Gartner são:

Nível 1: Desconhecido: Neste nível de maturidade, a exploração analítica ocorre de forma ad hoc, sem que exista um processo formal relacionado à práticas de análises e métricas de desempenho. Normalmente, executivos e gerentes solicitam informações para usuários que se esforçam para fornecê-las usando qualquer aplicativo operacional que esteja disponível. Na maior parte destas solicitações, os resultados são fornecidos através de planilhas projetadas para uso único. Nesta maturidade, a empresa não possui infraestrutura de informações para alavancar dados que sejam coletados de sistemas transacionais para tomada de decisão ou análise. Neste estágio, não há processos definidos para análise ou tomada de decisão ou ainda métricas de desempenho. Esta abordagem é utilizada porque não necessita de grandes investimentos iniciais, porém de uma forma ampla, gera um grande esforço de trabalho com sucessivas ações

de retrabalho e acabam por não fornecer informações consistentes ou precisas e que raramente são auditáveis.

Nível 2: Oportunista: Neste nível de maturidade, as áreas de negócios realizam os projetos analíticos de forma isolada com a intenção de otimizar um processo ou tomar decisões táticas. Cada projeto possui sua própria infraestrutura de informações, ferramentas, aplicativos e indicadores de desempenho. Desta forma, diferentes aplicativos proliferam por toda a organização, cada qual guiado por sua própria equipe de profissionais de TI, usuários de negócios e gerentes operacionais. Cada unidade de negócio pode usar diferentes capacidades analíticas, bancos de dados, ferramentas de integração de dados e plataformas de BI, enquanto os resultados obtidos são através de relatórios, consultas ad hoc e painéis. Para alimentar esses aplicativos, são criados DM's de assunto único e talvez com alguma tecnologia de qualidade de dados. Em implementações mais rudimentares, algumas tabelas transacionais podem ser copiadas para um banco de dados em que são gerados relatórios, isolando assim o sistema transacional do impacto de desempenho da geração e uso destes relatórios. Neste cenário, a empresa possui conjuntos múltiplos e fragmentados de dados e metadados armazenados em vários locais, incluindo a plataforma de BI, os painéis e as ferramentas de integração de dados. Esta abordagem oferece valor aos usuários rapidamente, com informações e análises relevantes. Os especialistas em BI têm liberdade para inovar, mas eles não precisam ter um alto nível de conhecimento técnico em BA. Esta filosofia cria silos de conhecimento juntamente com as aplicações e informações, de modo que a organização, de forma mais ampla, acaba não se beneficiando de conhecimentos compartilhados. Deste modo, gerentes em níveis superiores continuam a receber dados inconsistentes ou conflitantes dos vários sistemas e métricas que são empregados por departamentos diferentes.

Nível 3: Padronizado: Neste nível de maturidade, pessoas, processos e tecnologias começam a se coordenar por toda a empresa. Pelo menos um executivo sênior, geralmente do lado comercial, se torna o gestor da empresa para assuntos analíticos. Os gerentes de processo e os gestores de TI coordenam projetos que permeiam diferentes processos de negócios e que precisam compartilhar análises e decisões. A maioria das empresas implementa um centro de competências de BI (BICC) composto por usuários de negócios, profissionais de TI e analistas, com o intuito de compartilhar conhecimentos e melhorar a consistência em aplicações analíticas e no uso de informações.

Os padrões de tecnologia começam a surgir, inclusive aqueles para infraestrutura de informações, DW's e plataformas de BI. Esses padrões não são necessariamente obrigatórios, mas são preferidos para economias de escala e suporte aprimorado. No entanto, os projetos da empresa não compartilham dados ou modelos analíticos consistentemente. No máximo, um ou dois processos compartilham um modelo de dados mestre comum enquanto que os metadados são federados para cada tecnologia. Neste cenário, começa a surgir um pequeno compartilhamento de análises, processos de decisão, componentes e recursos. Alguns compartilhamentos de medidas e índices de desempenho acontecem nos processos, principalmente para ajudar áreas de negócios de forma individual, mas essas medidas e índices não estão vinculadas aos objetivos da empresa.

Pela primeira vez, a empresa começa a reduzir o custo total de seus esforços em ambiente analítico através de uma melhor coordenação e padronização de tecnologias. No entanto, a adaptabilidade dos sistemas analíticos continua baixa, de modo que a empresa ainda não consegue atingir fortes economias de escala.

Nível 4: Corporativo: Nesse nível de maturidade, os principais executivos tornaram-se patrocinadores de programas analíticos e a empresa definiu uma estrutura de métricas de desempenho que conecta vários processos a objetivos corporativos, além disso, essas métricas orientam a estratégia empresarial.

Os aplicativos analíticos suportam processos de decisão multifuncionais que atendem toda a empresa. Os executivos corporativos e operacionais podem ver as relações de causa e efeito entre as principais atividades. Os mesmos aplicativos são utilizados por analistas, gerentes de negócios e executivos seniores e às vezes até por parceiros, fornecedores, clientes e cidadãos comuns.

Existe uma arquitetura de informação empresarial que orienta o projeto de novos aplicativos analíticos. A empresa exibe um alto grau de disciplina em torno de projetos analíticos, com as versões de entrega orientadas através de técnicas de gerenciamento de projetos, desenvolvimento ágil e prototipação rápida. Existe um comprometimento com a qualidade e a integração dos dados para melhorar a precisão e a disponibilidade de informações por toda a empresa, mas há um pragmatismo quanto à qualidade dos dados em novas fontes de dados. Neste cenário, existe um financiamento substancial para o desenvolvimento destes projetos.

A utilização de modelos de dados comuns, regras e analíticos minimizam o número de versões de um determinado conjunto de informações. Em geral, o armazenamento e o

processamento de dados são pensados para ter eficiência de custo e performance, podendo ser utilizados conceitos de data lake e LDW. O departamento de TI fornece suporte escalável, atuando como um facilitador e não como o único executor.

Embora os esforços analíticos tenham se tornado mais eficientes, o uso cresce e, portanto, os custos permanecem elevados. A empresa deve manter pessoas com alto conhecimento técnico em muitas áreas diferentes. O BICC é confrontado com o desafio permanente de manter e melhorar gradualmente os sistemas analíticos atuais, enquanto continua evoluindo e inovando. A empresa aproveita os padrões em investimentos em tecnologia, mas aceita adequadamente a utilização de tecnologias novas e modelos de dados conforme cada unidade de negócios requer.

A maturidade em ambiente analítico ultrapassou os relatórios de BI descritivo e de self-service BI para análises de diagnóstico incluindo análises preditivas e avançadas. O papel do BICC muda para níveis de análise mais altos com uma abordagem mais consultiva, evoluindo, na sequência, para um centro de excelência de análise. Este centro de excelência trabalha ativamente com áreas de negócios individuais e partes interessadas sobre quando modelos de informação e padrões de tecnologia devem evoluir, em vez de visualizar os esforços departamentais como TI's paralelas.

Nível 5: Transformativo: Nesse nível de maturidade, o ambiente analítico torna-se uma iniciativa estratégica, gerida em conjunto pelas áreas de negócios e de TI, e apoiada e governada nos níveis mais altos da organização. A empresa olha para a informação como um ativo estratégico e usa analíticos para gerar receita, operar de forma eficiente ou fornecer o melhor serviço de atendimento ao cliente. Todas as partes interessadas usam a informação dos sistemas analíticos para coordenar uma resposta às mudanças das condições de negócios em toda a cadeia de valor e tomar decisões transformacionais.

Os usuários são oriundos de vários níveis dentro da organização, pertencentes a várias unidades de negócios, bem como de clientes e parceiros. Todos confiam na informação e análises que os sistemas geram como base para tomar decisões em busca dos objetivos estratégicos da empresa. Nesse nível, o uso de analíticos é muitas vezes incorporado à tarefas operacionais e modelos de decisão, trazendo percepções mais fortemente vinculados a ações. Todos os projetos usam processos e modelos padrões, com alguma customização para necessidades específicas. Todo o programa tornou-se ágil o suficiente para se adaptar às necessidades de mudança como uma rotina.

A empresa tem otimizado os custos ao compartilhar sistemas, processos e capacidades técnicas por toda a organização. O papel do BICC mudou completamente de reativo para proativo, incentivando a adoção e a inovação. Os usuários podem ver o desempenho da empresa e os fatores que contribuem para isso. A orientação do serviço oferece aos usuários mais controle sobre as atividades analíticas. No entanto, determinar a granularidade adequada dos serviços apresenta um desafio contínuo em uma organização com necessidades diferentes. Este programa requer investimento contínuo e atenção da alta administração. Futuras fusões ou aquisições podem reintroduzir muitos dos problemas que a empresa já superou.

5.3 MODELO TDWI

Segundo Halper e Stodder (2014), o Modelo de Maturidade e a Ferramenta de Avaliação da Matriz de Análise TDWI foram desenvolvidos em resposta à necessidade das organizações em entender como as suas implantações analíticas se comparam com as de outras empresas de nicho e em como fornecer as melhores práticas para seguir adiante. A avaliação mede a maturidade de um programa de análise objetivamente em muitas dimensões que são fundamentais para derivar o valor da análise.

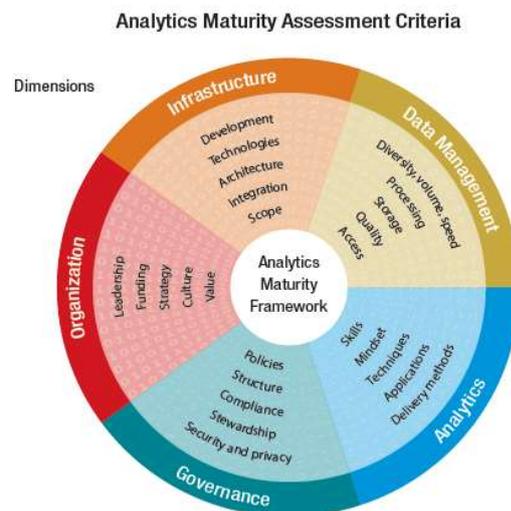


Figura 14 - Dimensões do modelo de maturidade do TDWI (HALPER e STODDER, 2014)

De acordo com a figura 14 as dimensões do modelo de maturidade do TDWI são:

- ✓ **Organização:** Define até que ponto a estratégia organizacional, a cultura, a liderança, as habilidades e os recursos financeiros suportam um programa

analítico bem-sucedido. Esta dimensão também verifica se o uso de analíticos é aplicado de maneira ampla nas decisões cotidianas da empresa.

- ✓ **Infraestrutura:** Verifica se a arquitetura está avançada e coerente para suportar uma iniciativa analítica, medindo até que ponto a infraestrutura suporta o consumo de analíticos, considerando a empresa como um todo e seus potenciais usuários. Esta dimensão também verifica quais tecnologias estão em vigor para apoiar estas iniciativas e como elas estão integradas no ambiente existente.
- ✓ **Gestão de dados:** Esta dimensão verifica se a variedade, o volume e a velocidade dos dados utilizados são vastos o suficiente para o uso analítico. Neste contexto, a qualidade, o processamento e a integração de dados bem como os problemas de acesso devem ser considerados.
- ✓ **Analíticos:** Esta dimensão mede o quanto a empresa está avançada no uso analítico, o que inclui os diferentes tipos de análises utilizadas e como estas são entregues na organização. Também inclui a cultura analítica, isto é, como a prática analítica contribui para decisões tomadas em toda a empresa.
- ✓ **Governança:** Verifica se a estratégia de governança de dados da empresa é coerente em apoio ao seu programa de BI e BA. Verifica também se a empresa é capaz de gerenciar a descoberta de dados de usuários e as explorações analíticas efetivamente, sem aplicar restrições em excesso que dificultem a geração de descobertas e novas percepções.

O modelo de maturidade em ambiente analítico do TDWI consiste em cinco estágios: nascente, pré-adoção, adoção precoce, adoção corporativa e maduro ou visionário. À medida que as organizações se movem através dessas etapas, elas devem obter maior valor de seus investimentos em analíticos conforme a figura 15.

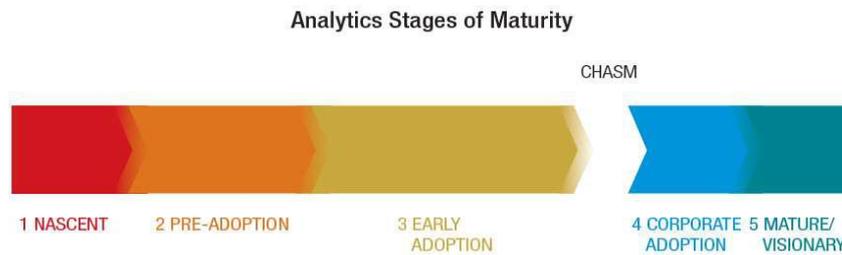


Figura 15 - Estágios de maturidade em BI e BA do TDWI (HALPER e STODDER, 2014)

Estágio 1: Nascente: O estágio nascente representa um ambiente pré-analítico, no qual a maioria das empresas não faz uso de práticas analíticas, exceto, talvez, por análises através de planilhas. Podem existir bolsões de usuários que possam estar interessados no valor potencial de um ambiente analítico e que possam estar testando algum software. Geralmente, no estágio nascente, a cultura não é analítica, em outras palavras, a cultura não é conduzida por dados e as decisões são tomadas mais com base no instinto do que no fato.

- ✓ **Organização:** Na organização nascente, a maioria dos executivos desconhece o poder de um ambiente analítico. As áreas de TI e negócios não trabalham em conjunto para viabilizarem a criação de um ambiente analítico. Normalmente, a TI não está particularmente interessada em ajudar o negócio a obter acesso aos dados, a menos que este esteja sob o controle da própria TI. Usuários de negócios desenvolvem seus próprios aplicativos para atividades que podem ser úteis para realizar análises, mas o usuário é geralmente um superutilizador de planilhas. Nessa cultura de não colaboração, podem haver grupos de usuários que tentam se educar sobre novos tipos de análise porque sabem que as formas antigas não os ajudarão a permanecerem competitivos.
- ✓ **Gestão de dados:** Normalmente, a organização nascente não possui estratégia de gerenciamento de dados. A TI pode acreditar que tem algum tipo de estratégia de gerenciamento de dados, talvez com alguns bancos de dados ou conjuntos de dados que ele usa para relatórios, mas não pensou em nomear padrões ou metadados. A qualidade e a consistência dos dados podem ser ruins, os volumes de dados geralmente são baixos e estão em silos desconectados, e geralmente são apenas estruturados. Do lado empresarial, as organizações estão pedindo informações ou relatórios à TI e quando acessam suas próprias fontes de dados, as análises são feitas em planilhas. Gasta-se muito tempo para formar um conjunto de dados para análise.

- ✓ **Analíticos:** Na organização nascente, o trabalho com uso de planilhas é uma regra normalmente. Os painéis e análises podem ser desenvolvidos em uma planilha e enviados por e-mail para aqueles que precisam. Se a empresa possui um grupo especializado em análises, geralmente, é um departamento e é direcionado a uma função específica, como o marketing. Em outras palavras, a análise está sendo feita em silos na organização e um grupo não sabe o que o outro está fazendo. As melhores práticas não são compartilhadas e não há esforços para criar uma cultura analítica ampla que envolva usuários não tradicionais. No entanto, em muitas organizações nascentes, os funcionários estão cansados de não ter acesso aos dados para tomar decisões, querem ferramentas analíticas para que possam responder à questões comerciais importantes. Eles não querem confiar na TI, pois esta demora muito para disponibilizar as informações.

Estágio 2: Pré-Adoção: Membros de equipes começam a ler sobre o tópico e talvez participar de seminários na internet ou conferências. As organizações podem investir em alguma tecnologia analítica, como instâncias únicas de uma ferramenta de self-service BI, ou ainda criando relatórios gerenciados a partir de um DM ou DW. As equipes estão começando a entender o poder da análise para melhorar as decisões e assim evoluir o resultado comercial.

- ✓ **Organização:** Na fase de pré-adoção, muitas vezes um patrocinador executivo intensifica a discussão sobre ambiente analítico na empresa. O patrocinador pode estar frustrado porque as decisões estão sendo tomadas na ausência de dados, ou o executivo vem de uma organização mais orientada a dados. É comum o uso de ferramentas freeware para fazer algumas análises. Uma discussão dentro da organização sobre o uso de analíticos começa a ganhar corpo.
- ✓ **Gestão de dados:** Os gestores de negócios e ou TI imaginam o potencial valor de combinar múltiplas fontes de dados para análise. A empresa percebe que precisa de uma infraestrutura de dados para apoiar seus esforços em análise e investimentos são feitos para a criação de DM e DW. A liderança da empresa está começando a pensar de forma holística sobre os dados, que ainda são tipicamente apenas estruturados, e das fontes diferentes em torno da empresa. A TI precisa avançar e entender os problemas de qualidade dos dados, porque a empresa agora está

pensando em reunir e compartilhar dados, tomando medidas para garantir que os usuários tenham dados relevantes, consistentes e oportunos para análise.

- ✓ Analíticos: A análise ainda é rudimentar, mas os grupos estão começando a se desenvolver e os analistas de negócios estão percebendo o poder da visualização, e podem estar investindo em uma ferramenta de visualização de baixo custo para tentar entender seus dados. Marketing ou outros departamentos estão começando a fazer perguntas que precisam de respostas. É necessário um entendimento que a TI e o negócio precisam se juntar para apoiar a criação de analíticos, evitando um cenário de caos.

Estágio 3: Adoção Precoce: Durante a fase de adoção precoce, a empresa investe em ferramentas analíticas e metodologias. O pensamento gira em gerenciamento de dados, relatórios e dashboards. Os usuários costumam passar muito tempo nesta fase de adoção precoce.

- ✓ Organização: Nesta fase, a TI e o negócio começam a trabalhar juntos como equipe em projetos selecionados. A empresa começa a se beneficiar do uso de analíticos como parte de seu processo de tomada de decisão em vários departamentos, mas a própria empresa ainda não é necessariamente orientada a analíticos. Projetos analíticos começam a gerar frutos e há interesse em construir protótipos para ver como a análise pode ser avançada e ampliada em toda a empresa. À medida que a organização se entusiasma com os resultados obtidos, novas áreas de negócio começam a se interessar pelo ambiente analítico em seus processos.
- ✓ Infraestrutura: Neste estágio, pode haver alguma tecnologia analítica local, mas normalmente não é uma plataforma dedicada para análise. A empresa pode estar pensando em uma arquitetura unificada que pode permitir aos usuários acessar múltiplas fontes de dados e diferentes tipos de dados, incluindo conteúdo não estruturado. A empresa pode considerar ou mesmo avaliar provas de conceito (POCs) para implantar sistemas analíticos na nuvem ou desenvolver sistemas com tecnologia big data para aumentar a flexibilidade e reduzir custos para algumas das suas iniciativas de análise.

- ✓ **Gestão de dados:** Nesta fase, a organização normalmente possui um DW ou DMs locais, tratando apenas dados estruturados. Do ponto de vista da estratégia de dados, metadados são atribuídos ao nível departamental, enquanto o gerenciamento definido do ciclo de vida dos dados, a auditabilidade e a linhagem dos dados são insuficientes ou inexistentes. A governança está começando a ser implantada para garantir que os dados confidenciais sejam protegidos e que os usuários tenham dados relevantes e apropriados para análise.
- ✓ **Analíticos:** Nesta estágio, existem grupos na organização que utilizam ferramentas de BI, realizando descoberta de dados e criando analíticos. Estas ferramentas podem ser de self-service BI, permitindo a exploração e visualização de dados. Podem acontecer iniciativas ligadas a analíticos avançados se existirem profissionais com conhecimento mais profundo em análise de dados. Estas iniciativas podem partir de necessidade de áreas específicas e serem implementadas por uma consultoria externa.
- ✓ **Governança:** Nesta fase de maturidade, algumas organizações podem ter um comitê diretor que supervisiona o programa sob perspectiva de governança, com representantes de departamentos que fornecem relatórios de progresso e conformidade.

Abismo: À medida que as organizações tentam passar da adoção precoce para a adoção corporativa e assim estender o valor do uso do ambiente analítico para mais usuários e departamentos, as empresas devem superar uma série de obstáculos. Neste momento, existe o desafio óbvio de obter a proficiência técnica necessária para novas tecnologias, como bancos de dados NoSQL, Hadoop Self-Service BI, bem como analíticos avançadas. Para tentar superar esta barreira, algumas empresas vão ao mercado de trabalho para tentar contratar profissionais com os perfis requeridos, contratam treinamentos específicos nas tecnologias desejadas, enviam profissionais para congressos e utilizam grupos de funcionários mais capacitados para difundir o conhecimento. Com relação a gestão e governança de dados, é necessário que um sólido plano deva ser implantado. O gerenciamento de dados deve se ajustar a diferentes e dinâmicos requisitos, com disponibilidade de dados e desempenho necessários para suportar análises democratizadas, o que significa que os controles e práticas de qualidade de dados e controle de acesso devem ser implementados e governados. As empresas precisam construir uma

arquitetura de informação unificada com uma plataforma analítica corporativa garantindo coerência e integridade para o uso de analíticos. Questões políticas e culturais devem ser resolvidas, como por exemplo, a visão de apropriação de dados por algum departamento, o que torna mais difícil a criação de um programa amplo de exploração de dados. Não menos importante, é necessário garantir o aporte financeiro para aquisições e treinamentos de novas tecnologias, para isso, uma estratégia pode ser através da implantação de pequenos projetos iniciais que agreguem valor para o negócio. Por estas razões, para atravessar com sucesso este abismo é que se gasta uma grande quantidade de tempo nesta fase.

Estágio 4: Adoção Corporativa: A adoção corporativa é a fase mais desafiadora no processo de consolidação na jornada analítica de uma organização. Durante a adoção corporativa, os usuários finais estão bem envolvidos e o uso do ambiente analítico transforma o modo de trabalho das áreas de negócio. Os usuários alteraram a forma como as decisões são tomadas com a implantação de analíticos. Neste estágio, estarão usando tipos diferentes de dados, big data e até dados não estruturados em seus esforços de análise. As organizações que atingem esta fase de maturidade já endereçaram alguns problemas na organização, infraestrutura, gerenciamento de dados, analíticos e governança.

- ✓ Organização: Nesta fase, percebe-se que o uso do ambiente analítico é um diferencial competitivo e existem projetos de sucesso ou protótipos maduros que comprovem isso. A inovação de dados e analíticos é vista com bons olhos pois já existe uma cultura em análises estabelecida. A estratégia de negócios é uma mistura de top-down e bottom-up, com uma infraestrutura de dados que possa suportá-la. As áreas de negócios e a de TI trabalham como uma única equipe, compartilhando projetos de análise em conjunto, ou seja, percebe-se uma crescente cultura e mentalidade analítica permeando a organização.
- ✓ Infraestrutura: Neste estágio de maturidade, uma gama de tecnologias podem estar sendo usadas, como bancos de dados NoSQL, Hadoop, ferramentas de self-service BI enterprise e um DW. A infraestrutura é predominante on premises, mas podendo ter implantações em nuvem. A arquitetura da informação é unificada de uma forma que sustenta a análise. A empresa trabalha em um ecossistema analítico que incentiva a inovação e permite aos usuários explorar tipos novos de dados em uma variedade de plataformas.

- ✓ Gerenciamento de dados: Quanto mais madura é uma empresa, melhor ela pode gerenciar e usar seus dados especificamente para análises avançadas, e não apenas para relatórios de BI. Nesta fase de maturidade, as organizações podem fazer uso de muitas formas de dados e absorver novas fontes de dados à medida que elas emergem. O compartilhamento de dados é uma atividade colaborativa que é bem gerenciada através de políticas sólidas de governança de dados. A posse de dados em silos individuais de usuários ou departamentos é considerado negativo nessa fase, pois a empresa está buscando consolidar ou empregar federação de dados e virtualização para unificar visualizações de dados de todos os componentes da infraestrutura de dados. A empresa está sempre à procura de novos dados, tanto interna como externamente, com o intuito de proporcionar melhores análises. Existe um gerenciamento de ciclo de vida e uma auditabilidade de dados e processos definidos, o que permite a existência de uma estrutura de linhagem de dados. Vale destacar que existem metadados atribuídos e até uma possível camada semântica de informações.

- ✓ Analíticos: Neste estágio de maturidade, os dados novos que entram na organização podem ser analisados rapidamente e passam a fazer parte da infraestrutura lógica, com isso, o ambiente analítico consegue suportar a organização. Normalmente, uma empresa também possui um centro de excelência (COE) em ambiente analítico que atende a diferentes partes da organização. O COE inclui a equipe de ciência dos dados, que pode até treinar outros grupos no uso de análises em diferentes formas. Neste ponto, o uso de analíticos é operacionalizado como parte dos processos de negócios, aonde são analisados diferentes tipos de dados, incluindo dados não estruturados e dados geoespaciais. Essas empresas geralmente usam capacidades novas e não apenas a infraestrutura de BI existente.

- ✓ Governança: Uma empresa nesta fase de maturidade entende que o ambiente analítico, com todos os seus benefícios, pode ser uma grande responsabilidade, especialmente se existirem dados de fontes externas a serem considerados. Nesta fase de maturidade, as organizações terão a governança com a orientação do PMO e um comitê diretor que supervisiona o programa em uma perspectiva para toda a empresa. Neste cenário, um patrocinador executivo geral está envolvido em

atualizações mensais. O programa é executado como uma iniciativa corporativa orçamentada e planejada e tratado a par com outros programas de integração de dados. O controle da governança não será muito pesado e equilibrará as necessidades quem envolvem políticas de privacidade da informação com requisitos dos usuários para análises relevantes, oportunas e consistentes.

Estágio 5: Madura ou Visionária: Apenas algumas empresas podem ser consideradas visionárias quando o assunto é ambiente analítico. Nesta fase, as organizações estão executando programas analíticos sem problemas ou dificuldades e usando uma infraestrutura altamente ajustada com estratégias bem estabelecidas de gerenciamento de programas e dados. O acesso de dados bem governado e flexível está disponível para os usuários realizarem exploração e desenvolverem visualizações através de self-service BI e de forma independente da área de TI. Muitos programas são executados como iniciativas orçamentadas e planejadas em uma perspectiva organizacional. No estágio visionário, há entusiasmo e energia para com o tema e a existência de uma cultura analítica flexível e ágil beneficia os usuários de todos os escalões da empresa.

- ✓ Organização: Neste segmento, os executivos investem em um ambiente analítico e consideram isso como crítico, pois são vistos como uma arma competitiva. O uso de analíticos não é visto apenas para direcionar a estratégia a ter novas percepções, em vez disso, a mentalidade criativa está sempre procurando por oportunidades para usar analíticos de novas maneiras. Aqui tem-se a percepção de que o uso de analíticos está se tornando mais abrangente e eles precisam estar pensando fora da caixa. As empresas maduras e visionárias geralmente conseguem acessar análises de vários dispositivos.
- ✓ Infraestrutura e gerenciamento de dados: A forma de gerenciar a complexidade é a chave para a maturidade da análise. A empresa visionária implantou uma infraestrutura analítica coerente que está totalmente operacional e pode ser usada nas missões críticas de negócios. Parte da infraestrutura inclui a capacidade de integrar novas fontes de dados para análise, sejam elas internas ou externas à empresa. A infraestrutura usa o que funcionou no passado, incluindo um DW, e pode aproveitar a tecnologia mais recente, como bancos de dados NoSQL, Hadoop além da utilização de nuvem, normalmente de forma híbrida. A governança está bem

estabelecida e o self-service BI é implementado com supervisão, a partir de uma estratégia de acesso a dados bem gerenciada.

- ✓ Analíticos: A empresa visionária desenvolve continuamente seu ambiente analítico. Normalmente, a empresa faz uso de todos os tipos de dados, incluindo dados não estruturados e dados em tempo real, para tomada de decisão e incorporação nos processos de negócios. A empresa visionária pode conectar os pontos entre os dados novos e os ativos existentes. A empresa estabeleceu COEs, e as equipes estão trabalhando para oferecer novas e atraentes formas de análises. Algumas organizações visionárias criam uma equipe conjunta de negócios e TI que inova nas tecnologias, as traz de volta ao negócio e os leva à produção.

5.4 MODELO HP

A HP desenvolveu um modelo de forma a ilustrar a evolução das capacidades de BI dos seus clientes. Neste sentido, o modelo proposto baseia-se na sua experiência com o cliente e estrutura-se em três capacidades: gerenciamento da informação, habilitação de negócio e gerenciamento de estratégia e programas. Este modelo é composto por cinco níveis de maturidade e apresenta as principais características a ter para poder evoluir em maturidade, conforme é visível na figura abaixo (HP, 2009).

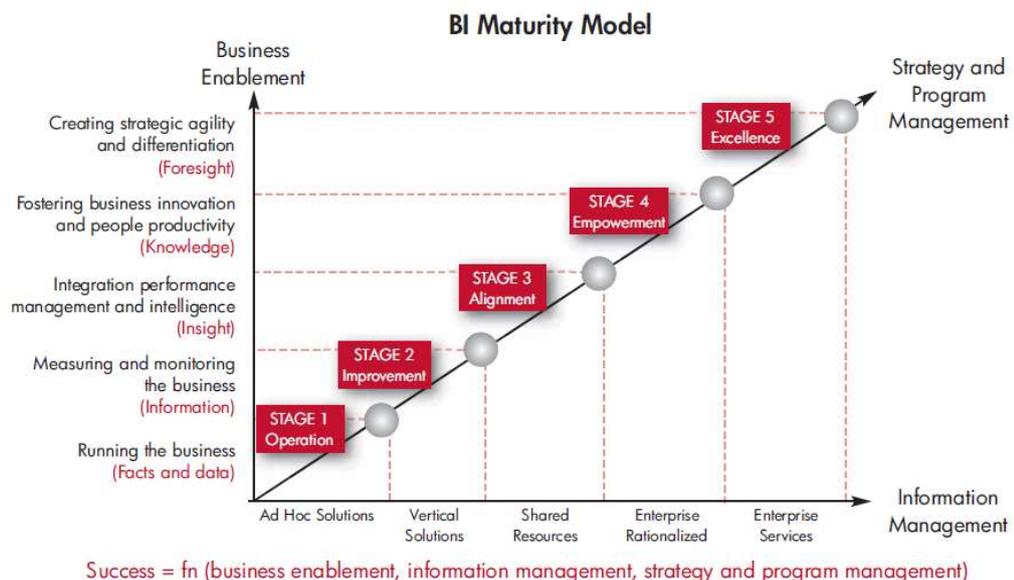


Figura 16 – Modelo de maturidade para BIA da HP (HP, 2009)

Na figura 16, a dimensão de habilitação de negócios descreve o avanço natural dos tipos de necessidades e problemas do negócio resolvidos com soluções de BI. A dimensão de gerenciamento da informação descreve o avanço natural das soluções de informação que uma empresa adota para atender a uma variedade de necessidades de negócios. A dimensão de gerenciamento de estratégia e programas descreve o natural avanço da habilidade de gerenciamento como um facilitador e catalisador chave para o sucesso do BI (HP, 2009).

Estágio 1: Soluções Ad hoc: Encontram-se empresas que estão apenas começando a pensar sobre BI, mas também podem englobar organizações que já passaram por iniciativas sem muito sucesso e estão revisitando sua estratégia básica de BI. Este é o estágio inicial em que quadros estão sendo estabelecidos e os conceitos estão sendo formados, muitas vezes no nível departamental da empresa.

- ✓ Habilidade de negócios: As necessidades de negócios são focadas no aprimoramento de recursos básicos de relatórios e análises, com planilhas frequentemente usadas como ferramenta de BI. Os consumidores de BI concentram-se principalmente entre executivos e gerentes, com um pequeno grupo de analistas ou usuários de operações fazendo o trabalho manual de reunir dados de várias fontes para criar relatórios e análises básicos para alimentar o gerenciamento.
- ✓ Gerenciamento da informação: Concentram em obter acesso básico a dados, geralmente através de esforços locais e ad hoc focados em atender uma demanda departamental. Caracteriza-se por esforços iniciais em ETL, DW e OLAP, além disso, não há integração entre departamentos ou compartilhamento de repositórios de dados.
- ✓ Gerenciamento de estratégia e programa: As atividades possuem foco limitado no gerenciamento de projetos como disciplina. Os projetos são pequenos em escala e principalmente intradepartamentais. A experiência em BI é concentrada em um pequeno grupo de indivíduos na organização.

Estágio 2: Soluções Verticais: Os esforços de BI são focados em áreas de assunto localizadas ou soluções verticais, como a implementação de análise de Recursos Humanos ou o início da integração de dados de clientes. Neste estágio, as soluções ad-hoc dão lugar a

estratégias mais planejadas de como não apenas medir o que está acontecendo com o negócio, mas também como monitorá-lo e fornecer um nível de visibilidade para o gerenciamento usar em uma maneira mais oportuna.

- ✓ **Habilitação de negócios:** O foco está no aprimoramento dos recursos de relatórios e análise, mas já verifica-se a criação de dashboards e scorecards. Os consumidores de BI ainda se concentram principalmente no nível executivo, mas há menos esforço manual envolvido na preparação e apresentação dos dados. A latência da informação está diminuindo, com relatórios e análises entregues semanalmente ou mensalmente.
- ✓ **Gerenciamento da informação:** Caracteriza-se pelo uso de repositórios como DW, DM e ODS que podem ser alavancados por vários departamentos ou unidades de negócios, mas cada um permanece focado em uma única área de assunto sem integração entre eles. A qualidade dos dados está se tornando mais importante, mas os esforços nesta área são em grande parte ad hoc e manual, envolvendo principalmente limpeza baseada em ETL.
- ✓ **Gerenciamento de estratégia e programa:** Começa o reconhecimento da criticidade do gerenciamento de projetos como uma disciplina. Os papéis e habilidades baseados em projetos para BI são formalmente identificados e os gerentes de projetos de BI têm responsabilidades interdepartamentais.

Estágio 3: Soluções Compartilhadas: Empresas no estágio 3 estão começando a alcançar uma visão integrada do assunto através das soluções verticais de informação que implementaram no estágio 2. Essa integração permite o aumento do valor do negócio através de métricas e análises mais complexas. Nesta etapa, o valor comercial do BI começa a ser percebido.

- ✓ **Habilitação de negócios:** É caracterizado por uma maior capacidade de integração e alinhamento, o que aumenta o valor de negócio. Neste ponto, há um foco maior no valor das análises, muitas vezes incorporadas em outros sistemas ou fluxos de trabalho comerciais. As empresas implementaram relatórios integrados e estão se movendo em busca de indicadores de desempenho. Os consumidores de BI estão se expandindo para incluir mais usuários de negócios, não apenas executivos e

gerentes. A disponibilidade de dados é maior, bem como a sua qualidade e utilização de acordo com as necessidades de negócio.

- ✓ Gerenciamento da informação: A integração de informações através das áreas de negócio e em soluções é a chave neste estágio. Os DW e DM departamentais são consolidados em um armazenamento de dados corporativo. Surgem os primeiros esforços para alcançar uma visão única da informação. A qualidade dos dados e os esforços de governança de dados também se tornam mais importantes e mais sofisticados nesta fase.
- ✓ Gerenciamento de estratégia e programa: As organizações evoluíram do gerenciamento de projetos de BI para o gerenciamento de programas de BI. Neste estágio, começam a aparecer os centros de competências de BI.

Estágio 4: Empresa Conformada: As empresas neste estágio estão transformando a informação em um ativo poderoso. Elas estão enfrentando os desafios em torno do gerenciamento de dados mestre, governança de dados e popularizando o BI. O uso do ambiente analítico é agora um componente chave da estratégia e é incorporado aos processos de negócios. A informação está disponível quando necessário, onde é necessário. A informação é centralizada e o ambiente é flexível e pode se adaptar, facilmente, às mudanças nas questões empresariais.

- ✓ Habilidade de negócios: Caracteriza-se pela alavancagem da inteligência para transformar a forma como os processos de negócios são projetados e a forma como as pessoas trabalham. As práticas analíticas são automatizadas e incorporadas aos processos de negócios para permitir a tomada de decisões. Os usuários têm novas expectativas em torno da entrega de informações e capacidades que lhes permitem ser mais sensíveis aos seus ambientes de negócios, prevendo os resultados e aprofundando a análise. Embora os usuários de negócios recebam maior foco como usuários de BI, executivos e gerentes também adotam BI e continuam a usá-lo para definir e gerenciar a estratégia corporativa.
- ✓ Gerenciamento da informação: Uma versão única da verdade é disseminada por toda a organização. Existe uma abordagem de gerenciamento de dados mestre mais

avançada. O programa de governança de dados tem o apoio de profissionais seniores de TI e de negócios sendo capaz de impor padrões e políticas por toda a organização. Um programa robusto de qualidade de dados está em vigor e as áreas de negócios e a TI se comprometem a lidar com problemas de qualidade de dados diretamente na fonte. Estão em curso esforços para integrar conteúdos não estruturados com dados estruturados para ampliar a capacidade analítica.

- ✓ Gerenciamento de estratégia e programa: A gestão de portfólio torna-se de maior importância à medida que o BI se torna um componente natural de todas as iniciativas estratégicas. A empresa possui um modelo de governança avançado para BI e uma abordagem sofisticada para identificar e realizar os benefícios comerciais dos programas de BI.

Estágio 5: Serviços Corporativos: Este é o estágio final em que o BI é integrado em todos os níveis de uma organização e as análises preditivas são usadas em todos os níveis para a maioria das decisões de negócios. Existe uma abordagem totalmente funcional baseada em uma arquitetura orientada a serviços que traz uma flexibilidade ao modelo de negócios através da geração ágil de informações. Sistemas, usuários e ferramentas estão integrados, e estão fornecendo orientação, suporte e impulso para o portfólio de BI.

- ✓ Habilidade de negócios: A organização se beneficia de um ambiente de informação ágil. A modelagem de negócios sistêmica e dinâmica torna-se uma realidade trazendo vantagem competitiva. O ambiente analítico é visto como um diferencial chave para a organização, não apenas como uma atividade de valor agregado. A inovação em BI torna-se um tema central no portfólio de investimentos da organização. Os usuários em todos os níveis da organização têm acesso a percepções e descobertas que os ajudam a trabalhar com mais eficiência e a otimizar os resultados comerciais.
- ✓ Gerenciamento da informação: As capacidades de gerenciamento de informações atingem seu nível mais alto, onde as informações são entregues através de um modelo de serviço. As informações integradas estão disponíveis de forma perfeita, independentemente da fonte de dados ou da tecnologia de integração utilizada. Os programas de governança de dados e qualidade continuam em pleno vigor, apoiados

pelo patrocínio de mais alto nível. Dados não estruturados e estruturados são totalmente integrados. O BI avançado é totalmente incorporado nos processos, sistemas e fluxo de trabalho. Os esforços de entrega de informações são caracterizados pela agilidade e novos analíticos são facilmente desenvolvidas. Os usuários podem acessar as informações de que precisam, quando precisam, entregues de forma a suportar perfeitamente seu papel na organização.

- ✓ Gerenciamento de estratégia e programa: Caracteriza-se pela percepção do completo valor de BI. O BI funciona como uma alavanca para mudanças estratégicas. Um novo papel, do Chief Analytics Officer (CAO), pode até surgir nesta fase para orientar a organização no uso de seus ativos de informação para superar a concorrência. O gerenciamento avançado de portfólio de BI ajuda a organização a alcançar o valor estratégico completo de seus investimentos em BI e a empresa tem o compromisso de incluir o BI na agenda de sua organização de pesquisa e desenvolvimento.

6 ESTUDO DE CASO

Este estudo de caso tem como objetivo de fazer uma avaliação do estágio atual em que uma empresa se encontra no que se refere a ambiente analítico, assim como identificar as estratégias e ações necessárias para a sua evolução. Como visto anteriormente, uma boa forma de realizar um estudo como este é através da aplicação de um modelo de capacidade e maturidade. O modelo de avaliação escolhido para em trabalho foi o do TDWI, pois a pesquisa necessária para realizar esta avaliação encontra-se disponível de forma gratuita na internet².

Esta pesquisa faz perguntas sobre as estratégias atuais da organização em relação à BI e ao uso de dados na geração de analíticos. Ao completar a avaliação, será fornecido um conjunto de pontuações que indica a maturidade da empresa para ambiente analítico através das cinco dimensões chave: organização, infraestrutura, gestão de dados, analíticos e governança, sendo ainda possível comparar a pontuação obtida com algumas referências de mercado, o que pode ajudar a planejar o futuro de forma mais eficaz. Para os fins desta pesquisa, serão avaliados tanto o BI tradicional, quanto análises mais avançadas, como análises preditivas e prescritivas.

As perguntas do questionário são fornecidas individualmente ou agrupadas em uma matriz e cada dimensão pode atingir uma pontuação de até 20 pontos. Como as organizações podem estar em diferentes níveis de maturidade nas cinco dimensões, cada uma será classificada separadamente. Na pesquisa, há perguntas que não são computadas, mas sim usadas para orientações sobre práticas recomendadas. O quadro 7 traz a referência para cada dimensão com base na pontuação obtida.

Quadro 7- Interpretação da pontuação da avaliação de maturidade do TDWI

Pontuação por dimensão	Estágio
4,00 a 7,10	Nascente
7,20 a 10,10	Pré-adoção
10,20 a 13,30	Adoção precoce
13,40 a 16,60	Adoção corporativa
16,70 a 20,00	Madura ou Visionária

Este estudo de caso começa com a apresentação do cenário em que a empresa em questão se encontra para em seguida termos a avaliação com uma análise de cada uma das cinco

² <https://tdwi.org/analyticsMM>

dimensões do modelo TDWI e por fim, traremos uma comparação com avaliações de referências de mercado. Na avaliação e análise de cada dimensão, será mostrada a percepção existente sobre a dimensão, a pontuação obtida com a sua análise, o seu grau de maturidade, um quadro de referências de mercado e um conjunto de recomendações para atingir o próximo estágio.

6.1 DESCRIÇÃO DO CONTEXTO

O estudo de caso foi realizado em uma empresa do setor elétrico brasileiro, ligada ao Governo Federal, que não desempenha atividade lucrativa e tem como missão planejar e operar o sistema elétrico do país. Esta organização não é proprietária de equipamentos ou instalações do setor, mas o seu grande ativo são as informações destes bem como de demais variáveis relativas à operação e planejamento de todo o sistema elétrico brasileiro

Atualmente, esta empresa está inserida em um cenário de transformação setorial com novos atores e tecnologias de energia que estão provocando uma revolução no modelo atual do setor elétrico brasileiro. Neste cenário, é nítido o aumento da complexidade da missão da empresa que passa a ter um aumento crescente no volume, diversidade e velocidade dos dados e processos do setor elétrico, o que traz a tona a necessidade de uma evolução tecnológica com ótica centrada a dados. Esta visão centrada em dados deve alavancar seu uso, como ativo da organização, e assim oferecer grandes possibilidades de exploração. Em face a isto, a empresa compreende que existe a necessidade de se construir uma estrutura para consumo de dados e capacidades analíticas, que habilite uma transformação digital com visão centrada a dados e prepare a organização para as mudanças do setor de energia.

Esta pesquisa foi realizada com a utilização do questionário em sua versão original, na língua inglesa, com usuários das áreas de estatísticas da pós-operação, estudos elétricos, contratos da transmissão e financeira da empresa.

6.2 DIMENSÕES APLICADAS SEGUNDO O MODELO DO TDWI

6.2.1 Dimensão Organização

✓ *Cenário percebido da dimensão*

A prática de BI na organização teve início com a implantação da plataforma Business Objects no ano de 2005, nesta época, os usuários das áreas de negócios começaram a desenvolver seus primeiros relatórios ad hoc a partir da camada semântica que o Business Objects disponibilizava. Nos anos seguintes, alguns cubos OLAP foram desenvolvidos aumentando assim a capacidade analítica disponível. Neste período, as necessidades de negócios estavam focadas em recursos básicos de relatórios e análises, para geralmente atender demandas departamentais de alguns executivos e gerentes.

Atualmente, existe por parte da maioria dos gestores e da diretoria a percepção de que o uso de analíticos poderá ser um diferencial na empresa, por esta razão um número crescente de áreas começa a trabalhar em conjunto com a TI em projetos específicos de analíticos. Estes projetos começam a gerar frutos e a despertar interesse de outras áreas. Recentemente, vem sendo discutida, na TI, a definição de uma infraestrutura de dados que possa suportar a estratégia analítica da corporação. Um ponto importante é que no planejamento estratégico dos últimos dois anos, foi alocada uma verba para a modernização de ferramentas, bem como a contratação de serviços para ambiente analítico.

✓ *Resultados do questionário e análises da avaliação pelas áreas*

O quadro 8 mostra que as quatro áreas avaliadas se encontram no estágio de adoção precoce para a dimensão organização. Este fato é um indicador de que as áreas de negócios e de TI começam a trabalhar como uma só equipe em projetos analíticos. Já é percebido que apesar da empresa não ser orientada como um todo para analíticos, alguns departamentos começam a se beneficiar do seu uso em seus processos de tomada de decisão.

Quadro 8 - Avaliação da dimensão organização

	Avaliação	Estágio
Pós-operação	13,00	Adoção precoce
Estudos elétricos	10,25	Adoção precoce
Contratos da transmissão	10,63	Adoção precoce
Financeira	13,13	Adoção precoce
Média	11,75	Adoção precoce

Considerando a média obtida de 11,75, a empresa pode ser classificada como em adoção precoce na dimensão organização conforme a figura 17.



Figura 17 - Estágio de maturidade na dimensão organização

De acordo com o quadro 9, as referências indicadas pela avaliação para a dimensão organização, percebemos que a maturidade da empresa encontra-se acima da média das empresas ligadas ao governo porém abaixo da empresa com maior maturidade. O mesmo ocorre quando comparamos com empresas que estão na mesma faixa de orçamento anual entre trezentos e cinquenta milhões de reais e 1 bilhão setecentos e cinquenta milhões de reais.

Quadro 9 – Referências para a dimensão organização

	Mínimo	Média	Máximo
Setor: Governo	4,00	10,30	16,75
Orçamento: R\$350M - R\$1750M	6,88	10,95	16,25

✓ *Recomendações*

- Criar um BICC, que poder começar através de um modelo virtual, não oficial e progredir para um modelo formal composto por profissionais das áreas de negócios e de TI.

- As área de negócios e a de TI devem trabalhar como uma única equipe, compartilhando projetos de análise em conjunto, desenvolvendo assim uma crescente cultura e mentalidade analítica que permeie toda a organização.
- Desenvolver uma estratégia e um roteiro de modernização e implantação para ambiente analítico a curto prazo e a longo prazo que possam ser utilizados como um guia por todos usuários de todos os departamentos da organização.

6.2.2 Dimensão Infraestrutura

✓ *Cenário percebido da dimensão*

Atualmente, a infraestrutura tecnológica da empresa para suportar a atividade analítica é predominante on premises, porém coexiste com algumas atividades de pesquisa e desenvolvimento na nuvem. O desenvolvimento e consumo de analíticos se sustenta sobre uma base de dados ODS, que consolida as informações dos diferentes sistemas transacionais. Esta base ODS encontra-se instalada em um SGBD Microsoft SQL Server 2012 e os processos de ETL que a alimentam são desenvolvidos na ferramenta também da Microsoft, o Integration Services. Para consumir os dados deste ODS, existe uma plataforma de BI tradicional instalada, o SAP Business Objects (BO). A plataforma BO foi adquirida em 2005 com o objetivo de facilitar a extração de informações das bases transacionais através de uma camada semântica, o universo do BO, que traduz toda a complexidade dos relacionamentos entre tabelas de uma base de dados para o linguajar do dia a dia da área de negócio, permitindo assim a criação de consultas ad hocs pela própria área usuária. Hoje em dia, existe um legado de cerca de 1750 documentos desenvolvidos no BO a partir de cerca de 50 universos. É importante destacar que, até hoje, ainda existem universos do BO que acessam bases transacionais. Posteriormente, foram implementados alguns cubos OLAP utilizando a ferramenta da Microsoft Analysis Services, porém esta tecnologia não obteve a adesão esperada na corporação como um todo, e acabou ficando restrita a apenas duas áreas de negócios. Ao longo dos anos, uma prática que sempre ocorreu e que ainda permanece forte é a criação de análises a partir de planilhas em Excel.

Recentemente, com intuito de modernizar o parque tecnológico e assim usufruir de todas as vantagens de um BA, foi adquirida a plataforma Tableau que em breve deverá substituir todas as análises existentes em BO.

✓ *Resultados do questionário e análises da avaliação pelas áreas*

O quadro 10 mostra que a área contratos da transmissão se encontra no estágio de maturidade de pré-adoção, enquanto as demais áreas se encontram no estágio de adoção precoce. Este resultado da área contratos da transmissão pode ser explicado pelo fato da plataforma Tableau ainda não ter sido implantada neste departamento, e a única forma de criação de análises e consultas é através dos universos do Business Objects, que para este departamento ainda apontam para bases transacionais. Nas demais áreas avaliadas, o Tableau já foi implantado e a prática do self-service BI começa a gerar produtos que agregam valor ao negócio.

Quadro 10- Avaliação da dimensão infraestrutura

	Avaliação	Estágio
Pós-operação	13,00	Adoção precoce
Estudos elétricos	12,50	Adoção precoce
Contratos da transmissão	9,50	Pré-adoção
Financeira	11,50	Adoção precoce
Média	11,63	Adoção precoce

Considerando a média obtida de 11,63, a empresa pode ser classificada como em adoção precoce na dimensão infraestrutura conforme a figura 18.



Figura 18 - Estágio de maturidade na dimensão infraestrutura

Considerando as referências mostradas no quadro 11 que foram indicadas pela avaliação para a dimensão infraestrutura, verificamos que a maturidade da empresa encontra-se acima da média das empresas ligadas ao governo porém abaixo

da empresa com maior maturidade. Tal fato se repete quando comparamos com empresas que estão na mesma faixa de orçamento anual entre trezentos e cinquenta milhões de reais e 1 bilhão setecentos e cinquenta milhões de reais.

Quadro 11- Referências para a dimensão infraestrutura

	Mínimo	Média	Máximo
Sector: Governo	4,00	10,13	14,50
Orçamento: R\$350M - R\$1750M	6,00	10,15	15,50

✓ *Recomendações*

- Implantar ferramentas modernas como o Tableau de forma corporativa, atendendo assim a todos os departamentos.
- Implantar bases de dados em memória (local ou na nuvem) para melhorar o desempenho de aplicativos analíticos.
- Incentivar a inovação através de novas tecnologias como bancos de dados NoSQL e Hadoop permitindo aos usuários explorar novos tipos de dados em uma variedade de plataformas.
- Adicionar à arquitetura vigente o componente de LDW, para acelerar desenvolvimentos sem ter a necessidade de criar processos de coleta de dados.

6.2.3 Dimensão Gestão de Dados

✓ *Cenário percebido da dimensão*

Uma característica do ambiente analítico desta empresa é que os dados que se encontram disponíveis no ODS, para o consumo de analíticos, não representa a totalidade dos dados gerados pelas aplicações transacionais. Estes dados representam apenas uma parcela que, em algum momento, foi solicitado por alguma área de negócios para atender a algum tipo de análise ou relatório. No dia a dia, um grande volume de manutenções evolutivas é solicitado à TI para acrescentar novos conjuntos de dados ao ODS, o que gera sensíveis alterações na modelagem de dados e nos processos de etl, o que eventualmente pode trazer um

certo grau de instabilidade no processo de carga como um todo. Uma particularidade que existe nesta gestão diz respeito a periodicidade de atualização da base ODS. Este processo de atualização é disparado por um job que executa uma carga completa do ODS duas vezes ao dia, a primeira começando às 03:00 e a segunda às 12:00, ou seja, não existe um processo de carga incremental e este processo de execução leva cerca de três horas. Esta particularidade traz inconvenientes para aplicações analíticas que possuam requisitos que necessitem de dados gerados próximos ao tempo real, o que pode inviabilizar o uso destes analíticos. Um ponto importante é que apenas dados estruturados estão disponíveis para criação e consumo de analíticos. Um outro ponto importante é que não existe um processo estabelecido para tratar qualidade dos dados nem a gestão de dados mestres.

✓ *Resultados do questionário e análises da avaliação pelas áreas*

O quadro 12 nos traz a informação que os departamentos de pós-operação e contratos da transmissão encontram-se no estágio de pré-adoção de maturidade. A explicação para este cenário é que os dados de diversos processos de negócios destas áreas não se encontram disponíveis para consumo no ambiente analítico. Este fato é mais grave na área de contratos da transmissão, aonde as análises são limitadas e feitas apenas em bases transacionais. Já os departamentos de estudos elétricos e financeiro encontram-se em um estágio de maturidade mais avançado, adoção precoce, pois boa parte dos dados de seus processos já se encontram disponíveis para consumo no ODS. Vale destacar que algumas aplicações analíticas já foram construídas com práticas de self-service através da ferramenta Tableau. Estas aplicações representam casos de sucesso que começam a despertar o interesse pelo tema em outras áreas da organização.

Quadro 12- Avaliação da dimensão gestão de dados

	Avaliação	Estágio
Pós-operação	9,00	Pré-adoção
Estudos elétricos	13,50	Adoção precoce
Contratos da transmissão	8,50	Pré-adoção
Financeira	10,50	Adoção precoce
Média	10,38	Adoção precoce

Considerando a média obtida de 10,38, a empresa pode ser classificada como em adoção precoce na dimensão gestão de dados conforme a figura 19.



Figura 19 - Estágio de maturidade na dimensão gestão de dados

Considerando as referências do quadro 13, informadas ao final da avaliação para a dimensão gestão de dados, constatamos que a maturidade da empresa encontra-se acima da média das empresas ligadas ao governo porém abaixo da empresa com maior maturidade. O mesmo ocorre quando comparamos com empresas que estão na mesma faixa de orçamento anual entre trezentos e cinquenta milhões de reais e 1 bilhão setecentos e cinquenta milhões de reais.

Quadro 13- Pontuação obtida na dimensão gestão de dados

	Mínimo	Média	Máximo
Sector: Governo	4,00	10,13	14,50
Orçamento: R\$350M - R\$1750M	5,00	8,72	13,00

✓ *Recomendações*

- Ampliar a quantidade de dados no ODS, com a maior abrangência possível, para que este possa atender a todos os interessados no uso do ambiente analítico.

- Começar um programa de qualidade de dados junto com as iniciativas de gerenciamento de dados mestres.
- Começar a tratar diferentes tipos de dados, incluindo dados não estruturados e dados geoespaciais.
- Federar metadados em todas as tecnologias.
- Os dados devem ser compartilhados através de uma atividade colaborativa bem gerenciada através de políticas sólidas de governança de dados.

6.2.4 Dimensão Analíticos

✓ *Cenário percebido da dimensão*

Apesar do principal ativo da empresa ser a informação, o uso de analíticos não faz parte da cultura da empresa e assim não é utilizado de forma sistêmica nos principais processos de tomada de decisão. Olhando para o espectro de capacidades analíticas, identificamos os analíticos descritivos e de diagnóstico como os mais disseminados por toda a corporação. Geralmente, são relatórios desenvolvidos em BO e que retratam o estado de um determinado processo ao longo de um período de tempo e ainda, em alguns casos, apresentam as suas causas e consequências. De um modo geral o uso destes relatórios servem de apoio à processos operacionais de negócio e a maior parte do legado foi desenvolvido pela equipe de TI. Com uma utilização mais intensa, as planilhas em Excel representam o maior legado de análises produzidas pelas próprias áreas usuárias, é comum encontramos documentos desenvolvidos em Business Objects que servem como geradores de dados para análises realizadas, posteriormente, em planilhas. Os analíticos preditivos e prescritivos são encontrados em pouquíssimas unidades de negócios e representam atividades isoladas das próprias áreas que compram suas próprias ferramentas e criam silos de conhecimento de tecnologia que não é compartilhada por toda a empresa. Recentemente, com a aquisição do Tableau, a prática de análise de descoberta começa de modo tímido a ganhar forma.

✓ *Resultados do questionário e análises da avaliação pelas áreas*

O quadro 14 retrata que a área de contratos da transmissão se encontra no estágio de pré-adoção de maturidade em analíticos, o que pode ser facilmente explicado pela dinâmica de trabalho da área, cujo foco principal sempre foi o operacional, e por conta disso, sempre pediu apoio à área de TI para construir relatórios que servissem de apoio às atividades do dia a dia. Atualmente, existem cerca de cinquenta documentos desenvolvidos no BO, sendo que a maioria está obsoleta ou caíram em desuso. Vale destacar que o próprio universo utilizado encontra-se defasado. Recentemente, a área de contratos mostrou interesse em modernizar suas análises. Já as áreas de pós-operação, estudos elétricos e financeira estão no nível de maturidade de adoção precoce pois já desenvolvem práticas de self-service com a utilização do Tableau, e ainda procuram fazer uso da visualização gráfica para otimizar processos e tomar decisões baseadas em dados. Neste contexto, a área financeira está um degrau a frente das demais, pois já desenvolveram um conjunto de painéis com a finalidade de ter um acompanhamento orçamentário preciso. Estes painéis já começam a ser utilizados por diversos executivos de diversas gerências, e inclusive por diretores.

Quadro 14- Avaliação da dimensão analíticos

	Avaliação	Estágio
Pós-operação	10,25	Adoção precoce
Estudos elétricos	10,75	Adoção precoce
Contratos da transmissão	8,75	Pré-adoção
Financeira	11,00	Adoção precoce
Média	10,19	Adoção precoce

Considerando a média obtida de 10,19, a empresa pode ser classificada como em adoção precoce na dimensão analíticos conforme a figura 20.



Figura 20 - Estágio de maturidade na dimensão analíticos

Comparando as referências informadas, vide quadro 15, ao final da avaliação para a dimensão de analíticos, constatamos que a maturidade da empresa encontra-se acima do mínimo das empresas ligadas ao governo porém abaixo média das empresa com maior maturidade. Porém quando comparamos com empresas que estão na mesma faixa de orçamento anual entre trezentos e cinquenta milhões de reais e 1 bilhão setecentos e cinquenta milhões de reais, verificamos que encontra-se acima da média e abaixo da maior avaliação encontrada.

Quadro 15 - Pontuação obtida na dimensão analíticos

	Mínimo	Média	Máximo
Setor: Governo	4,00	10,31	15,25
Orçamento: R\$350M - R\$1750M	5,25	9,52	14,00

✓ *Recomendações*

- Expandir o uso de ferramentas de self-service BI por toda a corporação, incentivando e facilitando a prática de descoberta de dados por usuários certos e com o grau apropriado de governança.
- Disseminar as iniciativas existentes de exploração da informação através do uso de analíticos avançados como preditivos e prescritivos.
- Tornar o uso de analíticos deve como parte dos processos de negócios, tornando este uso parte da cultura da empresa.

6.2.5 Dimensão Governança

✓ *Cenário percebido da dimensão*

Falar de governança na empresa é um assunto complexo, principalmente quando falamos de governança de dados. Não está claro para a corporação as motivações e os objetivos esperados com a adoção de um programa como este.

De uma forma geral, a necessidade de uso sempre surge através de determinados executivos ou diretores em situações que são influenciadas por questões de TI. Neste contexto, estas iniciativas são vistas como um modismo ou ainda encaradas como um trabalho imposto, sem sentido e burocrático. Percebe-se claramente a falta de engajamento e liderança na condução de um projeto de governança a nível corporativo. Alguns setores da empresa não possuem o sentimento que seu trabalho interfere diretamente em toda a cadeia de negócio, gerando frutos e retorno direto, e para piorar a situação, ainda existe uma cultura de propriedade da informação que permeia parte da organização, o que de maneira geral, é um complicador para práticas modernas de BI como o self-service, pois as áreas “donas” das informações dificultam o consumo destas informações por outras áreas. Falando de governança de aplicações analíticas, a área de TI está propondo uma governança para painéis desenvolvidos no Tableau, aonde cada painel desenvolvido pela área usuária poderá ter um selo que indicará se este é um painel corporativo, departamental ou apenas de uso do próprio usuário que o desenvolveu.

✓ *Resultados do questionário e análises da avaliação pelas áreas*

Com base no quadro 16, percebemos que as avaliações para a dimensão de governança são as piores de toda a pesquisa, e com exceção da área financeira em que seu nível de maturidade ficou em adoção precoce as demais áreas foram qualificadas como em um estágio de pré-adoção. A explicação para esta avaliação ruim se deve a alguns fatores como a falta de uma definição nas políticas de gestão, propriedade e responsabilidades sobre dados e informações, a não atuação de uma equipe de governança analítica composta de usuários chaves das áreas de negócio, a criação de mecanismos de segurança que dificultem o acesso ou compartilhamento de informações e ao não monitoramento à adesão das políticas analíticas quando criadas.

Quadro 16- Avaliação da dimensão governança

	Avaliação	Estágio
Pós-operação	8,50	Pré-adoção
Estudos elétricos	8,25	Pré-adoção
Contratos da transmissão	8,00	Pré-adoção
Financeira	10,75	Adoção precoce
Média	8,88	Pré-adoção

Considerando a média obtida de 8,88, a empresa pode ser classificada com o nível de maturidade em pré-adoção na dimensão governança conforme a figura 21.



Figura 21- Estágio de maturidade na dimensão governança

Utilizando as referências informadas ao final da avaliação, vide quadro 17, para a dimensão de governança, constatamos que a maturidade da empresa encontra-se um pouco acima da média das empresas ligadas ao governo porém abaixo média das empresa com maior maturidade. Porém quando comparamos com empresas que estão na mesma faixa de orçamento anual entre trezentos e cinquenta milhões de reais e 1 bilhão setecentos e cinquenta milhões de reais, verificamos que encontra-se abaixo da média e acima da menor avaliação encontrada.

Quadro 17- Pontuação obtida na dimensão governança

	Mínimo	Média	Máximo
Sector: Governo	5,50	8,84	12,00
Orçamento: R\$350M - R\$1750M	5,50	9,89	14,25

✓ *Recomendações*

- Criar uma equipe de governança analítica composta de usuários de diferentes unidades de negócio.
- Criar políticas de segurança para acesso e compartilhamento de informações que não engessem a dinâmica analítica e que sejam flexíveis a adaptações
- Criar mecanismos para monitorar o nível de adesão às políticas estabelecidas.
- Estender as iniciativas de gerenciamento de desempenho para processos que agregem os esforços em ambiente analítico e valor de negócio.
- As ações de governança devem contar com a orientação de um PMO e de um comitê diretor para supervisionar o programa em uma perspectiva para toda a empresa.

7 CONCLUSÕES

7.1 TENDÊNCIAS ANALÍTICAS

Neste trabalho, foi feita uma ampla revisão contemplando os conceitos tradicionais de BI, que vem sendo utilizado por muitas empresas há cerca de duas ou mais décadas. Foi mostrado que estes conceitos não mais atendem as necessidades e demandas das empresas e que para estas continuarem competitivas, tendo a informação como a principal fonte de tomadas de decisões é necessário uma modernização que engloba diversos aspectos como governança, infraestrutura, gestão de dados, organização e uso analítico. Neste contexto de modernização, a tendência é que as empresas busquem um equilíbrio entre as abordagens top-down e bottom-up de desenvolvimento analítico, aproveitando o que cada linha tem de melhor a oferecer, habilitando assim as áreas de negócios, em um time híbrido com a TI, a desenvolverem suas análises nas diferentes capacidades analíticas existentes, no tempo requerido pelo próprio negócio e com a governança necessária para se evitar um ambiente caótico.

7.2 IMPORTÂNCIA DA AVALIAÇÃO DA MATURIDADE ANALÍTICA

Este estudo demonstrou que a avaliação de maturidade é um quesito vital e inerente em um processo de modernização analítica. Esta avaliação torna-se o instrumento que possa ser o guia direcionador desta transformação dentro da empresa, capaz de confrontar os pontos fracos com os pontos fortes, bem como apresentar condições para análise de oportunidades e potenciais ameaças e assim desenvolver uma melhor estratégia para evolução. Uma vez apurado o nível de maturidade em ambiente analítico da organização, pode-se partir para etapas mais práticas de implantação da nova estrutura. Contudo, o processo de avaliação é algo complexo dada a transversalidade, subjetividade e natureza estratégica associados ao ambiente analítico.

Esta monografia deixa visível que a utilização de um modelo de capacidade e maturidade pode ajudar uma organização a evoluir de forma sistemática na sua capacidade de construir aplicativos analíticos de forma eficaz e eficiente, pois fornece informações que orientam na definição do plano de melhoria da qualidade e produtividade, permitindo ainda o aproveitamento dos processos e recursos já existentes, além de direcionar e justificar investimentos.

7.3 PROPOSTAS DE TRABALHOS FUTUROS

Com a intenção de ultrapassar o escopo do estudo desenvolvido neste trabalho, pode-se recomendar algumas sugestões para trabalhos futuros:

1. Desenvolver um estudo comparativo sobre maturidade analítica entre diferentes segmentos do setor elétrico como por exemplo distribuição e geração de energia.
2. Utilizar um outro modelo de maturidade, como por exemplo, o MPS.BR para realizar um estudo comparativo com o CMMI de forma a se identificar vantagens e desvantagens de cada modelo em sua aplicabilidade a modernização de um ambiente analítico.
3. Desenvolver um estudo semelhante em outros setores da sociedade como saúde, segurança, transportes entre outros com a possibilidade de criar um mapa de maturidade analítica por setor, que possa servir de referências para empresas dos respectivos nichos.

REFERÊNCIAS BIBLIOGRÁFICAS

ADDRIANS, P.; ZANTINGE, D. **Data Mining**. Inglaterra: Addison-Wesley, 1996.

ALBRIGHT, C.; WINSTON, W. **Business Analytics: Data Analysis and Decision Making**. 5th Edition, 2013.

ALECRIM, Emerson. **O que é Big Data?** 2015. <https://www.infowester.com/big-data.php>
(Acessado em 28/04/2017)

ALMEIDA, Kleber. **Revista Eletrônica de Administração** – ISSN 1676-6822 Periodicidade semestral, Edição número 8 – junho de 2005

AMAZON - **What is a Data Lake?** <https://aws.amazon.com/pt/big-data/data-lake-on-aws/>
(acessado em 06/11/2017)

BALTZAN, P.; PHILLIPS, A. **Sistemas de Informação**. 2012.

BARBIERI, Carlos. **BI – Business Intelligence: modelagem e tecnologia**. Axcel Books, 2001.

BARBIERI, Carlos. **BI2 – Business Intelligence: Modelagem e Qualidade**. Elsevier Editora, 2011.

BARONY, Yuri. **Big Data: questão de moda ou necessidade?**
<https://imasters.com.br/desenvolvimento/big-data-questao-de-moda-ou-necessidade>
(Acessado em 28/04/2017)

BARTIÉ, Alexandre. **Garantia da Qualidade de Software**. 13ª Tiragem, Elsevier Editora, 2002.

BIGURU, **The Business Intelligence Blog**, <https://biguru.wordpress.com/2012/01/08/bi-maturity-models/> (Acessado em 30/09/2017)

COOK, Henry. **Solution Path for Planning and Implementing the Logical Data Warehouse**. Gartner (G00320563), 2017.

DA SILVA, Alessandro Lemes. **Já ouviu falar em Data Federation?** <http://www.infomev.com.br/ja-ouviu-falar-em-data-federation/> (Acessado em 11/11/2017)

DE BRUIN, T. **Understanding the Main Phases of Developing a Maturity Assessment Model.** Australasian Conference on Information Systems (ACIS), 2005.

DECKER, K. and FOCARDI, S. **Technological overview: a report on data mining.** CSCS — Swiss National Supercomputing Center, Technical Report, Zúrique, 1995.

DUAN, L., and XU, L. D. **Business Intelligence for Enterprise Systems: A Survey.** IEEE Transactions on Industrial Informatics, 2012.

ECKERSON, Wayne. **Five steps for delivering self-service business Intelligence to everyone.** Information Builders, 2014.

ECKERSON, Wayne. **From Reporting to Analytics: A Pathway to Greater BI Value.** Eckerson Group, 2015.

ECKERSON, Wayne. **Governed Data Discovery, Balancing Flexibility and Standards.** Eckerson Group, 2016.

ECKERSON, Wayne. **How Visual Design Standards Improve Analytical Literacy.** Eckerson Group, 2015.

ELENA, C. (2011). **Business intelligence.** Journal of Knowledge Management, Economics and Information Technology, 1(2). Retrieved from http://www.scientificpapers.org/wp-content/files/1102_Business_intelligence.pdf

ELIAS, Diego. **O que significa OLTP e OLAP na prática.** <https://canaltech.com.br/business-intelligence/o-que-significa-oltp-e-olap-na-pratica/> (Acessado em 31/10/2017), 2014.

ELMASRI, R.; NAVATHE, S. B. **Sistemas de banco de dados: fundamentos e aplicações.** 3. ed. Rio de Janeiro: LTC, 2002.

EXAME. **Conheça os benefícios do self-service corporativo.** 2016. <https://exame.abril.com.br/negocios/dino/conheca-os-beneficios-do-self-service-corporativo-dino89084396131/> (Acessado em 12/11/2017)

FAYYAD, U.M. et al. **Advances in Knowledge Discovery and Data Mining.** California: AAAI Press, 1996b.

FERRARI, Sandra. **Proposta de metodologia para controle de qualidade em uma fábrica de software.** Tese de doutorado na Universidade Federal de Santa Catarina. Programa de Pós-Graduação de Engenharia de Produção, 2007.

GARTNER. **Gartner ITGlossary: Advanced Analytics.** <http://www.gartner.com/it-glossary/advanced-analytics/> (Acessado em 25/04/2017).

GARTNER. **Gartner ITGlossary: Big Data.** <https://www.gartner.com/it-glossary/big-data> (Acessado em 11/11/2017).

GARTNER. **Gartner ITGlossary: Business analytics.** <https://www.gartner.com/it-glossary/business-analytics/> (Acessado em 15/11/2017).

GARTNER. **Gartner ITGlossary: Business Intelligence.** <http://www.gartner.com/it-glossary/business-intelligence-bi/> (Acessado em 25/10/2017).

GARTNER. **Gartner ITGlossary: Self-Service Business Intelligence.** <https://www.gartner.com/it-glossary/self-service-business-intelligence> (Acessado em 12/11/2017)

GROSSMAN, R. L.; HORNICK, M.; MEYER G. **Emerging KDD Standards.** In: Communications of the ACM, 2002

GUERRA, Bruno. **Business Analytics, Business Intelligence e Data Visualization: você sabe a diferença?** <http://blog.in1.com.br/business-analytics-business-intelligence-e-data-visualization-voce-sabe-a-diferenca> , 2017 (Acessado em 15/11/2017).

HALPER, F.; KRISHNAN, K. **TDWI Big Data Maturity Model Guide: Interpreting your assessment score.** 2013.

HALPER, F. e STODDER, D. **TDWI Analytics Maturity Model Guide**. 2014.

HEKIMA. **O guia definitivo de big data para iniciantes**. (www.hekima.com)

HOWSON, Cindi e DUNCAN, Alan. **ITScore for BI and Analytics**. Gartner (G00314086), 2016.

HP. **The HP Business Intelligence Maturity Model: Describing the BI Journey**. 2009.

IDOINE, Carlie. **Why Business Analytics Projects Succeed: Voices From the Field**. Gartner (G00255328), 2015.

IDOINE, Carlie e TAPADINHAS, João. **Citizen Data Science Augments Data Discovery and Simplifies Data Science**. Gartner (G00314599), 2016

INOVAÇÃO. **Virtualização de dados**.

<http://www.inovacaogis.com.br/index.php/produtos/informatica-powercenter/virtualizacao-de-dados> (Acessado em 11/11/2017)

INMON, W.H. **What is a Data Warehouse?** Prisma, Volume 1, Número 1. 1995

INMON, W.H. **Building the Data Warehouse**. Wiley Publishing Inc. Fourth Edition. 2005

KIMBALL, Ralph. **The Data Warehouse Toolkit**. 2ª edição. John Wiley and Sons, 2002.

KING, D. **Numerical machine learning**. Georgia: Tech College of Computing, 2003

KONASANI, Venkat and KADRE, Shailendra. **Business Analytics Using SAS**. 2015.

LEEK, Jeff. **The Elements of Data Analytic Style**. 2015.

LEWIS, Jason. **Planning Guide for Data and Analytics**. Gartner (G00331851), 2017.

LEWIS, Jason. **Solution Path for Planning and Implementing a Data and Analytics Architecture**. Gartner (G00324344), 2017.

LIBERATORE, Mathew e LUO, Wenhong. **The Analytics Movement**. 2010.

LUHN, H. P. **A Business Intelligence System**. IBM Journal of Research and Development, 1958.

MATOS, David. **Data Lake, a fonte do Big Data**. <http://www.cienciaedados.com/data-lake-a-fonte-do-big-data/>, 2015. (acessado em 06/11/2017)

NASCIMENTO, Rodrigo. **O que é um dashboard?** <http://marketingpardados.com/analise-de-dados/o-que-e-dashboard-%F0%9F%93%8A/> , 2017 (acessado em 03/11/2017).

NEOINFINITO. **O que é Business Analytics?** <http://www.neoinfinito.com.br/site/o-que-e-business-analytics/> , 2015 (acessado em 06/04/2017).

NUNES, Breno. **Representações CMMI**. <http://tiinteligente.blogspot.com.br/2013/04/cmми-representacoes-introducao.html> , 2013 (acessado em 10/09/2017).

PAULK, M.C. **Surviving the Quagmire of Process Models, Integrated Models, and Standards. Surviving the quagmire of process models, integrated models, and standards**. Proceedings of the Annual Quality Congress, 2004.

PETERSON, Timothy. **Microsoft Sql Server 2000 dts**. 2001.

PRESSMAN, R. S. **Engenharia de Software**. 5ª edição, McGraw-Hill, 2002.

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

Pyramid Analytics. **BI Maturity Models**. <http://www.pyramidanalytics.com/pages/blog/2015/07/bi-maturity-models.aspx> (Acessado em 12/08/2017)

RAINER, R. K., & CEGIELSKI, C. G. **Introdução a sistemas de informação**. (3ª ed.). Elsevier, 2011.

RESENDE, Dirceu. **Analysis Services – Como criar seu primeiro cubo multidimensional no modelo estrela (Star schema)**. <https://www.dirceuresende.com/blog/analysis-services-como-criar-seu-primeiro-cubo-multidimensional-no-modelo-estrela-star-schema/> (Acessado em 10/07/2017).

SAMARANI, Paulo Roberto de Miranda. **Um modelo de implementação do Capability Maturity Model Integration nível 2**. Programa de Pós-Graduação em Computação da UFRGS, 2005.

SANTOS, M. R. **Aplicação de Business Intelligence para Análise de Indicadores das Redes de Referência no Estado do Paraná**. Universidade Estadual de Londrina, 2009.

SAS. **Big Data, o que é e por que é importante?** https://www.sas.com/pt_br/insights/big-data/what-is-big-data.html (Acessado em 11/11/2017)

SCHEIBLER, Douglas. **Self-Service BI: poder e governança nas mãos do usuário**. 2017. <https://itforum365.com.br/gestao/self-service-bi-poder-e-governanca-nas-maos-usuario> (Acessado em 12/11/2017)

SCHLEGEL, Kurt. **The Fundamentals of a BI and Analytics Program**. Gartner Business Intelligence & Information Management Summit 23-24 de junho 2015 | São Paulo, Brasil

SEI. **CMMI for Development**. Version 1.3 (CMMI-DEV, V1.3) CMU/SEI-2010-TR-033, ESC-TR-2010-033). Engineering Institute, Carnegie Mellon University, November 2010. <http://www.sei.cmu.edu>

SELVAGE, Mei Yang e THOO, Eric. **How to Drive Self-Service in Data Integration Strategies**. Gartner (G00343448), 2017.

SERRA, Laércio. **A essência do Business Intelligence**. 1ª Edição. Berkeley, 2002.

SILVEIRA, Debora Priscila. **O que é Data Science?**

<https://www.oficinadanet.com.br/post/16919-o-que-e-data-science> , 2016 (Acessado em 15/11/2017).

STODDER, David. **Improving Data Preparation for Business Analytics**. TDWI, 2016.

STODDER, David. **Rethinking Enterprise BI to Fit a Self-Service World**. TDWI.

TAPADINHAS, João. **How to Architect the BI and Analytics Platform**. Gartner (G00265003), 2016.

TDWI. **Business Intelligence**. <http://tdwi.org/portals/business-intelligence.aspx> , 2013 (Acessado em 26/10/2017).

TECHNOPEDIA. **Advanced Analytics**.

<https://www.techopedia.com/definition/32370/advanced-analytics> (Acessado em 15/11/2017).

TURBAN, E., e VOLONIMO, L. **Business Intelligence e Suporte a Decisão**. 2013.

VERCELLIS, C. **Business Intelligence: Data Mining and Optimization for Decision Making**, 2009.

VERT. **Por que a virtualização de dados é um elemento estratégico para seu negócio?**

<http://www.vert.com.br/blog-vert/por-que-a-virtualizacao-de-dados-e-um-elemento-estrategico-para-seu-negocio/> (Acessado em 11/11/2017)

WELLS, DAVE. **A Capability Model for Business Analytics: Part 1 – Dimensions of Capability**. Eckerson Group, 2016.

WELLS, DAVE. **A Capability Model for Business Analytics: Part 2 – Assessing Analytic Capabilities**. Eckerson Group, 2016.

WELLS, DAVE. **A Capability Model for Business Analytics: Part 3 – Using the Capability Assessment**. Eckerson Group, 2016.

WELLS, DAVE. **Beyond Data Visualization: The Power of Data Storytelling.** Eckerson Group, 2017.

WELLS, DAVE. **The Other Side of Data Visualization.** Eckerson Group, 2017.

Wikipedia. **Capability Maturity Model.**

https://pt.wikipedia.org/wiki/Capability_Maturity_Model (acessado em 27/03/2017)

ANEXO A - AVALIAÇÃO DO MODELO DE MATURIDADE DO TDWI

A avaliação do modelo de maturidade em analytics do TDWI é composto por 35 questões distribuídas em cinco categorias que compõem as dimensões do modelo conforme visto em 5.4.2.

Organização:

Esta seção foca a organização relacionada aos seus esforços de análise.

Liderança / cultura

1. Existe patrocínio das áreas de TI e de negócios para iniciativas analíticas na empresa?

- Não há patrocínio na empresa para atividades analíticas
- Há patrocínio apenas da área de TI
- Há patrocínio apenas da área de negócios
- Há patrocínio das áreas de TI e de negócios
- Há patrocínio das áreas de TI e de negócios e elas trabalham em conjunto
- Não sei

2. Nós somos capazes de expressar os potenciais benefícios de um projeto de ambiente analítico em linguagem comercial para que os executivos entendam.

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

Estratégia

3. Possuímos um processo de captação orçamentária bem estabelecido para ambiente analítico que seja conduzido tanto por negócios como por TI.

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

4. **Nós temos um planejamento para ambiente analítico acordado em toda a empresa e disciplina para alterá-lo se necessário.**
- Discordo totalmente
 - Discordo parcialmente
 - Não concordo nem discordo
 - Concordo parcialmente
 - Concordo totalmente
 - Não sei
5. **Nós tomamos ações com base em analíticos, ou seja, usamos análises como parte de um processo de negócios.**
- Discordo totalmente
 - Discordo parcialmente
 - Não concordo nem discordo
 - Concordo parcialmente
 - Concordo totalmente
 - Não sei
6. **Dados e uso de analíticos guiam o negócio da empresa.**
- Discordo totalmente
 - Discordo parcialmente
 - Não concordo nem discordo
 - Concordo parcialmente
 - Concordo totalmente
 - Não sei

Conhecimentos

7. **Treinamos usuários para executar análises avançadas.**
- Discordo totalmente
 - Discordo parcialmente
 - Não concordo nem discordo
 - Concordo parcialmente
 - Concordo totalmente
 - Não sei
8. **Há pessoas na organização com conhecimentos em análises avançadas para apoiar as necessidades do negócio, como cientistas de dados?**

- Não
- Sim, a nível departamental ou unidade de negócio
- Sim, em toda a empresa

9. Existem usuários com conhecimentos no espectro de capacidades analíticas e que utilizam analíticos na organização.

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

Infraestrutura:

Esta seção foca na infraestrutura para ambiente analítico.

Desenvolvimento

1. Os projetos analíticos são conduzidos pela liderança empresarial e realizam entregas incrementais ao invés de uma única entrega no final de todo o processo de desenvolvimento?

- Não
- Estamos caminhando para isso
- Sim, para alguns projetos
- Absolutamente

2. Temos os conhecimentos adequados para abordar as tecnologias existentes na infraestrutura em nossos esforços de análise.

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

3. Nossos cientistas e analistas de dados trabalham em estreita colaboração com nossas equipes de BI e administração de dados garantindo que as cargas de trabalho analíticas tenham a infraestrutura de dados necessária.

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo

- Concordo parcialmente
- Concordo totalmente
- Não sei

Tecnologias

4. Quais tecnologias você utiliza, atualmente, em seus trabalhos analíticos?

- Usamos arquivos de texto e planilhas
- Temos um data warehouse ou data marts
- Usamos uma plataforma analítica
- Usamos diferentes tecnologias como data warehouse e Hadoop, porém estas estão isoladas
- Utilizamos uma variedade de abordagens que formam um ecossistema analítico
- Nenhuma

5. Você faz uso de tecnologias móveis para análise?

- Não, e não temos planos para tal
- Não, mas estamos pensando a respeito
- Sim, mas apenas em poucos usos
- Sim, para todas as necessidades

Arquitetura

6. Possuímos uma arquitetura de informações em toda a empresa para análise.

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

6. Criamos nossa arquitetura analítica para aproveitar os sistemas legados já implantados.

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente

- Concordo totalmente
- Não sei

7. Utilizamos algum serviço de nuvem pública em aplicações analíticas?

- Não, nunca usaremos serviços de nuvem pública
- Temos um serviço em nuvem privado
- Temos um modelo híbrido onde usamos serviços de nuvem públicos e privados
- Ainda não utilizamos nuvem para aplicações analíticas, mas caminhamos nesta direção

Gestão de Dados:

Esta seção foca em gestão de dados em suporte à ambiente analítico.

Tipos de dados

1. Que tipos de dados você atualmente coleta e gerencia como parte de seus esforços analíticos?

- Nenhum
- Dados estruturados somente de sistemas internos
- Dados estruturados e demográficos
- Dados não estruturados juntamente com nossos dados estruturados Nós coletamos e gerenciamos dados de várias fontes, tanto internas como externas à empresa. Isso inclui dados não estruturados, dados geoespaciais e muito mais

2. Quantos o volume de dados que você está analisando atualmente?

- Megabytes
- Terabytes
- Petabytes
- Não sei

3. Fazemos uso de múltiplas fontes de dados em uma única análise?

- Não
- Sim, com dados estruturados
- Sim, com dados estruturados e uma ou duas fontes externas, como dados demográficos
- Sim, com diferentes tipos de dados, incluindo dados não estruturados e outros dados não tradicionais, mas sem integrá-los
- Sim, com diferentes tipos de dados, incluindo dados não estruturados e outros dados não tradicionais integrados

4. **Os funcionários da minha empresa podem facilmente encontrar os dados que precisam quando precisam.**

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

Integração

5. **Como os dados são integrado na empresa?**

- Ainda não temos uma boa maneira de integrá-los
- Nós temos metadados que usamos para ajudar na integração de dados
- Utilizamos para de integração de dados um produto de virtualização de dados ou data warehouse lógico
- Nós empregamos rotinas ETL para centralizar o máximo de dados possível em um data Warehouse

6. **Nossos dados são armazenados em silos.**

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

7. **Se usuários quiserem acesso a self-service para compartilhar recursos de dados eles conseguem.**

- Não, os usuários estão limitados aos seus próprios silos de dados e não têm acesso a recursos de dados compartilhados
- Sim, aplicamos técnicas para permitir o acesso via self-service a dados integrados de múltiplas fontes
- Usuários com determinados perfis possuem acesso self-service somente a um data warehouse centralizado
- Nenhum acesso de self-service é permitido, somente através da área de TI é possível acessar alguns recursos de dados compartilhados.

Qualidade de dados

8. **Temos um processo para tratar de qualidade de dados**

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

Analíticos:

Esta seção foca no ambiente analítico da sua empresa.

Escopo

1. Que tipos de técnicas analíticas sua empresa usa para analisar dados?

- Nenhuma ainda
- Ferramentas de BI/OLAP, dashboards, relatórios e a consultas ad hoc
- Todas acima assim como de visualização de dados
- Todas acima assim como de análises preditivas
- Todas acima assim como de mineração de dados ou de técnicas estatísticas
- Utilizamos todas as técnicas descritas acima, bem como técnicas como análise de redes sociais, análise geoespacial, análise de texto, análise de rede ou mineração de fluxo

2. As análises são muitas vezes automatizadas como parte dos processos de negócios da minha empresa.

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

Cultura

3. Nós temos uma boa idéia de quais questões de negócios estamos tentando resolver com dados em minha empresa.

- Ainda não
- Estamos trabalhando nisso
- Sim, e nós estamos tentando torná-las parte da nossa cultura
- Sim, as questões são orientadas para os negócios

4. **Temos tolerância para falhas precoces com novas tecnologias analíticas na minha empresa.**
- Discordo totalmente
 - Discordo parcialmente
 - Não concordo nem discordo
 - Concordo parcialmente
 - Concordo totalmente
 - É muito cedo para saber
5. **O ambiente analítico é visto como um diferencial competitivo na minha empresa.**
- Discordo totalmente
 - Discordo parcialmente
 - Não concordo nem discordo
 - Concordo parcialmente
 - Concordo totalmente
 - Não sei
6. **Podemos implementar análises para suportar métricas de gerenciamento de desempenho para que os usuários possam analisar mais profundamente os dados associados às métricas pelas quais eles são responsáveis.**
- Discordo totalmente
 - Discordo parcialmente
 - Não concordo nem discordo
 - Concordo parcialmente
 - Concordo totalmente
 - Não sei

Métodos de entrega

7. **Os usuários de negócios que desenvolvem visualizações de dados ou análises avançadas são capazes de trabalhar com a TI para garantir o orçamento e concentrar os recursos na implantação de analíticos para outros departamentos internos.**
- Discordo totalmente
 - Discordo parcialmente
 - Não concordo nem discordo
 - Concordo parcialmente

- Concordo totalmente
- Não sei

8. Como os analíticos são entregues em sua empresa?

- Nós não os entregamos ainda
- Estatísticos ou cientistas de dados os preparam e os entregam
- A área de TI ou estatísticos constroem um painel ou outro método de entrega interativo
- Analistas de negócios criam e distribuem
- Eles são operacionalizados como parte de um processo de negócio
- Usamos uma variedade de métodos de distribuição, incluindo a operacionalização e embutindo analíticos em um processo comercial

9. Qual a porcentagem de analistas de negócios, analistas de dados, cientistas de dados e usuários de sua organização que possuem ferramentas e conhecimento para analisar dados em um ambiente self-service sem um envolvimento da área de TI?

- Nenhum
- 1-20%
- 21-50%
- 51-75%
- 76-100%

Governança:

Esta seção foca em governança relacionada a dados e analíticos.

Políticas

1. Políticas de gestão de dados, de propriedade e responsabilidade da informação estão definidas e documentadas na minha empresa?

- Não
- Não tivemos tempo para isso, mas sabemos que precisamos fazer
- Estamos trabalhando nisso em nível de unidade de negócio
- Sim, em nível de unidade de negócio
- Sim, em nível corporativo

Estrutura

2. Temos uma equipe de governança analítica com representantes de toda a empresa, incluindo usuários chaves de áreas de negócios. As funções e responsabilidades estão claramente definidas.

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente

Conformidade

3. Estamos monitorando a adesão às nossas políticas analíticas?

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

Administração

4. O papel formalizado do administrador analítico está em curso com os papéis e as responsabilidades claramente identificados?

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

5. Definições de dados e metadados estão claramente estabelecidos.

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

Segurança

6. Políticas de segurança estão em vigor e aplicadas para todas as formas de dados na empresa.

- Discordo totalmente
- Discordo parcialmente
- Não concordo nem discordo
- Concordo parcialmente
- Concordo totalmente
- Não sei

Demográficos:

Esta seção foca em dados demográficos de pesquisa para ajudar a avaliar as pontuações.

1. Em que ramo de negócio sua empresa se enquadra? Selecione apenas um.

- Mercado financeiros
- Seguros
- Consultoria
- Software/Internet
- Telecomunicações
- Saúde
- Indústria, exceto de computadores
- Varejo / atacado / distribuição
- Governo
- Educação
- Farmacêutica
- Mídia / publicação / entretenimento
- Serviços
- Turismo
- Transporte/logística
- Produção de computadores
- Alimentação
- Vendas online
- Outros

2. Selecione a resposta que melhor descreve sua função em sua empresa.

- Gerente de TI
- Diretor de TI
- Executivo de TI
- Analista de negócios
- Cientista dados

- Gerente de negócios
- Diretor de negócios
- Executivo de negócios
- Outro

3. Qual o tamanho da sua empresa, por orçamento?

- Menor que R\$35M
- R\$35M-R\$175M
- R\$175M-R\$350M
- R\$350M-R\$1750M
- R\$1750M-R\$3.5B
- R\$3.5B-R\$17.5B
- R\$17.5B-R\$35B
- Maior que R\$35B
- Não sei

4. Onde a sua organização está localizada?

- África
- Ásia/Ilhas do Pacífico
- Austrália/Nova Zelândia
- Canadá
- América Latina
- Europa
- Oriente Médio
- Estados Unidos
- Outro