



**VII SINGEP**

Simposio Internacional de Gest3o de Projetos, Inova3o e Sustentabilidade  
International Symposium on Project Management, Innovation and Sustainability

ISSN: 2317-8302

## **TRANSFORMAÇÃO DIGITAL NA SAÚDE: O CASO DA IMPLANTAÇÃO DE UM DATA WAREHOUSE EM UMA REDE DE HOSPITAIS**

**HUGO CAVALCANTE DE OLIVEIRA**  
UNINOVE

**CRISTIANE DREBES PEDRON**  
UNINOVE – Universidade Nove de Julho

Agradeço a uninove pela oportunidade de cursar ao mestrado, agradeço também a minha orientadora pela construção em conjunto deste trabalho, e agradeço a minha família pela paciência nestes momentos difíceis.



## TRANSFORMAÇÃO DIGITAL NA SAÚDE: O CASO DA IMPLANTAÇÃO DE UM DATA WAREHOUSE EM UMA REDE DE HOSPITAIS

### Resumo

As organizações de saúde vivem a revolução e inovação tecnológicas do século XXI, era que apresenta tecnologias como: internet das coisas (*IoT*), interconectividade, inteligência artificial, inteligência preditiva e cognitiva, *big data*, implantação de ERP (*Enterprise Resource Planning*), entre outras. Com essa mudança tecnológica e a implantação de diversos sistemas, surge um grande volume de dados descentralizados e muitas vezes desestruturados. Considerando esse cenário, este relato técnico apresenta etapas para a criação e implantação de um *data warehouse* na saúde. A pesquisa segue uma abordagem qualitativa, tendo sido conduzido um estudo de caso. O relato contribui com uma discussão sobre as etapas para a criação e implantação de um *data warehouse* na área da saúde, na qual são destacadas as dificuldades por conta da complexidade específica deste segmento.

**Palavras-chave:** *Data warehouse*; *Analitics*; *Business Intelligence – BI*; Tecnologia na saúde.

### Abstract

The health organizations are living the revolution and technological innovations of the XXI century, which presents technologies such as Internet of Things (IoT), interconnectivity, artificial intelligence, predictive and cognitive intelligence, big data, Implementation of many ERP (Enterprise Resource Planning). This technological change and the deployment of many systems bring a large volume of decentralized and often unstructured data. Considering this scenario, this technical report presents steps for creation and implementation of a data warehouse in health area. The research follows a qualitative approach, and a case study was conducted. The report contributes with a discussion about the steps for the creation and implementation of a data warehouse in the health area, in which the difficulties due to the specific complexity of this segment are highlighted.

**Keywords:** *Data warehouse*; *Analitics*; *Business Intelligence – BI*; Health technology.



## 1 Introdução

Desde o início do século XXI, uma grande revolução tecnológica vem intrigando e desafiando as organizações da saúde a se movimentarem de forma mais estratégica no âmbito da tecnologia. A implantação de sistemas ERP (*Enterprise Resource Planning*) e a automatização de processos deixaram de ser diferenciais e se tornaram obrigatórios para que o setor da saúde não fique atrás de outros segmentos.

Um estudo realizado por Burke, Wang, Wan e Diana (2002) mostra que entre a década de 1990 e 2000 o uso da tecnologia da informação nos hospitais triplicou, demonstrando que dos 4000 hospitais avaliados no estudo, 400 hospitais informatizaram mais de 90% de suas operações, outros 1900 estão entre 60% e 70% de sua operação informatizada, enquanto o restante se encontra abaixo dos 50%. Apesar do constante crescimento, ainda existe muita dificuldade na aderência de tecnologia pelos profissionais da saúde, o que dificulta muito a implantação de tecnologia e avanço no setor da saúde (Burke et al., 2002).

Esta revolução não para, e o grande volume de dados gerados e espalhados nos diversos sistemas e tecnologias obrigam as organizações a qualificarem e tornarem esses dados acessíveis para a tomada de decisão (Favaretto, 2007).

Além da transformação digital, o aspecto humano deve ser considerado quanto ao impacto dessa transformação, e o envolvimento e a aderência de toda a organização é importante para o sucesso na implantação de tecnologia na área da saúde (Cardoso & Pedron, 2013).

Considerando tal cenário, este relato técnico apresenta uma situação problema quanto à implantação de um *data warehouse* na saúde, especificamente em uma rede de hospitais que está passando por uma grande transformação digital. Neste relato a organização objeto de estudo será chamada de rede *Hosp*.

A rede *Hosp* atuava em cada unidade através de um modelo de operacionalização totalmente independente e descentralizada da administração corporativa. Para resolver esse problema, a rede *Hosp* realizou um projeto de transformação tecnológica e de processos para que todas as unidades funcionassem de forma padronizada e integrada. Mesmo com todo esse movimento, era inviável assumir uma única base de dados e sistema que atendesse a todos os setores da organização, sendo necessária a implantação de diversos sistemas para que cobrissem toda a operação da rede.

Desta forma, surgiu a necessidade de estruturar uma base de dados única que agrupasse, organizasse e qualificasse os dados coletados das diversas fontes de dados descentralizadas. Para resolver esse problema, a abordagem escolhida foi a implantação de um *data warehouse* juntamente com suas ferramentas de apoio, como ETL (*Extract, Transform, Load*), *data marts*, OLAP (*Online Analytical Processing*), entre outras. Para Inmon (1997), o *data warehouse* é uma coleção de dados orientados por assunto, integrado, variável com o tempo e não volátil, que tem por objetivo dar suporte aos processos de tomada de decisões.

Esta decisão foi tomada com base nos benefícios apresentados pela implantação do *data warehouse*, por meio do qual, segundo Inmon, Terdeman e Imhoff (2001), seria possível acessar os dados mais facilmente e rapidamente, tornando-os mais seguros e exatos, além de fornecer estruturas para estudos estratégicos, como: estreitar as análises de produtos e serviços, descobrir novos mercados, dinamizar operações, realizar análises históricas, fazer previsões através de *mining*, entre outros.

Este relato técnico tem como objetivo apresentar as etapas para a criação e implantação de um *data warehouse* na saúde, que segundo Kimball e Ross (2002) é tratado do com as seguintes fases: bases de dados de origem, área de *data staging*, área de apresentação de dados, ferramentas de acesso de dados, além da conceituação de *data warehouse*, ETL,



data mart e OLAP, BI (*Business Intelligence*) e avaliação do estudo de caso aplicado a rede Hosp.

O relato está estruturado em cinco partes: introdução, referencial teórico, metodologia, estudo de caso e considerações finais.

## 2 Referencial Teórico

O referencial teórico está estruturado em dois tópicos, onde o primeiro trata da definição de um *data warehouse*, abordando conceitos e definições dos principais autores sobre *data warehouse*. No segundo tópico, de ferramentas e fases para a criação e implantação de um *data warehouse*, são avaliados especificamente os conceitos propostos por Kimball, que servirão como lente teórica para a análise do estudo de caso proposto nesta pesquisa.

### 2.1 Definição de um *data warehouse*

Dentre os bens mais preciosos e importantes de qualquer empresa estão os dados gerados durante o tempo em seu negócio (Kimball & Ross, 2002). As informações podem ser geradas das mais diversas fontes sistêmicas, processos e tecnologia, tais como sistemas de estoques, sistemas de vendas, sistemas de atendimento ao cliente, planilhas, mapeamento de processos, e outros (Kimball & Ross, 2002). Com essa massificação de dados e informações, é inevitável a falta de acessibilidade e qualidade dessas fontes.

Neste cenário complexo de tratamento de dados, as ferramentas de um *data warehouse* podem ajudar as empresas a estruturar, qualificar e centralizar os seus dados, tornando-os facilmente acessíveis e íntegros para a tomada de decisão no momento correto (Favaretto, 2007).

Com o advento da microinformática na década de 1970, as empresas iniciaram o processo de informatização nos mais variados setores e indústrias, gerando a implantação de diversos sistemas de informação e dados descentralizados. Ainda na década de 1970 surge o primeiro conceito de armazém de dados, porém com o foco somente no armazenamento centralizado dos mesmos. Na década de 1990 esse conceito de armazém de dados passa por uma grande transformação, e assim surge o *data warehouse* com o foco em ferramentas de processamento de grande volume de dados, qualificação e estruturação do banco de dados por assuntos (Inmon, 1997).

Para Kimball e Ross (2002), o *data warehouse* possibilita a coleta e o agrupamento de dados dos mais diversos sistemas transacionais de uma empresa. Os dados de um *data warehouse* não são voláteis, ou seja, essa base deve estar intacta com os dados de séries históricas coletadas dos sistemas transacionais, não permitindo qualquer alteração e assim garantindo sua integridade.

Por meio de uma visão de tomada de decisão é possível realizar uma comparação de uma base de dados tradicional e um *data warehouse*. A base de dados tradicional trabalha com dados e tabelas relacionais, onde contém dados voláteis que são alterados a todo o momento pelos sistemas ERPs e CRMs (*Customer Relationship Management*). O processamento do banco de dados concorre com as operações dos sistemas transacionais, podendo gerar falhas de performance na geração de relatórios e indicadores devido às diversas regras e vínculos entre as tabelas do banco, gerando um impacto direto tomada de decisão (Inmon, Terdeman & Imhoff, 2001).

No caso do *data warehouse* é utilizada uma modelagem dimensional das tabelas e organizados por assuntos, a base de dados é totalmente integrada, além de não volátil, o que garante a integridade dos dados, e desta forma não existe transações dos sistemas ERPs e



CRMs concorrendo com a geração de relatórios e indicadores, essa estrutura de dados torna a busca e disponibilização da informação muito mais rápida para o suporte a tomada de decisão (Turban, Shard, Aronson & King, 2009).

Conforme Barbieri (2001), é possível observar o *data warehouse* de acordo com um conjunto de características trazido pelos autores Inmon e Kimball, que estão apresentadas na figura 1.

Características	RALPH KIMBALL	WILLIAM H. INMON
Abordagem	Bottom-up / Star Schema	Top-Down
Nível de atuação	Tático	Estratégico
Requisitos de integração de dados	Requisitos de negócios individuais	Integração em toda a empresa
A estrutura dos dados	KPI, medidas de desempenho de negócios, scorecards	Dados que atendem a necessidades múltiplas do negócio
Conjunto de habilidades	Pequena equipe de generalistas	Maior equipe de especialistas
Restrição de tempo	Necessidades urgentes para o primeiro data warehouse	É permitido que mais tempo atenda às necessidades dos negócios.
Complexidade ETL	Modelo requer um processo ETL único que carrega o modelo de dados final usado para relatórios	O modelo requer 2 processos de ETL: - carregamento de sistemas de origem - construção de relatórios de data marts
Desempenho e performance	A arquitetura suporta o carregamento de várias áreas de assunto em paralelo (dimensões primeiro, depois fatos). A arquitetura de 2 camadas requer uma camada única de ETL que ofereça processamento mais rápido	A arquitetura suporta o carregamento de várias fontes de dados em paralelo. Dentro de cada tabela de origem, os carregamentos precisam ser sequenciados com base nas dependências.
Modelagem	A complexidade da modelagem varia de acordo com a indústria / área de assunto. Metodologia bem estabelecida com diretrizes e estruturas de modelagem para cada tipo de área de assunto por setor.	A modelagem não é complexa, pois a estrutura essencialmente copia o sistema de origem com algumas desnormalizações e a adição de registros de data e hora para cada chave primária para rastrear alterações. Requer modelagem adicional para datamarts dimensionais.

**Figura 1.** Principais autores sobre *data warehouse*.

Fonte: Adaptado de Barbieri (2001) e Ahmed, Islam e Karim (2017).

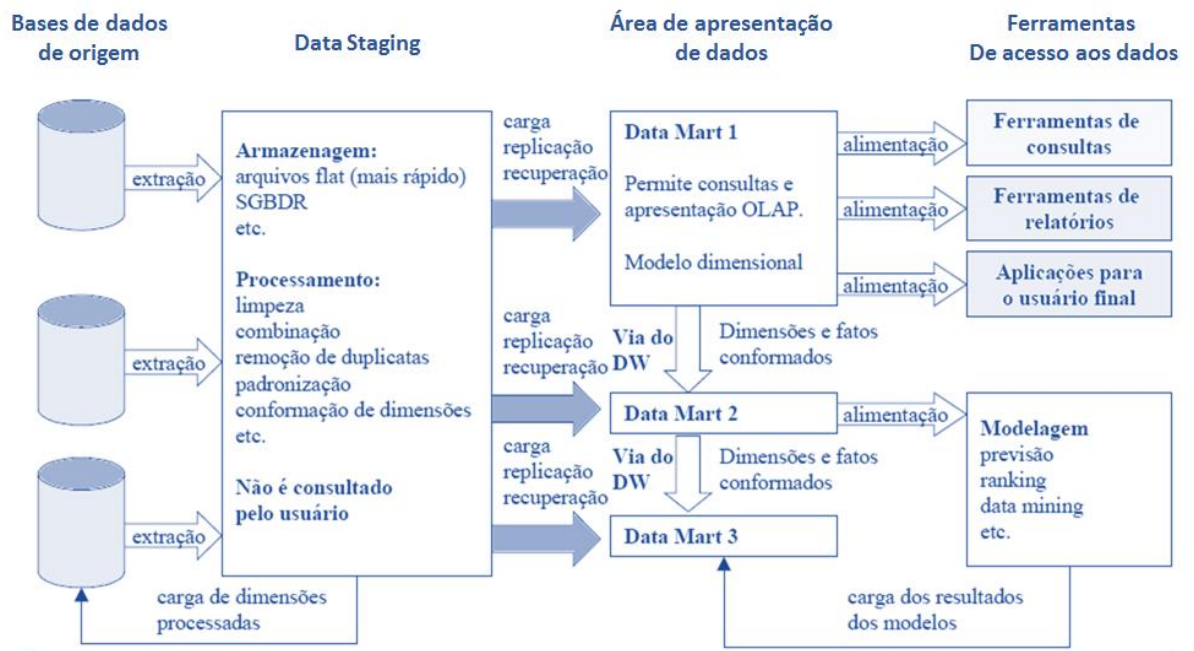
Com base no modelo de *data warehouse* proposto por Kimball e Ross (2002) e denominado *star schema*, que trabalha de forma *bottom-up* e propõe a criação do *data warehouse* de forma evolutiva, iniciando o mapeamento dos dados pelos relatórios e indicações primários de cada setor, criando as dimensões e complementando as ETLs conforme novos mapeamentos. O ponto positivo é que as entregas iniciais são realizadas de forma mais rápida, não sendo necessária a espera da finalização do *data warehouse*. Em contrapartida, o risco de retrabalho é enorme, pois as cargas são revistas a todo momento.

No modelo *top-down* proposto por Inmon os dados são mapeados por assunto e estruturados a partir das principais entidades, como por exemplo, clientes e fornecedores. As cargas são construídas através do mapeamento de todos os dados que compõem essa entidade. Neste modelo as cargas são mais demoradas, pois é necessário o mapeamento e *input* de toda a estrutura que compõe aquela entidade, porém, após construído, o nível de retrabalho é muito menor, pois inicialmente já foi realizado o mapeamento da estrutura de forma completa (Ahmed, Islam, & Karim, 2017).



## 2.2 Ferramentas e fases para a criação e implantação de um data warehouse

Partindo da definição básica do *data warehouse*, existem diversos componentes para sua construção, como bases de dados de origem, área de *data staging*, área de apresentação de dados e ferramentas de acesso de dados (Kimball & Ross, 2002).



**Figura 2.** Componentes para a construção do *data warehouse*.  
Fonte: Kimball e Ross (2002, p. 09).

Segundo Kimball e Ross (2002), a base de dados de origem é composta pelos dados operacionais e transacionais da empresa, sendo que o acesso a esta base é realizado por ferramentas de *mining* que possuem mecanismos de consultas específicos que não oneram o desempenho das bases transacionais. A base de dados de origem possui o papel de fornecer os dados para as cargas realizadas no *data warehouse*.

A área de *data staging* é uma das fases mais importantes para a criação e implantação do *data warehouse*, pois nesta fase está contido todo o processo de busca, tratamento e qualificação dos dados que serão imputados no *data warehouse* (Kimball & Ross, 2002). O *data staging* é basicamente responsável por toda comunicação com as bases de dados transacionais, sendo esse processo realizado através das ferramentas de ETL. Essa rotina é responsável por criar as tabelas de fatos e dimensões que estruturam e geram valor para a composição do *data warehouse*, essa área é totalmente técnica e não fica disponível para a visualização e uso dos usuários finais (Turban et al., 2009).

Ainda dentro da fase de *data staging* existe o processo de modelagem de dados, esse processo é denominado como criação das dimensões, e neste momento os dados são segmentados em cubos por assunto, e com esse processo são gerados os *data marts* (Favaretto, 2007). Os *data marts* possuem o papel de agrupar os dados por assunto, tornando-os mais confiáveis e acessíveis para os usuários que consomem as informações (Favaretto, 2007).

A fase seguinte é a preparação para a apresentação dos dados, na qual é realizado o refinamento dos cubos e preparação para trabalhar com as ferramentas OLAP (Kimball & Ross, 2002).



As ferramentas OLAPs permitem que os usuários transitem entre os *data marts* e gerem diversas visões dos dados, transformando-os em informações e conhecimento para a tomada de decisão. As ferramentas possuem recursos como *drill up* ou *drill down* que permitem ao usuário partir de uma visão macro ou gerencial, que contém pouca informação, e chegar na granularidade operacional para identificar qualquer *GAP* (Santos, 2002).

Após toda estrutura criada, restam apenas as ferramentas de acesso e visualização de dados, que buscam a exploração das mais diversas visões criadas nos *data marts*, e geram todo o conhecimento para a tomada de decisões nas organizações.

### 3 Metodologia

Para a realização do estudo proposto, a abordagem metodológica utilizada é qualitativa e exploratória. Segundo Creswell (2010), a pesquisa qualitativa é uma abordagem interpretativa, em que os pesquisadores analisam o que enxergam, ouvem e entendem as interpretações e não podem separar de suas histórias, contextos e entendimentos. Através da abordagem qualitativa é possível realizar uma investigação científica mais aberta, explorando o objeto de forma mais subjetiva (Flick, 2009).

O método utilizado foi o estudo de caso, sendo que os dados foram coletados por meio de entrevistas semiestruturadas, análise de documentos e observação participante. Segundo Yin (2015), o estudo de caso é utilizado para a investigação de um fenômeno contemporâneo em seu contexto real, ajudando a avaliar as relações entre o fenômeno e o contexto do projeto, ou a responder o porquê da utilização de tal ferramenta ou método. O estudo de caso pode ser conduzido através das fases de plano, desenho, preparação, coleta, análise e compartilhamento (Yin, 2015).

O estudo de caso é um método que busca estudar o fenômeno em profundidade (Creswell, 2010). Para tanto foram realizadas entrevistas semiestruturadas seguindo um roteiro anexado no Apêndice. O perfil dos entrevistados está apresentado na Figura 3. Finalmente, os dados foram coletados segundo uma observação participante, na qual um dos pesquisadores esteve acompanhando o projeto e teve a oportunidade de coletar dados ao longo de todo o processo. Estas diferentes formas de coleta de dados permitiram o que Yin (2010) chama de triangulação.

<u>Cargo</u>	<u>Abreviatura</u>	<u>Setor</u>
Analista TI	ATI	TI
Especialista <i>Business Intelligence</i>	EBI	Núcleo de informação
Analista de processos	ADP	Qualidade e processos
Especialista Corporativa de Negócios	ECN	Diretoria de operações
Observador: Gerente corporativo de TI	GTI	TI

**Figura 3.** Participantes das entrevistas semiestruturadas.

Os dados foram analisados seguindo as técnicas de análise de conteúdo propostas por Bardin (2011), que são apresentadas através das seguintes etapas: pré-análise, exploração do material, tratamento dos resultados obtidos e interpretação.

### 4 Estudo de Caso

Esta seção está estruturada em 2 grandes tópicos. O primeiro é a caracterização do estudo de caso da rede hosp, onde é apresentado a dimensão da rede hosp. O segundo capítulo apresentada um pouco sobre a complexidade da transformação digital na rede hosp, e em seguida é apresentado o estudo de caso juntamente com a análise do material para a



triangulação, tais como documentos do projeto, principais recortes das entrevistas semiestruturadas e as colocações do observador participante.

#### **4.1 Caracterização do estudo de caso rede Hosp**

A rede *Hosp* é uma rede filantrópica que atua no segmento da saúde, educação e assistência social que possui 120 anos de existência, e está presente em sete estados do Brasil: São Paulo, Espírito Santo, Mato Grosso, Goiás, Minas Gerais, Santa Catarina e Rio de Janeiro. Possui 24 unidades próprias, dentre elas: 15 unidades na área da saúde, 7 unidades na educação, 2 unidades de assistência social e 18 unidades em parceria com o governo que são administradas pela OS – Organização Social da rede *Hosp*.

A rede *Hosp* conta com 12.5 mil colaboradores na saúde, 700 colaboradores na educação e 100 colaboradores na assistência social. Possui um faturamento anual de 1.8 bilhões de reais, os quais são 100% revertidos para ações sociais. Fornecendo atendimentos que abrangem desde a atenção primária, secundária e terciária, a rede *Hosp* atua na estratégia de saúde da família, unidade básica de saúde, ambulatório médico de especialidades pronto socorro, pronto atendimento, hospitais de média e alta complexidades, além de atuar nas etapas da educação infantil, creche e pré-escola, ensino fundamental, ensino médio, educação para jovens e adultos e cursos técnicos profissionalizantes.

Através dessa estrutura são atendidos mais de 630 mil pacientes por mês, 4800 alunos e 700 pacientes acolhidos em situação de vulnerabilidade social. 74% dos serviços são realizados em convênio com o SUS – Sistema Único de Saúde e representação de 22% de bolsas de estudo nos serviços na educação.

#### **4.2 Análise do problema**

Devido à complexidade e descentralização da estrutura administrativa e localização das unidades, as unidades da rede *Hosp* trabalhavam de forma independente e totalmente descolada das diretrizes da administração corporativa, assumindo processos individuais em cada unidade e sistemas e tecnologia totalmente descentralizados. Esta situação estava gerando problemas de falta de informação para a alta gestão que necessitava de uma visão integrada para a tomada de decisões estratégicas.

Para resolver esse problema, como primeiro passo a rede *Hosp* formalizou a estrutura de administração corporativa e, na sequência, criou um projeto de reestruturação de tecnologia e processos para todas as unidades funcionarem de forma padronizada e totalmente conectadas em sistemas e bancos de dados únicos. Mesmo com todo esse movimento de reestruturação e inovação da tecnologia, é inviável assumir um único sistema de informação que atenda a todos os setores e segmentos da organização, sendo necessária a implantação de diversos sistemas para que cubram toda a operação da rede.

Diante dessa estrutura e complexidade, os processos foram organizados e diversos sistemas foram padronizados e implantados para que trabalhassem de forma única em todas as unidades, reduzindo boa parte da descentralização de processos e tecnologia. Mas ainda havia a necessidade, por parte da administração corporativa, de se obter informação centralizada de todos os sistemas, para que fosse possível a tomada de decisão de forma mais rápida e precisa.

#### **4.3 Intervenção e mecanismos adotados**

O projeto de implantação de um *data warehouse* na rede *Hosp* seguiu a proposta de Kimball e Ross (2002) que apresenta a implantação do *data warehouse* através de quatro





etapas: bases de dados de origem, área de *data staging*, área de apresentação de dados e ferramentas de acesso de dados.

A fim de segmentar o escopo do projeto em ondas, foi escolhida a implantação do *data warehouse* no segmento da saúde e como piloto o maior hospital da rede *Hosp*.

Para o início do projeto foi criado um grupo de condução que tem como objetivo a avaliação e padronização de todos os dados que serão disponibilizados em relatórios e indicadores. Esse grupo é composto pelas áreas de tecnologia da informação que fornece toda a estrutura técnica para o projeto, um setor de gestão da informação que cuida da padronização dos dados, o setor de processos e qualidade que é responsável por padronizar a forma de trabalho e atender os padrões de certificação vigente, e por último, por se tratar da área da saúde, foi necessária a inclusão do setor de práticas assistenciais que é responsável por garantir o cumprimento das normas técnicas.

#### **4.3.1 Bases de dados de origem**

Avaliando os documentos do projeto, identificou-se que a primeira etapa seria o levantamento da estrutura atual de disponibilização de relatórios e indicadores da rede *Hosp*. Através do levantamento foi possível identificar a grande dificuldade dos setores em gerar a informação para a tomada de decisão, pois os sistemas não eram únicos, e para a mesma informação gerada eram identificados uma particularidade e um padrão próprio de acordo com o sistema ou setor. Desta forma, dados que deveriam seguir uma visão única, como por exemplo, o reconhecimento de receita, possuíam valores diferentes, pois cada sistema ou setor calculava com suas convicções. Estas dificuldades também foram evidenciadas na entrevista com o Especialista *Business Intelligence* (EBI): “Principais problemas encontrados: áreas sem indicadores mapeados, indisponibilidade de agenda, falta de interesse ou descaso na área, (...), falta de conhecimento para tratar indicadores”.

De acordo com os documentos do projeto, foram avaliados mais de 90 setores e levantados mais de 400 indicadores e relatórios que compõem a lista de informações para apoiar a tomada de decisão da rede *Hosp*. Dado a observação participante, nesta fase de levantamento e análise da necessidade de cada setor, foi possível identificar a dificuldade dos setores em discutir sobre os indicadores de apoio à decisão. Muitos participantes possuíam conhecimentos limitados e básicos quanto à utilização das informações para a tomada de decisão: para os analistas que levantavam as informações, em diversos momentos houve dificuldades para a conclusão do levantamento em determinados setores. Estas informações também são enfatizadas na fala do Analista de Processos (ADP) entrevistado: “As maiores dificuldades foram doutrinar os colaboradores dentro do método estipulado e receber as fichas dos indicadores no prazo adequado para realizar as análises”.

Além do limitador do conhecimento, houve alguns setores que foram reativos na disponibilização dos dados, pois se sentiam confortáveis com as ferramentas e processos atuais e não gostariam de mudar, porém dada a observação, foi identificado que esses setores utilizavam ferramentas manuais e não confiáveis para a geração e consolidação dos dados, podendo ocasionar em uma informação incorreta para a tomada de decisão. Por último, também foi observado alguns setores avessos à inovação tecnológica, pois enxergavam esse processo apenas como mais uma ferramenta para utilizar e atrapalhar o trabalho principal que é o cuidado ao paciente, sendo esse grupo composto principalmente pela parte clínica e assistencial da rede *Hosp*.

Dado a documentação do projeto, para lidar com esse problema foi necessário um trabalho de conscientização desde a alta gestão até os participantes do projeto, onde foi discutido sobre a importância e envolvimento de todos para a geração de valor em cima da informação. Além da informação obtida na documentação do projeto, é possível confrontar



com a fala da Especialista Corporativa de Negócios (ECN), que diz o seguinte: “Durante as fases que eu participei, extremamente bem estruturado. Porém, precisa de um *sponsor* alto, porque existem definições que precisam de cadeira para serem tomadas com agilidade. Tivemos isso em nossas áreas, e percebi que isso facilitou bastante”.

#### 4.3.2 Data Staging

Antes de iniciar o processo de criação das cargas, foi realizada uma avaliação de todos os dados por parte do grupo do projeto, onde foi discutido o conceito de cada necessidade mapeada no levantamento. Conforme o exemplo apresentado anteriormente, alguns setores tratavam o reconhecimento de receita de formas diferentes, e dado a observação, o grupo do projeto ajudou a padronizar a forma em que os dados seriam disponibilizados para todos que consumiam a mesma informação, como exemplo, quem utiliza o valor de receita reconhecida, após essa padronização, vai enxergar o mesmo valor, independente do indicador ou relatório que está utilizando.

Dado a observação, foi possível identificar que a fase de padronização dos dados gerou muitos atritos entre os setores e equipes do projeto, pois algumas visões ou cálculos dos dados possuem respaldo de órgãos que regulamentam aquela forma de trabalho, porém outros dados são construídos de formas totalmente interpretativas, ocasionando diversas discussões sobre a forma correta de gerar e interpretar determinadas informações.

As documentações do projeto demonstram que os problemas foram resolvidos com rodadas de discussão entre os envolvidos e conciliação de conceitos para que todos estivessem confortáveis com a forma em que o dado seria disponibilizado. Dado a observação, em alguns momentos foi necessária a intervenção da alta administração para decidir a forma de apresentação de dados específicos. Este momento foi tido como um dos mais importantes do projeto, pois essas definições ditaram a forma na qual os dados seriam estruturados no *data warehouse*.

A etapa de padronização dos dados levou mais tempo que o esperado e comprometeu em partes o início da próxima atividade. O Analista de TI entrevistado (ATI) também comprova essa dificuldade em sua fala: “Devido à imaturidade no segmento da saúde para um modelo de dados voltado para BI, alta demanda de indicadores e curto prazo para elaboração das definições de dados, métricas, dimensões e origens, ocasionaram atrasos também nesta fase”. Os documentos do projeto demonstram que foi executado um paralelismo entre as etapas de padronização de dados e criação das cargas ETLs, adicionando uma complexidade maior ao projeto.

Conforme o trabalho de análise das bases e padronização dos dados era finalizado, os dados eram mapeados e fornecidos à equipe técnica para a criação das cargas de ETLs, que basicamente extraem os dados das bases de origem, transformam e carregam os mesmos para o *data warehouse*.

Dada a observação, o processo de carga foi considerado como relativamente tranquilo, pois se trata de um trabalho totalmente técnico do setor de tecnologia da informação. Os documentos do projeto apresentam uma observação importante, de que houve momentos de retrabalho devido ao paralelismo realizado com a padronização dos dados, pois em determinados momentos as cargas eram montadas de uma forma, e como a padronização ainda estava acontecendo, o mesmo dado que estava contido na carga montada voltava em evidência para a discussão do grupo, mudando o conceito daquele dado, gerando assim um retrabalho para a equipe técnica de recriar ou complementar a mesma carga.



### 4.3.3 Área de apresentação de dados

Após todo o mapeamento e criação das cargas, iniciou-se o processo de criação dos *data marts*. Avaliando a documentação do projeto pode se perceber que a etapa foi conduzida conforme o programado, pois toda a criação foi realizada seguindo a carga já mapeada e imputada no *data warehouse*. Os *data marts* montados seguiram o padrão de negócio classificado por assunto ou setor, e as dimensões foram criadas conforme o direcionamento de cada área.

Dado a observação, ficou claro um problema de disponibilidade da equipe técnica de tecnologia da informação devido a concorrência com outros projetos. Desta forma foi possível observar certa sobrecarga por parte da equipe para atender o escopo no prazo determinado

### 4.3.4 Ferramentas de acesso de dados

E como última fase, se inicia o momento de utilizar os dados de forma visual, gerando valor em cima da estrutura de dados. Através da coleta de dados do observador, ficou eminente a satisfação de várias áreas quanto aos resultados do projeto, tendo em vista que foram gerados *dashboards* que facilitaram o dia a dia dos setores. O ATI entrevistado também reforça esse ponto de conclusão: “Projeto fantástico que contribuiu com os dados para a gestão dos processos da empresa por indicadores”. Devido à complexidade do projeto, foram necessárias algumas alterações em cargas ETL e *data marts*, pois no momento de geração dos *dashboards*, foi identificada a necessidade de complemento de algumas informações, gerando um retrabalho.

## 5 Considerações Finais

Neste estudo foram apresentadas as etapas para a implantação de um *data warehouse* em uma rede de hospitais, sendo que a lente teórica utilizada para a avaliação do projeto foi baseada no estudo realizado por Kimball e Ross (2002), que defendem a implantação do *data warehouse* com quatro etapas: bases de dados de origem, área de *data staging*, área de apresentação de dados e ferramentas de acesso de dados.

Através do estudo de caso da rede *Hosp* foi possível identificar a complexidade e dificuldade do projeto devido o momento de transformação digital que a rede *Hosp* estava envolvida. As entrevistas demonstraram que a concorrência de recursos e a alta necessidade de disponibilização de uma estrutura de dados pressionaram o projeto a tomar decisões errôneas, mas que foram corrigidas a tempo para colocar o projeto em curso novamente.

Tratando as principais etapas do modelo proposto, as fases de bases de dados de origem e *data staging* geraram diversos conflitos, pois as áreas criavam seus indicadores de forma setorial e sem interagir umas com as outras, criando diversos padrões de indicadores e dados mestres individuais por área. Além da falta de interação entre indicadores comuns, essa fase evidenciou uma falta de conhecimento e maturidade dos setores para se trabalhar orientado a indicadores.

Nestas etapas ficam evidentes a necessidade de um aprimoramento conceitual para todos os participantes do projeto, visto que assim os levantamentos seriam muito mais produtivos e direcionados para os objetivos do projeto. A falta de padronização de dados comuns entre as áreas também prejudicou bastante o andamento desta etapa, pois foram necessárias diversas rodadas de discussão para a definição dos dados que seriam considerados para a criação das cargas ETL.



Nas fases mais técnicas o trabalho foi realizado sem muitas discussões e envolvimento com as áreas, e desta forma o andamento das atividades aconteceu de forma mais natural e evolutiva conforme o cronograma.

Apenas nas etapas finais de área de apresentação de dados e ferramentas de acesso de dados que ocorreram mais interações com os setores e geraram alguns retrabalhos devido à construção das cargas enquanto alguns indicadores ainda não estavam finalizados.

As recomendações após a conclusão deste estudo se estruturam em três principais pontos: primeiramente, recomenda-se um planejamento mais realista do projeto, avaliando fortemente a concorrência de demandas entre os participantes do projeto, como no caso da rede *Hosp*, que estava passando por uma transformação digital e existiam muitos projetos acontecendo em paralelo.

A segunda recomendação é a necessidade de preparar melhor os participantes do projeto quanto ao uso mais inteligente dos dados e indicadores de seus respectivos setores, pois por se tratar de uma instituição da saúde, os colaboradores focam mais em habilidades clínicas, assistenciais e atendimento ao paciente, deixando o aspecto administrativo e estratégico um pouco de lado. As dificuldades para discutir sobre dados, indicadores e indisponibilidades em alguns setores assistenciais e clínicos prejudicaram bastante o andamento do projeto.

A última recomendação é contar com o apoio e envolvimento maior da alta gestão, para que desta forma as atividades e entregas aconteçam mais no prazo, pois a alta gestão promove fortemente a participação e contribuição de todos.

Apesar das dificuldades apresentadas neste estudo, o modelo de implantação proposto por Kimball e Ross (2002) se mostrou funcional e contribuiu significativamente para a conclusão do projeto da rede *Hosp*.

Como sugestão para trabalhos futuros deixo a oportunidade de explorar melhor as fases de bases de dados de origem e *data staging*, que foram bem críticas. A avaliação de um modelo mais funcional para o levantamento dos dados faria toda a diferença técnica para a construção das ETL e daria muito mais agilidade para a condução das outras fases.

**Referências**

- Ahmed F., Islam S., Karim R. (2017). Comparison Between Inmon and Kilball Methodology for the Purpose of Designing, Constructing and Testing of a Commercial BIDW Project. *International Journal of Computer Graphics*, v. 8, n. 1, p. 11-20
- Barbieri, C. (2012). *BI2 – Business intelligence: modelagem e qualidade*. Editora Campus
- Bardin, L. (2011). *Análise de Conteúdo*. Editora Bookman. Editora ArtMed
- Burke, E. D., Wang, L. B., Wan, H. T. & Diana, L. M. (2002). Exploring Hospitals' Adoption of Information Technology. *Journal of Medical Systems*, v. 26, n. 4, Aug / 2002
- Cardoso, R. I. & Pedron, D. C. (2013). Users' commitment in information system implementation: the role of top-management commitment and organizational context. *Revista de Gestão em Sistemas de Saúde - RGSS*, São Paulo, v. 2, n. 1, p. 03-25, 2013.
- Creswell, W. J. (2010). *Projeto de pesquisa: Métodos Qualitativos, Quantitativos e Misto*. Editora Bookman. Editora Edições 70.
- Favaretto, F. (2007). Melhoria da qualidade da informação no controle da produção: estudo exploratório utilizando Data Warehouse. *Revista Produção*, v. 17, n. 2, p. 343-353
- Flick, U. (2009). *Introdução a pesquisa qualitativa*. 3ed. Editora Bookman. Editora ArtMed.
- Inmon, H. W. (1997). *Building the data warehouse*. Editora Wiley Computer, 3ed
- Inmon, H. W., Terdeman, H. R., & Imhoff C. (2001). *Data Warehousing: como transformar informações em oportunidades de negócios*. Editora Berkeley
- Kimball, R. & Ross, R. (2002). *The Data Warehouse Toolkit: Guia completo para a modelagem dimensional*. Editora Campus
- Santos, A. F. R. (2002). Metodologia e uso de técnicas de exploração e análise de dados na construção de um data warehouse. *Dissertação de mestrado*. Universidade Federal de Pernambuco.
- Turban, E., Shard, R., Aronson, E. J., & King, D. (2009). *Business Intelligence: Um enfoque gerencial para a inteligência do negócio*. Editora Bookman.
- Yin, K. R. (2015). *Estudo de caso: Planejamento e métodos*. Editora Bookman.

**Apêndice - Entrevistas semiestruturadas****Entrevista 1: Cargo Analista de TI (ATI)**

Pergunta 1: Na sua percepção, como foi a fase de planejamento do projeto?

Pergunta 2: Explane um pouco sobre a fase de levantamento das informações, como foi condução dessa fase? Quais as dificuldades?

Pergunta 3: Após o levantamento das informações, iniciou-se o processo de validação dos dados, como foi essa fase?

Pergunta 4: Pensando no âmbito técnico, como foi a construção das cargas e *data marts*?

Pergunta 5: As formas de apresentação ou visões que foram construídas, foram satisfatórias? Comente um pouco.

Pergunta 6: Em um âmbito geral, como você classificaria o projeto de *data warehouse* e BI?

**Entrevista 2: Cargo de Especialista *Business Intelligence* (EBI)**

Pergunta 1: Na sua percepção, como foi a fase de planejamento do projeto?

Pergunta 2: Explane um pouco sobre a fase de levantamento das informações, como foi condução dessa fase? Quais as dificuldades?

Pergunta 3: Após o levantamento das informações, iniciou-se o processo de validação dos dados, como foi essa fase?

Pergunta 4: Pensando no âmbito técnico, como foi a construção das cargas e *data marts*?

Pergunta 5: As formas de apresentação ou visões que foram construídas, foram satisfatórias? Comente um pouco.

Pergunta 6: Em um âmbito geral, como você classificaria o projeto de *data warehouse* e BI?

**Entrevista 3: Especialista Corporativa de Negócios (ECN)**

Pergunta 1: Na sua percepção, como foi a fase de planejamento do projeto?

Pergunta 2: Explane um pouco sobre a fase de levantamento das informações, como foi a condução dessa fase? Quais as dificuldades?

Pergunta 3: Após o levantamento das informações, iniciou-se o processo de validação dos dados, como foi essa fase?

Pergunta 4: Pensando no âmbito técnico, como foi a construção das cargas e *data marts*?

Pergunta 5: As formas de apresentação ou visões que foram construídas, foram satisfatórias? Comente um pouco.

Pergunta 6: Em um âmbito geral, como você classificaria o projeto de *data warehouse* e BI?

**Entrevista 4: Cargo de Analista de processos (ADP)**

Pergunta 1: Na sua percepção, como foi a fase de planejamento do projeto?

Pergunta 2: Explane um pouco sobre a fase de levantamento das informações, como foi condução dessa fase? Quais as dificuldades?

Pergunta 3: Após o levantamento das informações, iniciou-se o processo de validação dos dados, como foi essa fase?

Pergunta 4: Pensando no âmbito técnico, como foi a construção das cargas e *data marts*?

Pergunta 5: As formas de apresentação ou visões que foram construídas, foram satisfatórias? Comente um pouco.

Pergunta 6: Em um âmbito geral, como você classificaria o projeto de *data warehouse* e BI?