



EMPRESA BRASILEIRA DE SERVIÇOS HOSPITALARES

SCS Quadra 9, Edifício Parque Cidade Corporate, Torre C, 1° ao 3° andares - Bairro Asa Sul

Brasília-DF, CEP 70308-200

(61) 3255-8900 - <http://www.ebserh.gov.br>

Norma Operacional - SEI nº 7/2022/SGTI/DTI-EBSEH

Brasília, data da assinatura eletrônica.

NORMA OPERACIONAL PARA DESENVOLVIMENTO DE SOLUÇÕES DE ANÁLISE DE DADOS CORPORATIVOS DA REDE EBSEH

A **Diretoria de Tecnologia da Informação** da Empresa Brasileira de Serviços Hospitalares - Ebserh, no uso de suas atribuições institucionais que lhe foram conferidas pelo art. 103 do Regimento Interno da Administração Central da Ebserh, aprovado na 137ª Reunião Extraordinária do Conselho de Administração, realizada no dia 14 de junho de 2022, RESOLVE:

Divulgar a presente Norma Operacional que dispõe sobre as diretrizes acerca da utilização de soluções de tecnologia da informação para análise de dados corporativos no âmbito da Rede Ebserh, com intuito de definir padrões para seu uso, fomentar a cultura de compartilhamento de dados e disseminar o conhecimento gerado pela análise de dados.

CAPÍTULO I - DAS DISPOSIÇÕES PRELIMINARES

Seção I - Do Objeto e Âmbito de Aplicação

Art. 1º Esta Norma Operacional dispõe sobre procedimentos, regras, atribuições, responsabilidades e vedações quanto à gestão de dados, informação e conhecimento, bem como quanto à divulgação de informações corporativas relevantes no âmbito da Rede Ebserh.

Parágrafo único. A presente norma aplica-se aos usuários, empregados, servidores, colaboradores, terceirizados, estagiários, estudantes, docentes, gestores, ou quaisquer outros indivíduos que fazem uso de soluções de tecnologia da informação para desenvolvimento de soluções de análise de dados corporativos da Empresa.

Seção II - Dos Objetivos

Art. 2º Esta Norma Operacional tem como objetivo:

- I - estabelecer padrões para utilização da ferramenta de forma colaborativa e segura;
- II - fomentar a cultura de compartilhamento de dados;
- III - promover o gerenciamento de catálogo de soluções de análise de dados desenvolvidos na Rede Ebserh;
- IV - promover o monitoramento da disponibilidade das soluções de análise de dados desenvolvidos;
- V - disseminar o conhecimento gerado pela análise de dados;
- VI - promover qualificação e agregar valor aos dados negociais da empresa;

VII - organizar e sistematizar os dados oriundos das áreas de negócios;

VIII - promover autonomia no desenvolvimento e manutenção de soluções de análise de dados, em cada unidade hospitalar; e

IX - estabelecer procedimentos de distribuição das licenças para uso da ferramenta de análise de negócio e análise de dados.

Seção III - Das Definições

Art. 3º Para os fins desta norma operacional, considera-se:

I - área negocial: área da Ebserh responsável pela definição de processos de trabalho, requisitos, regras de negócio e níveis aplicáveis ao serviço de TI;

II - ativo de informação: qualquer dado, informação ou conhecimento a que esteja associado um valor para o negócio da organização;

III - *Business Intelligence*: conjunto de metodologias implementadas por meio de *softwares* que vão, em última análise, coletar informações e organizá-las em conhecimentos úteis para ajudar na tomada de decisão;

IV - contas de serviço: contas de acesso à rede corporativa de computadores necessários a um procedimento automático (aplicação, *script* etc.), sem qualquer intervenção humana no seu uso;

V - dados corporativos: todo e qualquer dado produzido, consumido, armazenado ou processado em meio digital, para execução das atividades da Empresa;

VI - ferramenta de análise de negócio e análise de dados: ferramenta adotada como padrão pela Ebserh, cuja principal finalidade é apoiar decisões utilizando o conceito de *Self-Service Business Intelligence*;

VII - segurança da informação: preservação da confidencialidade, da integridade e da disponibilidade da informação, conjunto de ações para proteção de um grupo de dados, salvaguardando seu valor, seja para um indivíduo específico, no âmbito de sua atividade laboral, seja para a organização;

VIII - *Self-Service Business Intelligence*: conceito de que o usuário não técnico possa criar e implantar suas próprias análises, apoiado na estruturação de uma arquitetura corporativa e de ferramentas que entreguem autonomia na concepção e modelagem para a implementação de relatórios e *dashboards*; e

IX - solução de análise de dados: ferramenta tecnológica de tratamento de dados que implemente processos de coleta, armazenamento, processamento, compartilhamento ou apresentação de dados corporativos, tais como painéis, relatórios ou planilhas eletrônicas.

CAPÍTULO II - DAS RESPONSABILIDADES

Seção I - Das Competências e Atribuições

Art.4º São competências das áreas negociais que desenvolvem ou utilizam soluções de análise de dados:

I - criar processo no SEI, observando o Anexo I desta norma, para os projetos desenvolvidos ou em desenvolvimento relacionados à análise de dados, indicando nome do projeto, finalidade da solução de análise de dados, fontes de dados, periodicidade de atualização, endereço de acesso e contato do responsável técnico;

II - encaminhar à Diretoria de Tecnologia da Informação (DTI) as informações utilizadas no desenvolvimento da ferramenta, incluindo as documentações, manuais e código-fonte, se for o caso;

III - definir o grau de confidencialidade dos dados exibidos na ferramenta desenvolvida pela unidade, sob sua responsabilidade;

IV - assegurar o controle de acesso aos dados, garantindo-o somente aos usuários autorizados;

V - esclarecer dúvidas técnicas ou negociais acerca da ferramenta desenvolvida ou das informações apresentadas, sempre que demandado;

VI - responsabilizar-se pelo controle da divulgação, compartilhamento, uso e proteção dos dados presentes na ferramenta desenvolvida; e

VII - dar publicidade a esta norma nas suas unidades, promovendo a realização de encontros e debates para esclarecimentos de sua relevância.

Parágrafo único. A manutenção da solução de análise de dados desenvolvida pela área negocial é de sua inteira responsabilidade.

Art. 5º São competências da Diretoria de Tecnologia da Informação:

I - definir diretrizes, políticas, normas e procedimentos operacionais acerca da utilização de soluções de análise de dados;

II - zelar, em conjunto com o Setor de Tecnologia da Informação e Saúde Digital das unidades hospitalares, pelo bom funcionamento da estrutura tecnológica da Rede Ebserh, providenciando mecanismos e ferramentas para conceder ou revogar acesso controlado aos dados;

III - monitorar o acesso aos dados e notificar o Comitê de Segurança da Informação (CSI), sempre que forem identificadas falhas de segurança;

IV - manter e disponibilizar informações sobre o catálogo de solução de análise de dados da Ebserh;

V - realizar cadastramento inicial das contas de serviços relacionadas à estrutura organizacional das unidades hospitalares da Rede Ebserh, bem como sua atualização;

VI - gerir o contrato da solução, suas licenças e a infraestrutura tecnológica que permita o seu uso; e

VII - intermediar a execução de desenvolvimento de projetos de análise de dados, realizada pela empresa contratada para este fim, prezando pela execução contratual e cumprimento dos requisitos legais.

Parágrafo único. A gestão contratual da solução, seu licenciamento e a infraestrutura tecnológica necessária serão realizadas pela DTI, com o apoio do Setor de Tecnologia da Informação e Saúde Digital dos hospitais da Rede Ebserh.

Art. 6º São atribuições dos usuários que utilizam soluções de análise de dados:

I - zelar pela confidencialidade dos dados e informações por eles acessados; e

II - observar os normativos da Rede Ebserh em relação ao uso de dados corporativos, em especial a Política de Segurança de Informação (PSI) e a Norma de Acesso e Disponibilização de Dados Digitais, que buscam preservar os ativos de informação da empresa mediante a implementação de controles de acesso e proteção aos dados digitais.

Seção II - Das Vedações

Art. 7º Para fins desta norma operacional é vedado:

I - utilizar a ferramenta para projetos alheios aos interesses da Ebserh, inclusive em projetos não relacionados às atividades laborais;

II - compartilhar senhas de acesso;

III - utilizar ferramenta de análise de negócio e análise de dados não padronizadas; e

IV - publicar solução de análise de dados ou relatórios com dados privativos às áreas negociais, sem a devida autorização do responsável pela área.

CAPÍTULO III - DAS REGRAS OPERACIONAIS

Seção I - Da Utilização da Solução

Art. 8º Os relatórios com visualização de dados corporativos devem ser desenvolvidos, preferencialmente, com a solução tecnológica de análise de negócio e análise de dados definida como padrão pela Ebserh.

Parágrafo único. A DTI divulgará, por meio de ofício circular, a adoção de ferramenta padrão de análise de negócio e análise de dados, bem como os procedimentos para uso e licenciamento quando necessário.

Seção II - Do Desenvolvimento de Solução de Análise de Dados

Art. 9 A construção de soluções de análise de dados poderá ser realizada por qualquer usuário da Rede Ebserh, relativamente às informações de sua área de atuação, com anuência do superior imediato, desde que observados a legislação, os normativos vigentes, em especial a Norma de Acesso e Disponibilização de Dados Digitais, e as boas práticas de desenvolvimento de soluções de análise de dados recomendadas pela DTI.

Parágrafo único. O uso das soluções de análise de dados, seu acesso público e divulgação deverão estar em conformidade com a Lei Geral de Proteção de Dados (LGPD).

Art. 10 A área negocial responsável pelo desenvolvimento de uma solução de análise de dados deverá:

I - estar apta a orientar a leitura, análise e suporte técnico da respectiva solução; e

II - responsabilizar-se pela extração e tratamento dos dados.

Art. 11 As soluções de análise de dados desenvolvidos na Rede Ebserh devem seguir:

I - preferencialmente, o padrão de arquitetura de BI definido no Anexo II desta norma; e

II - obrigatoriamente, o Manual de Identidade Visual da Ebserh, sob responsabilidade da Comunicação Social da Administração Central.

Seção III - Da Publicação da Solução de Análise de Dados

Art. 12. A publicação e disponibilização de soluções de análise de dados, resultante do uso de ferramenta padrão de análise de negócio e análise de dados, deverão observar o que segue:

I - realização de análise de riscos relacionada à publicação das soluções de análise de dados, de modo a evitar que informações sensíveis sejam indevidamente divulgadas em conformidade com as normas e políticas de segurança da informação da Ebserh;

II - aprovação por escrito do superior imediato no processo SEI respectivo;

III - liberação da funcionalidade de criação do endereço público pela área de TI após prévia submissão de solicitação por parte da área negocial.

Parágrafo único. A demanda de divulgação das soluções de análise de dados publicados poderá ser encaminhada à área de Comunicação Social da Administração Central e dos hospitais para análise da estratégia de comunicação e coordenação das atividades de divulgação.

Seção IV - Do Catálogo de Soluções de Análise de Dados

Art. 13 A DTI manterá o catálogo atualizado das soluções de análise de dados desenvolvidos pela Administração Central e pelos hospitais da Rede Ebserh, conforme formulário constante no Anexo III.

§1º As áreas responsáveis pelo desenvolvimento das soluções de análise de dados deverão encaminhar à DTI formulário preenchido:

I - para as soluções de análise de dados desenvolvidas, em até 120 (cento e vinte) dias após a publicação desta norma operacional; e

II - para novas soluções de análise de dados, em até 15 (quinze) dias após a publicação da respectiva solução.

§2º Somente será autorizada a publicação de soluções de análise de dados constantes no catálogo.

§3º O catálogo de soluções de análise de dados será divulgado na intranet da Administração Central.

CAPÍTULO IV - DAS DISPOSIÇÕES FINAIS

Art. 14. O esclarecimento de casos omissos e dúvidas técnicas relacionadas à aplicação desta norma serão realizados pela Diretoria de Tecnologia da Informação (DTI) da Ebserh.

Art. 15 Esta norma operacional será atualizada a cada dois anos ou a qualquer tempo, conforme necessidade.

Art. 16. O descumprimento das disposições desta norma poderá resultar em apuração de responsabilidade.

Art. 17 Esta norma operacional entra em vigor na data de sua publicação.

(assinado eletronicamente)

SIMONE HENRIQUETA COSSETIN SCHOLZE

Diretora de Tecnologia da Informação



Documento assinado eletronicamente por **Simone Henriqueta Cossetin Scholze, Diretor(a)**, em 22/12/2022, às 10:28, conforme horário oficial de Brasília, com fundamento no art. 6º, caput, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

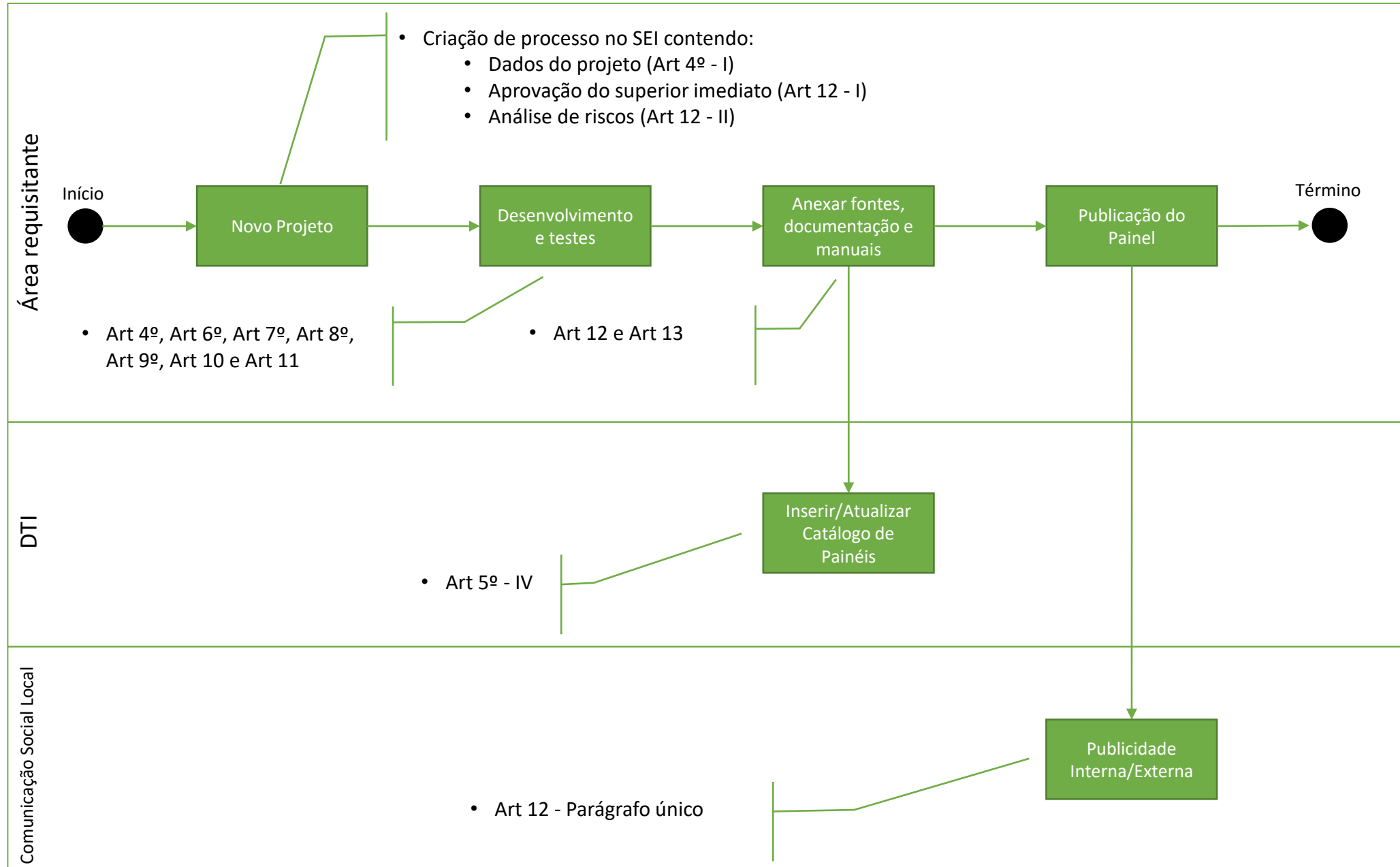


A autenticidade deste documento pode ser conferida no site https://sei.ebserh.gov.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **26501263** e o código CRC **F3F5053F**.

Referência: Processo nº 23477.016573/2022-78 SEI nº 26501263

Criado por [delana.silva](#), versão 5 por [simone.scholze](#) em 22/12/2022 10:28:23.

Macrofluxo Painéis BI





DOCUMENTO DE ARQUITETURA

DEFINIÇÃO DE PADRÕES DE PROJETO

Sumário

Histórico de Revisões	3
Introdução	4
Visão Geral	4
1. Definição e separação das camadas de um Projeto de BI	7
1.1. Aplicação de Estrutura de BI com soluções Open Source.....	7
1.2. Aplicação de Estrutura de BI com soluções Microsoft.....	8
2. Fontes de dados não estruturados (fontes de dados heterogêneas).....	9
3. Topologia de Rede x Disponibilidade de banco de dados e arquivos em FTP	11
4. Padronização de nomenclatura	13
5. Plataforma de versionamento de Código Fonte.....	14
Conclusão	15

Histórico de Revisões

Data	Versão	Descrição	Autor
16/08/2022	1.0	Documento Inicial	Cristiano Joaquim
23/08/2022	1.1	Inclusão de dados de Topologia de Rede e Plataforma de Sistema Operacional	Cristiano Joaquim
05/09/2022	1.2	Explicação do fluxo de ETL com as diversas fontes de dados aplicadas	Cristiano Joaquim
15/09/2022	1.3	Tratativa de fontes de dados heterogêneas	Cristiano Joaquim
26/09/2022	1.4	Inclusão do padrão de nomenclatura de objetos de banco de dados, Repositório de Fonte de Arquivos e respectivamente a proposta de controle de versão destes arquivos	Cristiano Joaquim

Introdução

Este documento tem o objetivo de apresentar dentro da Organização EBSERH, boas práticas e recomendações do time de Arquitetura de Business Intelligence e Soluções, para criação de Estruturas de Dados e o conceito de Data Driven, que envolvam as entregas de Produtos Analíticos, definindo padrões de aceitação de Projetos e recomendações para refactoring de projetos existentes.

Visão Geral

Os padrões apresentados neste documento, trazem conhecimento técnicas e recomendações utilizadas no mercado, especificamente para apresentar uma padronização em estrutura de construção de um ambiente de **Business Intelligence** (a partir de agora descrito apenas como BI), seu banco de dados de Data Warehouse (a partir de agora descrito como DW) e/ou Data Mart (a partir de agora descrito como DM), para compor um ambiente de Data Analytics. Ainda serão descritos neste documento todas as boas práticas envolvendo o uso de ferramentas, bem como fluxos que façam menção a entrega de soluções relacionadas ao ambiente de BI.

Obs.: Por definição, BI é um conjunto de ferramentas e técnicas para análise de dados.

Como ponto inicial, nossa abordagem de estrutura, Modelagem de Dados do DW, Técnicas e ferramentas, conceituam os padrões em materiais já utilizados e difundidos no mercado, tornando assim a **Proposta de Arquitetura** de fácil compreensão.

Com este conceito, abordaremos não somente as boas práticas de Desenvolvimento de um ambiente de BI (onde apresentaremos em um diagrama, ainda nesta Visão Geral), como também algumas regras principais do fluxo de desenvolvimento, descritas abaixo:

- A modelagem de dados para um ambiente de BI necessita seguir padrões, onde atualmente temos no mercado duas propostas de Modelagem comumente utilizadas, descritas
 - **Star Schema**, que é comumente descrita como uma tabela Fato centralizada e Tabelas de Dimensões que se relacionam com está Fato, trabalhando com uma proposta de ambiente desnormalizado (se diferenciando do padrão 3FN utilizado em uma modelagem de ambientes de Sistemas Produtivos); Como referência, o desenho da **Imagem 1** abaixo descreve este entendimento.

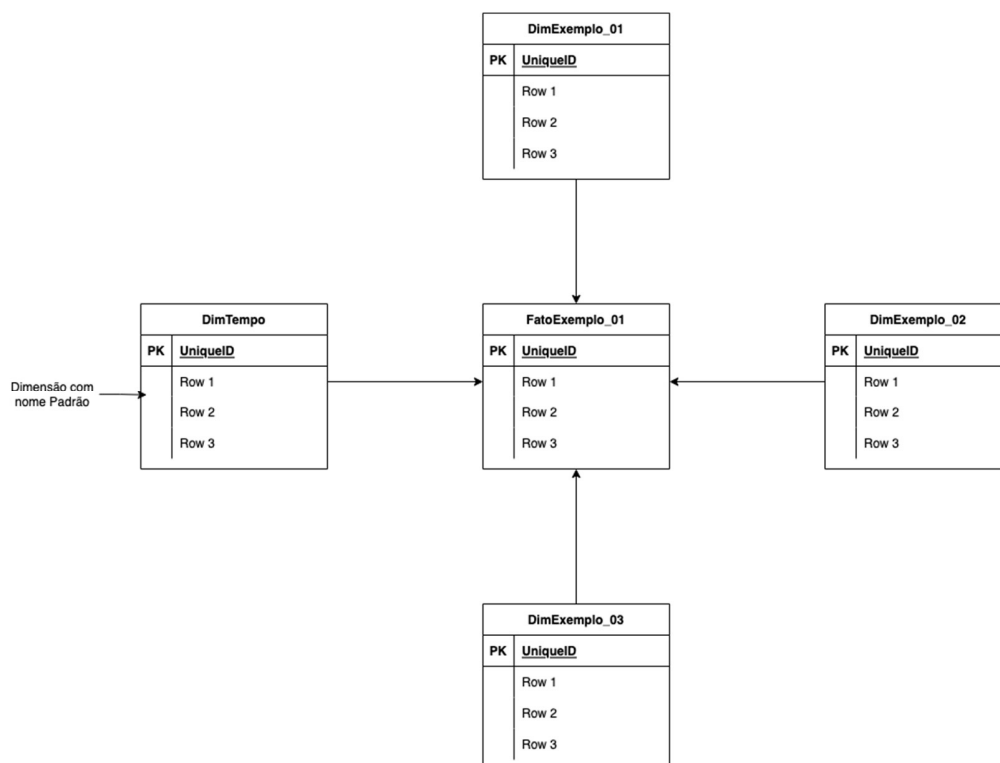


Imagem 1 – Exemplo de um Modelo de Dados Star Schema

- **Snowflake**, que trás a proposta de relacionamento entre Tabelas Fatos e Tabelas de Dimensões, porém possibilitando com que estas Dimensões possam se relacionar entre si, com a possibilidade de utilização do conceito da regra 3FN (também utilizada em modelagem de ambientes de Sistemas Produtivos). A **Imagem 2** abaixo apresenta o exemplo de uma Modelagem de Dados para BI no formato Snowflake, onde poderemos ver a possibilidade de relacionamento entre as Tabelas de Dimensões.

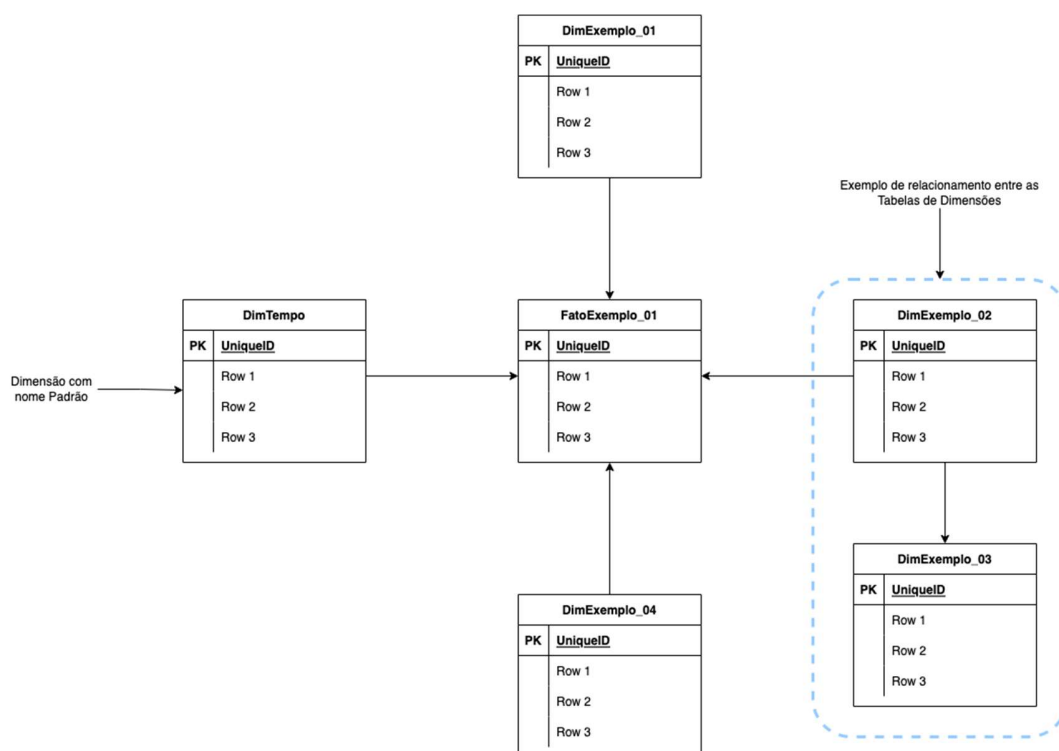


Imagem 2 – Exemplo de um Modelo de Dados Snowflake

- Todo processo de extração de dados é realizado por técnicas (e ferramentas) de ETL, que trás como significado de letras a nomenclatura Extract, Transform and Load (Extrair, Transformar e Carregar e que **neste momento referenciaremos somente como ETL**). Este é o processo vital de um projeto de BI, alinhado com a proposta do modelo de dados que será construído. A Imagem 3 abaixo apresenta uma visão de como um processo de ETL é definido.



Imagem 3 – Fluxo de um processo ETL

- Com a estrutura de Modelagem de Dados e processo de ETL definidos, será possível aplicar o conceito de Data Driven (Orientação a Dados). Para isso, podemos aplicar o uso da camada de Cubo Analítico como uma estrutura semântica. O uso deste ambiente nos favorece aumentar a disponibilidade do ambiente de DW, favorecendo que o processo de execuções de Cargas de ETL possa ocorrer em janelas pré-definidas.
- Ferramentas de Data Visualization, que nos favorecem visualizar os dados extraídos, tratados e carregados pelo nosso processo de ETL dentro de uma Modelagem de Dados para o DW e com a posterior carga do ambiente semântico do Cubo Analítico, em formatos de gráficos e Dashboards.

1. Definição e separação das camadas de um Projeto de BI

Para melhor entendimento da estrutura de BI, segue abaixo uma imagem que descreve a proposta geral de aplicação para cada camada.

Conforme poderá ser visualizado na Imagem 4 abaixo, cada camada está disponível em sua sequência de aplicação.

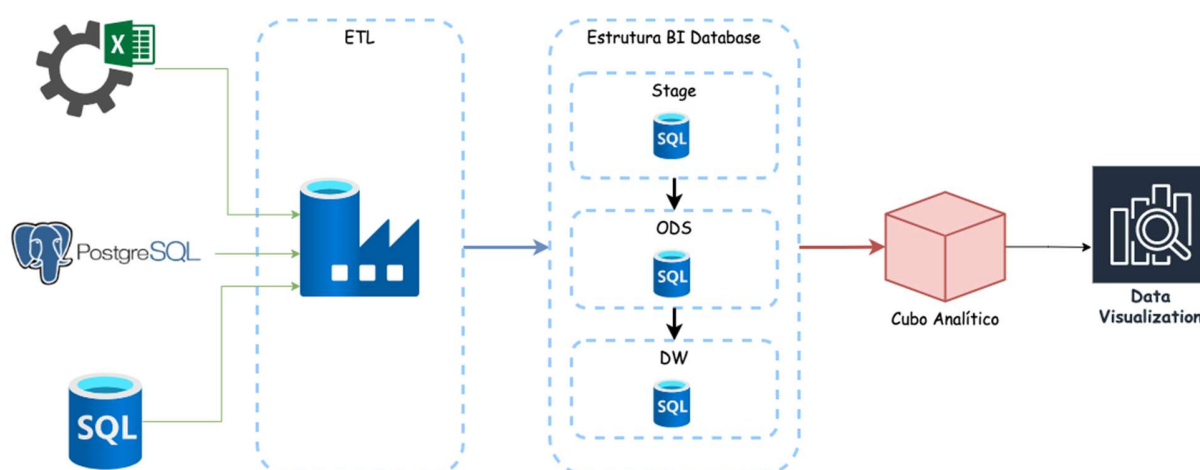


Imagem 4 – Aplicação de camadas em uma estrutura de BI

1.1. Aplicação de Estrutura de BI com soluções Open Source

Existem ferramentas Open Source que podem ser abordadas para a tratativa e aplicação de Projetos de BI, variando desde uma solução de banco de dados, passando por uma ferramenta de ETL.

Entretanto, alguns pontos podem não ser atendidos para que se obtenha uma estrutura orientada a Data Driven, no que diz respeito a aplicação de algumas camadas do Projeto de BI. Diante disto, segue o descrito de ferramentas com este formato Open Source:

- Database Engine:
 - PostgreSQL
 - MySql
 - MariaDB
- Ferramentas de ETL:
 - Pentaho Data Integration
 - Talend Open Studio
 - Apache NiFi

A Imagem 5 abaixo descreve uma estrutura em camadas, porém, sem a aplicação de uma Estrutura de Cubo Analítico especificamente definida. Como ponto de observação, vale ressaltar o

comentário da complexidade em se trabalhar com ferramentas **OLAP não tabulares**, aumentando assim a complexidade de uso com a linguagem MDX, proveniente de ferramentas neste perfil.

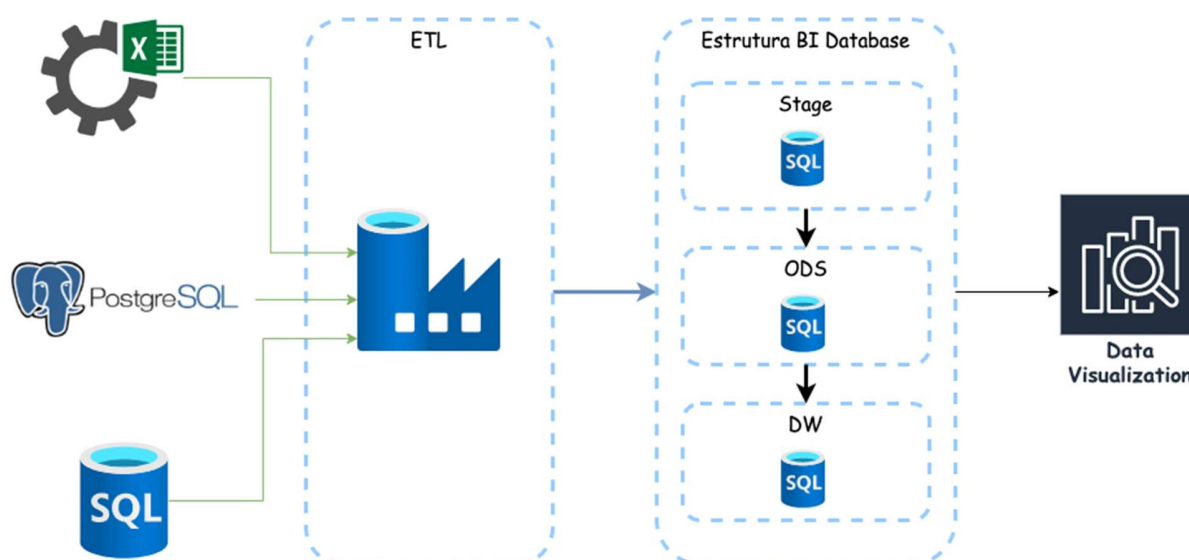


Imagem 5 – Solução de estrutura de BI com plataformas Open Source

De acordo com o que pode ser na Imagem 5 acima, temos no desenho como fontes de dados, tanto Origens heterogêneas que podem vir de arquivos estruturados e/ou não estruturados (como arquivos TXT, CSV neste exemplo).

Da mesma forma, também podemos ter diversos bancos de dados que estejam no formato SQL ANSI-92 e especificamente em bancos de dados PostgreSQL (que foi descrito como origem principal dentro de diversas fontes de Hospitais que serão as origens principais de extração de dados).

Todo o fluxo de ELT deverá ter mapeado esta estrutura de origem de dados, sendo possível realizar seu processo de extração com os devidos fluxos de Transformação aplicado em seu desenvolvimento.

1.2. Aplicação de Estrutura de BI com soluções Microsoft

Com a disponibilidade de uma versão do Microsoft SQL Server em edições Standard, Enterprise e/ou Business Intelligence, existe a possibilidade de se utilizar a suíte de BI da plataforma Microsoft, tendo como ferramentas específicas:

- Microsoft SQL Server Database Engine, solução de banco de dados On-Premise padrão da Microsoft;
- Microsoft SQL Server Integration Services, ferramenta padrão On-Premise de ETL;
- Microsoft SQL Server Analysis Services, que é a ferramenta de Cubo Tabular e Analítico para proposta de Data Driven;

- Microsoft Power BI, que permite a conexão com diversas fontes de dados, dentre as quais fazem parte da proposta de arquitetura desenhada na **Imagem 6** disponível abaixo.

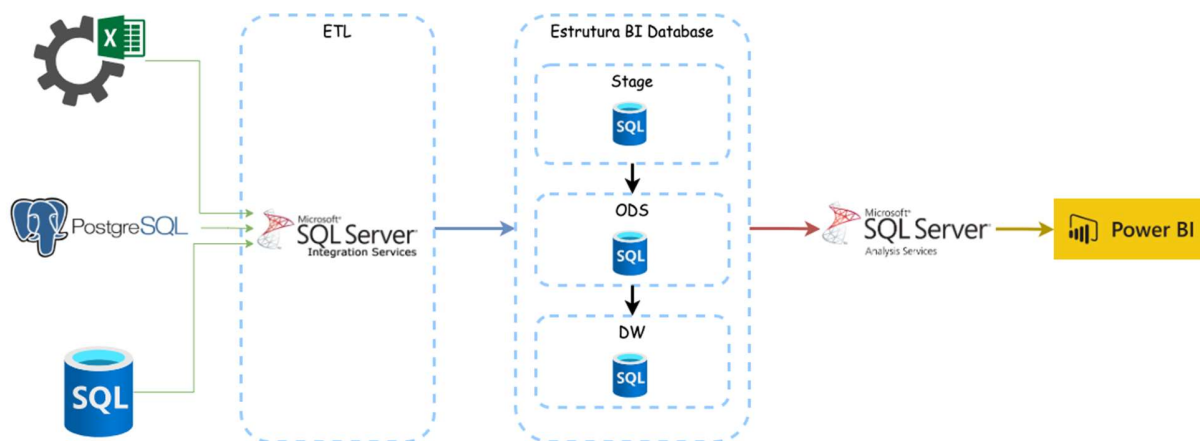


Imagem 6 – Soluções Microsoft com ferramentas de BI em ambiente On-Premise

2. Fontes de dados não estruturados (fontes de dados heterogêneas)

Conforme necessidade apresentada em reuniões com a Equipe de Projetos de BI, levaram em consideração extração de fontes de dados de origens heterogêneas, tendo estas origens as informações provenientes de Redes sociais (neste caso utilizando descrevendo redes como Twitter, Facebook, LinkedIn, entre outra). Para este caso, a estrutura se dará através de uso de ferramentas de controle de Streaming, como citado em reuniões o Apache Kafka, Apache Flumen e Apache Airflow.

O conceito de aplicação desta ferramenta vem de encontro a proposta de uso de estrutura de dados em Data Lake, onde a estrutura poderá estar centralizada em Cloud Provider e/ou OnPremisse.

A figura abaixo pode descrever como o fluxo com o início de dados em Streaming, tendo como middleware de estrutura Batch Layer e Serving Layer o fluxo de query, além de camadas de Speed Layer.

Tratar dados por streaming, ou seja, em tempo real, é uma das demandas mais requisitadas do momento, pois a necessidade de tomar decisões sobre dados que são muito voláteis ou em sistemas críticos, tem aumentado consideravelmente. Assim problemas que antes eram tratados apenas por processamento em batch, agora ganham a necessitada de se ter mais uma camada, a de streaming.

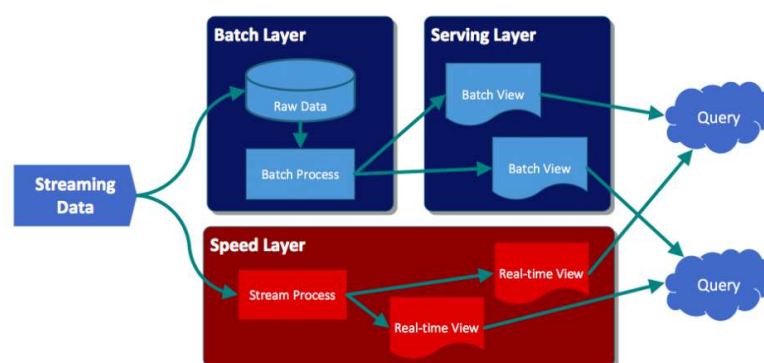


Imagem 7 – Fluxo de Streaming de Dados utilizado uma estrutura com Kafka

A imagem acima se trata da arquitetura de Big Data chamada Kappa, onde exemplifica o que foi mencionado anteriormente, existindo uma camada para tratar dados por batch (Batch Layer), ou seja, processamento de dados mais pesados, histórico e de forma agendada. E outra camada para processamento de dados de forma rápida e em tempo real, no caso streaming (Speed Layer), que permite processamento e análise dos dados com menor delay.

A proposta de uso de streaming de dados vem de encontro a necessidade de análise de dados em tempo real não poder ser medidos usando a arquitetura de BI. Com o aumento da adoção de IoT, computações e visualizações precisam ser feitas em dados de streaming (por exemplo, ter intervalo de tempo de 3 seg , 5 seg etc) para obter o valor a partir dele. Para analisar os dados de streaming, precisamos primeiro publicar os dados em sistemas de mensagens. Em seguida, todas as etapas de limpeza de dados e processamento de dados são feitas usando o uma Engine de Streaming de dados, que como proposta inicial citamos a ferramenta do Kafka.

Todos os dados processados são então armazenados no Data Lake, tendo também a possibilidade de serem replicados em uma estrutura de baseada na arquitetura oferecida pela suíte do Elasticsearch, que é então visualizado por Kibana ou Grafana, conforme imagem 8 apresentada abaixo.

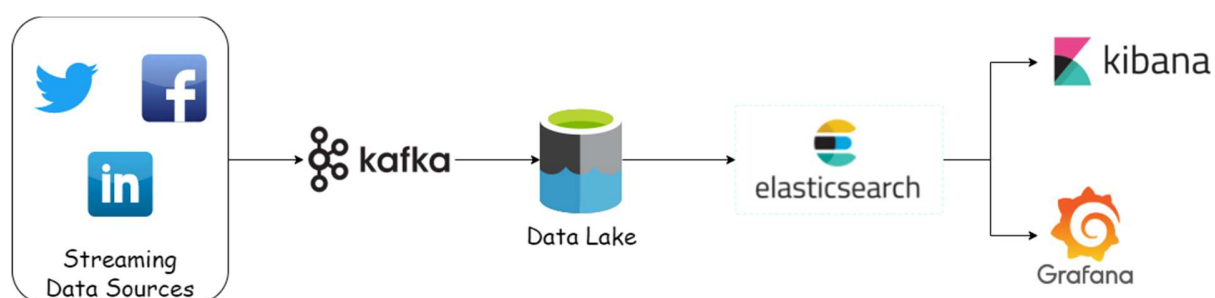


Imagem 8 – Conceito de uso de uma estrutura com proposta de Data Lake e ferramentas como Elasticsearch, Kibana e Grafana

De acordo com a necessidade, também poderemos ter o fluxo de streaming de dados conectando diretamente na estrutura do Elasticsearch, sendo este a origem de informações para a camada de apresentação do Kibana e Grafana, conforme imagem 9 apresentada abaixo.



Figura 9 – Estrutura de Streaming com tratativa direta na camada do Elasticsearch para apresentação de dados.

Para entendimento de cada camada descrita na imagem acima, segue detalhamento das ferramentas utilizadas (não levando em consideração o conceito e uso de um Data Lake).

Kafka: É um sistema de mensagens de publicação distribuída usado para construir pipelines de dados em tempo real e aplicativos de streaming.

Elasticsearch: É um mecanismo de pesquisa e análise de texto completo altamente escalável. Ele permite armazenar, pesquisar e analisar grandes volumes de dados de forma rápida e em tempo real.

Kibana: Uma ferramenta que facilita a compreensão de grandes volumes de dados. Sua interface simples e baseada no navegador permite que você crie e compartilhe rapidamente dashboards dinâmicos que exibem alterações nas consultas da Elasticsearch em tempo real.

Grafana: É uma solução que facilita a análise de dados e obtenção de métricas para análises e monitoramento de sistemas online. Contém opção para monitoramento de aplicativos e dados, com possibilidade de personalização de painéis.

3. Topologia de Rede x Disponibilidade de banco de dados e arquivos em FTP

De acordo com a estrutura apresentada, atualmente a EBSEH tem bancos de dados distribuídos em Hospitais, sendo que cada um destes possui sua estrutura de Rede e servidores exclusivos para sustentação da fonte de dados.

A proposta deste documento visa apresentar uma estrutura que integre a comunicação destas redes, facilitando assim a extração de dados em janelas que serão definidas, garantindo assim o

agendamento de rotinas que realizem este fluxo de extração de dados, mediante ao uso de ferramentas e técnicas de ETL.

Conforme poderá ser visualizado nas Imagens 10 e 11, são apresentados dois tipos de propostas para comunicação e extração destes dados, sendo que cada um possuem uma estratégia diferente, porém, com o mesmo resultado, que é encaminhar as informações para o Servidor de Datawarehouse disponível na Sede da EBSERH.

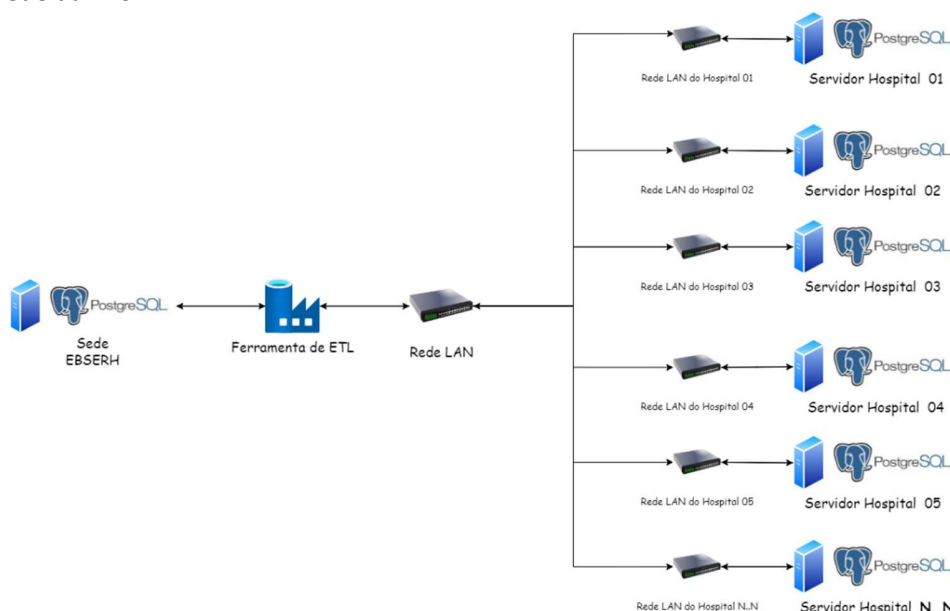


Imagem 10 – Proposta de Comunicação entre as Redes dos Hospitais e Sede da EBSERH para troca de informações entre bases de dados

Conforme pode ser visualizado na Imagem 7 abaixo, a premissa deste desenho é que as Redes de Hospitais possam estar ligadas com a Rede da Sede da EBSERH, favorecendo assim a comunicação da estrutura de banco de dados que será realizada mediante a tratativa da ferramenta e técnica de ETL aplicada.

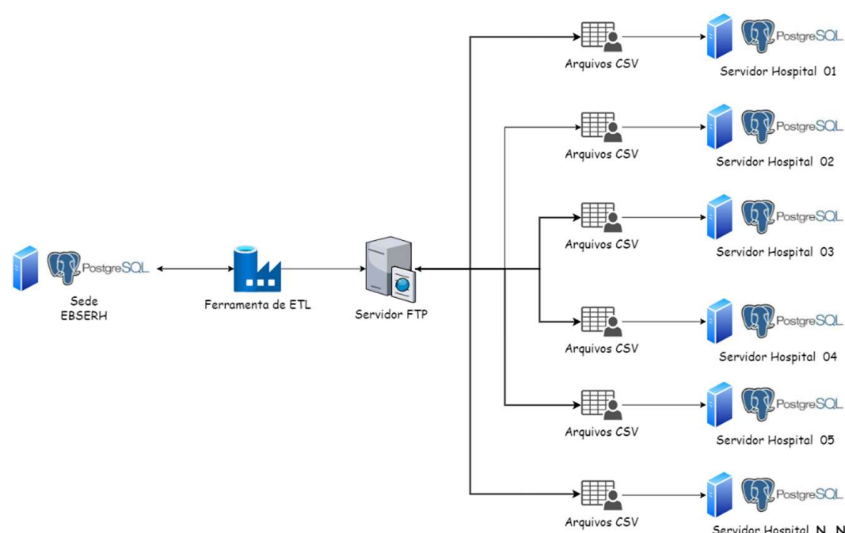


Imagem 11 – Proposta de Comunicação com geração de arquivos CSV com extração dos dados dos Hospitais e disponibilização destes dados via Servidor FTP

Conforme poderá ser visualizado na Imagem 8 acima, caso a proposta de comunicação apresentado na Imagem 7 não consiga ser realizada, existe também a possibilidade de proposta de criação de uma estrutura de Servidor FTP, onde as informações que serão extraídas para a estrutura de dados do Datawarehouse possam ser efetivadas com estruturas parametrizadas na proposta de uso deste servidor.

4. Padronização de nomenclatura

Para definir um padrão de nomenclatura de utilização de objetos no banco de dados, este documento trás como proposta o uso do padrão Pascal Case. Isto significa que os objetos de banco de dados, como exemplo de tabelas e suas colunas, serão entregues no formato de exemplo abaixo:

ExemploDeNomeDaColuna

Conforme poderá ser visto, o template de nomenclatura descrito acima apresenta um formato que, pelo texto inserido, demonstra a seguinte texto em explicação: Exemplo de nome da Coluna.

Obs.: A proposta de uso deste padrão, trás como benefício o entendimento do texto descrito, além de não ser necessário o uso de um **Catálogo de nomenclatura**.

Para as tabelas que serão utilizadas como propostas de modelos de um ambiente de BI, seguem os acrônimos a serem utilizados na estrutura:

- Tabelas de Dimensões deverão ser compostas pelo acrônimo Dim, somando o nome da Proposta de uso desta dimensão. Exemplo: DimContasMedicas, que descreverá que esta tabela se refere a Dimensão de Contas Médicas;

- Tabelas Fato deverão ser compostas pelo acrônimo Fato, somando o nome da proposta de uso desta tabela Fato. Exemplo: FatoContasMedicas, que descreverá que esta tabela se refere a Fato de Contas Médicas;
- Tabelas de Stage serão utilizadas pelo acrônimo Stg, somando o nome da proposta de uso da respectiva tabela;
Obs.: Como esta tabela poderá ser uma associação de várias tabelas da Origem, não foi pego um exemplo específico para apresentação. Contudo, entenda-se que sua nomenclatura se baseará no exemplo: Stg, somando o nme da proposta desta tabela de Stage.
- Tabelas de ODS serão utilizadas pelo acrônimo Ods, somando o nome da proposta de uso da respectiva tabela;
Obs.: Como esta tabela poderá ser uma associação de várias tabelas da Origem, não foi pego um exemplo específico para apresentação. Contudo, entenda-se que sua nomenclatura se baseará no exemplo: Stg, somando o nme da proposta desta tabela de Stage.

Da mesma forma, também teremos objetos de banco de dados, como Procedures e Index que poderão ser utilizados na estrutura do Projeto. Os padrões propostos para serem utilizados nestes objetos são descritos abaixo:

- Procedures, deverão iniciar com o acrônimo SP, somando com ação a ser executada na estrutura da tabela que se refere, como nos exemplos:
 - SP_Inserir_ContasMedicas, que executará o INSERT INTO;
 - SP_Atualiza_ContasMedicas, que executará o UPDATE;
 - SP_Deleta_ContasMedicas, que executará o DELETE;
 - SP_Obter_ContasMedicas, que poderá executar um SELECT na estrutura da tabela, tendo a opção de receber parâmetros para filtro na clausula WHERE ou não;
- Index, deverão iniciar com acrônimos IDX. Porém, poderemos utilizar dois tipos de INDEX na estrutura do banco de dados, sendo INDEX Clusterizados e não Clusterizados. Para cada um dos casos, recomenda-se o único formato de acrônimo, sendo que estes serão diferenciados especificamente pelos seus tipos;

5. Plataforma de versionamento de Código Fonte

Para a plataforma de versionamento de Código Fonte, utilizaremos o Azure DevOps, com projetos no formato de Repositório Git.

Fica como proposta que a estrutura de repositórios poderá seguir o formato de uso de fluxo do Git Flow, tendo como foco o uso de desenvolvimento de Features para as atividades que serão inseridas no Backlog de estrutura do Projeto.

Como linha principal de branch, utilizarmos o padrão MAIN, já comumente utilizado em estrutura de projetos que utilizam Git.. Entretanto, dependendo da cultura de desenvolvimento

utilizado na instituição, a entrega de código pode ser adaptada para uso de branches fixas ligadas a ambientes, com o processo de atualização de código sendo feito entre as branches utilizando Pull Request.

Para um bom entendimento de uso da estrutura do Git Flow, segue a Imagem 12 abaixo extraída de documentos baseados em uma plataforma Git, que descreve a aplicação e adoção do fluxo Git Flow.



Imagem 12 – Apresentação do uso do Git Flow como estrutura de atualização de Branchs em projetos que utilizam o Git como gestão de códigos fontes

Conclusão

Para uma proposta eficaz de solução de BI, é recomendável que todos os requisitos de ferramentas e técnicas sejam inclusos na estrutura de um Projeto de BI, a fim de padronizar toda a estrutura de entrega dos requisitos de dados que necessitam ser analisados. De acordo com o levantamento em reunião, a proposta de uma Solução baseada na estrutura de Suíte de BI da Microsoft, visa atender a entrega de requisitos a serem implementados. Contudo, tende-se a chegar ao uso de soluções Open Source, ficando ao encargo do Time de Desenvolvimento de Business Intelligence a apresentação de uma proposta que traga em visão geral, o uso da arquitetura por camada, levando em consideração o conjunto de técnicas para inicialmente se iniciar uma extração de dados do ambiente de Origem.



Modelo de Formulário para envio de painel para atualização do Catálogo de Painéis

Hospital filial:

Registrar informações de cada painel desenvolvido ou em desenvolvimento

[illegible]