



IV- 459 - DESAFIOS PARA USO DE BUSINESS INTELLIGENCE NA GESTÃO DE INFORMAÇÕES DE EMPRESAS DE SANEAMENTO

Jacqueline Nayara Ferraça Leite⁽¹⁾

Engenheira Agrônoma e Mestre em Ciência do Solo pela Universidade Estadual Paulista Júlio de Mesquita Filho (FCAV-UNESP). Graduanda em Ciências de Dados pela Universidade Virtual do Estado de São Paulo (UNIVESP). Assistente Administrativo na Sociedade de Abastecimento de Água (SANASA-Campinas/SP).

Diego de Oliveira Pinto⁽²⁾

Biólogo e Mestre em Parasitologia e Saúde Ambiental pela Universidade Estadual de Campinas (UNICAMP). Consultor Técnico da Organização Pan Americana de Saúde (OPAS), atuando pelo Ministério da Saúde e FUNASA. Atualmente Biólogo na Sociedade de Abastecimento de Água e Saneamento S. A. (SANASA Campinas), atuando no Programa Plano de Segurança da Água (PSA).

Adriana Angelica Rosa Vahteric Isenburg⁽³⁾

Engenheira Civil pela Pontifícia Universidade Católica (PUC-Campinas); Mestre em Geociências pela Universidade Estadual de Campinas (UNICAMP); MBA Administração de Projetos pela Fundação Instituto de Administração (FIA). Secretária de Obras e Serviços públicos e Projetos de Campinas, 1999 a 2000 e Diretora Técnica da Agência das Bacias PCJ, 2010 a 2012. Atualmente Gerente de Integração, Controle e Desenvolvimento Tecnológico, (SANASA-Campinas/SP).

Endereço⁽¹⁾: Avenida da Saudade, 500 – Ponte Preta - Campinas - SP - CEP: 13041-903 - Brasil - Tel: (19) 3735-5176 - e-mail: jacqueline.leite@sanasa.com.br

RESUMO

O objetivo desse trabalho é demonstrar a implementação de *Business Intelligence* (BI) ao Plano de Segurança da Água evidenciando os desafios para alcançá-la.

O projeto foi desenvolvido em três etapas. Na primeira etapa buscou-se o entendimento de negócio, a análise de requisitos e o mapeamento de fontes de dados. Na segunda etapa foi desenvolvida a solução BI para o conjunto de dados: Extração, Transformação, Carregamento (ETL), Modelagem, Cálculos, Validação e Visualização. Na terceira etapa foi realizada a implantação do projeto com automação, entrega, compartilhamento e documentação.

A solução entregue foi um aplicativo em Power BI composto pelos seguintes painéis operacionais (com métricas para acompanhamento) de dados históricos e disponibilizado aos consumidores: Monitoramento Hidrológico e Climático, Qualidade de Água Bruta; Qualidade de Água Tratada e Gestão de Risco.

A solução promoveu uma visão geral e integrada das informações de registros e indicadores que compõem o Plano de segurança da Água com maior rapidez, alcance, impacto e transparência. Além disso, promoveu gestão de risco mais assertiva possibilitando identificar riscos e oportunidades e apoiar a tomada de decisões estratégicas e operacionais.

PALAVRAS-CHAVE: Avaliação e Gestão de Risco, Banco de Dados, *Business Intelligence*, Gestão da Informação, Gestão Pública.

INTRODUÇÃO

A gestão integrada de informações requer atenção das empresas e órgãos públicos de saneamento no que se refere à demanda de uma base mínima necessária para Gestão de Recursos Hídricos eficiente e eficaz. Para isso, dados de monitoramento hidrometeorológico, qualidade da água, operação de reservatórios, sistemas operacionais de outorga e Licença de obra Hídrica, entre outros, são imprescindíveis. No entanto, a



SIMPÓSIO LUSO-BRASILEIRO
DE ENGENHARIA SANITÁRIA
E AMBIENTAL



grande quantidade e diversidade na geração, armazenamento e coleta de dados pode resultar na existência de dados conflitantes e dispersos que levam a informações inconsistentes e dificultam o processo de tomada de decisão. Dessa forma, é importante integrar as informações das diversas fontes geridas pelos *stakeholders* envolvidos na cadeia do processo de abastecimento de água e convertê-las em formato que possibilite sua análise e acesso.

Para isso, é necessário utilizar conjunto de processos para estruturação, acesso e exploração de informações direcionando a compreensão sobre as atividades estratégicas e operacionais do ambiente de negócios (*Business Intelligence* - BI).

Ressalta-se que, a aplicação desses processos BI em órgãos públicos enfrenta grandes desafios institucionais e técnicos, como infraestrutura inadequada, resistência cultural e, ainda, falta de recursos (FARIA; SANTOS; LIMA, 2023). Dessa maneira, em primeiro momento e abrangendo os dois primeiros desafios, é passo essencial a criação de uma estrutura de governança para tratar questões-chaves como a padronização da estrutura de dados que permitam interoperabilidade das informações, integração, armazenamento e acesso aos dados.

Do ponto de vista técnico, e aqui impacta muitas vezes o desafio da falta de recursos, seria ideal a combinação de algumas tecnologias (no *Back-end*), como sistemas *Enterprise Resource Planning* (ERP) e *Online Analytical Processing* (OLAP). O ERP trata-se de sistema modular integrado por banco de dados comum, caracterizado por sistema integrado e parametrizável que opera em tempo real ou próximo, capaz de integrar módulos, possibilitando otimização de processos, gestão integrada e provimento da capacidade de planejamento. OLAP, trata-se de um sistema responsável pelo armazenamento de dados multidimensionais representando informações por dimensões e categorias que coleta, organiza, agrega e analisa dados. O OLAP é composto ainda por *Data Warehouse* (DW) que armazena o conjunto de dados de apoio às decisões gerenciais, integrado, não volátil, variável em relação ao tempo e baseado em assuntos, obtido por processo de extração de informações de diferentes fontes, entre aplicações, arquivos e banco de dados, processamento e carregamento (ETL).

Com essas tecnologias, é possível executar processo de modelagem de dados OLAP, representação visual conhecida como esquema (estrela ou floco de neve), que define os sistemas de coleta e gerenciamento de informações. Isso significa dizer que, por meio da modelagem descreve-se quais dados a empresa coleta, estabelece-se as relações entre diferentes conjuntos de dados e determina-se métodos para armazenar e analisar esses dados de maneira a incidir em uma visão unificada das informações e permitir abranger o atendimento de diferentes partes interessadas, como analistas, cientistas e engenheiros de dados. Após modelagem, por meio de ferramentas analíticas OLAP, é possível executar operações analíticas como decomposição, fragmentação e pivotagem e, finalmente, com ferramentas de visualização de dados (no *Front-end*), criar elementos visuais como diagramas, gráficos e mapas que irão traduzir dados complexos para uma comunicação acessível e gerar informações.

Em contrapartida, apesar de limitações quanto ao acesso às ferramentas citadas, é possível construir uma solução com ferramentas geralmente mais usuais, *low code* e desconectadas, como aquelas que compõem o Office 365 (Excel, Power Query, Power BI, Power Apps, Forms, OneDrive e Sharepoint).

Todavia, é preciso salientar ainda que, nesse contexto, são maiores as dificuldades ao lidar com a integração de dados provenientes de diversas fontes, as quais abrangem formatos estruturados (tais como planilhas Excel e tabelas de bancos de dados), semiestruturados (como páginas da web e documentos XML) e não estruturados (incluindo imagens e arquivos PDF). Estes dados frequentemente exibem uma organização extremamente heterogênea, exigindo um processo minucioso de limpeza, padronização, organização, definição das informações para o modelo de negócio e construção de estruturas lógicas, como um modelo relacional.

Outrossim, a qualidade dos dados como Corretude, Completude, Consistência, Credibilidade, Atualidade, Acessibilidade, Conformidade, Confidencialidade, Eficiência, Precisão, Rastreabilidade, Compreensibilidade, Disponibilidade, Portabilidade e Recuperabilidade (dimensões conceituadas pela ISO/IEC 25012:2008) demanda bastante atenção nesse processo.

Contudo, este trabalho buscou a construção de gestão integrada de informações por meio de ferramentas de *Business Intelligence* combinando processos de extração, transformação e carregamento de dados (ETL) e visualização (*Dashboard*) para integrar informações para análise e acompanhamento de dados estratégicos do Plano de Segurança da Água.

MATERIAL E MÉTODOS

Para este trabalho será descrito o processo para implementação do projeto BI para Plano de Segurança da Água.

O projeto foi desenvolvido em três etapas. Na primeira etapa buscou-se o entendimento de negócio, a análise de requisitos e o mapeamento de fontes de dados por meio de entrevistas com coordenadorias e gerências internas além de leitura de documentos técnicos (WHO, 2022; ABNT NBR 17.080).

Na segunda etapa foi desenvolvida a solução BI para o conjunto de dados: Extração, Transformação, Carregamento (ETL), Modelagem, Cálculos, Validação e Visualização. Nessa etapa foram criadas pastas na rede em compartilhamento com coordenadorias e gerências internas e formulários online (ferramenta Forms) para acomodar, tornar acessíveis e tempestivos os dados geridos internamente. Para dados gerados em OLE DB sistema supervisorio (Historian Data GE - Sistema SCADA) foi estruturada credencial de acesso para consulta junto ao departamento de automação, definindo e elencando parâmetros controlados. Dados externos de monitoramento hidrogeológico foram extraídos via API SABESP e Agência das Bacias PCJ e os dados de cadastro técnico de georreferenciamento de sistemas e setores de abastecimento da empresa foram obtidos do banco de dados do software QGIS.

A transformação e limpeza dos dados internos (estruturados, semiestruturados e não estruturados) foi realizada com suplemento Query (Linguagem M) das ferramentas Excel e Power BI, dados de API foram extraídos e tratados diretamente na ferramenta Power BI. Já para os dados do Software QGIS foram extraídas feições de interesse em WKT formato CSV e transformado em Power BI.

A modelagem adotada para a maioria dos conjuntos de dados, devido grande variedade de dimensões, foi *Snowflake*, conforme demonstrado no exemplo da figura 1.

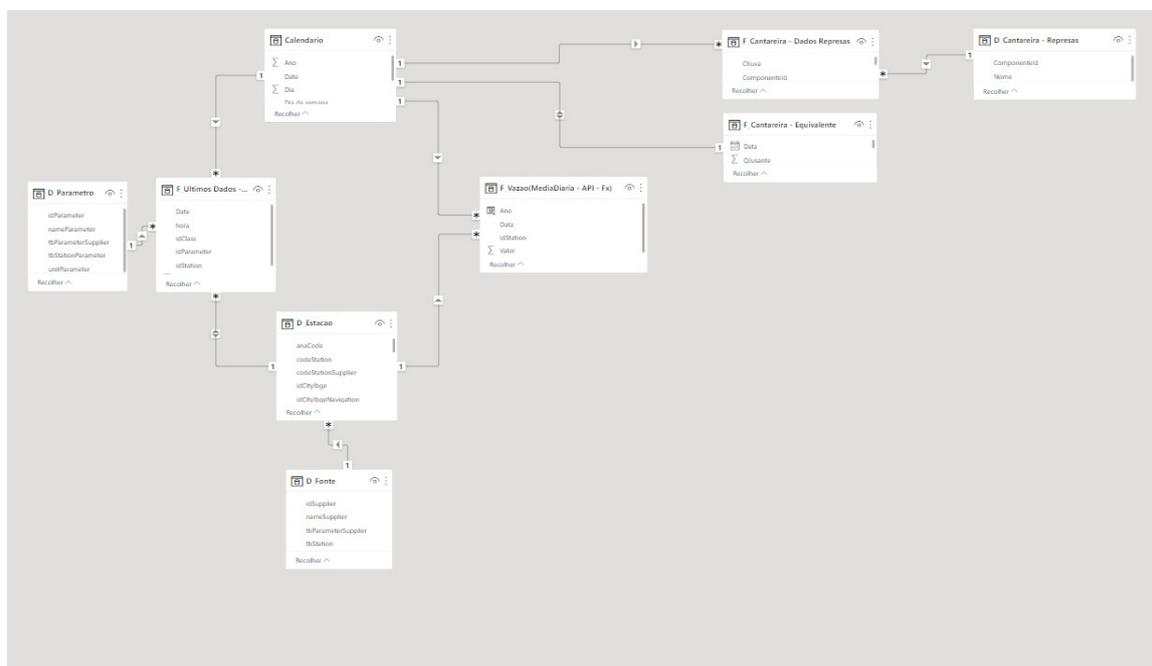


Figura 1 – Modelagem de dados *Snowflake* para painel de Monitoramento Hidrológico – consultas API e Historian

A temporalidade dos dados de API e Sistema supervisorio Historian variam em intervalos de 10 minutos, 60 minutos e diário. Foi estabelecido para o projeto cálculo de média diária para os parâmetros Vazão das estações Valinhos, Paulínia e Atibaia (dados API), Nível do Atibaia e Temperatura (dados Historian) com atualização periódica às 10 horas, compatibilizando temporalidade de atualização semelhante aos dados do Sistema Cantareira (Volume Equivalente, Qnat, QT5/PS-SC).



Após validação dos dados normalizados obtidos dos processos de ETL, modelagem e cálculo foram elaborados as visualizações com *Dashboards* Operacionais (Métricas para acompanhamento).

Na terceira etapa foi realizada a implantação do projeto com automação, entrega, compartilhamento e documentação.

RESULTADOS OBTIDOS OU ESPERADOS

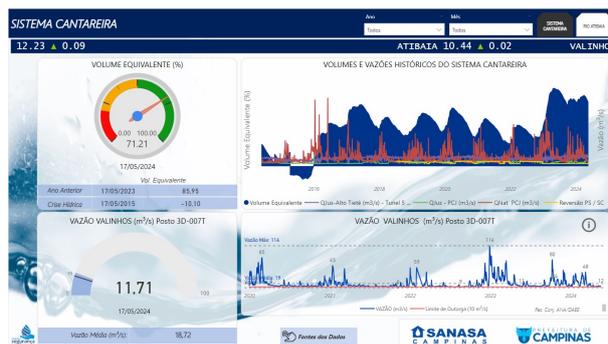
A solução entregue foi um aplicativo composto pelos seguintes painéis operacionais (com métricas para acompanhamento) de dados históricos e disponibilizado aos consumidores:

1 – Hidrológicos e Climáticos – Volumes e Vazões que compõem o Sistema Cantareira, Nível do Rio Atibaia, Temperatura e Precipitação (figuras 2a, 2b e 2c);

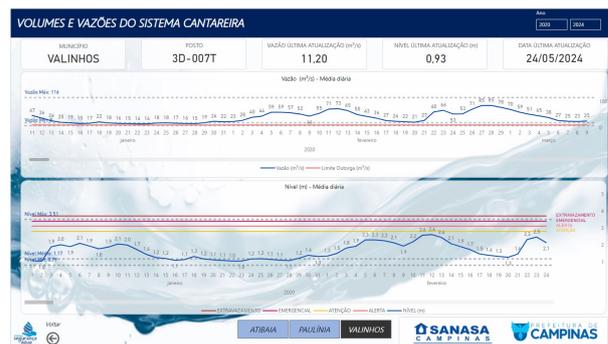
2 – Qualidade de Água Bruta – Parâmetros químicos (Orgânicos e Inorgânicos), físicos (Temperatura, Turbidez, Sólidos Totais e Resíduos) e biológicos (bacteriológicos) de qualidade e Índice de Qualidade da Água (IQA) para pontos de amostragem nos mananciais de abastecimento do município do projeto (figura 2d);

3 – Qualidade de Água Tratada – Contagem de parâmetros químicos (Orgânicos e Inorgânicos), biológicos (bacteriológicos e cianotoxinas) em conformidade com legislação pertinente, exibido em formato de relatório anual de qualidade de água (figura 2e);

4 – Gestão de Risco – Interativo para demonstração da estrutura do Plano de Segurança da Água - Gestão de Risco associado aos principais perigos e eventos perigosos identificados (figura 2f).



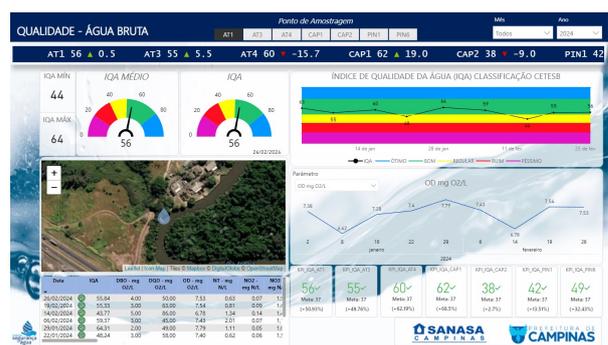
2a – Monitoramento Hidrológico – Sistema Cantareira



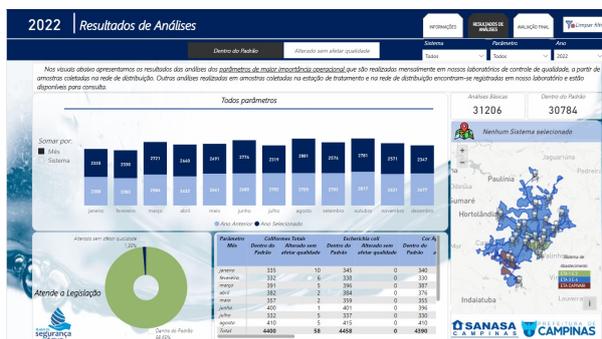
2b – Monitoramento Hidrológico – Volumes e Vazões



2c – Monitoramento Hidrológico-Níveis do Rio Atibaia, Temperatura e Precipitação



2d – Qualidade de Água Bruta



2e – Qualidade de Água Tratada



2f – Gestão de Risco

Com o desenvolvimento dessa solução, espera-se ainda, promover relativos impactos para inserção da cultura de dados nas empresas de saneamento, de modo a alavancar a gestão integrada de informação utilizando dados como centro das decisões corporativas e com isso demonstrar a importância para desenvolver conjunto de arquiteturas e/ou sistemas de informação orientados a assunto (*Data Lake* e *Warehouse*) que darão amparo ao tratamento, armazenamento e trânsito das informações e, dessa forma, oferecer a possibilidade de consultas intersetoriais com rapidez, integridade, eficiência e transparência.

Os painéis desenvolvidos promoveram uma visão geral e integrada das informações de registros e indicadores que compõem o Plano de Segurança da Água com maior alcance, impacto e transparência. Isso significa dizer que, controles antes restritos à uma gerência, puderam ser expandidos aos demais integrantes da cadeia de produção gerando maior entendimento e engajamento quanto aos desafios do Plano.

Além disso, o monitoramento desses indicadores permitiu uma gestão de risco mais assertiva e com tomadas de decisões mais ágeis, eficientes e eficazes no sentido de identificar riscos e oportunidades, seja por meio de análises de ponto crítico e controle, promovendo a alocação de KPIs em etapas chave da gestão, identificação de gargalos e do compartilhamento de soluções, gerando cadência e alinhamento de expectativas e entregas entre os stakeholders.

CONCLUSÕES

A gestão integrada de informação poderá simplificar ações, processos administrativos burocráticos, reduzir variabilidade, aumentar a qualidade e integridade na obtenção e gestão de dados garantindo planejamento estratégico de gestão forte e otimizada.

Embora tenha sido possível construir a visualização de dados operacionais por meio de plataformas *low code*, é importante ressaltar a necessidade da implantação de processos integrados (ERP) e banco de dados estruturados (*Data Lake* e *Warehouse*) para fundamentar a utilização de ferramentas de BI e consequentemente orientar tomada de decisão para planejamento e execução de ações.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ABNT - Associação Brasileira de Normas Técnicas. ABNT NBR 17.080: *Plano de Segurança da Água – Princípios e diretrizes para elaboração e implementação*. ISBN 978-85-07-09497-5. 2023.
2. Sady, F.S.A., Darcy, S.J. (2013). *Information architecture analysis using business intelligence tools based on the information needs of executives*, *Journal of Information Systems and Technology Management*, 2, 251-270.
3. FARIA, Osvaldo Pacheco de; SANTOS, Luciléia Soares dos; LIMA, Juocerlee Tavares Guadalupe Pereira de. O impacto das tecnologias de informação e telecomunicação no setor público brasileiro: uma revisão bibliográfica. *Ciências Humanas*, v. 27, ed. 128, nov. 2023. DOI: 10.5281/zenodo.10127140.
4. ISO25012 System and Software Engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) — Data quality model. ISO/IEC - International Organization for Standardization and International Electrotechnical Commission, Switzerland, 1 edition, 2008b
5. WHO - WORLD HEALTH ORGANIZATION. *Guidelines for drinking water quality*:. 4ª ed incorporating the first and second addenda. Geneva: WHO, 2022.