



**UNIVERSIDADE FEDERAL DO MARANHÃO
CURSO DE CIENCIA DA COMPUTAÇÃO**

MYLADY DOS SANTOS AMARAL

**IMPLEMENTAÇÃO DE UM PROCESSO ETL COM FERRAMENTAS *OPEN
SOURCE* APLICADO AO PORTAL DA TRANSPARÊNCIA DA ASSEMBLEIA
LEGISLATIVA DO ESTADO DO MARANHÃO**

**São Luís – MA
2025**

MYLADY DOS SANTOS AMARAL

IMPLEMENTAÇÃO DE UM PROCESSO ETL COM FERRAMENTAS *OPEN SOURCE* APLICADO AO PORTAL DA TRANSPARÊNCIA DA ASSEMBLEIA LEGISLATIVA DO ESTADO DO MARANHÃO

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão, como parte dos requisitos necessários para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Profa. Dra. Simara Vieira da Rocha

**São Luís - MA
2025**

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Diretoria Integrada de Bibliotecas/UFMA

dos Santos Amaral, Mylady.

IMPLEMENTAÇÃO DE UM PROCESSO ETL COM FERRAMENTAS OPEN
SOURCE APLICADO AO PORTAL DA TRANSPARÊNCIA DA ASSEMBLEIA
LEGISLATIVA DO ESTADO DO MARANHÃO / Mylady dos Santos
Amaral. - 2025.

68 f.

Orientador(a): Simara Vieira da Rocha.

Curso de Ciência da Computação, Universidade Federal do
Maranhão, São Luís, 2025.

1. Etl. 2. Transparência Pública. 3. Pentaho. 4.
Oracle. 5. Postgresql. I. Vieira da Rocha, Simara. II.
Título.

MYLADY DOS SANTOS AMARAL

IMPLEMENTAÇÃO DE UM PROCESSO ETL COM FERRAMENTAS *OPEN SOURCE* APLICADO AO PORTAL DA TRANSPARÊNCIA DA ASSEMBLEIA LEGISLATIVA DO ESTADO DO MARANHÃO

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão, como parte dos requisitos necessários para obtenção do grau de Bacharel em Ciência da Computação.

São Luís, 07 de agosto de 2025

BANCA EXAMINADORA

Profª. Dra. Simara Vieira da Rocha
Orientador

Prof. Dr. Carlos de Salles Soares Neto
Examinador

Prof. Me. Carlos Eduardo Portela Serra de Castro
Examinador

**São Luís – MA
2025**

Agradecimentos

Em primeiro lugar, agradeço a mim mesma pelo esforço, dedicação e perseverança que me permitiram concluir esta jornada.

A minha orientadora, Simara Vieira da Rocha, pela paciência, sabedoria, apoio indispensáveis durante a elaboração desta monografia e por ser inspiração para mim.

A minha família (Edna, Wilma, Lya, Augusto, Giovanni e Rosana), por estarem comigo, me apoiarem e me incentivarem em todos os momentos. Agradeço as minhas melhores amigas: Diandra e Andressa, por segurarem a minha mão durante esses 11 anos de amizade. Agradeço a Mariana, Ruan, Ruanita, Nathasha, Beckman, Beneilton, Hiago, Mike, José, Paulo, Ana, Arthur e Luigi que tornaram minha vida muito mais divertida.

Agradeço a LUF (Isanna, Jean, Sabrina, Diego, Guilherme, Cláudio, Kaynara, Lucas, Emanuelle, Mateus, Vitória, Tatiana, Soraya, Sanara e Vinicius Socram), que me proporcionaram experiências incríveis além da vida acadêmica.

Um agradecimento especial à Atlético Lorde, ao time feminino da Lorde em especial a Lisle, Mayrla, Alana, Laís e Mackele que além de minhas companheiras de quadra foram grandes amigas e incentivadoras.

Ao Departamento de Tecnologia da Assembleia Legislativa do Maranhão, por autorizar a utilização deste projeto como tema de pesquisa.

Por fim, aos meus professores, pelo conhecimento compartilhado, pelos conselhos valiosos e pelo incentivo ao longo desta graduação.

“Os sonhos das pessoas não têm fim.”
Eiichiro Oda, em *"One Piece"*

Resumo

A transparência pública é um pilar fundamental para a governança e a participação cidadã. A Lei de Acesso à Informação (LAI - Lei 12.527/2011) marcou uma mudança crucial no Brasil, transformando o Estado de detentor do monopólio de documentos oficiais para guardião de informações públicas. Isso promove a prestação de contas e permite o controle social sobre os gastos e ações do governo. A Assembleia Legislativa do Estado do Maranhão (ALEMA) enfrenta desafios significativos para garantir transparência, evidenciados pela classificação "C"(atendimento regular) no índice do Tribunal de Contas do Estado (TCE/MA) em 2021. As principais dificuldades são: a dependência e incompatibilidade técnica, pois a ALEMA depende de dados do sistema SIGEF, mantido pela SEPLAN em banco de dados Oracle. Porém, a ALEMA utiliza PostgreSQL. As diferenças estruturais, de tipos de dados, linguagem SQL e gerenciamento entre Oracle e PostgreSQL geram complexidade na integração e migração dos dados, que são fornecidos apenas via views; e, a fragmentação de dados, uma vez que, desde 2018, um novo banco de dados é criado a cada ano exercício. Essa fragmentação compromete a eficiência na publicação das informações no Portal da Transparência da ALEMA e dificulta o fornecimento claro e acessível de dados, especialmente sobre despesas e notas de empenho. Nesse sentido, o objetivo deste trabalho é propor a criação de uma solução de Extração, Transformação e Carga (ETL) com tecnologias open source, viabilizando a interoperabilidade entre os bancos de dados em Oracle e PostgreSQL, para que a ALEMA atenda aos requisitos legais vigentes em seu Portal da Transparência. A solução criada trouxe um impacto positivo para o nível de transparência do Portal da ALEMA, saindo de um conceito "C" em 2021 para o nível "Elevado"(nota 7,90) em 2024 (Relatório de Informação nº 253/2025 NUFIS I, de 11 de março de 2025).

Palavras-chave: ETL; transparência pública; Pentaho; Oracle; PostgreSQL; Portal da Transparência.

Abstract

Public transparency is a fundamental pillar for governance and citizen participation. The Access to Information Law (LAI - Law 12.527/2011) marked a crucial shift in Brazil, transforming the State from a monopolistic holder of official documents to a guardian of public information. This promotes accountability and enables social oversight over government spending and actions. The Legislative Assembly of the State of Maranhão (ALEMA) faces significant challenges in ensuring transparency, as evidenced by its "C" rating (regular compliance) in the State Audit Court's (TCE/MA) index in 2021. The main difficulties are: technical dependence and incompatibility, as ALEMA relies on data from the SIGEF system, maintained by SEPLAN in an Oracle database, while ALEMA uses PostgreSQL. The structural differences in data types, SQL language, and management between Oracle and PostgreSQL create complexity in data integration and migration, as the data is only provided via views; and data fragmentation, since a new database is created every fiscal year starting in 2018. This fragmentation undermines the efficiency of publishing information on ALEMA's Transparency Portal and hinders the clear and accessible provision of data, especially regarding expenses and payment orders. In this context, the objective of this work is to propose the creation of an open- source Extract, Transform, and Load (ETL) solution, enabling interoperability between Oracle and PostgreSQL databases and allowing ALEMA to comply with current legal requirements on its Transparency Portal. The implemented solution had a positive impact on the transparency level of ALEMA's Portal, raising it from a "C" rating in 2021 to a "High" level (score 7.90) in 2024 (Information Report No. 253/2025 NUFIS I, dated March 11, 2025).

Keywords: ETL; public transparency; Pentaho; Oracle; PostgreSQL; Transparency Portal.

Lista de ilustrações

Figura 1 – Exemplo de uma transformação simples no <i>Pentaho Kettle</i>	22
Figura 2 – Exemplo de um <i>job</i> no Pentaho Kettle.	24
Figura 3 – Tela de configuração de conexão com base de dados.	25
Figura 4 – Etapas seguidas para desenvolvimento do projeto.	31
Figura 5 – Exemplo do mapeamento da view nota de empenho do Oracle para a tabela nota de empenho no PostgreSQL.	36
Figura 6 – Grupo Orçamento.	37
Figura 7 – Grupo Despesas e empenhos.	38
Figura 8 – Grupo Pagamentos e processamentos.	40
Figura 9 – Grupo Programas e ações.	41
Figura 10 – Transformação da VSEPLANNELEG para a tabela nota_empenho.	43
Figura 11 – Table input: VSEPLANNELEG.	44
Figura 12 – Aba: <i>Select&Alter</i>	45
Figura 13 – Aba: <i>Meta-data</i>	46
Figura 14 – <i>Sort rows</i> : SR 11.....	47
Figura 15 – <i>Filter rows</i> : tipo_ne.....	47
Figura 16 – <i>Calculator</i> : Calc_NE.....	48
Figura 17 – Dados da coluna VLNOTAEMPENHO antes e coluna valor_nota_empenho após formatação	49
Figura 18 – <i>Calculator</i> : Calc_NE 2.....	49
Figura 19 – Coluna valor_nota_empenho com os dados antes e a coluna valor_nota_empenho_1 após formatação	50
Figura 20 – Insert/Update - update_NE e Insert/Update - update_NE_tipo_AX.....	51
Figura 21 – Comparativo da quantidade de registros do ano de 2019 na view VSEPLANNELEG e na tabela nota_empenho pós migração.	52
Figura 22 – Dados após migração.....	52
Figura 23 – Job: MIGRA_SIGEF_PROD.	53
Figura 24 – Transformation: SIGEF_cd.....	54
Figura 25 – Transformation: SIGEF notas empenho	55

Figura 26 – Transformation: SIGEF notas e rap.	55
Figura 27 – Transformation: UPDATE_TOTAIS.....	56
Figura 28 – Transformation: UPDATE_TOTAIS_RAP	56
Figura 29 – SQL versionamento.	56
Figura 30 – Tela de Pesquisa de Notas de Empenho.....	62

Lista de tabelas

Tabela 1 – Detalhes dos requisitos mapeados na etapa de Análise.....	32
Tabela 2 – Mapeamentos de tipos de dados Oracle para tipos equivalentes no PostgreSQL.....	35

Lista de abreviaturas e siglas

ALEMA	<i>Assembleia Legislativa do Maranhão</i>
BD	<i>Banco de Dados</i>
ETL	<i>Extract, Transform, Load</i> (Extrair, Transformar, Carregar)
IN	<i>Instrução Normativa</i>
JVM	<i>Java Virtual Machine</i> (Máquina Virtual Java)
LAI	<i>Lei de Acesso à Informação</i>
NUFIS	<i>Núcleo de Fiscalização</i>
PDI	<i>Pentaho Data Integration</i> (Integração de Dados Pentaho)
SEPLAN	<i>Secretaria de Estado de Planejamento e Orçamento</i>
SGBD	<i>Sistema de Gerenciamento de Banco de Dados</i>
SGBDR	<i>Sistema Gerenciador de Banco de Dados Relacional</i>
SICONV	<i>Sistema de Convênios e Transferências da União</i>
SIGEF	<i>Sistema Integrado de Planejamento e Gestão Fiscal</i>
SQL	<i>Structured Query Language</i>
STC	<i>Secretaria de Transparência e Controle</i>
TAG	<i>Termo de Ajustamento de Gestão</i>
TCE/MA	<i>Tribunal de Contas do Estado do Maranhão</i>

Sumário

1	INTRODUÇÃO	13
1.1	Objetivos.....	14
1.1.1	Objetivos Específicos.....	15
1.2	Organização do Trabalho	15
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	Assembleia Legislativa do Estado do Maranhão(ALEMA)	16
2.1.1	Constituição e atribuições	16
2.1.2	Papel atual.....	17
2.2	Evolução das Normas de Transparência no TCE/MA.....	18
2.2.1	Instrução Normativa - TCE/MA nº 59/2020	18
2.2.2	Instrução Normativa-TCE/MA nº 81/2024	19
2.2.3	Importância das Mudanças	20
2.3	Processo ETL	21
2.3.1	Extração	21
2.3.2	Transformação	21
2.3.3	Carga	21
2.4	Pentaho Data Integration (PDI ou Kettle)	22
2.4.1	<i>Spoon</i>	22
2.4.1.1	Transformações.....	22
2.4.1.2	<i>Jobs</i> (Trabalhos).....	23
2.4.1.3	<i>Database Connections</i> (Conexões de Base de dados).....	25
2.4.2	Transformação de dados em tipos Java.....	26
2.5	<i>Pentaho Server</i>	27
2.5.1	<i>User Console</i> (Interface do Usuário)	27
2.5.2	<i>Scheduler</i> (Agendador)	27
2.6	Sistema gerenciador de banco de dados(SGBD)	27
2.6.1	SGBD Oracle	28
2.6.2	SGBD PostgreSQL	28

3	METODOLOGIA	29
3.1	Descrição do problema	29
3.2	Etapas da Metodologia	30
3.2.1	ANÁLISE	32
3.2.2	ESPECIFICAÇÃO	33
3.2.2.1	Mapeamento e criação de tabelas:	34
3.2.3	IMPLEMENTAÇÃO.....	42
3.2.3.1	Extração dos Dados	43
3.2.3.2	Transformação dos Dados.....	44
3.2.3.3	Carga dos Dados no PostgreSQL.....	50
3.2.3.4	Validação e Testes.....	51
3.2.3.5	Job Pentaho - MIGRA_SIGEF_PROD.....	53
3.2.4	HOMOLOGAÇÃO	57
3.2.4.1	Configuração do Pentaho Server.....	57
3.2.4.2	Configuração do Agendamento do Job MIGRA_SIGEF	57
3.2.5	PUBLICAÇÃO.....	58
3.2.5.1	Configuração do Ambiente de Produção	59
3.2.5.2	Integração com o Portal da Transparência	59
3.2.5.3	Resultados da Publicação.....	59
3.2.6	MANUTENÇÃO.....	59
3.3	Discussão de resultados	60
3.3.1	Desempenho em 2021	61
3.3.2	Desempenho em 2023.....	61
3.3.3	Desempenho em 2024.....	61
3.3.4	Comparativo das Classificações.....	63
3.3.5	Impacto e Benefícios	63
4	CONCLUSÃO	64
	REFERÊNCIAS	66

1 INTRODUÇÃO

A transparência pública é um dos pilares fundamentais para garantir a governança e a participação cidadã em processos de tomada de decisão. No Brasil, no ano de 2011, foi criada a Lei 12.527, que ficou conhecida como lei de acesso à informação – LAI (Ato institucional Nº. Lei nº 12.527, de 18 de novembro de 2011) (BRASIL, 2011), com sua implementação o Estado passou da posição de detentor do monopólio de "documentos oficiais" para guardião de "informações públicas" (MICHENER; CONTRERAS; NISKIER, 2018). Essa mudança trouxe avanços importantes, mas também desafios, especialmente no que diz respeito à integração e disponibilização de dados provenientes de diferentes sistemas.

Os impactos da LAI foram sentidos em todo o território nacional. Porém, no estado do Maranhão, sua implementação efetiva ocorreu somente em 2015, com a posse do então governador eleito, Flávio Dino. Quando assumiu o poder ele abriu as portas da administração estadual, criando a Secretaria de Transparência e Controle (STC), aprovando a regulamentação estadual da LAI (Decreto nº 10.217/2015) e lançando o portal da transparência do governo do estado (MICHENER; CONTRERAS; NISKIER, 2018).

A Lei Ordinária nº 10.217, de 23 de Março de 2015, dispõe sobre as regras específicas para garantir o acesso a informações públicas e tem como subordinados todos os órgãos públicos integrantes da administração direta dos Poderes Executivo, Legislativo e Judiciário, do Tribunal de Contas do Estado e do Ministério Público do Estado do Maranhão (LEI ORDINÁRIA Nº 10.217. 2015. CAP I, Art. 1º) (MARANHÃO, 2015).

Neste contexto, sendo a Assembleia Legislativa do Estado do Maranhão (ALEMA), o poder legislativo deste estado, teria portanto a responsabilidade e o dever de cumprir o dito na referida lei. Assim em 2016 foi lançado o primeiro portal da transparência direcionado especificamente aos gastos e informações públicas deste órgão.

Em 2020, o Tribunal de Contas do Estado do Maranhão (TCE/MA) publicou a Instrução Normativa (IN) nº 59/2020 (TCE/MA, 2020), que estabeleceu critérios rigorosos para a avaliação da transparência, classificando os órgãos públicos com base

no percentual de cumprimento das exigências, utilizando uma escala que vai do índice A (transparência excelente) ao índice C- (transparência insuficiente). E em 2021, no Relatório de acompanhamento N° 1056/2021 – NUFIS 2 publicado em 13 de Dezembro, o portal da ALEMA foi avaliado com o índice C, indicando um atendimento regular aos critérios de transparência (TCE/MA, 2021).

Afim de elevar essa classificação foram iniciados esforços para adequar o portal aos critérios da IN nº 59/2020. No entanto, a ALEMA enfrenta desafios técnicos significativos para garantir a transparência de seus dados. O órgão depende de informações financeiras e orçamentárias coletadas por meio do Sistema Integrado de Planejamento e Gestão Fiscal (SIGEF), cujos dados são armazenados em banco de dados (BD) no Sistema Gerenciador de Banco de Dados (SGBD) Oracle e mantidos pela Secretaria de Estado do Planejamento e Orçamento (SEPLAN). Esses dados são disponibilizados por meio de um esquema de views, o que gera complexidades na integração e migração destes, pois, o SGBD Oracle, usado pela SEPLAN, e o SGBD PostgreSQL, utilizado pela ALEMA, apresentam diferenças significativas em termos de estrutura de dados, tipos de dados, linguagem SQL e gerenciamento de transações.

Adicionalmente, desde o ano de 2018, existe a problemática da criação de um BD por ano de exercício o que resulta em fragmentação de dados. Comprometendo a eficiência na publicação de informações no Portal da Transparência, bem como dificultando o fornecimento de dados relacionados a despesas e notas de empenho de forma clara e acessível ao público de maneira geral.

A partir desses fatores surgiu a questão: "Como unificar os dados disponibilizados pela SEPLAN e exibir esses dados em tempo real ou com o mínimo possível de atraso, dentro do portal da transparência da ALEMA, utilizando ferramentas open source?". Este trabalho busca responder a essa problemática, propondo uma solução viável para otimizar a integração e a acessibilidade das informações públicas.

1.1 OBJETIVOS

O objetivo geral deste trabalho é criar uma solução de Extração, Transformação e Carga (ETL) com tecnologias open source, viabilizando a interoperabilidade entre os bancos de dados em Oracle e PostgreSQL, para que a ALEMA atenda aos requisitos legais vigentes em seu Portal da Transparência.

1.1.1 Objetivos Específicos

- Mapear as diferenças entre os modelos de dados do Oracle e PostgreSQL.
- Configurar e utilizar o Pentaho Data Integration para realizar o ETL dos dados.
- Implementar uma solução que permita atualizações diárias e automáticas.
- Validar e testar o processo desenvolvido para garantir a consistência e integridade dos dados.

1.2 Organização do Trabalho

Este trabalho está estruturado em quatro capítulos. O Capítulo 1 apresenta a introdução ao tema e os objetivos. O Capítulo 2, traz a fundamentação teórica necessária para compreensão dos procedimentos que foram adotados. O Capítulo 3 trata da metodologia propriamente dita. E o Capítulo 4 traz as considerações finais e trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo será apresentada a fundamentação teórica usada no desenvolvimento desta monografia, importante para compreensão dos procedimentos utilizados para alcançar os objetivos esperados. Para tanto, abordaremos a Assembleia Legislativa do Estado do Maranhão(ALEMA), a Evolução das Normas de Transparência no TCE/MA, o Processo ETL, o Pentaho Data Integration (PDI ou Kettle), o Pentaho Server e Sistema gerenciador de banco de dados(SGBD).

2.1 Assembleia Legislativa do Estado do Maranhão(ALEMA)

A Assembleia Legislativa do Estado do Maranhão (ALEMA) é o órgão máximo do Poder Legislativo estadual, constituindo-se como espaço de representação popular e de exercício democrático no âmbito do Estado do Maranhão (ALEMA, 2025a). Sua criação e evolução acompanham o desenvolvimento político-institucional brasileiro desde o período imperial.

Segundo documentos oficiais (ALEMA, 2025b), a ALEMA é: o parlamento estadual do Maranhão, instituição política com origens no Conselho Geral da Província (1829). E composta atualmente por 42 deputados estaduais eleitos e organizada em Plenário, Comissões Temáticas e Mesa Diretora

2.1.1 Constituição e atribuições

A ALEMA foi constituída para:

- a) Representar o povo maranhense no âmbito estadual (ALEMA, 2025a)
- b) Estabelecer o equilíbrio entre os Poderes do Estado
- c) Garantir a participação popular nas decisões políticas
- d) Promover o desenvolvimento legislativo regional (ALEMA, 2025b) E tem como atribuições constitucionais (ALEMA, 2025a):

a) **Legislar:**

- Elaborar leis estaduais
- Propor emendas à Constituição Estadual
- Aprovar orçamentos e planos de governo

b) **Fiscalizar:**

- Controlar os atos do Governador e Secretários
- Acompanhar a execução orçamentária
- Realizar Comissões Parlamentares de Inquérito

c) **Representar:**

- Ser canal de expressão das demandas populares
- Mediar conflitos políticos e sociais

2.1.2 Papel atual

Atualmente, conforme registros oficiais (ALEMA, 2025a; ALEMA, 2025b), a ALEMA desempenha os seguintes papéis:

- a) **Democrático:** Garantir a participação popular através de audiências públicas e canais de comunicação
- b) **Transparência:** Manter o Portal da Transparência e a TV Assembleia
- c) **Modernização:** Implementar novas tecnologias no processo legislativo
- d) **Capacitação:** Manter a Escola do Legislativo para formação política
- e) **Integração:** Promover a Jornada de Integração Legislativa com municípios

2.2 Evolução das Normas de Transparência no TCE/MA

A transparência na gestão pública é um princípio fundamental para o controle social e a promoção da integridade administrativa. No âmbito do Tribunal de Contas do Estado do Maranhão (TCE/MA), esse princípio foi regulamentado por meio de normativas específicas que estabelecem critérios para avaliação e classificação da transparência nos municípios e órgãos públicos do estado. A evolução dessas normas reflete a necessidade de adaptação às demandas da sociedade por maior acesso à informação e à prestação de contas.

2.2.1 Instrução Normativa - TCE/MA nº 59/2020

Em abril de 2020, o TCE/MA publicou a Instrução Normativa-TCE/MA nº 59/2020 (TCE/MA, 2020), que definiu os critérios para avaliação da transparência, classificando os entes públicos com base no percentual de cumprimento das exigências. Conforme essa norma, os níveis de transparência eram classificados em:

- **A (Excelente):** igual ou superior a 90% de atendimento às exigências.
- **B (Bom):** igual ou superior a 70%.
- **C (Regular):** igual ou superior a 40%.
- **C- (Insuficiente):** entre 0% e 39,90%.

Além disso, os critérios de avaliação eram divididos em três tipos, com pesos distintos para o cálculo do percentual atendido:

- **Essencial (Peso 3):** critérios fundamentais para garantir a transparência mínima exigida.
- **Obrigatória (Peso 2):** critérios importantes, mas com menor impacto que os essenciais.
- **Recomendada (Peso 1):** critérios complementares, que agregam valor à transparência.

Essa classificação permitia ao TCE/MA identificar os órgãos que necessitavam de melhorias e adotar medidas corretivas, como a emissão de recomendações, a celebração de Termos de Ajustamento de Gestão (TAG) e, em casos de descumprimento, a aplicação de multas e o registro no portal do Sistema de Convênios e Transferências da União (SICONV) (TCE/MA, 2020; TCE/MA, 2021).

2.2.2 Instrução Normativa-TCE/MA nº 81/2024

Em 2024, a Instrução Normativa-TCE/MA nº 59/2020 foi revogada e substituída pela Instrução Normativa-TCE/MA nº 81/2024, que trouxe mudanças significativas nos critérios de avaliação e na classificação dos índices de transparência. A nova norma estabeleceu uma escala de classificação mais detalhada, que inclui os seguintes níveis (TCE/MA, 2024a):

- **Diamante:** atendimento a 100% dos critérios essenciais e índice de transparência igual ou superior a 95%.
- **Ouro:** atendimento a 100% dos critérios essenciais e índice de transparência igual ou superior a 85% e inferior a 95%.
- **Prata:** atendimento a 100% dos critérios essenciais e índice de transparência igual ou superior a 75% e inferior a 85%.
- **Elevado:** atendimento a menos de 100% dos critérios essenciais e índice de transparência igual ou superior a 75%.
- **Intermediário:** índice de transparência igual ou superior a 50% e inferior a 75%.
- **Básico:** índice de transparência igual ou superior a 30% e inferior a 50%.
- **Inicial:** índice de transparência superior a 0% e inferior a 30%.
- **Inexistente:** índice de transparência igual a 0%.

Além disso, os critérios de pontuação foram atualizados, priorizando aspectos

como:

- **Disponibilidade (30%):** garantia de que as informações estejam acessíveis ao público.
- **Atualidade (30%):** atualização frequente dos dados disponibilizados.
- **Série histórica (20%):** disponibilização de dados históricos para análise comparativa.
- **Gravação de relatórios (10%):** geração automática de relatórios para auditoria e controle.
- **Filtro de pesquisa (10%):** facilidade de pesquisa e filtragem de informações no portal.

2.2.3 Importância das Mudanças

A evolução das normas de transparência no TCE/MA reflete a necessidade de adaptação às demandas da sociedade por maior acesso à informação e à prestação de contas (TCE/MA, 2020). A nova classificação, mais detalhada e abrangente, permite uma avaliação mais precisa do nível de transparência dos órgãos públicos, incentivando a melhoria contínua e a adoção de boas práticas (TCE/MA, 2024a).

A mudança para uma escala mais granular, com níveis como **Diamante**, **Ouro** e **Prata**, também visa reconhecer e premiar os órgãos que se destacam na transparência, enquanto os níveis **Básico** e **Inicial** servem como alerta para aqueles que precisam de melhorias. Além disso, a priorização de critérios como **Disponibilidade** e **Atualidade** reforça a importância de manter os dados atualizados e acessíveis, garantindo que a transparência seja efetiva e útil para o cidadão (TCE/MA, 2024a).

Nesse novo contexto, a nota C obtida pela ALEMA em 2021 equivale ao nível **Básico**, que representa um percentual de atendimento entre 40% e 69,99%. Essa mudança reflete a evolução dos critérios de avaliação e a necessidade de adaptação dos órgãos públicos para alcançar padrões mais elevados de transparência. A nova classificação também reforça a importância de melhorias contínuas nos portais de transparência, visando atender às expectativas da sociedade e às exigências do controle externo (TCE/MA, 2024a).

2.3 Processo ETL

ETL vem do inglês *Extract, Transform, Load*, em português, Extração, Transformação e Carga. Segundo Ferreira et al. (2010), "é um processo para extrair dados de um sistema de Banco de Dados (BD), sendo esses dados processados, modificados, e posteriormente inseridos em outro BD". Esse processo é essencial em projetos de integração de dados, como o apresentado neste trabalho.

O processo ETL, é composto por fases principais:

2.3.1 Extração

A extração é a primeira fase do processo ETL, nesta etapa podem ser coletados dados de diversas fontes, sejam estas bancos de dados, APIs etc. Segundo, Kimball e Ross (2002) extrair significa ler e compreender os dados de origem e copiar os dados necessários para o sistema ETL para posterior manipulação. Esta fase deve garantir que os dados sejam capturados de forma consistente e sem perdas, mesmo quando houver um grande volume de informações.

2.3.2 Transformação

Após a extração, os dados passam por um processo de transformação, onde são limpos, validados e formatados para atender aos requisitos do sistema de destino. Isso inclui a remoção de duplicidades, a correção de erros e a conversão de tipos de dados (KIMBALL; ROSS, 2002). A transformação é essencial para garantir a qualidade dos dados, principalmente em ambiente com múltiplas fontes de informação.

2.3.3 Carga

Na fase de carga, os dados transformados são inseridos no sistema de destino. A carga pode ser feita de forma total (substituindo todos os dados existentes) ou incremental (adicionando apenas os dados novos) (KIMBALL; ROSS, 2002). Ferreira et al. (2010) destaca que esse processo é um dos mais difíceis e complexos devido à variedade e diferenças entre os bancos de dados envolvidos.

2.4 Pentaho Data Integration (PDI ou Kettle)

O Pentaho Data Integration (PDI), também conhecido como Kettle, é um software para processos de extração, transformação e carregamento (ETL) de dados. Sendo uma solução baseada em Java, ele roda em qualquer computador que tenha uma *Java Virtual Machine* (JVM) instalada (CASTERS; BOUMAN; DONGEN, 2010). O PDI permite a criação de fluxos de trabalho visuais por meio de sua interface gráfica, chamada *Spoon*.

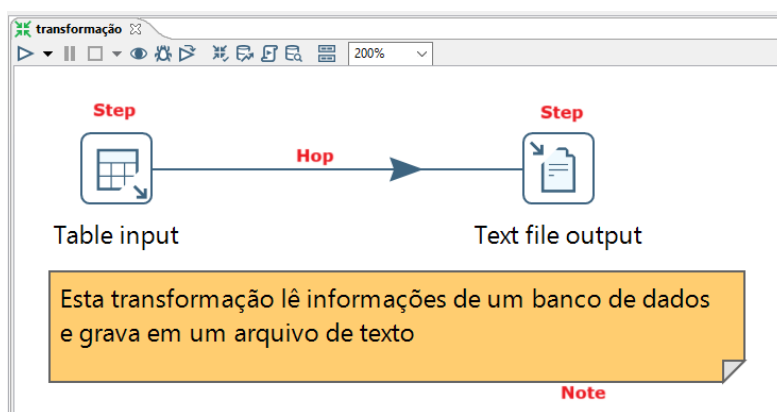
2.4.1 *Spoon*

Casters, Bouman e Dongen (2010) definem a *Spoon* como uma interface gráfica do usuário que permite projetar e gerenciar rapidamente fluxos complexos de processos ETL. Na *Spoon*, no contexto deste trabalho, teremos contato com os seguintes componentes:

2.4.1.1 Transformações

As transformações são consideradas o núcleo de uma solução ETL, elas são compostas por *steps* (etapas), *hops*(conexões) e *notes*(notas), que manipulam dados e definem o fluxo de informações (CASTERS; BOUMAN; DONGEN, 2010) (tradução nossa). A Figura 1 ilustra um processo de transformação de leitura de uma *Table input*(tabela) e escrita em um *Text file output*(arquivo de texto).

Figura 1 – Exemplo de uma transformação simples no *Pentaho Kettle*.



Fonte: Acervo do autor

1. **Steps (Passos):** um *step* é uma das etapas fundamentais de uma transformação. De acordo com Casters, Bouman e Dongen (2010), um *step* pode ser descrito como um bloco básico que realiza operações específicas, como leitura, escrita e processamento de dados. Ele se conecta a outros *steps* por meio de *hops*, definindo o fluxo de dados. Neste trabalho, serão abordados os seguintes tipos de *steps*:

- **Input (Entrada):** responsável pela leitura de dados de fontes externas, como bancos de dados ou arquivos.
- **Output (Saída):** responsável por escrever os dados processados em um destino, como um banco de dados ou arquivo.
- **Transform (Transformação/Conversão):** realiza operações de limpeza, filtragem ou conversão de dados.
- **Flow (Fluxo/Classificação):** define a lógica de fluxo de dados, como condições e bifurcações.

2. **Transformation Hops (Conexões de Transformação):** um *hop* é representado por uma seta que conecta dois *steps*, definindo o caminho dos dados entre eles. Segundo Casters, Bouman e Dongen (2010, p. 26), "um *hop* também representa um *buffer* de linhas chamado *row set*, que armazena temporariamente os dados em trânsito entre os *steps*"(tradução nossa). O tamanho do *row set* pode ser configurado para otimizar o desempenho da transformação.

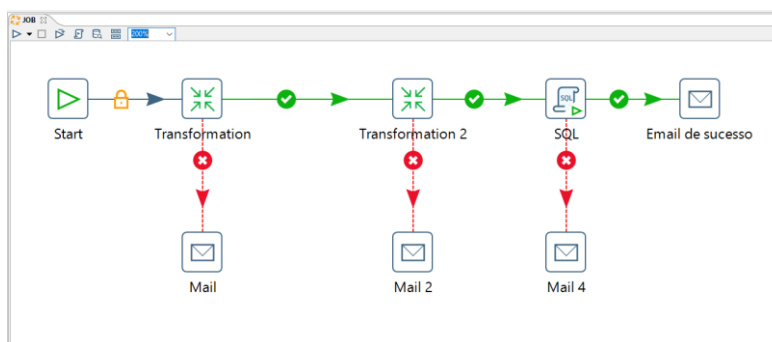
3. **Notes (Notas):** as *notes* são elementos opcionais que permitem adicionar comentários ou documentação diretamente na transformação. Elas são úteis para explicar a lógica de um *step* ou de um conjunto de *steps*, facilitando a compreensão e a manutenção do processo ETL.

2.4.1.2 Jobs (Trabalhos)

De acordo com Casters, Bouman e Dongen (2010), um *job* no *Pentaho Kettle* é um conjunto de tarefas (chamadas de *job entries*) que são executadas em uma ordem específica. Essas tarefas podem incluir ações como enviar e-mails, executar *scripts* ou carregar dados em um banco de dados(por meio de transformações). O *job* garante que essas tarefas sejam

realizadas de forma organizada e sequencial, o que é essencial para processos de manutenção e automação. A Figura 2 ilustra um exemplo de como um *job* pode ser estruturado.

Figura 2 – Exemplo de um *job* no Pentaho Kettle.



Fonte: Acervo do autor

Os principais componentes de um *job* são:

1. **Job Entries (Tarefas de Job):** Uma *job entry* é uma unidade básica de um *job*, responsável por realizar uma tarefa específica. Por exemplo, uma *job entry* pode ser configurada para enviar um e-mail ou executar um comando no sistema. Segundo Casters, Bouman e Dongen (2010, p. 31), "é possível criar cópias de uma *job entry*, chamadas de *shadow copies* (cópia sombra), que permitem reutilizar a mesma tarefa em diferentes partes do *job* sem precisar configurá-la novamente. Além disso, cada *job entry* produz um *result object* (objeto de resultado), que contém informações sobre o sucesso ou falha da execução. Todo *job* começa com uma tarefa especial chamada *Start*, que define o ponto de partida do processo" (tradução nossa).
2. **Job Hops (Conexões de Job):** Os *job hops* são as setas que conectam as *job entries*, definindo a ordem em que as tarefas serão executadas. Casters, Bouman e Dongen (2010), destacam que *job hops* podem ter diferentes regras de execução, como:
 - **Incondicional:** a próxima tarefa é executada independentemente do resultado da anterior.
 - **Seguir quando o resultado for verdadeiro:** a próxima tarefa é executada apenas se a anterior for bem-sucedida.

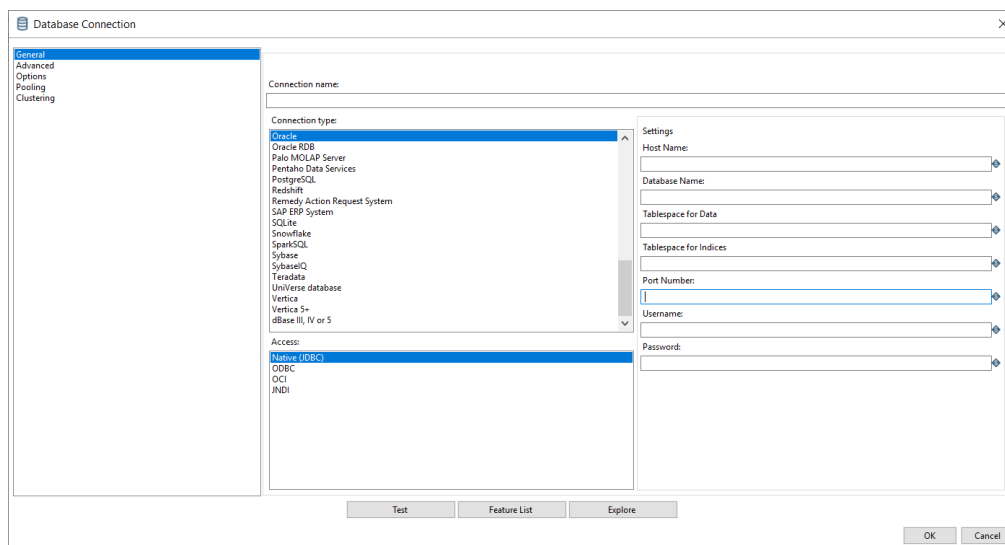
- **Seguir quando o resultado for falso:** a próxima tarefa é executada apenas se a anterior falhar.

Essas regras permitem criar fluxos de trabalho complexos, onde o sistema pode tomar decisões com base no resultado de cada tarefa.

2.4.1.3 Database Connections (Conexões de Base de dados)

As conexões de banco de dados *Kettle* são usadas por transformações e *jobs* para conecta-los a bancos de dados relacionais. De acordo com Casters, Bouman e Dongen (2010, p. 37), "as conexões de banco de dados no *Kettle* são descritores que permitem abrir conexões com bancos de dados relacionais em tempo de execução, com opções específicas para diferentes SGBDs" (tradução nossa). A Figura 3 mostra a tela de configuração de conexão de banco de dados no *Pentaho Kettle*.

Figura 3 – Tela de configuração de conexão com base de dados.



Fonte: Acervo do autor

Os principais aspectos das conexões de banco de dados são:

1. **Configuração de Parâmetros:** Cada conexão de banco de dados requer a definição de parâmetros específicos, como o tipo de SGBD, endereço do servidor, nome do banco de dados,

usuário e senha. Essas configurações permitem que o *Kettle* se conecte a diferentes bancos de dados de forma dinâmica, adaptando-se às necessidades do ambiente de execução (CASTERS; BOUMAN; DONGEN, 2010).

2. **Reutilização de Conexões:** As conexões de banco de dados podem ser reutilizadas em diferentes transformações e *jobs*, evitando a necessidade de reconfiguração. Uma vez que a conexão é definida, esta pode ser referenciada em múltiplos pontos do fluxo de trabalho, o que garantindo consistência e facilitando a manutenção de transformações e *jobs* (CASTERS; BOUMAN; DONGEN, 2010).
3. **Teste de Conexão:** O *Kettle* oferece a funcionalidade de teste de conexão, que verifica se os parâmetros configurados estão corretos e se a conexão com o banco de dados pode ser estabelecida. Esta funcionalidade se mostra essencial para garantir que os erros de conexão sejam identificados antes da execução do fluxo de trabalho (CASTERS; BOUMAN; DONGEN, 2010).

2.4.2 Transformação de dados em tipos Java

O PDI transforma os dados em tipos Java antes de realizar o input em uma tabela ou em qualquer outra operação. Isso ocorre porque o PDI é implementado em Java, e todas as operações internas são baseadas nos tipos de dados da linguagem Java. Ao converter os dados para tipos Java, o PDI garante que o processamento seja uniforme e consistente em todo o fluxo ETL, independentemente do formato original dos dados nas fontes (CASTERS; BOUMAN; DONGEN, 2010).

Essa abordagem oferece várias vantagens:

- **Consistência:** Ao utilizar tipos Java, o PDI padroniza o tratamento dos dados, evitando inconsistências que poderiam surgir devido a diferenças entre formatos de origem.
- **Portabilidade:** Como Java é uma linguagem multiplataforma, os dados podem ser processados em qualquer ambiente que suporte uma *Java Virtual Machine* (JVM), garantindo flexibilidade e portabilidade.

- **Integração:** A conversão para tipos Java facilita a integração com outras bibliotecas e ferramentas baseadas em Java, ampliando as possibilidades de processamento e análise.

Essa transformação é fundamental para o funcionamento eficiente e confiável do PDI em cenários complexos de integração de dados.

2.5 *Pentaho Server*

O *Pentaho Server* é o ambiente responsável pela execução automatizada de *jobs* e pela publicação de *dashboards*, garantindo que os dados sejam atualizados periodicamente. Para o projeto, é fundamental compreender os seguintes elementos que compõem o *Pentaho Server*:

2.5.1 *User Console* (Interface do Usuário)

A *User Console* é a interface gráfica web do *Pentaho Server*, que permite o acesso via *login* a funcionalidades como o agendador de tarefas, o repositório de arquivos e a visualização de *dashboards*. Essa interface é intuitiva e centraliza todas as operações do servidor, facilitando o gerenciamento de *jobs* e transformações.

2.5.2 *Scheduler* (Agendador)

O *Scheduler* (agendador) é uma ferramenta essencial para automatizar a execução de *jobs* e transformações. De acordo com Casters, Bouman e Dongen (2010, p. 322), "o *Quartz Scheduler*, integrado ao *Pentaho BI Server*, pode ser usado para executar sequências de ações que incluem a execução de transformações e *jobs* do *Kettle*"(tradução nossa). O agendador permite configurar horários e frequências para a execução de tarefas, como atualizações diárias de dados ou geração de relatórios periódicos.

2.6 Sistema gerenciador de banco de dados (SGBD)

Um sistema gerenciador de banco de dados (SGBD) é um *software* que controla o armazenamento, organização e recuperação de dados (ORACLE, 2025).

2.6.1 SGBD Oracle

O *Oracle Database* é um Sistema Gerenciador de Banco de Dados Relacional (SGBDR) desenvolvido pela *Oracle Corporation*. Ele é amplamente utilizado em ambientes corporativos devido à sua escalabilidade, segurança e suporte a grandes volumes de dados. O Oracle é frequentemente utilizado em aplicações críticas, como sistemas financeiros e de gestão empresarial (ORACLE, 2025).

2.6.2 SGBD PostgreSQL

O PostgreSQL é um SGBD desenvolvido na Universidade da Califórnia, no Departamento de Ciência da Computação de *Berkeley*. O PostgreSQL foi pioneiro em muitos conceitos que só se tornaram disponíveis em alguns sistemas de banco de dados comerciais muito mais tarde (POSTGRESQL, 2025).

PostgreSQL é um descendente de código aberto deste código original de *Berkeley*. Por ser *open source*, pode ser usado, modificado e distribuído por qualquer pessoa gratuitamente para qualquer finalidade, seja privada, comercial ou acadêmica (POSTGRESQL, 2025).

Este capítulo apresentou os principais conceitos e ferramentas que fundamentam a proposta de integração de dados entre sistemas heterogêneos, com destaque para o processo ETL e as particularidades dos SGBDs Oracle e PostgreSQL. A discussão sobre a evolução das normas de transparência no TCE/MA e a utilização do Pentaho Data Integration (PDI) como ferramenta de integração de dados proporcionou uma base sólida para o entendimento dos desafios técnicos envolvidos no projeto. Com essa fundamentação teórica estabelecida, o próximo capítulo, Metodologia, detalhará as etapas práticas adotadas para a implementação da solução, desde a análise do problema até a migração e publicação dos dados no Portal da Transparência da ALEMA, demonstrando a aplicação dos conceitos discutidos neste capítulo.

3 METODOLOGIA

Neste capítulo serão descritos os procedimentos realizados pela metodologia proposta para criação da solução de ETL para migração e disponibilização dos dados no Portal da Transparência da ALEMA. Para tanto, serão apresentados a descrição do problema, as etapas da metodologia propriamente dita e os resultados alcançados.

3.1 Descrição do problema

O problema central deste trabalho está relacionado à integração e migração de dados financeiros e orçamentários entre dois sistemas heterogêneos: o banco de dados Oracle, utilizado pela Secretaria de Estado de Planejamento e Orçamento (SEPLAN), e o banco de dados PostgreSQL, adotado pela Assembleia Legislativa do Estado do Maranhão (ALEMA). Como discutido no Capítulo 2, Fundamentação Teórica, a transparência pública é um princípio fundamental para a governança e o controle social, e a Lei de Acesso à Informação (LAI) exige que os órgãos públicos disponibilizem dados de forma clara, acessível e atualizada. No entanto, a ALEMA enfrenta desafios técnicos significativos para atender a essas exigências, principalmente devido às diferenças entre os sistemas de banco de dados utilizados.

O Sistema Integrado de Planejamento e Gestão Fiscal (SIGEF), mantido pela SEPLAN, armazena os dados financeiros e orçamentários em um banco de dados Oracle. Esses dados são disponibilizados por meio de views, que são estruturas de consulta que simplificam o acesso às informações. No entanto, a ALEMA utiliza um banco de dados PostgreSQL para gerenciar suas informações, o que gera incompatibilidades técnicas. Como destacado na Fundamentação Teórica, o Oracle e o PostgreSQL possuem diferenças significativas em termos de tipos de dados, estrutura de tabelas, linguagem SQL e gerenciamento de transações. Por exemplo, tipos de dados como VARCHAR2 no Oracle precisam ser convertidos para VARCHAR no PostgreSQL, e campos numéricos como NUMBER exigem ajustes para serem compatíveis com o tipo FLOAT do PostgreSQL. Essas diferenças tornam complexa a migração e a integração dos dados.

Além disso, a SEPLAN adota a prática de criar um novo banco de dados a cada ano, o que resulta em uma fragmentação dos dados. Essa fragmentação dificulta a consulta e a análise de informações históricas, já que os dados estão distribuídos em múltiplos bancos de dados. Para o Portal da Transparência da ALEMA, isso significa que os dados precisam ser consolidados e integrados de forma consistente, o que aumenta a complexidade do processo de migração. A fragmentação também impacta a eficiência do portal, que deve fornecer dados atualizados e de fácil acesso ao público.

Outro desafio técnico é a necessidade de garantir a consistência e a integridade dos dados durante o processo de migração. Dados financeiros e orçamentários são críticos, e qualquer erro ou inconsistência pode comprometer a confiabilidade das informações disponibilizadas. Como discutido na Seção 2.3 Processo ETL, na Fundamentação Teórica, o processo de ETL (Extração, Transformação e Carga) é essencial para garantir a qualidade dos dados, especialmente em ambientes com múltiplas fontes de informação. Portanto, é fundamental desenvolver um fluxo de ETL robusto, que não apenas migre os dados, mas também os valide e ajuste para garantir que estejam corretos e prontos para uso no PostgreSQL.

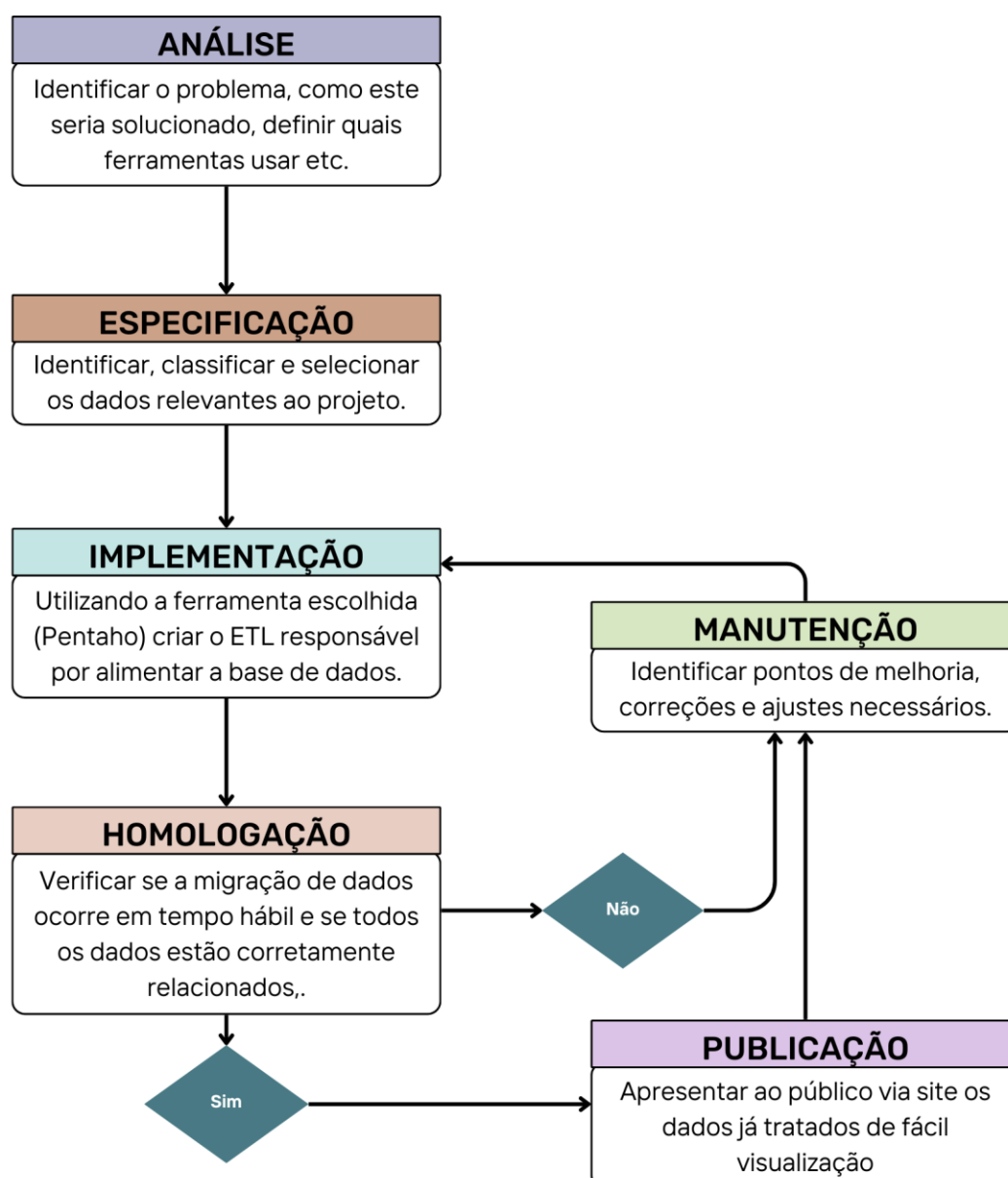
A questão "Como unificar os dados disponibilizados pela SEPLAN, armazenados em Oracle, e exibi-los de forma clara e atualizada no Portal da Transparência da ALEMA, utilizando ferramentas de código aberto?", levantada no Capítulo de Introdução deste trabalho envolve a superação de desafios técnicos relacionados à integração de sistemas heterogêneos, a fragmentação de dados e a garantia de consistência e atualização contínua das informações. A solução proposta, baseada no uso do Pentaho Data Integration (PDI) e em técnicas de ETL, visa criar um fluxo automatizado que permita a migração eficiente dos dados do Oracle para o PostgreSQL, garantindo que o Portal da Transparência da ALEMA possa exibir informações atualizadas e confiáveis ao público, em conformidade com as exigências de transparência estabelecidas pelo TCE/MA (TCE/MA, 2024a).

3.2 Etapas da Metodologia

Em junho de 2022 se iniciaram os esforços para a resolução do problema apresentado na seção anterior. Contudo, antes de fazer qualquer migração era necessário identificar, classificar e selecionar quais dados eram necessários e como eles se

relacionavam, para que fosse criado o esquema dentro da base de dados da ALEMA, que seria usado para as consultas no portal. Para isto foram seguidas as seguintes etapas da metodologia proposta, ilustradas na Figura 4 que são: Análise, Especificação, Implementação, Homologação, Publicação, Manutenção.

Figura 4 – Etapas seguidas para desenvolvimento do projeto.



Fonte: Acervo do autor.

3.2.1 ANÁLISE

A primeira etapa da metodologia foi voltada a compreender do problema apresentado e identificar ferramentas e tecnologias que poderiam ser utilizadas para solucioná-lo. Além de fazer o levantamento e documentação dos requisitos necessários para o desenvolvimento do projeto.

Durante esta etapa foi identificada a necessidade de desenvolver um processo ETL capaz de unificar todas as views do Oracle disponibilizadas pelo TCE/MA em uma base de dados PostgreSQL. Essa integração foi considerada essencial para que o módulo de consulta de despesas do Portal da Transparência pudesse consumir os dados de forma eficiente e disponibilizá-los ao público.

Os requisitos do ETL mapeados nesta etapa são apresentados na Tabela 1.

Tabela 1 – Detalhes dos requisitos mapeados na etapa de Análise

Requisito	Descrição
Fonte de Dados	Os dados financeiros e orçamentários são provenientes do Sistema Integrado de Planejamento e Gestão Fiscal (SIGEF), mantido pela Secretaria de Estado de Planejamento e Orçamento (SEPLAN), e estão armazenados em um banco de dados Oracle. As informações são disponibilizadas por meio de <i>views</i> .
Frequência de Atualização	Processo diário com atualizações programadas para às 00:00h, garantindo informações atualizadas no Portal da Transparência.
Mapeamento de Dados	Identificação de tabelas e <i>views</i> relevantes no Oracle e mapeamento detalhado dos relacionamentos para estrutura PostgreSQL
Compatibilidade entre Sistemas	Conversão de tipos de dados (ex: VARCHAR2 para VARCHAR, NUMBER para FLOAT) e ajuste de formatos.

Automatização do Processo	ETL totalmente automatizado utilizando <i>Pentaho Data Integration</i> (PDI) para atualizações sem intervenção manual.
Validação e Consistência	Verificação de contagem de registros, consistência de valores e integridade referencial após migração.
Desempenho e Escalabilidade	Solução capaz de lidar com grandes volumes de dados e fragmentação de bancos anuais, com otimização para consultas rápidas.
Documentação e Rastreabilidade	Documentação completa do processo e sistema de rastreabilidade para auditoria e transparência.

Fonte: Elaborada pelo autor (2025).

Esses requisitos foram fundamentais para orientar o desenvolvimento do projeto, garantindo que a solução proposta atendesse às necessidades técnicas e operacionais da ALEMA, além de cumprir com as exigências de transparência estabelecidas pelo TCE/MA. A próxima etapa, Especificação, detalhará as estratégias e ferramentas escolhidas para atender a esses requisitos.

3.2.2 ESPECIFICAÇÃO

A segunda etapa da metodologia teve como objetivo a definição das soluções e ferramentas que atendem aos requisitos identificados na fase de análise. Nesta etapa, foram especificadas as estratégias e tecnologias a serem utilizadas.

Dessa maneira, optou-se pela ferramenta Pentaho BI, uma suíte *open source* de *Business Intelligence*. Mais especificamente os módulos *Pentaho Data Integration* (PDI ou *Kettle*) e *Pentaho Server*.

A escolha do *Pentaho BI* foi baseada nas seguintes razões: conectividade com várias fontes de dados; interface gráfica amigável; automação eficiente; transformações flexíveis e personalizáveis.

Além de escolher a ferramenta para o processo ETL, essa etapa também incluiu o desenho do esquema do banco de dados do módulo de consulta de notas de empenho do Portal da

Transparência nomeado sigef(em minúsculas), mantendo a grafia original do Sistema Integrado de Planejamento e Gestão Fiscal (SIGEF), fonte primária dos dados, e seguindo convenções técnicas que preferem nomes em minúsculas para esquemas de banco de dados.

Tendo em vista o requisito de mapeamento de dados apresentado na Subseção

3.2.1 *Álise*, os seguintes processos foram executados:

3.2.2.1 Mapeamento e criação de tabelas

Nesta subetapa foram identificadas as *views* necessárias ao sigefe foram criadas tabelas correspondentes no PostgreSQL, para o armazenamento e posterior consulta dos dados pelo portal da transparência. O esquema de cada tabela foi definido com base na estrutura da *view* Oracle, mas com ajustes para atender às necessidades do projeto. Exemplos de mapeamento:

- VSEPLANNELEG(*view*) → nota_empenho(tabela).
- VSSEPLANLLEG(*view*) → nota_lancamento(tabela).
- VPORTALCREDOR(*view*) → credor(tabela).
- VSEPLANRAPNAOPROCLEG(*view*) → restos_a_pagar_nao_processados (tabela).

Dentro do mapeamento das tabelas houveram ainda as seguintes subetapas adicionais:

a) **Mapeamento de Colunas:** nesta etapa as colunas da *view* Oracle foram mapeadas para as colunas correspondentes na tabela do PostgreSQL. Os nomes das colunas foram ajustados para seguir a convenção *snake_case* do PostgreSQL. Exemplos de mapeamento de nomes de colunas da *view* VSEPLANNELEG para a tabela nota_empenho:

- NUNOTAEMPENHO(*view*) → num_nota_empenho(tabela): campo que guarda o número de uma nota de empenho;
- DEHISTORICO(*view*) → historico(tabela): campo que armazena uma breve descrição do que se trata a nota de empenho;

- VLNOTAEMPENHO(*view*) → valor_nota_empenho(tabela): campo que armazena o valor total da nota de empenho;
- CDCREDOR(*view*) → id_credor(tabela): campo que guarda o id do credor referente a tabela credor;
- DTLANCAMENTO(*view*) → data_lancamento(tabela): campo que guarda a data que a nota de empenho foi lançada no SIGEF;
- NUPROCESSO(*view*) → num_processo(tabela): campo que guarda o número do processo que gerou a nota de empenho.

b) **Mapeamento dos tipos de dados:** nesta etapa os tipo de dados Oracle foram mapeados para os equivalentes no PostgreSQL, Tabela 2:

Tabela 2 – Mapeamentos de tipos de dados Oracle para tipos equivalentes no PostgreSQL

Atributo Oracle	Atributo PostgreSQL	Observação
VARCHAR2	VARCHAR	Campos de texto
NUMBER	FLOAT	Campos referentes a valores monetários
DATE	TIMESTAMP	Campos de datas

Fonte: Elaborada pelo autor (2025).

A Figura 5 ilustra a *view* VSEPLANNELEG antes do processo de mapeamento e criação de tabela e o produto deste processo, a tabela nota_empenho.

Figura 5 – Exemplo do mapeamento da view nota de empenho do Oracle para a tabela nota de empenho no PostgreSQL.

VSEPLANNELEG		nota_empenho	
CDUNIDADEGESTORA	t	id	#
CDGESTAO	t	num_nota_empenho	t
NUNOTAEMPENHO	t	valor_nota_empenho	#
NUNEORIGINAL	t	num_nota_empenho_original	t
CDCREDOR	t	tipo	c
DTLANCAMENTO	t	id_subacao	↗
CDEVENTO	#	id_natureza	↗
TIPO	c	id_unidade_orcamentaria	#
CDSUBACAO	t	id_fonte	↗
CDUNIDADEORCAMENTARIA	t	id_modalidade_empenho	↗
CDNATUREZADESPESA	t	id_modalidade_licitacao	↗
CDPROGRAMA	t	id_programa	↗
CDFONTE	t	nome_modalidade_empenho	t
CDMODALIDADELICITACAO	t	num_processo	t
CDMODALIDADEEMPENHO	t	historico	t
DEMODALIDADEEMPENHO	t	id_unidade_gestora	↗
NUPROCESSO	t	id_gestao	#
VLNOTAEMPENHO	t	id_credor	t
NUQUANTIDADE	t	data_lancamento	d
DEHISTORICO	t	quantidade	#
		id_evento	↗

(View : VSEPLANNELEG)

(Tabela : nota_empenho)

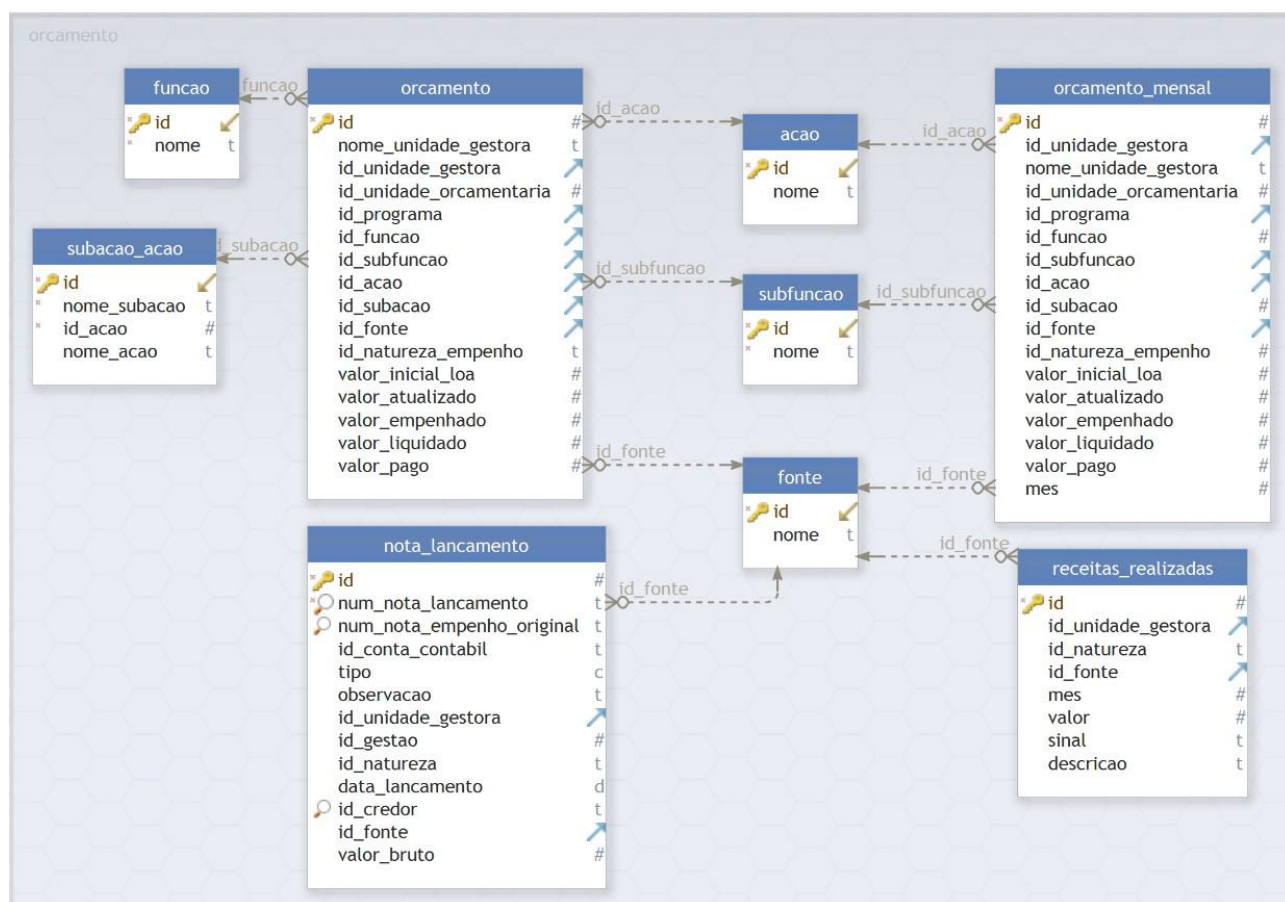
Fonte: acervo do autor.

c) **Criação do esquema - sigef PostgreSQL:** após a criação de todas as tabelas PostgreSQL correspondentes as *views* Oracle o desenho do esquema ficou como ilustrado abaixo nas figuras 6 a 9, para contexto de explicação separamos o esquema nos seguintes grupos:

- **Grupo Orçamento:** o "Grupo Orçamento"(Figura 6) refere-se à gestão e execução orçamentária das unidades gestoras. As tabelas desse grupo armazenam informações sobre o planejamento financeiro, execução mensal e classificações orçamentárias. O objetivo deste grupo de tabelas é armazenar dados de monitoramento e controle dos valores orçamentários planejados ao longo do

exercício fiscal.

Figura 6 – Grupo Orçamento.



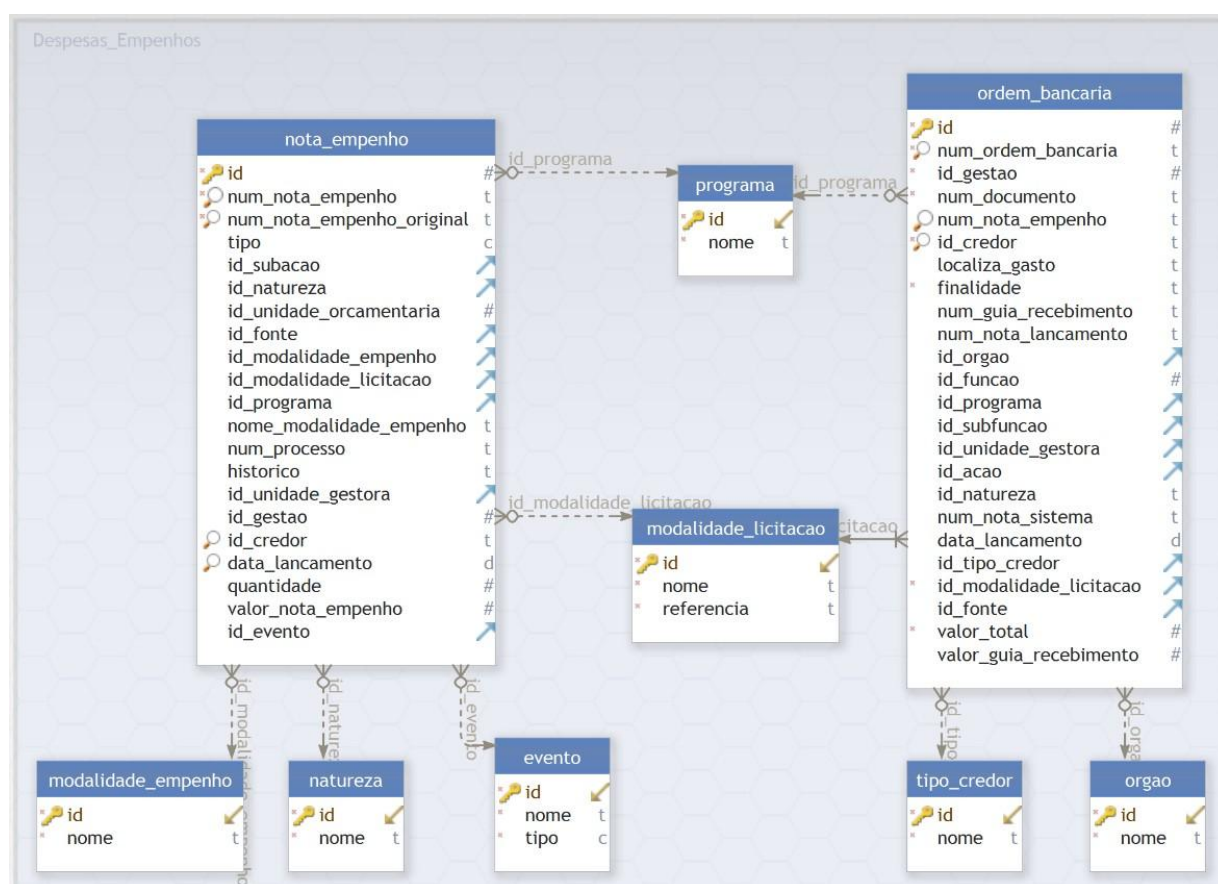
Fonte: Acervo do autor.

As seguintes tabelas compõem este grupo:

- **Tabela orçamento**: representa o orçamento de unidades gestoras. Contém informações como: Unidade gestora e orçamentária (id_unidade_gestora, id_unidade_orcamentaria); Programa, subfunção e ação vinculados ao orçamento; Valores orçamentários, como o inicial(valor_inicial_loa), atualizado(valor_atualizado), empenhado, liquidado e pago.
- **Tabela orçamento_mensal**: mostra a execução do orçamento em períodos mensais. Relaciona-se à tabela orçamento através da unidade gestora.
- **Tabela nota_lancamento**: armazena detalhes sobre lançamentos específicos, como: Notas de empenho relacionadas(num_nota_empenho_original);Data e valor do lançamento(data_lancamento, valor_bruto).

- **Tabelas auxiliares:** incluem as tabelas funcao, subacao, acao, subfuncao, e fonte, que fornecem suporte ao orçamento com informações classificatórias, como funções e subfunções relacionadas.
- **Grupo Despesas e empenhos:** o "Grupo Despesas e empenhos" (Figura 7) é focado na gestão das despesas realizadas e dos empenhos emitidos pelas unidades gestoras, abrangendo desde o registro de lançamentos até o controle de restos a pagar. Tem como objetivo garantir o acompanhamento detalhado dos gastos públicos e compromissos financeiros.

Figura 7 – Grupo Despesas e empenhos.



Fonte: Acervo do autor.

Esse grupo é composto pelas seguintes tabelas:

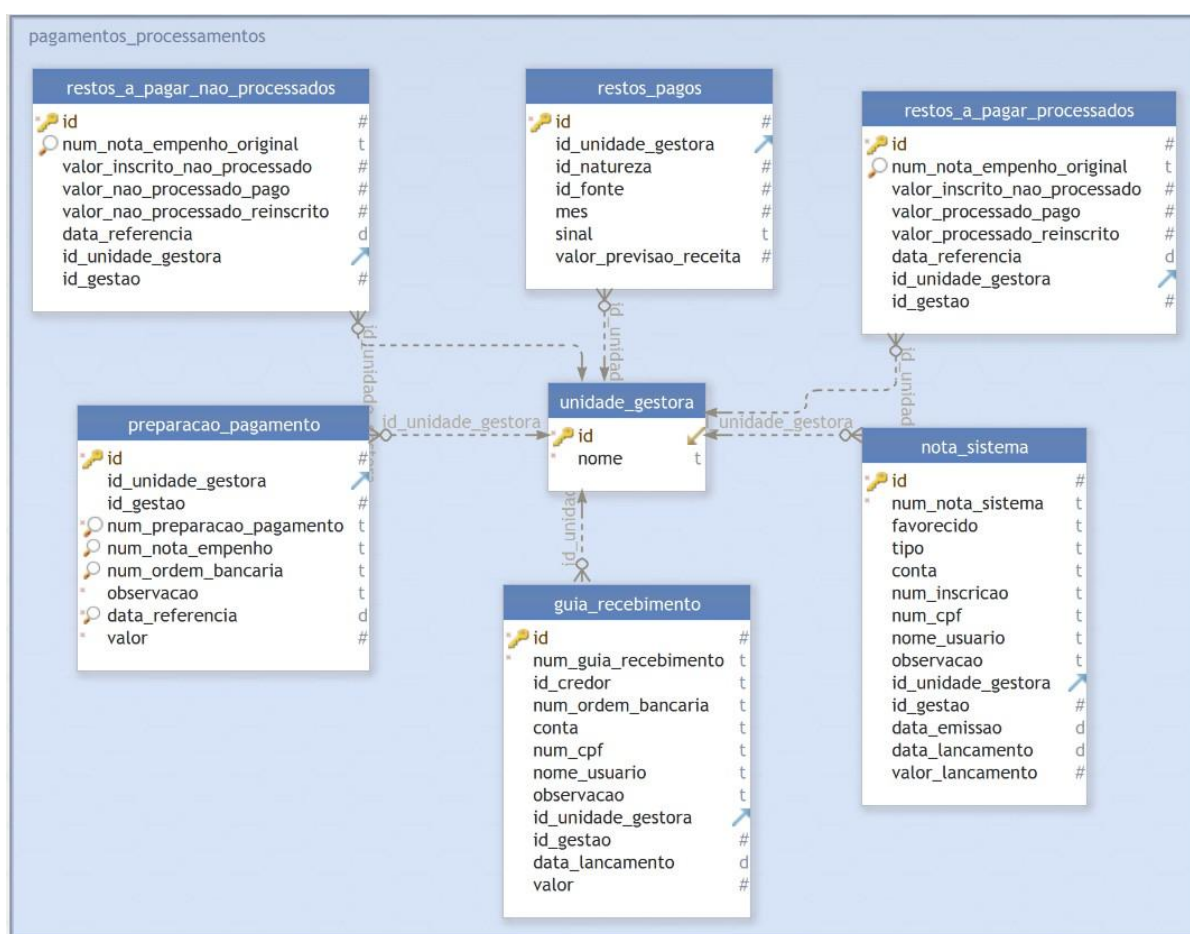
- **Tabela nota_lancamento:** armazena os registros das notas de lançamento de despesas. Tem como principais atributos: guarda Número identificador da nota de lançamento(num_nota_lancamento); Referência ao empenho relacionado(num_nota_empenho_original); Unidade gestora responsável(id_unidade_gestora); Data em que o lançamento foi

realizado(data_lancamento); Valor total do lançamento(valor_bruto).

- **Tabela restos_a_pagar_nao_processados:** registra despesas empenhadas que ainda não foram processadas. Seus principais atributos são: Referência ao empenho não processado(num_nota_empenho_original); Valor inscrito como restos a pagar não processados(valor_inscrito_nao_processado); Valor efetivamente pago de restos não processados (valor_nao_processado_pago); Data de referência do registro(data_referencia).
- **Tabela restos_a_pagar_processados:** contém informações de despesas processadas, mas ainda não pagas. Tem como principais atributos: Valor processado do empenho(valor_processado); Valor reinscrito como restos a pagar processados(valor_reinscrito); Data de referência(data_referencia).
- **Tabela restos_pagos:** registra os pagamentos efetuados relacionados a restos a pagar. Seus principais atributos são: Unidade gestora responsável pelo pagamento(id_unidade_gestora); Valor previsto para cobertura das despesas(valor_previsao_receita); Mês em que o pagamento foi realizado(mes).
- **Tabela preparacao_pagamento:** gerencia o fluxo de preparação de pagamentos. Os principais atributos são: Número da ordem bancária vinculada ao pagamento(num_ordem_bancaria); Número do empenho relacionado(num_nota_empenho); Valor a ser pago(valor); Observações adicionais(observacao).
- **Tabela credor_nota_empenho:** relaciona os credores com os empenhos a eles vinculados. Tem como principais atributos: Identificação do credor(id_credor); Valor original do empenho(empenho_originario); Valor já liquidado do empenho(valor_liquidado); Valor efetivamente pago(valor_pago).

- **Grupo Pagamentos e processamentos:** o "Grupo Pagamentos e processamentos" (Figura 8) foca no processamento e controle de pagamentos realizados, abrangendo etapas desde restos a pagar até a execução de pagamentos finais. Seu objetivo é controlar o fluxo de pagamentos desde a preparação até a quitação, garantindo rastreabilidade e eficiência.

Figura 8 – Grupo Pagamentos e processamentos.



Fonte: Acervo do autor.

As tabelas listadas abaixo compõe o grupo:

- **Tabela restos_a_pagar_nao_processados:** detalha valores empenhados, mas ainda não processados, sendo importante para identificar despesas pendentes.
- **Tabela restos_a_pagar_processados:** contém os valores que já foram empenhados e processados, mas ainda aguardam pagamento.
- **Tabela restos_pagos:** lista os valores já liquidados, e inclui dados como: Natureza da receita e fonte; e previsão e valores pagos mensalmente.

- **Tabela preparacao_pagamento:** registra os detalhes de preparação do pagamento, incluindo observações e valores a serem pagos.
- **Tabela nota_sistema:** Abrange informações sobre as notas emitidas no sistema.
- **Grupo Programas e ações:** o "Grupo Programas e ações" (Figura 9) organiza e classifica os dados relacionados a programas, ações, projetos e lançamentos contábeis, além de gerenciar informações de credores. Tem como objetivo classificar e organizar as despesas públicas de acordo com objetivos orçamentários e registrar informações detalhadas sobre credores e lançamentos.

Figura 9 – Grupo Programas e ações.



Fonte: Acervo do autor.

Como ilustrado na Figura 9 ele agrupa as seguintes tabelas:

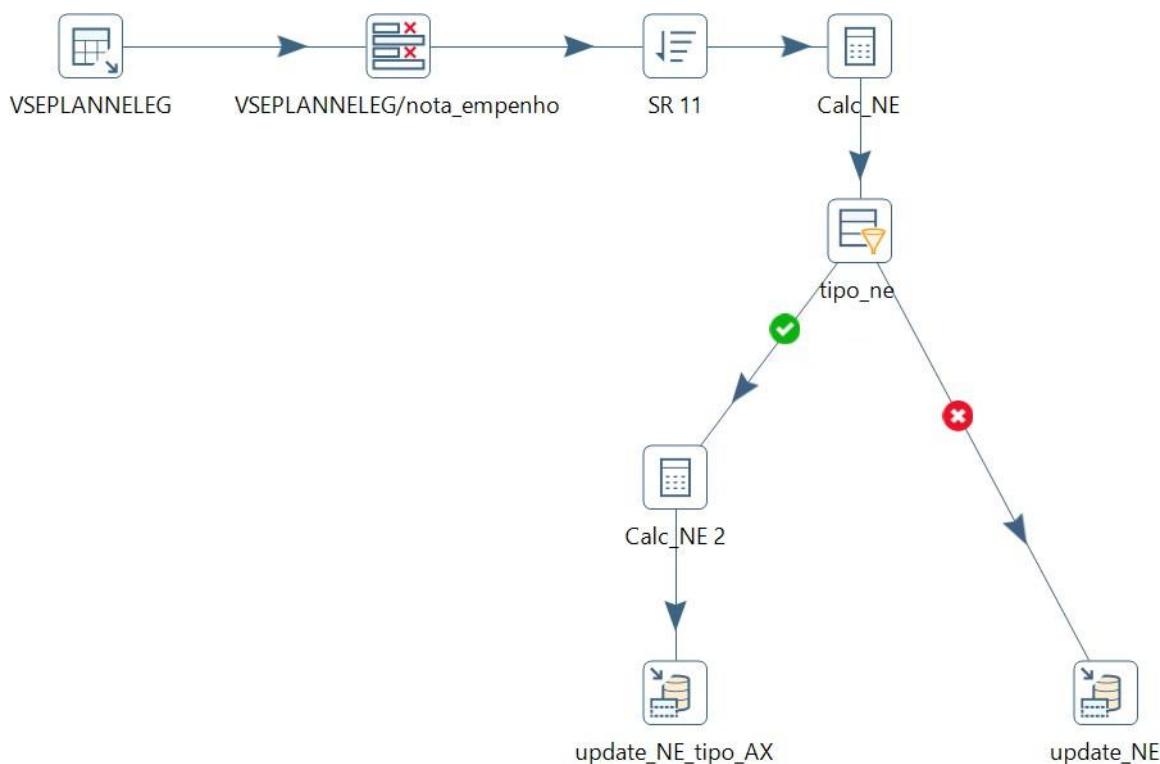
- **Tabela lançamento_contabil:** representa os lançamentos contábeis, inclui: Evento associado e conta contábil (id_evento, id_conta_contabil); Valores lançados e origem dos recursos.
- **Tabela credor:** contém informações detalhadas sobre os credores (nome, CPF/CNPJ, endereço, etc.).
- **Tabelas projeto e subacao:** classificam os projetos e subações vinculados ao orçamento.
- **Tabela credor_nota_empenho:** liga os credores aos empenhos e registra valores empenhados, liquidados, pagos e pendentes.

3.2.3 IMPLEMENTAÇÃO

Esta etapa marca o início da construção prática do projeto. Aqui, o processo de ETL, as integrações, e todas as funcionalidades previamente especificadas começam a ser desenvolvidas. As configurações do Pentaho Data Integration e do Pentaho Server são realizadas, além da criação de *scripts* e fluxos necessários para atender aos requisitos de mapeamento, transformação e carregamento dos dados.

Usaremos como exemplo a transformação da *view* VSEPLANNELEG para a tabela nota_empenho (Figura 10) a fim de explicar as etapas seguidas para a criação do fluxo ETL e migração dos dados das views SEPLAN para o esquema PostgreSQL.

Figura 10 – Transformação da VSEPLANNELEG para a tabela nota_empenho.



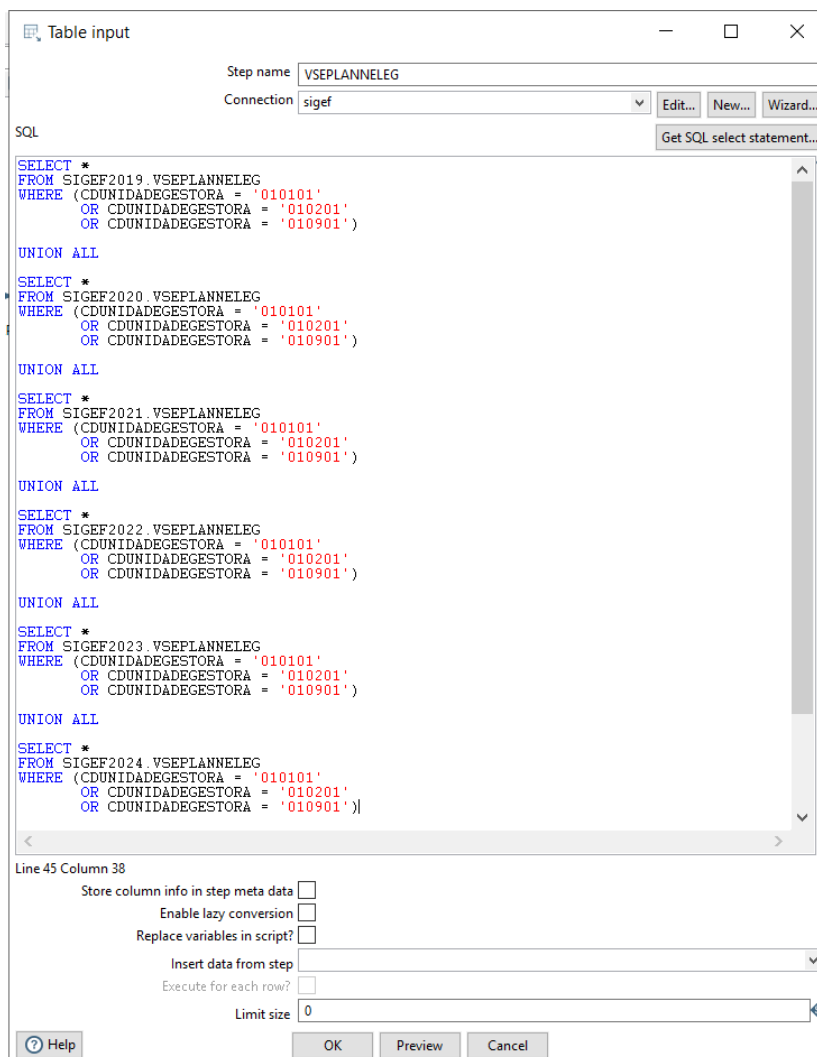
Fonte: Acervo do autor.

3.2.3.1 Extração dos Dados

A extração dos dados foi realizada utilizando o *Pentaho Data Integration (PDI/Kettle)*, que se conectou a um banco de dados do SIGEF para extrair informações de *views* predefinidas. O processo empregou a *Step Input "Table Input"*, responsável por executar consultas SQL diretamente nas *views* e capturar todos os campos necessários, incluindo metadados, para alimentar as etapas seguintes de transformação.

A Figura 11 ilustra a configuração da Table Input - VSEPLANNELEG, que estabelece a conexão com o Oracle e extrai dados da *view* VSEPLANNELEG, responsável por armazenar registros de notas de empenho. Para garantir a cobertura temporal, a consulta SQL utilizada combinou dados de múltiplos anos (2019 a 2024) por meio de cláusulas UNION ALL, consolidando os resultados em um único fluxo de dados para processamento posterior.

Figura 11 – Table input: VSEPLANNELEG.



Fonte: Acervo do autor.

3.2.3.2 Transformação dos Dados

A etapa de transformação é essencial para garantir a compatibilidade dos dados entre os sistemas Oracle (origem) e PostgreSQL (destino). O Pentaho Data Integration (PDI) adota uma abordagem intermediária onde converte primeiro os dados para tipos Java antes da conversão final para os tipos do banco de destino. Neste projeto, essa conversão foi feita na subetapa de mapeamento e conversão de tipos de dados, para posteriormente seguirem para as subetapas de classificação e de formatação.

- a) **Mapeamento e conversão de tipos de Dados:** esta etapa teve como objetivo mapear e converter os tipos de dados do Oracle para tipos compatíveis com o PostgreSQL. Para isto foi usada a *Step Transform "Select values"* como intermediadora, nomeada

"VSEPLANNELEG/nota_empenho", o mapeamento e seleção das colunas é feito na aba "Select&Alter" Figura 12.

Figura 12 – Aba: *Select&Alter*.

Select values

Step name: VSEPLANNELEG/nota_empenho

Select & Alter Remove Meta-data

Fields:

#	Fieldname	Rename to	Length	Precision
1	CDUNIDADEGESTORA			
2	CDGESTAO			
3	NUNOTAEMPENHO			
4	VLNOTAEMPENHO			
5	NUNEOORIGINAL			
6	CDCREDOR			
7	DTLANCAMENTO			
8	TIPO			
9	CDSUBACAO			
10	CDUNIDADEORCAMENTARIA			
11	CDNATUREZADESPESA			
12	CDPROGRAMA			
13	CDFONTE			
14	CDMODALIDADELICITACAO			
15	CDMODALIDADEEMPENHO			
16	DEMODALIDADEEMPENHO			
17	NUPROCESSO			
18	NUQUANTIDADE			
19	DEHISTORICO			
20	CDEVENTO			

Get fields to select
Edit Mapping

Include unspecified fields, ordered by ☐

Help OK Cancel

Fonte: Acervo do autor.

A conversão dos tipos de dados é feita na aba "Meta-data", Figura 13, da seguinte forma:

- Valores monetários (ex: VLNOTAEMPENHO) é convertido na coluna *Type* para o tipo *BigNumber*, visto que esse valores são extensos, e com configuração de *locale* (pt_BR) para garantir a correta interpretação dos decimais e visando garantir a precisão nos cálculos;
- A opção "Binary to Normal?" desativada indica que não há necessidade de conversão de dados binários;
- Campos de data foram padronizados para o tipo *Timestamp* no formato ddMMyyyy, aceito pelo PostgreSQL e configuração de *locale* (pt_BR) para

assegurar a consistência cultural do formato.

Figura 13 – Aba: *Meta-data*.

#	Fieldname	Rename to	Type	Length	Precision	Binary to Normal?	Format	Date Format Lenient?	Date Locale
1	CDUNIDADEGESTORA		Integer	6		N		N	
2	CDGESTAO		Integer	5		N		N	
3	CDSUBACAO		Integer	6		N		N	
4	CDUNIDADEORCAMENTARIA		Integer	5		N		N	
5	CDNATUREZADESPESA		BigNumber	9		N		N	
6	CDFONTE		String	10		N		N	
7	CDMODALIDADELICITACAO		Integer			N		N	
8	CDMODALIDADEEMPENHO		Integer			N		N	
9	NUQUANTIDADE		Integer	6		N		N	
10	DTLANCAMENTO		Timestamp			N	ddMMyyyy	N	pt_BR
11	VLNOTAEMPENHO		Number			N	###,###,###	N	pt_BR
12	CDPROGRAMA		Integer			N		N	

Fonte: Acervo do autor.

b) **Classificação:** os objetivos da etapa de classificação são: Tratamento de valores nulos para evitar erros na carga; Correção de inconsistências nos dados, como remoção de espaços extras ou ajuste de códigos inválidos; Ordenação de dados, para auxiliar na etapas seguintes; Filtragem de dados, para formatação, descarte ou classificação. Nesta etapa foram utilizadas as *Step Flow* "*Sort rows*" e "*Filter rows*", para ordenação e agrupamento, e definição do fluxo de tratamento dos dados respectivamente.

- *Sort rows*: a "*Sort rows*", nomeada SR 11, Figura 14, é usada para ordena dados afim de facilitar a filtragem, agregação e/ou junção com outros conjuntos de dados. Nesta transformação essa *Step* foi usada para ordenação dos dados para processamento no *Filter rows*.
- *Filter rows*: a *Step* "*Filter rows*" nomeada tipo_ne, é utilizada para filtrar notas de empenho entre Originário(O) e Reforço(R), que possuem valores positivos, e Anulação(A) e Devolução(X), pois as notas dos tipos A e X devem ser processadas para dar entrada no banco PostgreSQL com valores negativos. Na tipo_ne colocamos a condição: tipo = A OR tipo = X, como ilustrado na Figura 15. Caso a nota processada se encaixe na condição ela segue o fluxo positivo(check verde)e passa pela formatação da Step Calculator(Calc_NE 2), caso a nota não se encaixe na condição ela segue o fluxo negativo(X vermelho) direto para a etapa de carga de dado, fluxo ilustrado na Figura 10.

Figura 14 – *Sort rows*: SR 11.

Step name: SR 11

Sort directory: %%java.io.tmpdir%% Browse...

TMP-file prefix: out

Sort size (rows in memory): 1000000

Free memory threshold (in %):

Compress TMP Files? ☒

Only pass unique rows? (verifies keys only) ☐

Fields:

#	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?
1	NUNOTAEMPENHO	Y	N	N	0	N
2	TIPO	N	N	N	0	N

Help OK Cancel Get Fields

Fonte: Acervo do autor.

Figura 15 – *Filter rows*: tipo_ne.

Step name: tipo_ne

Send 'true' data to step: Calc_NE 2

Send 'false' data to step: update_NE

The condition:

☐ +

OR

tipo = [A]

tipo = [X]

Help OK Cancel

Fonte: Acervo do autor.

c) **Formatação:** na etapa de Formatação, os valores numéricos monetários foram ajustados, pois ao observar os dados na saída da etapa "VSEPLANNELEG/nota_empenho", verificou-se que os valores, mesmo após a transformação vem sem vírgula separando corretamente as casas decimais e todos os valores positivos. Portanto se fez necessário aplicar uma etapa que ajustasse as casas decimais e fizesse a conversão de valores das notas de empenho do tipo A e X em valores negativos. Nesta *Transformation* foram usadas duas *Steps Calculator*:

- *Calculator* - Calc_NE: a Calc_NE (Figura 16) transforma valor do campo VLNOTAEMPENHO no campo valor_nota_empenho fazendo o cálculo para posicionamento correto da vírgula após transformação do campo de *String* para *number*.

Figura 16 – *Calculator*: Calc_NE.

Calculator

Step name: Calc_NE

☒ Throw an error on non existing files

Fields:

#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask
1	div	Set field to constant value A	100			Integer			N	
2	valor_nota_empenho	A / B	VLNOTAEMPENHO	div		Number		2	N	###,###,###.#

Help OK Cancel

Fonte: Acervo do autor.

A Figura 17 mostra a coluna VLNOTAEMPENHO com os dados antes da formatação e a coluna valor_nota_empenho mostra os dados após passarem pela Calc_NE.

Figura 17 – Dados da coluna VLNOTAEMPENHO antes e coluna valor_nota_empenho após formatação

Examine preview data

Rows of step: Select values (1000 rows)

#	CDUNIDADEGESTORA	CDGESTAO	NUNOTAEMPENHO	VLNOTAEMPENHO	valor_nota_empenho	NUNEORIGINAL	TIPO
1	010101	00001	2019NE000001	16,994,572	169,945,7	2019NE000001	O
2	010901	01901	2019NE000001	558,000	5,580	2019NE000001	O
3	010101	00001	2019NE000002	13,006,950	130,069,5	2019NE000002	O
4	010901	01901	2019NE000002	1,795,000	17,950	2019NE000002	O
5	010901	01901	2019NE000003	236,000	2,360	2019NE000003	O
6	010101	00001	2019NE000003	16,994,572	169,945,7	2019NE000001	A
7	010901	01901	2019NE000004	150,000	1,500	2019NE000004	O
8	010101	00001	2019NE000004	13,006,950	130,069,5	2019NE000002	A
9	010101	00001	2019NE000005	15,000	150	2019NE000005	O
10	010901	01901	2019NE000005	150,000	1,500	2019NE000005	O
11	010101	00001	2019NE000006	24,878,400	248,784	2019NE000006	O
12	010901	01901	2019NE000006	186,000	1,860	2019NE000006	O

Fonte: Acervo do autor.

- *Calculator* - Calc_NE 2: a Calc_NE 2 (Figura 18) converte o valor do campo valor_nota_empenho fazendo a multiplicação do valor por -1 para negatização de valores de notas do tipo A e X.

Figura 18 – *Calculator*: Calc_NE 2.

Calculator

Step name
Calc_NE 2

☒ Throw an error on non existing files

Fields:

#	New field	Calculation	Field A	Field B	Field C	Value type	Length	Precision	Remove	Conversion mask
1	neg	Set field to constant value A	-1			Integer			N	
2	valor_nota_empenho_1	A * B	valor_nota_empenho	neg		Number		2	N	###,###,###

Help OK Cancel

Fonte: Acervo do autor.

A Figura 19 mostra a coluna valor_nota_empenho com os dados antes da formatação e a coluna valor_nota_empenho_1 mostra os dados após passarem pela Calc_NE 2.

Figura 19 – Coluna valor_nota_empenho com os dados antes e a coluna valor_nota_empenho_1 após formatação

Examine preview data

Rows of step: Select values (1000 rows)

#	CDUNIDADEGESTORA	CDGESTAO	NUNOTAEMPENHO	VLNOTAEMPENHO	valor_nota_empenho	valor_nota_empenho_1	NUNEOORIGINAL	TIPO
1	010101	00001	2019NE000003	16,994,572	169,945,7	-169,945,7	2019NE000001	A
2	010101	00001	2019NE000004	13,006,950	130,069,5	-130,069,5	2019NE000002	A
3	010901	01901	2019NE000020	85,000	850	-850	2019NE000019	A
4	010101	00001	2019NE000027	4,888,889	48,888,9	-48,888,9	2019NE000025	A
5	010101	00001	2019NE000041	19,143,432	191,434,3	-191,434,3	2019NE000024	A
6	010101	00001	2019NE000052	395,000	3,950	-3,950	2019NE000048	A
7	010101	00001	2019NE000061	29,200,000	292,000	-292,000	2019NE000056	A
8	010101	00001	2019NE000067	64,187,201	641,872	-641,872	2019NE000065	A
9	010101	00001	2019NE000086	4,177,983	41,779,8	-41,779,8	2019NE000076	A
10	010101	00001	2019NE000088	4,177,983	41,779,8	-41,779,8	2019NE000073	A
11	010101	00001	2019NE000090	4,155,923	41,559,2	-41,559,2	2019NE000071	A
12	010101	00001	2019NE000095	18,026,318	180,263,2	-180,263,2	2019NE000044	A

Fonte: Acervo do autor.

3.2.3.3 Carga dos Dados no PostgreSQL

Após as transformações, os dados foram carregados nas tabelas correspondentes do PostgreSQL utilizando a *Step output "Table output"*. A *Table output* é responsável por conectar-se ao banco PostgreSQL e inserir os dados na tabela notas_empenho. Para que os dados sejam inseridos corretamente é necessário fazer a correlação das colunas vindas das etapas anteriores com as colunas correspondentes na tabela nota_empenho no BD sig.

Como mostrado na Figura 20, na primeira sessão é feita uma comparação entre a coluna *"Table field"*, que diz respeito as colunas da tabela notas_empenho, e a coluna *"Stream field1"*, que corresponde as colunas vindas das etapas anteriores, para saber se os dados se tratam de dados novos que serão inseridos no banco ou de dados que devem ser atualizados. Abaixo, na sessão *"Update fields"*, é feita a correlação das colunas, onde os campos da coluna *"Stream field"* correspondem aos dados que devem dar entrada nos campos da coluna *"Table field"*.

Por possuir dois fluxos de alimentação foi necessário ter duas *Steps* de inserção/atualização:

- Insert/Update - update_NE: insere ou atualiza dados de notas de empenho do tipo O e usa o campo valor_nota_empenho vindo da Calc_NE para alimentar a coluna valor_nota_empenho da tabela nota_empenho.
- Insert/Update - update_NE_tipo_AX: insere ou atualiza dados de notas de empenho do tipo A e X e usa o campo valor_nota_empenho_1 vindo da Calc_NE 2 para alimentar a coluna valor_nota_empenho da tabela nota_empenho.

Figura 20 – Insert/Update - update_NE e Insert/Update - update_NE_tipo_AX.

Left Screenshot: update_NE

Step name: update_NE
 Connection: alema_desenv
 Target schema: sigef
 Target table: nota_empenho
 Commit size: 100
 Don't perform any updates: ☒
 The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1
1	num_nota_empenho	=	NUNOTAEMPENHO
2	num_nota_empenho_original	=	NUNEORIGINAL
3	id_unidade_gestora	=	CDUNIDADEGESTORA

Update fields:

#	Table field	Stream field
1	id_unidade_gestora	CDUNIDADEGESTORA
2	id_gestao	CDGESTAO
3	num_nota_empenho	NUNOTAEMPENHO
4	num_nota_empenho_original	NUNEORIGINAL
5	id_credor	CDCRETOR
6	data_lancamento	DTLANCAMENTO
7	tipo	TIPO
8	id_subacao	CDSUBACAO
9	id_unidade_orcamentaria	CDUNIDADEORCAMENTARIA
10	id_natureza	CDNATUREZADESPESA
11	id_programa	CDPROGRAMA
12	id_fonte	CDFONTE
13	id_modalidade_licitacao	CDMODALIDADELICITACAO
14	id_modalidade_empenho	CDMODALIDADEEMPENHO
15	nome_modalidade_empenho	DEMODALIDADEEMPENHO
16	num_processo	NUPROCESSO
17	valor_nota_empenho	valor_nota_empenho
18	quantidade	NUQUANTIDADE
19	historico	DEHISTORICO
20	id_evento	CDEVENTO

Right Screenshot: update_NE_tipo_AX

Step name: update_NE_tipo_AX
 Connection: alema_desenv
 Target schema: sigef
 Target table: nota_empenho
 Commit size: 100
 Don't perform any updates: ☒
 The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1
1	num_nota_empenho	=	NUNOTAEMPENHO
2	num_nota_empenho_original	=	NUNEORIGINAL
3	id_unidade_gestora	=	CDUNIDADEGESTORA

Update fields:

#	Table field	Stream field
1	id_unidade_gestora	CDUNIDADEGESTORA
2	id_gestao	CDGESTAO
3	num_nota_empenho	NUNOTAEMPENHO
4	num_nota_empenho_original	NUNEORIGINAL
5	id_credor	CDCRETOR
6	data_lancamento	DTLANCAMENTO
7	tipo	TIPO
8	id_subacao	CDSUBACAO
9	id_unidade_orcamentaria	CDUNIDADEORCAMENTARIA
10	id_natureza	CDNATUREZADESPESA
11	id_programa	CDPROGRAMA
12	id_fonte	CDFONTE
13	id_modalidade_licitacao	CDMODALIDADELICITACAO
14	id_modalidade_empenho	CDMODALIDADEEMPENHO
15	nome_modalidade_empenho	DEMODALIDADEEMPENHO
16	num_processo	NUPROCESSO
17	valor_nota_empenho	valor_nota_empenho_1
18	quantidade	NUQUANTIDADE
19	historico	DEHISTORICO
20	id_evento	CDEVENTO

(update_NE)

(update_NE_tipo_AX)

Fonte: acervo do autor.

3.2.3.4 Validação e Testes

Após a carga dos dados, foram realizados testes para assegurar a integridade e consistência das informações:

- Verificação de contagem de registros:** comparação entre o número de registros na VSEPLANNELEG (Oracle) e na tabela nota_empenho (PostgreSQL). A Figura 21 mostra a quantidade de registros referentes a notas de empenho da ALEMA e unidades gestoras relacionadas a mesma na view VSEPLANNELEG e na tabela nota_empenho.

Figura 21 – Comparativo da quantidade de registros do ano de 2019 na *view* VSEPLANNELEG e na tabela *nota_empenho* pós migração.

<pre>SELECT COUNT(*) FROM SIGEF2019.VSEPLANNELEG G WHERE G.NUNOTAEMPENHO LIKE '2019%' AND (CDUNIDADEGESTORA = '010101' OR CDUNIDADEGESTORA = '010201' OR CDUNIDADEGESTORA = '010901'); ... COUNT(*) ===== 3374 </pre> Read 1 rows	<pre>SELECT COUNT(*) FROM sigef.nota_empenho o WHERE o.num_nota_empenho LIKE '2019%' count ===== 3374 </pre> Read 1 rows
---	--

(Quantitativo View)

(Quantitativo Tabela)

Fonte: acervo do autor.

- b) **Validação de valores críticos:** conferência de somatórios de valores monetários (como *valor_nota_empenho*). A Figura 22 mostra os valores salvos no banco PostgreSQL formatados e transformados de acordo com o tipo.

Figura 22 – Dados após migração.

	num_nota_empenho	valor_nota_empenho	num_nota_empenho_original	tipo
1	2019NE000001	169.945,72	2019NE000001	O
2	2019NE000003	-169.945,72	2019NE000001	A
3	2019NE000001	5.580,00	2019NE000001	O
4	2019NE000004	-130.069,50	2019NE000002	A
5	2019NE000002	130.069,50	2019NE000002	O
6	2019NE000002	17.950,00	2019NE000002	O
7	2019NE000003	2.360,00	2019NE000003	O
8	2019NE000004	1.500,00	2019NE000004	O
9	2019NE000005	150,00	2019NE000005	O
10	2019NE000005	1.500,00	2019NE000005	O
11	2019NE000006	248.784,00	2019NE000006	O
12	2019NE000006	1.860,00	2019NE000006	O

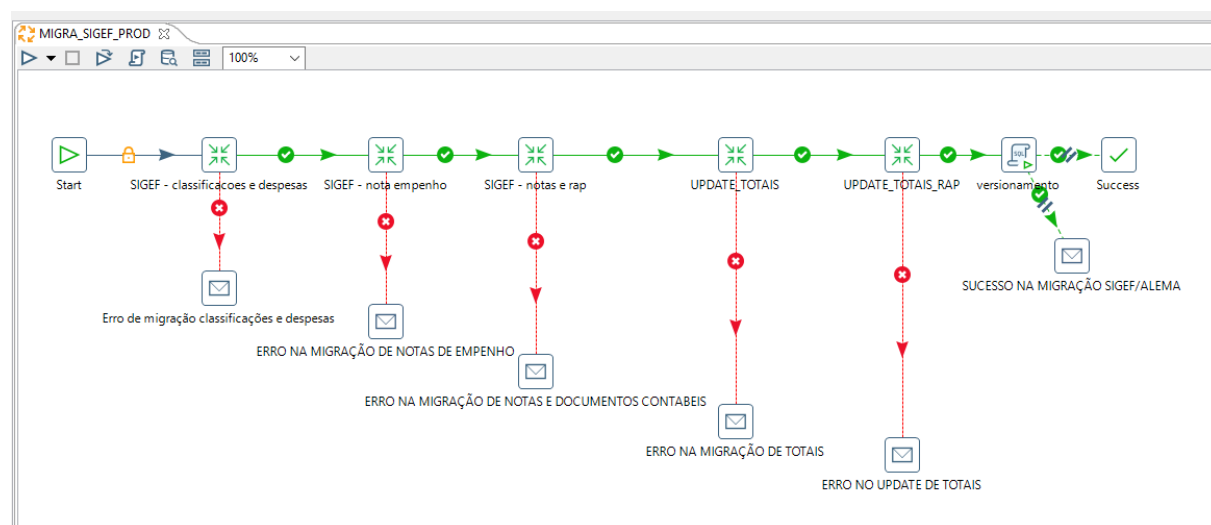
Fonte: Acervo do autor.

- c) **Testes de consultas:** garantia de que os dados estão acessíveis e consistentes no PostgreSQL.

3.2.3.5 Job Pentaho - MIGRA_SIGEF_PROD

O *Job* MIGRA_SIGEF_PROD se trata do conjunto de transformações responsáveis pela migração total de dados para a base *sig*, estas que seguem uma ordem de migração de dados dependentes um do outro como ilustrado na Figura 23).

Figura 23 – Job: MIGRA_SIGEF_PROD.



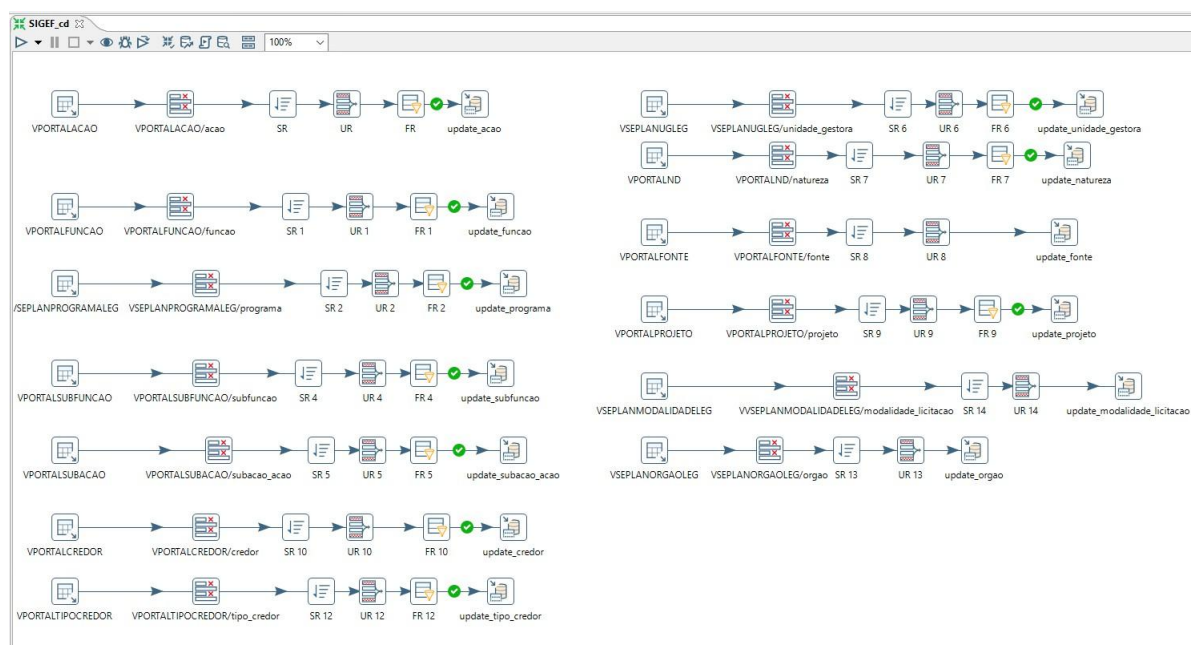
Fonte: Acervo do autor.

As *Transformations* seguem a seguinte ordem de migração de dados:

- SIGEF - classificacoes e despesas (Figura 24): migração de dados para tabelas auxiliares necessárias para posterior migração de notas de empenho como dados de fonte, credor, unidade gestora etc.
- SIGEF - nota empenho (Figura 25): migração de notas de empenho é a segunda transformação a ser executada e é responsável pela inserção de novas notas vindas das views, além da atualização de notas que possam ter sido editadas anteriormente.
- SIGEF - notas e rap (Figura 26): migração responsável por preparações de pagamentos, ordens bancárias, restos a pagar e restos pagos.

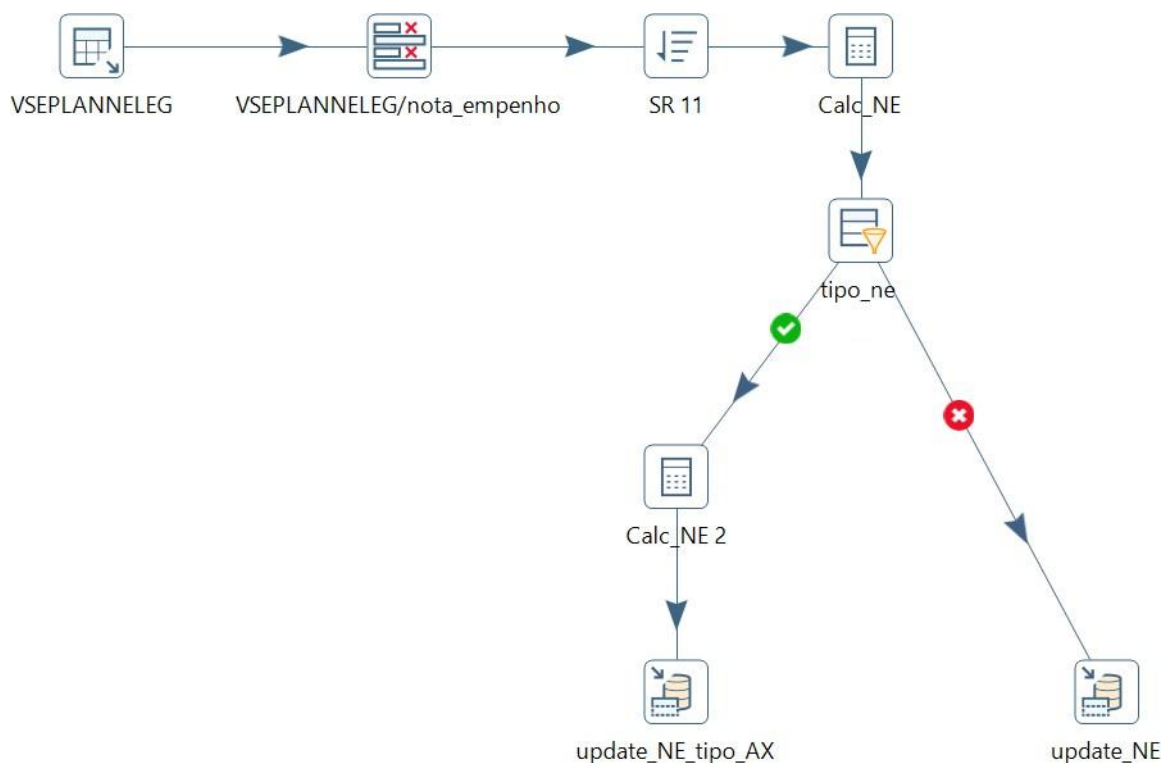
- d) UPDATE_TOTAIS (Figura 27): é responsável pelos cálculos de totais de valores pagos, liquidados, empenhados e cancelados de cada nota de empenho.
- e) UPDATE_TOTAIS_RAP(Figura 28): responsável pelos cálculos de totais de restos a pagar de cada nota de empenho.
- f) Versionamento(Figura 29): script SQL responsável por salvar na base PostgreSQL a data e hora da última atualização feita pelo Job.

Figura 24 – Transformation: SIGEF_cd.



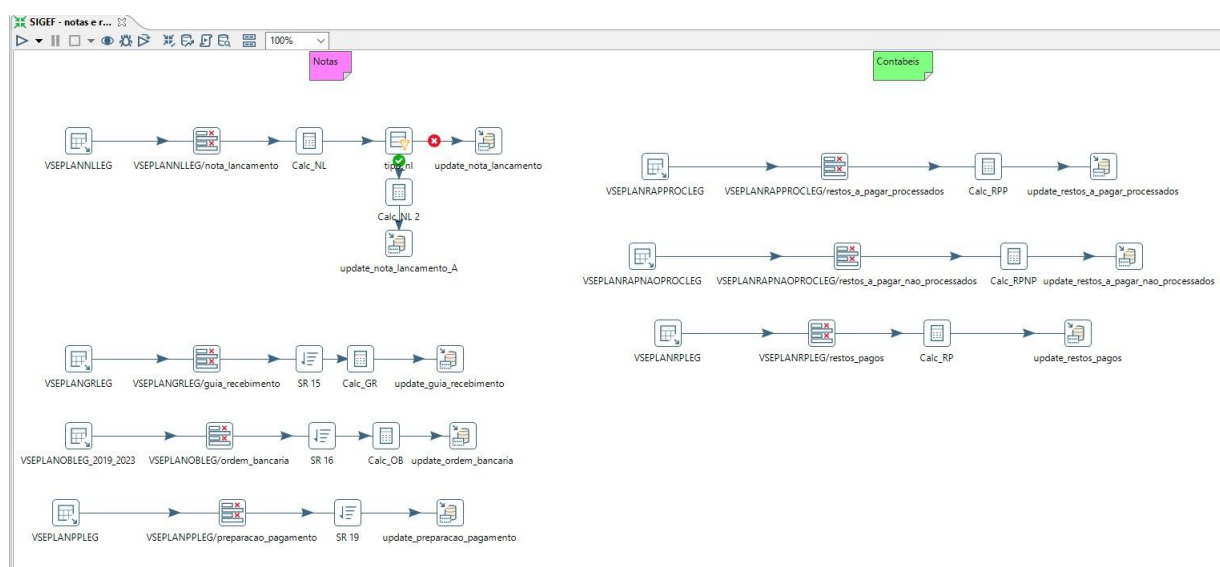
Fonte: Acervo do autor.

Figura 25 – Transformation: SIGEF notas empenho



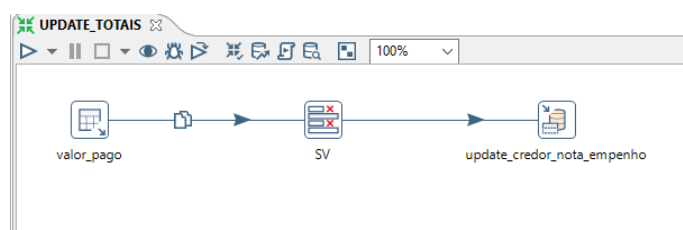
Fonte: Acervo do autor.

Figura 26 – Transformation: SIGEF notas e rap.



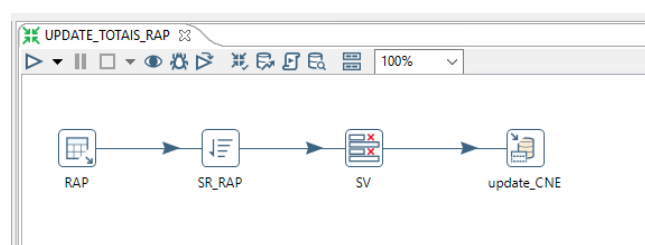
Fonte: Acervo do autor.

Figura 27 – Transformation: UPDATE_TOTAIS.



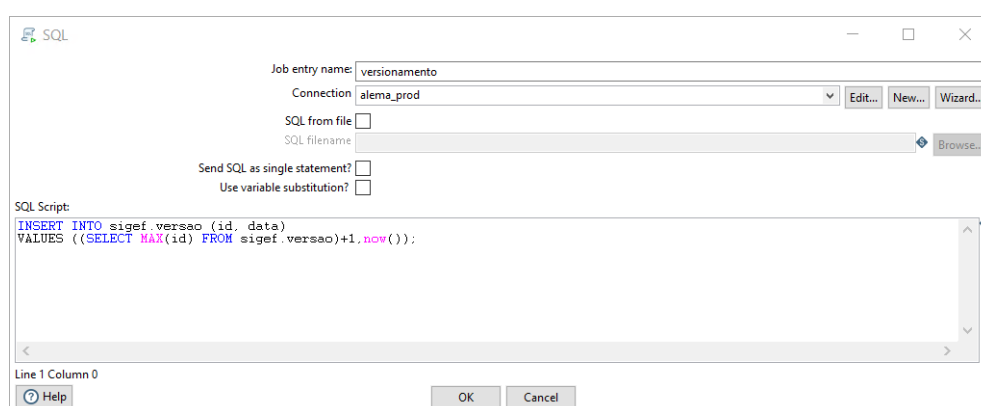
Fonte: Acervo do autor.

Figura 28 – Transformation: UPDATE_TOTAIS_RAP.



Fonte: Acervo do autor.

Figura 29 – SQL versionamento.



Fonte: Acervo do autor.

3.2.4 HOMOLOGAÇÃO

O objetivo desta fase é validar o funcionamento do projeto desenvolvido. Realizam-se testes para assegurar que os dados estão sendo extraídos, transformados e carregados corretamente. A etapa também envolve verificações de integridade, desempenho e conformidade com os requisitos levantados anteriormente.

A etapa de Homologação foi subdividida em subetapas voltadas as atividades citadas anteriormente, abaixo descreveremos cada uma delas:

3.2.4.1 Configuração do Pentaho Server

A versão utilizada foi o Pentaho Server 9.3.0.0, hospedada em uma máquina virtual disponibilizada, pelo setor de redes da ALEMA, exclusivamente para este fim. O ambiente foi configurado com os seguintes componentes e parâmetros:

- a) **Container Tomcat:** o Pentaho Server roda sobre um container Tomcat, que gerencia a execução do sistema web.
- b) **Repositório de Arquivos:** os jobs e transformações desenvolvidos no Pentaho Data Integration (PDI) foram publicados no repositório do Pentaho Server, organizados em pastas específicas para facilitar o gerenciamento.
- c) **Conexões de Banco de Dados:** as conexões com os bancos de dados Oracle (origem) e PostgreSQL (destino) foram configuradas diretamente nas transformations e jobs no PDI, utilizando os drivers JDBC apropriados.
- d) **Usuários e Permissões:** foram criados usuários com permissões específicas para acessar e executar os jobs, garantindo que apenas pessoal autorizado pudesse modificar ou agendar processos.

3.2.4.2 Configuração do Agendamento do Job MIGRA_SIGEF

O agendamento do job MIGRA_SIGEF_PROD foi configurado utilizando o agendador integrado ao Pentaho Server, baseado no Quartz Scheduler. As etapas para a configuração foram as seguintes:

- a) **Acesso à Interface de Agendamento:** através da User Console do Pentaho Server, acessou-se a seção "Agendamento"(Schedule), onde é possível criar e gerenciar agendamentos para jobs e transformações.
- b) **Definição da Frequência:** o job foi configurado para executar diariamente às 00:00h, garantindo que os dados fossem atualizados no Portal da Transparência com o mínimo de atraso.
- c) **Parâmetros de Execução:** foram definidos parâmetros específicos para o job, como o diretório de logs e opções de notificação em caso de falha. O agendador foi configurado para gerar logs detalhados de cada execução, facilitando a identificação de problemas.
- d) **Monitoramento:** a interface do Pentaho Server permite visualizar o histórico de execuções, incluindo status (sucesso ou falha), tempo de execução e mensagens de erro. Alertas por e-mail foram configurados para notificar a equipe responsável em caso de falha na execução do job.
- e) **Testes Iniciais:** antes de colocar o agendamento em produção, foram realizados testes manuais para validar a execução do job e a integridade dos dados após a migração. Esses testes incluíram a verificação da contagem de registros e a consistência dos valores monetários.

Com essa configuração, o processo ETL passou a ser executado automaticamente todas as noites, garantindo que os dados do Portal da Transparência da ALEMA estejam sempre atualizados e em conformidade com as exigências de transparência do TCE/MA.

3.2.5 PUBLICAÇÃO

A etapa de publicação marca a disponibilização do sistema em ambiente de produção, tornando-o acessível aos usuários finais. Nesta fase, o processo ETL desenvolvido passou a operar de forma automatizada, garantindo que o módulo de consulta de despesas do Portal da Transparência da ALEMA fosse alimentado com dados atualizados diariamente. A implementação do Job MIGRA_SIGEF_PROD no Pentaho Server permitiu que as atualizações ocorressem sem intervenção manual, assegurando a consistência e a pontualidade das informações disponibilizadas ao público.

3.2.5.1 Configuração do Ambiente de Produção

O Pentaho Server foi configurado para executar o Job MIGRA_SIGEF_PROD automaticamente todas as noites às 00:00h, garantindo que os dados fossem atualizados antes do início do dia útil. Essa configuração foi realizada por meio da interface do agendador (Scheduler) do Pentaho, onde foram definidos os parâmetros de execução, como frequência, horário e prioridade. Além disso, foram estabelecidos mecanismos de monitoramento para verificar o sucesso da execução do Job e alertar em caso de falhas.

3.2.5.2 Integração com o Portal da Transparência

Os dados migrados para o PostgreSQL foram integrados ao módulo de consulta de despesas do Portal da Transparência da ALEMA. O portal foi adaptado para consumir as informações diretamente do esquema sigef, exibindo-as de forma clara e organizada aos cidadãos. Foram implementadas funcionalidades de filtro e busca, permitindo que os usuários localizassem informações específicas, como notas de empenho por credor, período ou valor.

3.2.5.3 Resultados da Publicação

Com a publicação do sistema, o Portal da Transparência da ALEMA passou a atender aos critérios de atualidade e disponibilidade exigidos pelas normas do TCE/MA, como a Instrução Normativa nº 81/2024(TCE/MA, 2024a). A automatização do processo ETL eliminou atrasos na disponibilização dos dados, enquanto a consolidação das informações em um único esquema no PostgreSQL facilitou a consulta e a análise de dados históricos.

Esta etapa concluiu o ciclo de desenvolvimento, demonstrando a viabilidade da solução proposta e seu alinhamento com os objetivos de transparência e eficiência na gestão pública.

3.2.6 MANUTENÇÃO

A etapa de manutenção visa garantir a operação contínua e eficiente do processo ETL, implementando melhorias e correções necessárias. A partir da implantação do Job MIGRA_SIGEF_PROD, no ano de 2024, foram realizados ajustes periódicos para manter a qualidade dos dados e a estabilidade do sistema.

- **Melhorias Implementadas:**

- **Notificação por E-mail:** foi adicionado um sistema de alerta que envia notificações automáticas por e-mail ao final de cada execução do *Job*, informando status (sucesso ou falha) e tempo de processamento, permitindo monitoramento proativo.
- **Atualização Anual do Schema:** como a SEPLAN cria um novo banco de dados Oracle a cada ano (ex.: SIGEF2023, SIGEF2024), é necessário atualizar manualmente as consultas SQL no início de cada exercício fiscal para incluir o schema do ano corrente. Essa atualização é realizada diretamente nas transformações do Pentaho.
- **Otimização de Desempenho:** foram revisadas transformações críticas como *Sort rows* e *Filter rows*, com ajustes nos *buffers (row set)* e criação de índices no PostgreSQL para melhorar performance.
- **Tratamento de Exceções:** implementou-se validações adicionais para:
 - * Dados nulos/inconsistentes nas *views*;
 - * Divergências de tipos de dados.
 - * Correções Realizadas:
- **Campos Monetários:** corrigido truncamento de decimais mediante revisão das configurações de precisão nas *Steps Calculator* e *Select values*.
- **Integridade Referencial:** reordenadas transformações para garantir migração sequencial de tabelas relacionadas, evitando falhas por chaves estrangeiras.

A manutenção contínua assegura a adaptação do sistema às mudanças nos bancos- fonte e demandas do Portal.

3.3 Discussão de resultados

Os resultados obtidos pelo Portal da Transparência da Assembleia Legislativa do Estado do Maranhão (ALEMA) demonstram uma evolução significativa no atendimento aos critérios

de transparência estabelecidos pelo Tribunal de Contas do Estado do Maranhão (TCE/MA).

3.3.1 Desempenho em 2021

No ano de 2021 conforme o Relatório de Acompanhamento nº 1056/2021 – NUFIS 2 (TCE/MA, 2021), publicado em 13 de Dezembro de 2021, o portal da ALEMA obteve um índice de atendimento nota C na avaliação estadual, indicando um atendimento regular aos critérios de transparência.

3.3.2 Desempenho em 2023

Conforme o Relatório de Acompanhamento nº 595/2023 – NUFIS I (TCE/MA, 2023), de 21 de Junho de 2023, o portal da ALEMA obteve um índice de atendimento de 75,30%, classificando-se no nível B (Bom) na avaliação estadual.

3.3.3 Desempenho em 2024

Em 2024 a ALEMA alcançou um avanço ainda maior, conforme publicado no Relatório de Informação nº 253/2025 NUFIS I, de 11 de Março de 2025 (TCE/MA, 2024b). O portal foi classificado com nota 7,90 (equivalente a 79% de índice geral), atingindo o nível Elevado na nova escala de transparência introduzida pela Instrução Normativa TCE/MA nº 81/2024(TCE/MA, 2024a), que substituiu a anterior.

A classificação no nível Elevado refletiu melhorias na disponibilidade e organização dos dados financeiros, especialmente após a implantação do processo ETL com ferramentas *open source*, que permitiu a consolidação e atualização diária das informações. A Figura 30 mostra como esta sendo apresentada uma consulta de notas de empenho atualmente no Portal da Transparência da ALEMA.

Fonte: Captura de tela do Portal da transparência da ALEMA (2025c)

3.3.4 Comparativo das Classificações

- **2021 (RELATÓRIO DE ACOMPANHAMENTO Nº 1056/2021 – NUFIS 2):** nota C (Regular) - equivalente ao atual nível Básico; atendimento parcial aos critérios essenciais.
- **2023 (RELATÓRIO DE ACOMPANHAMENTO Nº 595/2023 – NUFIS I):** nota B (Bom) - equivalente ao atual nível Prata; 75,30% de conformidade, com destaque para a disponibilidade e atualização dos dados.
- **2024 (RELATÓRIO DE INFORMAÇÃO Nº 253/2025 NUFIS I):** nota 7,90 - nível Elevado; 90,48% dos critérios essenciais alcançados e de 79,02% de índice geral.

3.3.5 Impacto e Benefícios

A solução apresentada trouxe os seguintes benefícios:

- **Transparência ampliada:** Os cidadãos passaram a ter acesso a informações financeiras atualizadas e completas, incluindo notas de empenho, restos a pagar e pagamentos realizados.
- **Conformidade com as normas:** O portal atingiu níveis mais elevados de classificação no TCE/MA, conforme os critérios de disponibilidade, atualidade e série histórica.
- **Eficiência operacional:** A automação do processo reduziu a carga de trabalho manual e minimizou erros, garantindo a confiabilidade dos dados.

Essa progressão evidencia o impacto positivo das soluções técnicas implementadas, como a automação do ETL e a padronização dos dados, que garantiram maior confiabilidade, atualização frequente e acesso facilitado às informações.

Este capítulo tratou dos procedimentos adotados para a criação da solução de ETL, o próximo capítulo abordará a conclusão e sugestão de trabalhos futuros.

4 CONCLUSÃO

O objetivo deste trabalho foi criar uma solução de Extração, Transformação e Carga (ETL) com tecnologias *open source*, viabilizando a interoperabilidade entre os bancos de dados em Oracle e PostgreSQL, para que a ALEMA atenda aos requisitos legais vigentes em seu Portal da Transparência.

Para tanto, foi feito um levantamento teórico sobre os fundamentos do processo ETL, as ferramentas disponíveis, como o Pentaho Data Integration, e as particularidades dos sistemas Oracle e PostgreSQL. Além disso, foram analisadas as normas de transparência do TCE/MA, como a Instrução Normativa nº 59/2020 e sua evolução para a IN nº 81/2024, que estabelecem critérios rigorosos para a disponibilização de dados públicos. Essa fundamentação teórica foi essencial para entender os desafios técnicos e legais envolvidos no projeto.

Em seguida, propôs-se uma metodologia dividida em seis etapas: Análise, Especificação, Implementação, Homologação, Publicação e Manutenção. Essa abordagem permitiu desenvolver um fluxo ETL robusto, capaz de extrair dados das *views* do Oracle, transformá-los para compatibilidade com o PostgreSQL e carregá-los de forma automatizada no Portal da Transparência. A metodologia também incluiu a criação de um esquema de banco de dados no PostgreSQL, organizado em grupos lógicos (Orçamento, Despesas e Empenhos, Pagamentos e Processamentos, Programas e Ações), garantindo a integridade e acessibilidade dos dados.

Essa metodologia propiciou a modernização do Portal da Transparência da ALEMA, trazendo benefícios significativos, como a elevação da classificação do portal de "C"(Regular) em 2021 para "Elevado"(nota 7,90) em 2024, conforme Relatório de Informação nº 253/2025 do TCE/MA. Além disso, a solução garantiu maior eficiência operacional, redução de erros manuais e atualização diária dos dados, atendendo aos critérios de disponibilidade, atualidade e série histórica exigidos pelas normas vigentes.

Apesar dos benefícios alcançados, a solução tem a limitação de ter sido desenvolvida especificamente para o módulo de despesas, abrangendo principalmente a consulta de notas de empenho. Neste sentido, como trabalhos futuros, sugere-se:

- A expansão do processo ETL para outros módulos do Portal da Transparência, como receitas e contratos;
- A integração de ferramentas de visualização mais avançadas, como *dashboards*

interativos, para facilitar o entendimento e a participação cidadã.

Essas melhorias contribuiriam ainda mais para a transparência e o controle social na gestão pública.

REFERÊNCIAS

- ALEMA. *Função e definição*. 2025. Disponível em: <https://sistemas.al.ma.leg.br/transparencia/pagina.html?p=funcao-e-definicao&dswid=8523>. Acesso em: 28 jul. 2025.
- ALEMA. *História*. 2025. Disponível em: <https://sistemas.al.ma.leg.br/transparencia/pagina.html?p=historia&dswid=-116>. Acesso em: 28 jul. 2025.
- ALEMA. *Portal da Transparência da ALEMA - Módulo de Pesquisa de Notas de Empenho*. 2025. Disponível em: <https://sistemas.al.ma.leg.br/transparencia/despesas/notas-empenho/busca.html?dswid=6648>. Acesso em: 29 jul. 2025.
- BRASIL. *Lei nº 12.527, de 18 de novembro de 2011*. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal. *Diário Oficial da União*, Brasília, DF, 18 nov. 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm. Acesso em: 26 jul. 2025.
- CASTERS, M.; BOUMAN, R.; DONGEN, J. van. *Pentaho Kettle Solutions: Building Open Source ETL Solutions with Pentaho Data Integration*. Wiley, 2010. ISBN 978-0-470-63517-9. Disponível em: <http://repo.darmajaya.ac.id/4983/1/PentahoKettleSolutionsBuildingOpenSourceETLSolutionsWithPentahoDataIntegration.pdf>. Acesso em: 10 out. 2023.
- FERREIRA, J. et al. O processo ETL em sistemas data warehouse. In: *INForum 2010 - II Simpósio de Informática*. Braga, Portugal: [s.n.], 2010. Disponível em: https://www.researchgate.net/publication/265195317_O_Processo_ETL_em_Sistemas_Data_Warehouse. Acesso em: 26 jul. 2025.
- KIMBALL, R.; ROSS, M. *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. Wiley, 2002. ISBN 978-0-471-20024-8. Disponível em: https://www.r-5.org/files/books/computers/databases/warehouses/Ralph_Kimball_Margy_Ross-The_Data_Warehouse_Toolkit-EN.pdf. Acesso em: 26 jul. 2025.
- MARANHÃO. *Lei Estadual do Maranhão nº 10.217/2015*. 2015. Disponível em: <https://app.stc.ma.gov.br/legisla/consulta/publicacao/3809>. Acesso em: 26 jul. 2025.
- MICHENER, G.; CONTRERAS, E.; NISKIER, I. From opacity to transparency? Evaluating access to information in Brazil five years later. 2018. Disponível em:

<https://doi.org/10.1590/0034-761220170289>. Acesso em: 26 jul. 2025.

ORACLE. *Oracle Database Documentation*. 2025. Disponível em: <https://docs.oracle.com/en/database/oracle/oracle-database/23/cncpt/introduction-to-oracle-database.html#GUID-166C1E31-CDBC-47D9-867A-3D4C9AAC837D>. Acesso em: 26 jul. 2025.

POSTGRESQL. *PostgreSQL Documentation*. 2025. Disponível em: <https://www.postgresql.org/docs/>. Acesso em: 25 fev. 2025.

TCE/MA. *Instrução Normativa TCE/MA nº 59/2020*. 2020. Disponível em: <https://app.tcema.tc.br/publicacao/#!/documentohtml/2801>. Acesso em: 12 fev. 2025.

TCE/MA. *Instrução Normativa TCE/MA nº 81/2024*. 2024. Disponível em: <https://app.tcema.tc.br/publicacao/#!/documentohtml/23582>. Acesso em: 11 fev. 2025.

TCE/MA. *Relatório de Acompanhamento TCE/MA nº 595/2023*. 2023. Disponível em: <https://sistemas.al.ma.leg.br/transparencia/pagina.html?p=avaliacao-do-tce&dswid=1325>.

Acesso em: 28 jul. 2025.

TCE/MA. *Relatório de Acompanhamento TCE/MA nº 1056/2021*. 2021. Disponível em: <https://sistemas.al.ma.leg.br/transparencia/pagina.html?p=avaliacao-do-tce&dswid=1325>.

Acesso em: 28 jul. 2025.

TCE/MA. *Relatório de Informação nº 253/2025 NUFIS I*. 2024. Disponível em: [https://metabase.app2.tcema.tc.br/public/dashboard/5da70550-25fe-4512-8a72-](https://metabase.app2.tcema.tc.br/public/dashboard/5da70550-25fe-4512-8a72-afd980ef96e1?tab=21-)

[afd980ef96e1?tab=21-avalia%C3%A7%C3%B5es&ente_da_federa%C3%A7%C3%A3o=&esfera=Estadual&poder%252F%C3%B3rg%C3%A3o=Legislativo&faixa_de_transpar%C3%Aancia=&exerc%C3%ADcio=](https://metabase.app2.tcema.tc.br/public/dashboard/5da70550-25fe-4512-8a72-afd980ef96e1?tab=21-avalia%C3%A7%C3%B5es&ente_da_federa%C3%A7%C3%A3o=&esfera=Estadual&poder%252F%C3%B3rg%C3%A3o=Legislativo&faixa_de_transpar%C3%Aancia=&exerc%C3%ADcio=). Acesso em: 28 jul. 2025.