



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários

**ANP**  
**TRILHOS TIC TRENS**

**32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA**  
**13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS**  
**CATEGORIA 3**

**TEXT-TO-SQL NA DEMOCRATIZAÇÃO DE DADOS**  
**FERROVIÁRIOS**

## **INTRODUÇÃO**

A cultura corporativa moderna, fortemente orientada por dados, impulsiona a necessidade crescente de ferramentas que democratizem o acesso à informação. Decisões estratégicas baseadas em dados estruturados têm se mostrado significativamente mais assertivas e consistentes do que aquelas fundamentadas apenas na intuição gerencial. Bancos de dados relacionais continuam a ser a espinha dorsal para o armazenamento e a análise dessas informações vitais, formando a base operacional de sistemas complexos de inteligência de negócios e ferramentas de gestão. No entanto, persiste uma barreira técnica histórica e limitante que impede a exploração plena e ágil desses dados: a exigência de proficiência técnica na linguagem *Structured Query Language* (SQL). Para a grande maioria dos usuários de negócio — como diretores, gestores e analistas do setor metroferroviário —, a falta de domínio sobre a sintaxe e a lógica relacional cria um gargalo operacional, gerando uma dependência contínua de equipes especializadas de engenharia de dados para a extração de *insights* básicos.



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



## **32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA**

### **13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS**

Para superar essa barreira, as soluções de tradução de linguagem natural para SQL (*Text-to-SQL*) surgem como um mecanismo para traduzir perguntas cotidianas em consultas de banco de dados executáveis. A evolução desse campo tecnológico transitou de abordagens rígidas baseadas em regras e gramáticas para métodos sofisticados de aprendizado profundo, consolidando-se no uso de Grandes Modelos de Linguagem (LLMs - *Large Language Models*). Embora os LLMs contemporâneos possuam habilidades notáveis para interpretar nuances e a intenção semântica das perguntas dos usuários, a aplicação prática dessas tecnologias em ambientes corporativos de missão crítica enfrenta desafios substanciais. Os principais obstáculos incluem o elevado custo computacional, o alto consumo de *tokens* ao lidar com bancos de dados de esquemas extensos, a propensão a gerar *hallucinations* (consultas sintaticamente incorretas ou semanticamente falhas) e a dificuldade de adaptação a dialetos específicos de SQL (como Microsoft SQL Server ou PostgreSQL) sem um dispendioso processo de *fine-tuning*.

Além disso, em cenários onde a língua de interação não é o inglês, como na operação brasileira da TrensRJ, a complexidade aumenta consideravelmente. Estudos recentes demonstram que modelos de linguagem multilíngues frequentemente sofrem quedas sistêmicas de acurácia (*Execution Accuracy - EX*) ao traduzirem intenções do português para estruturas SQL, devido à menor sobreposição lexical e maior variabilidade morfológica em comparação com os *corpora* de treinamento centrados no inglês.

Diante deste cenário desafiador, este artigo apresenta o desenvolvimento, a arquitetura e a validação de um sistema avançado de *Text-to-SQL* implementado na TrensRJ. A solução orquestra a integração de LLMs por meio de *prompt engineering* avançado, mecanismos rígidos de validação de segurança (*Gatekeeper*) e interfaces de visualização interativas, equilibrando a acessibilidade da linguagem natural com o rigor técnico, o controle de custos



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



## 32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA 13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS

computacionais e a estabilidade exigida na inteligência de mercado do transporte de passageiros de massa

### DIAGNÓSTICO

O desenvolvimento do sistema de Inteligência de Mercado da TrensRJ baseou-se em uma arquitetura de *software* moderna e containerizada, utilizando Next.js 16, React 19, Tailwind CSS para estilização e Node.js no *backend*. O acesso aos dados é realizado em um banco de dados Microsoft SQL Server utilizando a biblioteca *mssql* com um *pool* de até 10 conexões simultâneas gerenciadas e controle de *timeout*. A inteligência artificial é provida através da integração com a API da Groq (compatível com o SDK da OpenAI), utilizando o modelo *llama-3.3-70b-versatile*.

O diagnóstico do problema revelou que delegar a geração de SQL a um LLM genérico (*frontier model*), fornecendo-lhe o esquema completo do banco de dados, é ineficiente e arriscado. Trabalhos paralelos que investigam leis de escala para a tarefa de *Text-to-SQL*, como os realizados com a família de modelos Qwen2.5, demonstram que, embora modelos massivos de 32 bilhões de parâmetros atinjam altas taxas de precisão (até 83,5% de EX em consultas complexas), seus custos de inferência são até oito vezes maiores que os de modelos compactos de 3 bilhões de parâmetros. Da mesma forma, a biblioteca *open-source* LexiQL demonstrou que a orquestração inteligente do fluxo de geração — através de introspecção de esquema e seleção semântica de contexto — permite que modelos de linguagem compactos (como o Qwen3-4B) operem de forma competitiva, reduzindo o consumo de *tokens* e evitando *hallucinations* ao restringir o contexto ao estritamente necessário.

Inspirado nesses princípios de orquestração restrita e eficiência de contexto, o sistema da TrensRJ não depende da capacidade bruta do LLM de inferir o



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



## 32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA 13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS

esquema global. Em vez disso, foi projetado com um *SYSTEM\_PROMPT* monumental atuando como a única fonte da verdade (*Single Source of Truth*) para as regras de negócio. Este *prompt*, com mais de 1.200 linhas de código injetadas no contexto, contém as seguintes diretrizes arquitetônicas que isolam riscos:

- **Glossário Corporativo e Regras de Negócio Inflexíveis:** O modelo é instruído rigidamente a traduzir jargões do setor metroferroviário. Por exemplo, a sigla "PAX" (Total de Passageiros) deve ser obrigatoriamente traduzida para SUM(p.qtd) e nunca para COUNT(\*). A "MDU" (Média Diária em Dias Úteis) exige uma lógica matemática específica (SUM(p.qtd) / COUNT(DISTINCT c.data)) atrelada ao filtro c.util = 'Sim'.
- **Schema Whitelist:** Para garantir o isolamento de dados sensíveis e economia de *tokens*, o sistema restringe a operação a exatas 9 tabelas e *views* predefinidas no *prompt*, eliminando o risco do LLM acessar tabelas financeiras ou de recursos humanos não autorizadas.
- **Habilidades Matemáticas SQL:** O *prompt* ensina ativamente o LLM a construir *Common Table Expressions* (CTEs) para cálculos complexos, como a exibição de gráficos de "Base 100" (índices comparativos temporais) e variações percentuais entre dois períodos.

Para garantir a segurança sistêmica contra ataques de *prompt injection* ou a geração de consultas excessivamente custosas (*full-table scans* que poderiam travar o SQL Server), implementou-se um mecanismo de *Gatekeeper* bidirecional:

1. **Gatekeeper Nível LLM (Detecção de Período Temporal):** A primeira validação ocorre antes de qualquer código SQL ser gerado. O *SYSTEM\_PROMPT* obriga o LLM a analisar a entrada do usuário em busca de um recorte temporal (ex: "últimos 15 dias", "março"). Se o



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



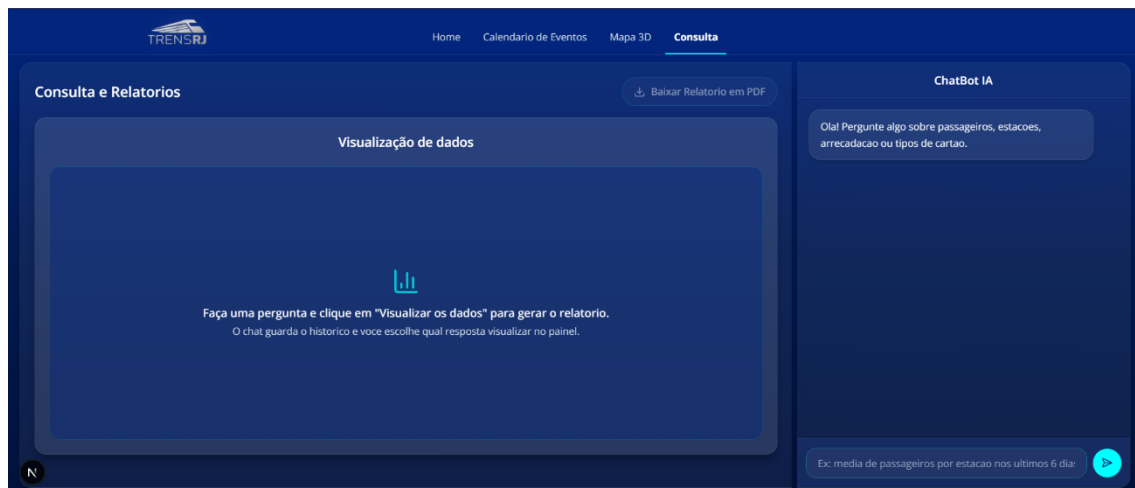
## 32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA 13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS

período estiver ausente, o modelo é instruído a não gerar SQL, mas a retornar um aviso padronizado.

2. **Gatekeeper Nível Backend (Enforcement):** A função `isMissingPeriodWarning()` no *backend* intercepta a resposta. Adicionalmente, a função `ensureReadOnlySql()` varre o código gerado bloqueando sumariamente qualquer palavra-chave de modificação de dados (INSERT, UPDATE, DELETE, DROP), garantindo operação estritamente *Read-Only*.

A interface do usuário consolida o ecossistema fornecendo um painel de consulta analítico e um mapa espacial 3D interativo construído com *Deck.gl* e *Mapbox*, capaz de cruzar os dados gerados com representações geográficas das operadoras (TrensRJ, MetrôRio, Mobi-Rio).

**Figura 1 – Interface de Consulta e Relatórios do Sistema TrensRJ.**



**Fonte: Painel de inteligência da aplicação**

(Refere-se à arquitetura visual onde o usuário interage via linguagem natural, com a área "Visualização de dados" à esquerda e o ChatBot IA à direita, conforme "Captura de tela")



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



## 32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA

### 13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS

**Tabela 1 – Sumário de validações de segurança da arquitetura Text-to-SQL.**

| Nível          | Validação Técnica            | Implementação Lógica                                          | Resultado Prático e Impacto                                                                |
|----------------|------------------------------|---------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| LLM / Contexto | Período temporal obrigatório | Diretriz prioritária no SYSTEM_PROMPT                         | Bloqueia a execução se a intenção do usuário for vaga; economia de CPU do banco de dados.  |
| LLM / Negócio  | Dicionário corporativo       | Tradução semântica de métricas (PAX, MDU, Impacto) via regras | Previne <i>hallucination</i> semântica e garante SQL válido atrelado à regra de negócio.   |
| LLM / Sintaxe  | Integridade do GROUP BY      | Regra Rígida de estruturação T-SQL                            | Evita erros sistêmicos na sumarização de dados relacionais                                 |
| Backend        | Read-only enforcement        | Função ensureReadOnlySql() antes da execução                  | Bloqueio imediato de tentativas de <i>Data Manipulation Language</i> (DML) não autorizadas |



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



## 32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA 13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS

|          |                            |                                                         |                                                                                                  |
|----------|----------------------------|---------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| Frontend | Formatação de padrão local | Função <code>formatCellValue()</code> na view da tabela | Conversão inteligente de notação americana para padrão brasileiro (ex: 156.348).                 |
| Frontend | Base 100 Detection         | Renderização condicional <code>isBase100Chart</code>    | Plotagem automática de <i>ReferenceLine</i> em gráficos <i>Recharts</i> para índices econômicos. |

Fonte: Elaboração dos autores baseada na implementação TrensRJ

### ANÁLISE DOS RESULTADOS

O fluxo de execução validado em produção demonstrou melhorias notáveis tanto na performance de acesso aos dados quanto na robustez da infraestrutura de *software*. Abaixo, detalha-se a análise sistemática do comportamento da solução em operação:

1. **Sucesso do Modelo de Orquestração com "Temperature Duality":** A estratégia de usar duas chamadas ao LLM com temperaturas distintas provou-se altamente eficaz. Na primeira chamada (geração do SQL), o modelo opera com temperatura zero, priorizando respostas determinísticas, precisão estrutural e zero criatividade. Após a execução do código no SQL Server, os metadados (linhas retornadas e colunas) e o resultado tabular são enviados em uma segunda requisição ao LLM com temperatura levemente superior (0.2), que interpreta os achados matemáticos e redige um sumário natural, interpretando quedas de



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



## 32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA 13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS

passageiros devido a chuvas ou picos operacionais de forma humanizada.

2. **Mitigação do Efeito Multilíngue (Cross-lingual Gap):** Conforme apontado pelo estudo no *benchmark* BIRD\_PT, transitar do inglês para o português na geração de SQL impõe um custo sistemático na Acurácia de Execução (EX), frequentemente devido ao fenômeno de *language mixing* (onde o modelo "pensa" internamente em inglês mesmo ao processar português) e à dificuldade de vinculação semântica (*schema linking*). O sistema TrensRJ contornou este problema com uma engenharia de *prompt* híbrida: o usuário escreve e recebe respostas em português, mas o *SYSTEM\_PROMPT* retém os elementos essenciais da arquitetura relacional (esquemas de banco e cláusulas T-SQL) nativamente em inglês, neutralizando assim a confusão morfológica apontada na literatura acadêmica e minimizando filtros errôneos causados por alterações de caixa alta e baixa.
3. **Desempenho da Detecção de Omissões (Gatekeeper):** A análise do "Exemplo 2: Detecção de Falta de Período" demonstrou a eficácia do mecanismo de recusa. Ao receber a pergunta ambígua "Quantos passageiros embarcaram?", o *Gatekeeper* do LLM, aliado à verificação *isMissingPeriodWarning()* do *backend*, retornou HTTP 200 com uma mensagem educativa exigindo maior especificidade. Isso erradicou a incidência de falhas críticas por *timeout* (acima de 30 segundos) que ocorreriam caso bilhões de linhas sofressem varredura (*full-table scan*).
4. **Resiliência através de Fallback Textual e Renderização:**
  - No *backend*, se a segunda chamada do LLM (encarregada da redação da resposta analítica) falhar por problemas na rede ou sobrecarga na API, uma heurística matemática de *fallback* entra em ação instantaneamente, calculando os totais, médias, mínimos



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



## 32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA 13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS

e máximos diretamente dos dados *raw* (brutos), formatando-os para exibição sem prejudicar a experiência do usuário.

- No *frontend*, os dados retornados passam pela normalização da função `toChartData()`, alimentando gráficos construídos com a biblioteca *Recharts*. Uma função de detecção `getDynamicBarColor()` lê a chave do dado retornado e aplica dinamicamente a identidade visual da empresa (ex: roxo para o ramal Belford Roxo, vermelho para Deodoro, verde para Santa Cruz), enriquecendo o nível tático da análise de Business Intelligence.

5. **Acurácia Estrutural Focada e Custo-Benefício:** Corroborando com os estudos aplicados sobre os modelos Qwen e a biblioteca LexiQL — que indicam que a acurácia no processamento de esquemas densos melhora substancialmente ao focar o contexto semântico —, a *Schema Whitelist* implementada na TrensRJ garantiu alta estabilidade na predição de consultas. Ao invés de usar *retrieval-augmented generation* (RAG) dinâmico, o uso de um contexto estático altamente curado com 9 entidades no *prompt* estabilizou a execução, permitindo que a Inteligência de Mercado consultasse *rankings* de "Farol" operacional e indicadores diários com latências razoáveis (média de 5 segundos, majoritariamente devida ao processamento duplo do modelo de 70 bilhões de parâmetros).
6. **Efetividade nas Decisões Operacionais (Micro) e estratégicas (Macro):** A democratização do acesso aos dados complexos de transporte resultou em melhorias diretas na qualidade do serviço ferroviário, atuando em duas frentes decisórias fundamentais:
7. **Decisões Micro (Tático-Operacionais):** No dia a dia, a operação ferroviária exige respostas rápidas. O sistema permitiu aos gestores monitorar indicadores cruciais de forma *ad-hoc*, como o "Farol das 09h" e a "Parcial das 16h", obtendo o ranqueamento imediato das estações mais



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



## **32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA**

### **13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS**

carregadas. A inteligência do sistema processa o impacto de variáveis externas em tempo real, permitindo aos gestores interrogar a IA sobre a perda de passageiros devido a chuvas fortes ou avaliar o impacto e o carregamento esperado de trens gerado por eventos específicos (como jogos de futebol e shows) para dimensionar o efetivo das estações e a oferta de trens.

8. **Decisões Macro (Estratégicas e Inteligência de Mercado):** A longo prazo, diretores passaram a realizar análises comparativas sofisticadas, como o cálculo da Média Diária em Dias Úteis (MDU) e variações percentuais multi-período, sem depender de scripts SQL complexos. Além disso, a integração com um mapa espacial 3D possibilita o cruzamento de dados de volume de passageiros da SuperVia com outras operadoras (MetrôRio e Mobi-Rio). Essa visualização espacial e comparativa das rotas permite à diretoria entender o fluxo metropolitano de passageiros frente a modais concorrentes (ônibus, BRT e carros), embasando decisões estratégicas sobre investimentos na malha e negociações de subsídios de forma visual e orientada a dados.

O resultado dessas melhorias é a transição de um modelo de gestão ferroviária tradicionalmente reativo para uma operação preditiva e responsiva, onde a visualização de gráficos dinâmicos acelerou a interpretação visual dos gargalos da rede.

## **CONCLUSÕES**

O projeto se provou um mecanismo altamente inovador e pragmático para a modernização do Business Intelligence no contexto de operações metroferroviárias e mobilidade urbana.

Tradicionalmente, as empresas enfrentam um paradoxo: disponibilizar dashboards estáticos e engessados (que oferecem segurança, mas zero



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



## 32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA 13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS

flexibilidade) ou fornecer acesso irrestrito ao banco de dados (que garante flexibilidade, mas compromete a governança, a segurança e requer alto conhecimento técnico).

O sistema implementado pela TrensRJ soluciona esse dilema. A solução concede flexibilidade analítica absoluta por meio de interfaces de linguagem natural, ao mesmo tempo que mantém rigor técnico draconiano utilizando validações arquitetônicas (Gatekeepers de backend e prompt engineering corporativo).

Ao definir o SYSTEM\_PROMPT como um "código" independente, estabeleceu-se uma separação clara entre a infraestrutura estrutural em Next.js/SQL e a lógica de negócios transitória, facilitando a escalabilidade.

Nota: O roadmap futuro para a evolução da aplicação (versão 2.0) buscará enfrentar a principal limitação observada: a latência causada pelas requisições sequenciais à API do LLM. Prevê-se a paralelização de chamadas dos modelos para redução de tempo de resposta, a implementação de caching para queries frequentes, o fine-tuning de um modelo menor e open-source (visando reduzir os custos computacionais conforme demonstrado viável pela literatura em LLMs compactos) e a ampliação do suporte nativo a múltiplos idiomas com painéis robustos para auditoria das queries geradas.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Campos, G. S.; Amaral, L. S. (2025). **Otimização de Text-to-SQL com Modelos de Linguagem Compactos**. Resumo executivo oriundo de TCC de Especialização em Data Science e Analytics do MBA USP/Esalq.



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



## 32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA 13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS

- Hosen, M. S.; Islam, R.; Naeem, Z.; Folorunso, E. O.; Chu, T. S.; Mamun, M. A. A.; Orunbon, N. O. (2024). Data-driven decision making: advanced database systems for business intelligence. Nanotechnology Perceptions 20 (3): 687-704.
- Ionescu, S.-A.; Radu, A.-O. (2024). **Assessment and integration of relational databases, big data, and cloud computing in financial institutions: performance comparison**. In: International Conference on Innovations in Intelligent Systems and Applications [INISTA], Craiova, Dolj County, România.
- Kanburoğlu, A. B.; Tek, F. B. (2024). **Text-to-SQL: A methodical review of challenges and models**. Turkish Journal of Electrical Engineering and Computer Sciences 32 (3): 403-419.
- Katsogiannis-Meimarakis, G.; Xydias, M.; Koutrika, G. (2023). **Data democratisation with deep learning: the anatomy of a natural language data interface**. In: ACM INTERNATIONAL CONFERENCE ON WEB SEARCH AND DATA MINING [WSDM].
- Nobre, L.; Sousa, S.; Teles, S.; Soares, A. (2025). **Language Effects in Text-to-SQL Across English and Portuguese**. Instituto de Informática - Universidade Federal de Goiás.
- Silva, L. O.; Silva, P. H. C.; Silva, F. A. (2025). **Leis de Escala para Text-to-SQL: Um Estudo sobre a Relação entre Tamanho e Desempenho de Modelos de Linguagem**. In: Proceedings of the 40th Brazilian Symposium on Data Bases, Fortaleza, CE, Brazil.



Prêmio  
Tecnologia &  
Desenvolvimento  
Metroferroviários



## **32ª SEMANA DE TECNOLOGIA METROFERROVIÁRIA**

### **13º PRÊMIO TECNOLOGIA E DESENVOLVIMENTO METROFERROVIÁRIOS**