IA para gestão de OAEs: do PDF ao mapa de criticidade, com trilha para vida útil remanescente

INSTITUTO MILITAR DE ENGENHARIA
Praça Gen. Tibúrcio, 80 - Urca, Rio de Janeiro – RJ

Luiz Araujo de Souza Junior| Rio de Janeiro | Brasil | luiz.junior@ime.eb.br
Tiago Tamagusko, | Rio de Janeiro | Brasil | tiago.tamagusko@ime.eb.br
Filipe A. C. Nascimento | Rio de Janeiro | Brasil | filipe.nascimento@ime.eb.br
Robson Luiz Gaiofatto IM | Rio de Janeiro | Brasil | robson.gaiofatto@ime.eb.br

Resumo

Este trabalho apresenta um pipeline orientado por agente de inteligência artificial para transformar relatórios de inspeção de Obras de Arte Especiais (OAEs) em uma base tabular e em visualizações de apoio à manutenção. A contribuição é a integração entre extração de informações, tratamento das notas de inspeção e localização geográfica, reunindo registros dispersos em um fluxo de consulta por ativo e corredor. Na arquitetura proposta, um modelo de linguagem lê relatórios no padrão DER-ES e produz registros estruturados; o processamento em Python consolida os campos e prepara histogramas e mapas. O desenvolvimento contempla também o aproveitamento de planilhas cadastrais privadas e arquivos KMZ, mantendo a origem e os critérios de compatibilização de cada conjunto. Os produtos do MVP ilustram a leitura conjunta da distribuição das notas e da localização dos registros, com revisão humana dos dados incompletos. A evolução prevista integra históricos de inspeção e medições de vibração ou telemetria para subsidiar estimativas de vida útil remanescente. O fluxo oferece uma base para organizar informações e discutir prioridades de manutenção, cuja escolha depende da avaliação de engenharia.

Palavras-chave: obras de arte especiais; agentes de inteligência artificial; geovisualização; manutenção; vida útil remanescente.

1 Introdução

A gestão de Obras de Arte Especiais (OAEs) tem dependido, historicamente, de ciclos de inspeção que produzem relatórios em formato PDF e demandam leitura humana para classificar condições, registrar anomalias e recomendar intervenções. No Brasil, a ABNT NBR 9452:2023 é uma referência para a inspeção de pontes, viadutos e passarelas e a apresentação das avaliações de condição (ABNT, 2023). A consolidação operacional dessas informações demanda reunir documentos, notas e localização, preservando os critérios adotados em cada inspeção. A dispersão dos registros dificulta a consulta conjunta e o planejamento de intervenções.

Nos últimos anos, duas frentes tecnológicas passaram a oferecer oportunidades para encurtar esse ciclo. A primeira é a aplicação de agentes de linguagem, que processam documentos semiestruturados para extrair campos e normalizar registros em bases tabulares. A segunda é o avanço de ferramentas geoespaciais em Python, como Pandas e GeoPandas, combinadas a bibliotecas de mapas de base como o contextily, que facilitam a produção de visualizações por trecho de rodovia ou corredor ferroviário e, com isso, apoiam a leitura territorial das prioridades de manutenção (GeoPandas, s.d.; OpenStreetMap, s.d.).

Em paralelo, o monitoramento da saúde estrutural por respostas dinâmicas oferece informações complementares às inspeções. A análise de vibração pode identificar alterações do comportamento estrutural quando consideradas as condições de medição, os efeitos ambientais e uma resposta de referência (Saidin et al., 2022). Redes de sensores sem fio oferecem recursos para organizar medições continuadas (Sonbul; Rashid, 2023). A associação desses dados ao histórico de inspeções pode subsidiar prognósticos, desde que se definam o mecanismo de deterioração, os limites de desempenho e os modelos a calibrar.

Este trabalho se insere nesse contexto e propõe um pipeline orientado por agente para transformar relatórios de inspeção no padrão DER-ES em uma base tabular e em um mapa de criticidade para apoio à manutenção. O fluxo localiza os documentos, identifica suas seções e extrai campos de identificação, localização, extensão e notas, registrando arquivo e data de inspeção. Em seguida, os registros são organizados em Python e preparados para duas saídas principais: um histograma das notas e uma representação geográfica por corredor. Neste artigo, o mapa de criticidade representa as notas de condição registradas; o cálculo probabilístico de risco permanece fora do escopo.

Como desenvolvimento inicial, o MVP reúne produtos ilustrativos para discutir a consulta conjunta das condições e da localização dos registros. O arranjo busca reduzir o trabalho de consolidação, melhorar a padronização e facilitar a comunicação entre engenharia, planejamento e operação. A ampliação para planilhas de OAEs ferroviárias e arquivos KMZ permite incorporar outras fontes ao mesmo esquema de informação, preservando suas chaves e diferenças de classificação. A integração de vibração constitui uma etapa futura, voltada à evolução da manutenção preditiva.

O objetivo é apresentar um fluxo integrado de extração assistida por IA, consolidação de dados e geovisualização para apoio à gestão de manutenção de OAEs. Sua contribuição está na articulação entre a leitura dos relatórios, o tratamento explícito das notas e o vínculo geográfico, e no produto que essa integração oferece ao gestor. O agente é um componente do fluxo. A proposta abrange a passagem do PDF à tabela e ao mapa, a incorporação de fontes cadastrais e uma trilha para associar históricos de inspeção a informações dinâmicas em estudos de vida útil remanescente.

2 Dados e métodos

2.1 Fontes de dados e escopo

A fonte documental compreende relatórios de inspeção no padrão DER-ES, nas modalidades cadastral, rotineira e especial, produzidos em atividades profissionais da Encopetro para o DER-ES. Os arquivos apresentam variação de layout; a arquitetura prevê leitura da camada textual e reconhecimento óptico de caracteres (OCR) quando o PDF contém somente imagens. O desenvolvimento inclui uma planilha privada com 120 registros cadastrais rodoviário e ferroviários e arquivos KMZ de pontes, viadutos e túneis, utilizados como fontes de localização.

Os conjuntos rodoviário e ferroviário são tratados como entradas distintas. A compatibilização consiste em organizar campos comuns e explicitar o vínculo entre identificação, inspeção e posição geográfica, mantendo o identificador de origem. O compartilhamento de um esquema tabular não comprova que todas as fontes estejam conciliadas por ativo nem autoriza somar seus registros como uma amostra única.

A demonstração gráfica utiliza a tabela auxiliar do MVP, que contém 54 registros. O conjunto tem finalidade ilustrativa: os valores de notas e as coordenadas precisam ser confirmados nas fontes antes de uso operacional. Os dados privados são apresentados de forma agregada, sem divulgação de identificadores ou de informações contratuais.

Os atributos-alvo são código e tipo da OAE, rodovia ou linha ferroviária, trecho, quilometragem, município e UF, extensão, tipo e data de inspeção, notas estrutural, funcional e de durabilidade e nota geral explícita. O esquema também contempla restrições, latitude, longitude e arquivo de origem. A norma ou o critério de classificação, a página de evidência e a qualidade da localização devem acompanhar os valores para permitir sua interpretação e revisão.

2.2 Agente de IA e extração estruturada

A Figura 1 apresenta a arquitetura do fluxo, desde as fontes de entrada até os produtos para apoio à manutenção. Para os relatórios PDF, o agente baseado em modelo de linguagem localiza as seções relevantes e extrai os atributos definidos no item 2.1. A leitura distingue identificação e cadastro, resultados de avaliação e recomendações do relatório. A entrada por planilha dispensa a etapa de extração textual, mas segue as mesmas verificações de campos e proveniência.

Na extração proposta, o agente recebe um esquema de saída, regras de formatação por campo e exemplos. Os valores são emitidos como registros JSON vinculados ao documento e à inspeção. A validação verifica formato de data, unidades, domínio das notas conforme a fonte e coerência entre identificação e evidência. Ausências, conflitos ou informações sem suporte no documento são encaminhados à revisão humana, sem completar valores por suposição.

A arquitetura prevê registrar a versão das instruções de extração, o modelo utilizado e as ocorrências por arquivo. Esses elementos permitem distinguir o dado transcrito das transformações posteriores. O artigo descreve a arquitetura e apresenta os produtos de demonstração; a medição da precisão do agente e do tempo total de revisão integra o protocolo de avaliação proposto no item 4.5.

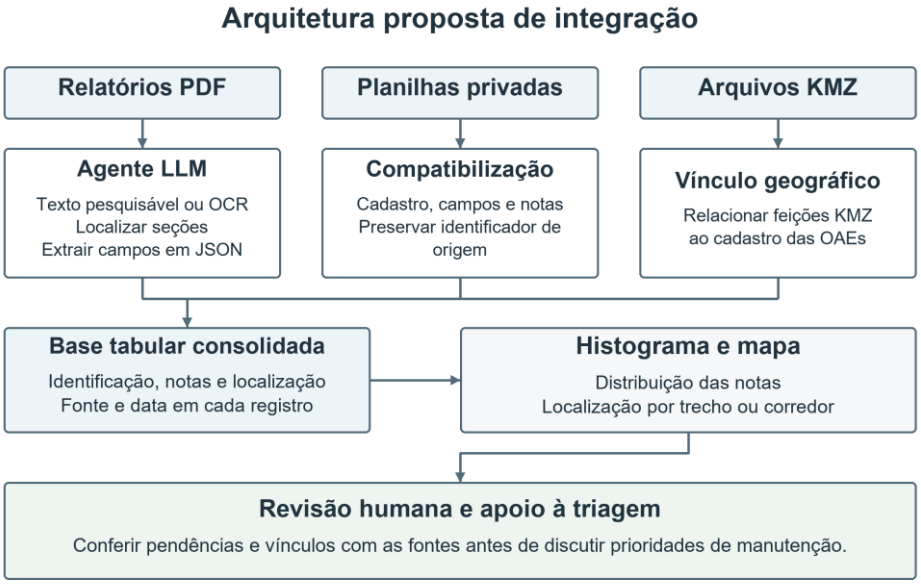


Figura 1 – Fluxo proposto de integração entre relatórios, dados de inspeção e localização.

Fonte: elaboração do autor.

2.3 ETL e regras de negócio da Nota Geral

Os registros são consolidados em CSV por um processo de extração, transformação e carga (ETL) em Python com Pandas. A Figura 2 mostra o tratamento da nota geral. O valor explícito no relatório é preservado com seu critério de origem. Quando faltam dados, a linha recebe uma indicação de pendência para revisão.

Na ausência de nota geral, o mínimo entre as notas estrutural, funcional e de durabilidade, quando completas e comparáveis, pode ser registrado em um campo separado de síntese conservadora. Essa é uma regra operacional proposta, que não substitui a nota emitida pelo responsável pela inspeção. Quando alguma dimensão está ausente, mantém-se a síntese sem valor. A ausência de informação não deve ser convertida automaticamente em nota de condição.

As classificações de documentos e edições diferentes mantêm sua identificação original. Valores fracionários ou fora do domínio exigem conferência na fonte, sem arredondamento ou reclassificação silenciosos. Toda correção registra valor anterior, justificativa e versão. Restrições operacionais são armazenadas em campo próprio e submetidas à avaliação de engenharia.

O esquema tabular contempla identificação do ativo e da inspeção, notas observadas, eventual síntese derivada, indicador de pendência, arquivo, data e página de evidência. A ligação às coordenadas inclui sua origem e qualidade. Essa estrutura permite consultar o registro e retornar ao documento que o sustenta.

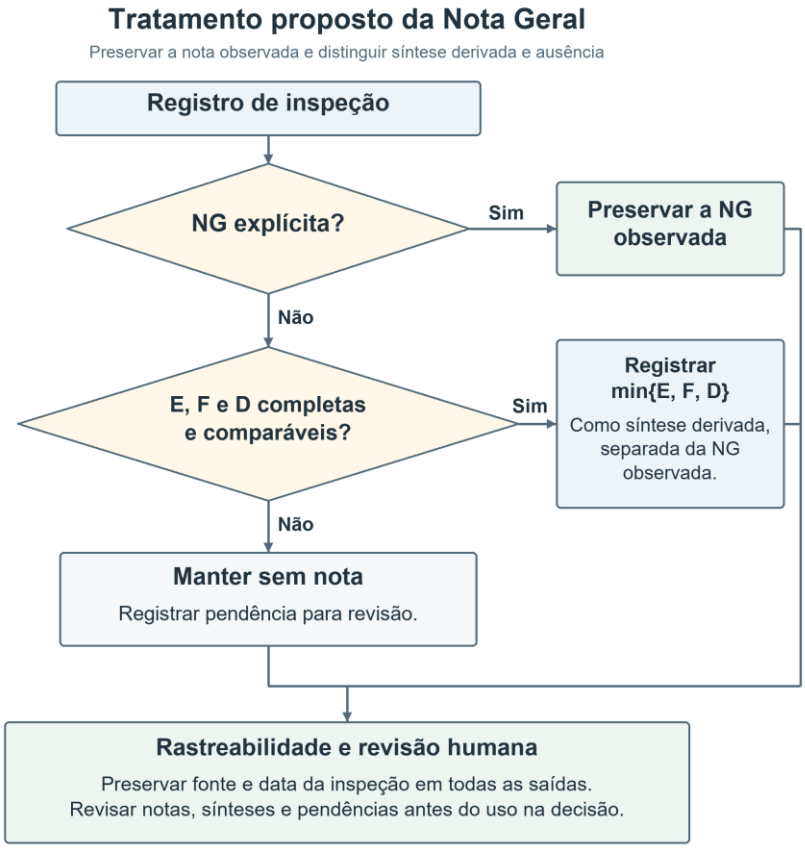


Figura 2 – Tratamento proposto da nota geral e das informações ausentes.

Fonte: elaboração do autor.

2.4 Geoprocessamento e visualizações

A visualização territorial proposta utiliza GeoPandas e contextily para organizar as coordenadas e apresentá-las sobre uma base cartográfica. As posições em latitude e longitude são associadas ao sistema de referência informado pela fonte; para visualização sobre tiles OSM, o fluxo pode reprojetá-las para Web Mercator (EPSG:3857). A legenda numérica e as cores devem permanecer consistentes no mapa e no histograma.

Os filtros permitem consultar registros por rodovia, linha ferroviária, trecho e nota de condição. A distribuição por nota ajuda a dimensionar grupos de interesse, enquanto o mapa mostra como esses registros se distribuem no território. As cores são um recurso de comunicação; a legenda deve indicar a classificação aplicável, sem substituir o critério técnico do relatório. O detalhamento da paleta foi reunido no Apêndice A.

Quando faltam coordenadas da obra, o centróide municipal pode ser usado somente para apresentação agregada e deve ser identificado como localização aproximada. Essa posição não substitui o georreferenciamento da OAE nem sustenta distâncias entre estruturas. O refinamento por quilometragem e a incorporação de arquivos KMZ exigem verificar as chaves de correspondência e registrar a origem do vínculo.

2.5 Reprodutibilidade e ambiente computacional

O desenvolvimento original foi organizado em Google Colab e Python, com etapas de consolidação e geovisualização. O ambiente e os elementos necessários para reprodução do fluxo são apresentados no Apêndice A. A reprodução ponta a ponta requer conservar o código, as instruções do agente, a versão do modelo, as entradas e os registros de execução.

2.6 Governança de dados e ética

Os relatórios e cadastros profissionais são utilizados com preservação das condições de acesso e de confidencialidade. As saídas devem manter a proveniência por arquivo e inspeção, suprimindo ou agregando informações sensíveis para divulgação. Os mapas precisam identificar a fonte cartográfica e a qualidade da posição representada. Registros incompletos e casos com impacto potencial em segurança permanecem sujeitos à revisão de engenharia.

2.7 Síntese do fluxo

O fluxo da Figura 1 liga relatórios, planilhas e dados de localização ao mesmo esquema de consulta. A etapa de extração converte o conteúdo documental em campos estruturados; o ETL verifica e organiza os registros; o tratamento da Figura 2 preserva as notas e sinaliza pendências; e a geovisualização apresenta a distribuição e a localização. O resultado é uma base de trabalho cuja origem pode ser consultada e que permite discutir prioridades sem perder o vínculo com a inspeção.

3 Resultados

3.1 Base de demonstração

Os produtos do MVP são uma base tabular, o histograma das notas e um mapa dos registros. Para apresentar os dois produtos no mesmo recorte, foram consideradas as 54 linhas da tabela auxiliar disponível, com sua nota geral e coordenadas. O conjunto ferroviário de 120 registros permanece como fonte adicional para a arquitetura de integração e não compõe o denominador dessas figuras.

A base demonstrativa apresenta três registros com nota 0, cinco com nota 1, nove com nota 2, dezesseis com nota 3, vinte com nota 4 e um com nota 5. As notas 3 e 4 reúnem 36 dos 54 registros. Esses valores descrevem a demonstração do fluxo; sua utilização como diagnóstico de uma rede depende da revisão das informações de origem.

3.2 Histograma das notas de condição

A Figura 3 apresenta o histograma das notas armazenadas na base demonstrativa. O produto permite reconhecer a distribuição dos registros e selecionar grupos para exame mais detalhado. O denominador está explícito para que filtros por corredor, período ou disponibilidade de informação possam ser comparados.

A concentração nas notas 3 e 4 orienta a discussão sobre acompanhamento e planejamento no conjunto representado. Os registros com notas menores merecem verificação técnica prioritária, considerando o significado da escala usada no relatório. O histograma não define, isoladamente, o tipo de intervenção nem a probabilidade de falha.

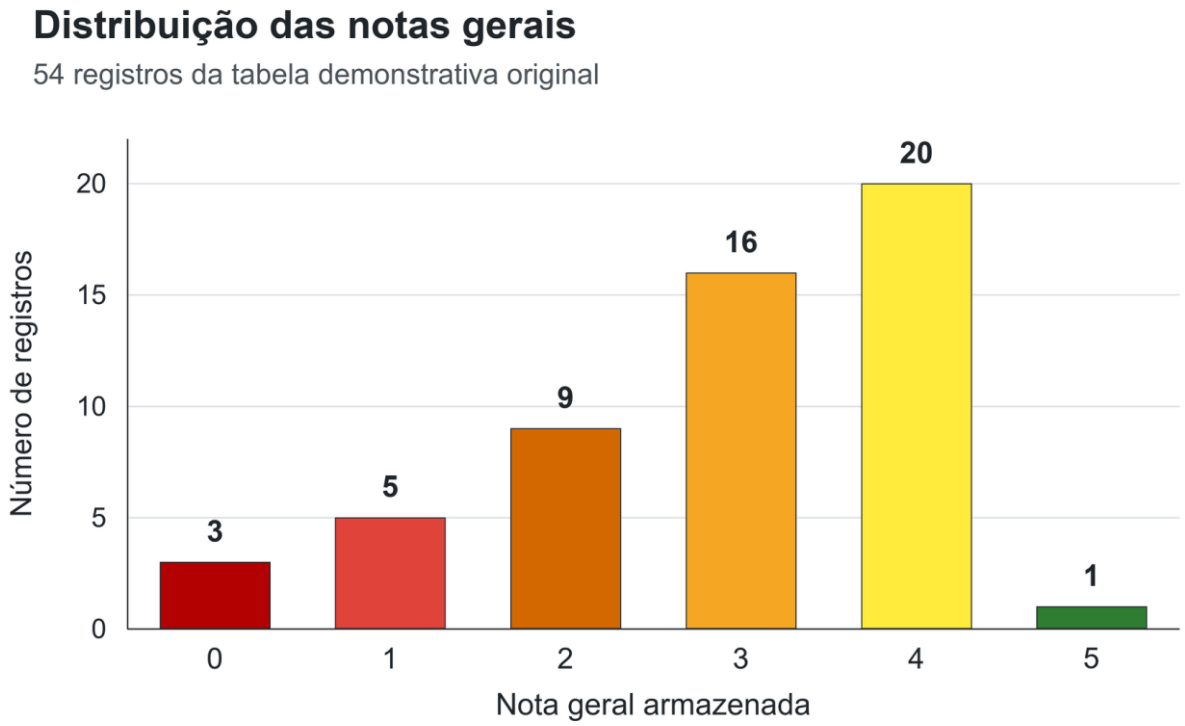


Figura 3 – Distribuição das notas de condição na base demonstrativa do MVP (54 registros).

Fonte: tabela auxiliar do MVP.

3.3 Mapa por trecho

A Figura 4 mostra a distribuição geográfica dos registros da demonstração, com a mesma paleta numérica do histograma. O mapa permite localizar grupos de interesse e associar sua leitura às consultas por corredor. A representação utiliza um contorno territorial público do IBGE para apresentar as coordenadas disponíveis; a arquitetura do MVP mantém a opção de mapas de base OSM com contextily (IBGE, s.d.).

No planejamento, a visualização pode auxiliar a organização espacial das equipes e a discussão de janelas de inspeção. A elaboração de roteiros e a interpretação de agrupamentos

dependem de coordenadas verificadas, acesso viário e condições operacionais. A representação de pontos sobrepostos não comprova um agrupamento de ativos independentes.

As coordenadas da tabela demonstrativa foram mantidas sem deslocamentos artificiais. Parte delas pode representar posições aproximadas. O mapa comunica a aplicação do fluxo e precisa ser atualizado com o cadastro geográfico validado antes de servir ao planejamento de campo.

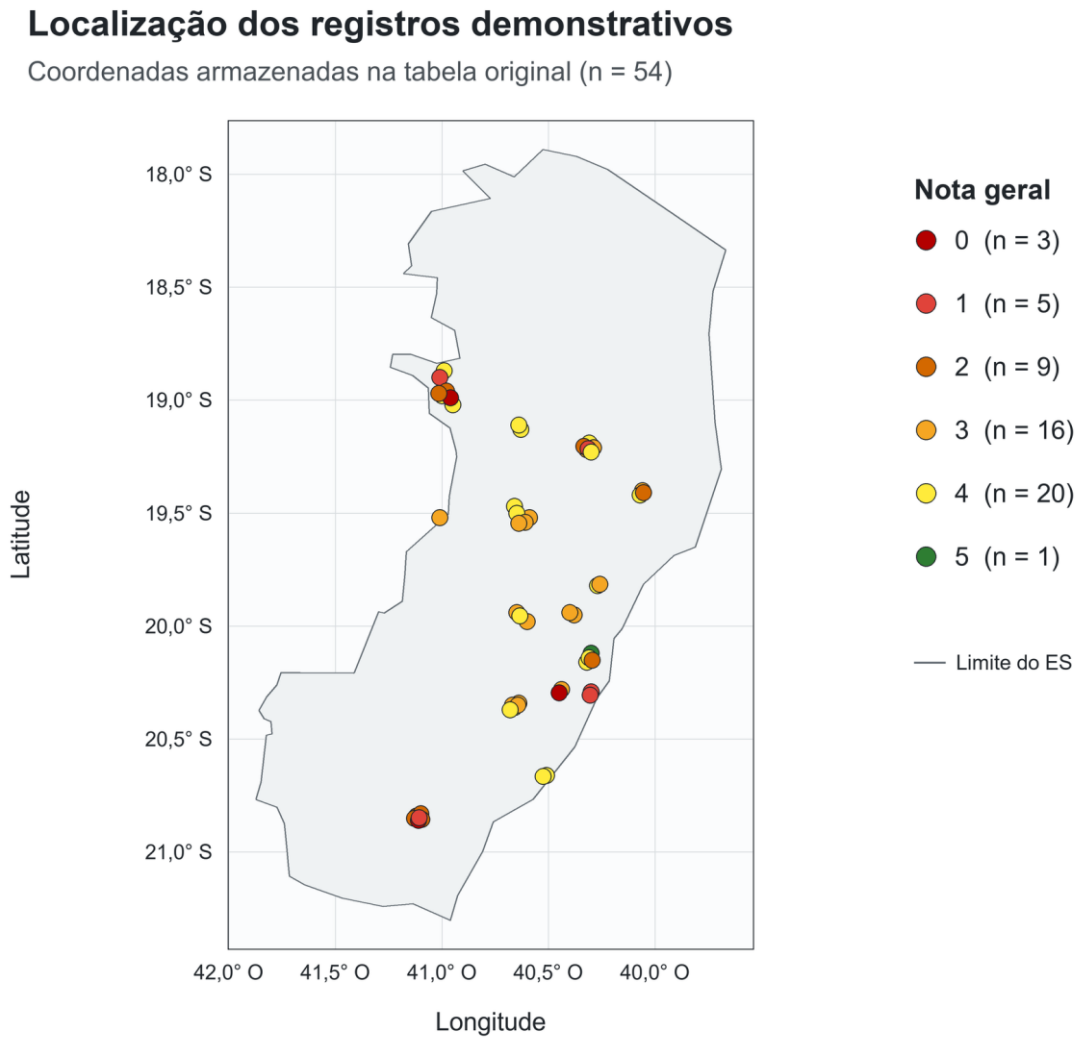


Figura 4 – Distribuição geográfica dos registros demonstrativos por nota geral.

Fonte: tabela auxiliar do MVP e contorno territorial do IBGE. As posições são ilustrativas e podem estar sobrepostas ou aproximadas.

3.4 Qualidade e rastreabilidade

A base inclui campos destinados a registrar arquivo, inspeção e localização, permitindo consultas e correções por registro. A revisão deve conferir a evidência das notas e manter valores ausentes separados de condições críticas. A origem documental e a qualidade da localização acompanham a informação usada para interpretar histogramas e mapas.

3.5 Implicações práticas

A manutenção corretiva é discutida quando a avaliação identifica dano ou perda de funcionalidade que exige intervenção. Notas baixas podem orientar a triagem para revisão

imediata, mas a decisão considera o mecanismo de dano, as restrições existentes e as consequências para os usuários.

O planejamento preventivo utiliza os registros de inspeção para programar ações e acompanhar a evolução das condições. O agrupamento das obras por corredor pode apoiar a organização das equipes, desde que o cadastro e a localização sejam conferidos.

A manutenção preditiva depende do acompanhamento temporal e de informação sobre deterioração. Vibração e telemetria podem complementar as inspeções em diferentes classes de condição; a escolha da instrumentação não deve ser limitada automaticamente às notas mais altas.

3.6 Síntese

O MVP ilustra a passagem de informações documentais para uma tabela e dois produtos de consulta complementares: o histograma resume as notas e o mapa apresenta a distribuição territorial. A contribuição do fluxo é articular essas etapas e preservar sua ligação com as fontes, oferecendo um suporte comum para a discussão de prioridades entre engenharia e operação.

4 Discussão, limitações e trabalhos futuros

4.1 Principais achados

A integração entre extração assistida por IA, consolidação tabular e geovisualização organiza o percurso do relatório ao produto de gestão. O agente trata o conteúdo textual; o ETL torna os campos comparáveis dentro de critérios explícitos; e os produtos gráficos facilitam a consulta do conjunto. Essa articulação é o foco do trabalho.

O histograma e o mapa respondem a questões diferentes: quantos registros pertencem a cada faixa e onde estão localizados. Sua leitura conjunta pode apoiar a preparação de inspeções e a discussão de prioridades, sem dispensar a avaliação individual das estruturas.

A identificação da fonte, da inspeção e das pendências permite revisar o dado e atualizar a base. O benefício esperado em tempo de consolidação e qualidade da decisão deve ser medido em uma aplicação acompanhada, incluindo o esforço de revisão humana.

4.2 Limitações do MVP

A extração depende do layout e da qualidade textual do PDF. Digitalizações com baixa resolução podem produzir erros numéricos e exigir OCR e conferência. As instruções do agente precisam distinguir valores presentes no documento de inferências e registrar os campos que não puderam ser extraídos.

A ausência de coordenadas restringe a interpretação espacial. Centróides e posições aproximadas preservam a comunicação agregada, mas não substituem o cadastro da obra. As correspondências entre planilha e KMZ precisam de verificação da identidade do ativo.

A síntese pelo mínimo das notas parciais é uma regra conservadora de apresentação, dependente de comparabilidade entre as dimensões. Ela não estima probabilidade de falha nem substitui julgamento técnico. Valores ausentes exigem pendência, em vez da atribuição de uma condição.

O conjunto demonstrativo não foi selecionado para representar a distribuição de condições de uma rede. As notas e coordenadas são tratadas como entradas para mostrar os produtos do

fluxo. Não foram demonstrados ganhos quantitativos de tempo, precisão do agente ou estimativas de vida útil remanescente.

4.3 Ameaças à validade

Os registros disponíveis podem concentrar regiões, tipos estruturais e critérios recorrentes no acervo profissional. Diferenças de edição normativa, equipe e data da inspeção afetam a comparação entre notas. A organização dos dados precisa preservar essas informações para separar diferenças de critério de mudanças reais de condição.

Agrupamentos aparentes podem decorrer de coordenadas aproximadas ou de pontos sobrepostos. Associações entre localização, idade, ambiente e deterioração exigem dados complementares e análise própria; não podem ser concluídas apenas pela inspeção visual do mapa.

4.4 Riscos e ética

Os dados profissionais exigem controle de acesso e cuidado na divulgação. O arquivo-fonte e o histórico de revisões devem permanecer disponíveis no ambiente autorizado, enquanto figuras e tabelas de publicação utilizam somente as informações permitidas.

O processo precisa mostrar como cada campo foi obtido, transformado e associado ao ativo. A revisão humana atua nos conflitos, nos campos ausentes e nas situações com repercussão em segurança. O mapa e o histograma apoiam a decisão, cuja responsabilidade permanece com a equipe de engenharia.

4.5 Trabalhos futuros e evolução para VUR

No curto prazo, o desenvolvimento prevê ampliar o conjunto de relatórios e compatibilizar as planilhas e os arquivos de localização, com critérios documentados. A ingestão programada no Drive, a revisão das correspondências cadastrais e o refinamento por quilometragem podem atualizar a base por corredor. Informações de tipologia, idade e ambiente devem acompanhar esse enriquecimento.

Na evolução para VUR, os históricos de inspeção poderão ser associados a medições de vibração e telemetria. A estimativa exige definir evento e horizonte de previsão, modelo de deterioração e limite de desempenho, além de avaliar as incertezas (Saidin et al., 2022; Sonbul; Rashid, 2023). A análise de custos e do momento de intervenção poderá ser desenvolvida quando houver dados compatíveis, sem converter diretamente uma nota em vida remanescente.

A avaliação proposta inclui precisão da extração por campo, proporção de pendências, concordância das notas com uma referência revisada e tempo total de consolidação, incluindo a conferência. A utilidade dos mapas e histogramas deve ser examinada com a equipe usuária. Os dados de avaliação precisam ser separados dos exemplos usados para configurar o agente.

5 Conclusões

Este trabalho apresenta um fluxo orientado por agente de IA para organizar relatórios de inspeção de OAEs em uma base tabular e em visualizações de apoio à manutenção. A contribuição é a integração entre extração documental, tratamento das notas e associação geográfica, preservando a origem de cada registro ao longo do processo.

O conjunto demonstrativo ilustra dois produtos complementares: um histograma de notas de condição e um mapa dos registros por localização. Sua leitura permite selecionar grupos de interesse e estruturar a discussão de prioridades, sujeita à confirmação dos dados e à avaliação de engenharia.

A variação dos relatórios, as lacunas nas notas, as coordenadas aproximadas e as diferenças entre critérios limitam a interpretação dos resultados. O tratamento explícito dessas situações e a revisão humana são necessários para o uso do fluxo em aplicações operacionais. O ganho em tempo e o desempenho do agente permanecem como objetos de avaliação.

A associação dos históricos de inspeção a dados de vibração e telemetria constitui a trilha futura para VUR. Esse desenvolvimento requer modelos calibrados e limites de desempenho definidos, mantendo a ligação entre condição observada, evolução temporal e decisão de manutenção.

Agradecimentos

Aos professores Tiago Tamagusko e Filipe A. C. Nascimento pelas orientações e à equipe do Grupo Encopetro pelo acesso ao material de estudo.

Referências

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. ABNT NBR 9452:2023. Inspeção de pontes, viadutos e passarelas — Procedimento. Rio de Janeiro: ABNT, 2023.

GEOPANDAS. Adding a background map to plots. Documentação oficial, s.d. Disponível em: https://geopandas.org/en/stable/gallery/plotting_basemap_background.html. Acesso em: 19 nov. 2026.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Malhas territoriais. s.d. Disponível em: https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/15774-malhas.html. Acesso em: 19 nov. 2026.

OPENSTREETMAP. Copyright and License. s.d. Disponível em: https://www.openstreetmap.org/copyright. Acesso em: 19 nov. 2026.

SAIDIN, S. S.; JAMADIN, A.; ABDUL KUDUS, S.; MOHD AMIN, N.; ANUAR, M. A. An Overview: The Application of Vibration-Based Techniques in Bridge Structural Health Monitoring. International Journal of Concrete Structures and Materials, v. 16, artigo 69, 2022. https://doi.org/10.1186/s40069-022-00557-1.

SONBUL, O. S.; RASHID, M. Towards the Structural Health Monitoring of Bridges Using Wireless Sensor Networks: A Systematic Study. Sensors, v. 23, n. 20, artigo 8468, 2023. https://doi.org/10.3390/s23208468.

Apêndice A — Ambiente computacional e paleta de visualização

O MVP utiliza Google Colab e Python, com Pandas, GeoPandas, Matplotlib e contextily. Para reprodução, o notebook deve registrar as versões utilizadas, o modelo de linguagem, suas instruções, a data de execução e as entradas. O vínculo abaixo referencia o protótipo; a avaliação de desempenho exige arquivar sua execução e os registros produzidos.

Notebook de referência:
https://colab.research.google.com/drive/1LoCYPDwpbUjaJghDrwIhnD84pa9INPFE

Figura A1 – QR do notebook apresentado no MVP. Fonte: arquivo original do autor.

Paleta das notas 0–5, em ordem:

- #B30000,
- #E0433A,
- #D36800,
- #F5A623,
- #FFEB3B
- e #2E7D32.



As cores são convenções visuais; ausências permanecem em classe própria, sem nota atribuída.

Glossário

- **OAE** — Obra de Arte Especial (ponte/viaduto).
- **DER-ES** — Departamento de Edificações e de Rodovias do Espírito Santo (padrão dos relatórios que você usou).
- **NBR 9452** — Norma ABNT para inspeção de pontes/viadutos (classificação por parâmetros e notas).
- **Relatório Cadastral / Rotineiro / Especial**
  Tipos de inspeção: cadastro inicial, verificação periódica e avaliação aprofundada com ensaios.
- **E/F/D** — Notas de avaliação: Estrutural (E), Funcional (F), Durabilidade (D).
- **NG (Nota Geral)** — Nota operacional adotada; se não vier explícita no relatório, você usa NG = min{E,F,D} (regra conservadora).
- **Semáforo 0–5**
  Paleta de criticidade: 0 (vermelho crítico) → 5 (verde bom), padroniza leitura e ação.
- **LLM** — Large Language Model; aqui, o "agente" que lê PDFs padronizados do DER-ES e extrai campos/valores.
- **OCR** (Reconhecimento Óptico de Caracteres)
  Conversão de imagem em texto pesquisável quando o PDF não possui camada de texto.
- **JSON / CSV**
  Estruturas de dados usadas para registrar extrações por arquivo (JSON) e a base unificada auditável (CSV).
- **Agente (baseado em LLM)** — Pipeline de prompts + exemplos para localizar "Código OAE", "Rodovia", "Nota geral", etc., e gerar um JSON por documento.
- **ETL** — Extração/Transformação/Carga dos dados em Pandas (staging→bronze→silver, com esquema final definido).
- **GeoPandas / contextily (OSM)** — Ferramentas para georreferenciar e mapear as OAEs sobre mosaico OSM (EPSG:4326→3857).
- **Paleta 0–5 (vermelho→verde)** — Classes de criticidade e cores: 0–1 corretiva; 2–3 preventiva; 4–5 preditiva/instrumentada.
- **Hotspot**
  Agrupamento espacial de OAEs com maior criticidade, útil para roteirização e janela de obra.
- **Clusters de criticidade** — Agrupamentos espaciais de OAEs com NG semelhantes para priorização territorial (rotas, equipes).
- **Imputação (nota = 1)** — Quando E/F/D estiverem ausentes, regra conservadora para não subestimar risco.
- **Rastreabilidade** — Guardar fonte_arquivo, data e observações em cada linha para auditoria e reprocessamento.
- **VUR (Vida Útil Remanescente)** — Conceito central da sua pesquisa futura: integrar telemetria/vibração para reduzir incerteza e prever o t* de intervenção.