Aplicação da Metodologia DSRM no Desenvolvimento de um Assistente Virtual Baseado em RAG para Gestão de Contratos Públicos

Murilo M. Q. C. Santos[1], Daniel B. S. Sousa[1], Cassiano H. Albuquerque[2], Emily B. S. Pinto[3]

[1]Coordenação de Sistemas de Informação – Instituto Federal de Alagoas (IFAL) Campus Arapiraca – Rodovia AL-110, 359, Deputado Nezinho – 57317-291 – Arapiraca – AL – Brasil

[2]Instituto Federal de Alagoas (IFAL) - Campus Maragogi – Rodovia AL-101, 139, – 57955-000 – Maragogi – AL – Brasil

[3]Instituto Federal de Alagoas (IFAL) - Reitoria – Rua Odilon Vasconcelos, 103, – 57035-660 – Maceió – AL – Brasil

{mmqcs1,dbss1}@aluno.ifal.edu.br, {cassiano.henrique, emily.santos}@ifal.edu.br

Abstract. *Public contract management requires frequent consultation of fragmented regulations, imposing high cognitive load on inspectors. This paper applies the Design Science Research Methodology (DSRM) to guide the development and evaluation of GeFi, a virtual assistant using Retrieval Augmented Generation (RAG) for the Federal Institute of Alagoas. The artifact is a layered REST API integrating NestJS, Prisma, PostgreSQL with pgvector, LangChain, and Google Gemini. Demonstration indicated the artifact meets project requirements for moderate documents, while evaluation revealed that large notices require advanced chunking strategies.*

Resumo. *A gestão de contratos públicos exige consulta frequente a normativas dispersas, o que impõe elevada carga cognitiva aos fiscais e amplia o tempo de resposta. Este artigo aplica a Design Science Research Methodology para guiar a construção e avaliação do GeFi, um assistente virtual baseado em Retrieval Augmented Generation para o setor de contratos do Instituto Federal de Alagoas. O artefato foi construído como API REST em camadas, utilizando NestJS, Prisma, PostgreSQL com pgvector, LangChain e Google Gemini. A demonstração indicou o atendimento aos requisitos em documentos moderados, enquanto a avaliação revelou que editais volumosos demandam estratégias de segmentação avançadas.*

1. Introdução

A gestão de contratos no setor público envolve o tratamento de grande volume de informações dispersas em editais, contratos, termos aditivos, leis e normas técnicas, o que dificulta a localização rápida dos dispositivos normativos aplicáveis a cada situação. Essa dispersão, somada à complexidade da linguagem jurídica e ao formato não estruturado dos documentos, eleva o tempo de resposta e aumenta o risco de interpretações equivocadas, podendo gerar prejuízos ao erário e responsabilização administrativa dos agentes públicos [STJ 2023].

Nesse contexto, a gestão e a fiscalização de contratos assumem papel central, pois envolvem execução orçamentária, conformidade normativa e responsabilização de

agentes públicos [Matos and Reis 2023; Silva et al. 2024]. Em organizações multicampi, esse desafio se intensifica devido à necessidade de atender unidades geograficamente dispersas e interpretar documentos heterogêneos, como leis, normativos internos, pareceres, manuais e contratos [Barbosa et al. 2024; STJ 2023].

Diante desse contexto, formula-se o seguinte problema de pesquisa: como reduzir a carga cognitiva e o tempo despendido na consulta normativa sem comprometer a fundamentação legal das decisões. Mecanismos de busca por palavra-chave são insuficientes, pois exigem que o fiscal já conheça a terminologia exata empregada pelo legislador [Gao et al. 2023]. Assistentes conversacionais com IA generativa já vêm sendo aplicados ao atendimento em instituições públicas de ensino, seja para conteúdos acadêmicos [Layane et al. 2025], seja para documentos institucionais [Souza et al. 2025; Bomfim 2026], mas raramente ao domínio da fiscalização contratual, no qual a rastreabilidade da fonte é condição de uso e não atributo desejável e no qual os documentos apresentam estrutura hierárquica de artigos, incisos e anexos, com remissões internas frequentes. O projeto GeFi (Gestão e Fiscalização Inteligente) enfrenta esse problema com um assistente virtual que responde em linguagem natural, fundamentado exclusivamente em um acervo normativo indexado, por meio da técnica de Geração Aumentada por Recuperação (RAG, do inglês *Retrieval Augmented Generation*), que combina busca semântica vetorial com Modelos de Linguagem de Grande Escala (LLM) [Lewis et al. 2020].

O desenvolvimento de um artefato tecnológico para esse problema exige rigor metodológico. Por isso, este trabalho adota a Design Science Research Methodology (DSRM) como guia, articulando relevância prática e avaliação sistemática [Peffers et al. 2007; Hevner et al. 2004]. O artigo descreve a aplicação da DSRM ao desenvolvimento do GeFi, com foco na arquitetura em camadas do backend, na persistência com Prisma e na avaliação do artefato. As principais contribuições são: (i) a formalização dos requisitos do setor de contratos; (ii) o projeto e a construção de um sistema funcional integrando API, banco de dados vetorial e modelos de linguagem; e (iii) a identificação de limitações relacionadas à escalabilidade de dados.

2. Fundamentação Teórica

2.1 Gestão de Contratos Públicos e Inteligência Artificial

Sob a Lei nº 14.133/2021, o fiscal acumula atribuições antes distribuídas entre diferentes atores da contratação, o que amplia sua carga de trabalho e sua exposição à responsabilização pessoal [Almeida 2022]. Esse acúmulo ocorre em um contexto de adoção ainda desigual de Inteligência Artificial (IA) no setor público: Ribeiro e Segatto [2025] identificam que cerca de um quarto das organizações federais e estaduais brasileiras já utiliza alguma solução de IA, com concentração naquelas de maior capacidade instalada em tecnologia da informação.

A aplicação de IA a esse contexto pode transformar a rotina de fiscalização ao automatizar a recuperação de informação normativa. Diferentemente da busca por palavra-chave, que depende da correspondência exata de termos, representações vetoriais densas capturam o significado das consultas e recuperam trechos pertinentes ainda que nenhuma palavra da pergunta apareça no texto de origem [Karpukhin et al.

2020] propriedade relevante no domínio jurídico-administrativo, no qual um mesmo conceito é designado por expressões distintas em leis, instruções normativas e editais.

2.2 RAG e Modelos de Linguagem de Grande Escala

Modelos de Linguagem de Grande Escala (LLMs) são redes neurais treinadas sobre extensos corpos textuais, capazes de gerar texto coerente e de executar tarefas de processamento de linguagem natural [Brown et al. 2020]. Apesar de sua capacidade generativa, os LLMs apresentam limitações relevantes: o conhecimento parametrizado é estático e o modelo pode produzir afirmações incorretas, fenômeno conhecido como alucinação [Ji et al. 2023].

A técnica RAG [Lewis et al. 2020] mitiga essas limitações ao condicionar a geração a um conjunto de documentos recuperados dinamicamente. Em sua formulação canônica, um sistema RAG opera em dois estágios. No primeiro, de indexação, os documentos são segmentados em fragmentos, convertidos em vetores por um modelo de embeddings e armazenados em um banco vetorial. No segundo, de consulta, a pergunta do usuário é vetorizada, os fragmentos semanticamente mais próximos são recuperados e fornecidos ao modelo como contexto para a geração da resposta [Gao et al. 2023]. Essa abordagem permite respostas fundamentadas e auditáveis, propriedade essencial em aplicações do setor público.

2.3 A Metodologia DSRM

A Design Science Research Methodology, proposta por Peffers et al. [2007], estrutura a pesquisa orientada à construção de artefatos em seis fases: (1) identificação do problema e motivação; (2) definição dos objetivos da solução; (3) projeto e desenvolvimento; (4) demonstração; (5) avaliação; e (6) comunicação. A metodologia se fundamenta no paradigma da ciência do projeto em sistemas de informação [Hevner et al. 2004], que valoriza tanto a relevância prática quanto o rigor científico.

A DSRM é particularmente adequada ao desenvolvimento do GeFi por três razões. Primeira, o objeto central da pesquisa é um artefato de software cuja utilidade precisa ser demonstrada em um contexto organizacional real. Segunda, a metodologia prevê ciclos iterativos, o que acomoda a natureza incremental do projeto. Terceira, a fase explícita de avaliação permite identificar limitações do artefato antes de sua implantação definitiva, reduzindo o risco de falha operacional. A Tabela 1 mapeia as seis fases da DSRM às atividades realizadas no GeFi.

Tabela 1. Mapeamento das fases da DSRM às atividades de desenvolvimento do GeFi.

Fase da DSRM	Atividade no GeFi
1. Identificação do problema	Sobrecarga cognitiva e morosidade na consulta normativa do setor de contratos
2. Objetivos da solução	Definição dos requisitos funcionais e não funcionais (Tabelas 2 e 3)
3. Projeto e desenvolvimento	Arquitetura em camadas, API NestJS, ORM Prisma e pipeline RAG
4. Demonstração	Consulta bem sucedida sobre documento de extensão moderada

5. Avaliação	Análise do comportamento em editais volumosos e da parametrização do RAG
6. Comunicação	Artigos e documentação técnica do projeto

3. Modelagem e Desenvolvimento do Artefato

Esta seção descreve as etapas construtivas do artefato, englobando as três primeiras fases da metodologia DSRM. Inicialmente, relata-se a identificação do problema e a definição dos objetivos da solução (Fases 1 e 2) por meio do levantamento de requisitos. Em seguida, detalha-se o projeto da arquitetura de software, a modelagem de dados e a implementação do pipeline de Inteligência Artificial (Fase 3).

3.1 Identificação do Problema e Objetivos

A primeira fase da DSRM consolidou o problema central: o fiscal de contratos despende tempo excessivo na localização de dispositivos normativos pertinentes, em virtude do volume documental, da complexidade jurídica e da ausência de uma interface unificada de consulta. A segunda fase traduziu esse problema em objetivos mensuráveis de solução, expressos como requisitos funcionais (Tabela 2) e requisitos não funcionais (Tabela 3), extraídos das necessidades do setor de contratos do Instituto Federal de Alagoas (IFAL).

Tabela 2. Requisitos funcionais do artefato GeFi.

ID	Requisito Funcional
RF01	Cadastro e autenticação de usuários com perfis distintos (administrador e fiscal)
RF02	Envio de documentos normativos em formato PDF, restrito ao administrador
RF03	Extração e indexação automática do conteúdo textual dos documentos
RF04	Consulta em linguagem natural à base normativa
RF05	Geração de respostas fundamentadas exclusivamente no acervo indexado
RF06	Rastreabilidade das fontes normativas citadas em cada resposta
RF07	Gestão do histórico de conversas (criação, listagem, fixação e exclusão lógica)

Tabela 3. Requisitos não funcionais do artefato GeFi.

ID	Requisito Não Funcional
RNF01	Segurança: hashing de senhas com bcrypt e autenticação stateless via JWT
RNF02	Escalabilidade horizontal do backend
RNF03	Manutenibilidade por separação em camadas e baixo acoplamento
RNF04	Testabilidade por injeção de dependências
RNF05	Desempenho: indexação assíncrona para não bloquear o envio do documento
RNF06	Integridade e auditabilidade dos dados (exclusão lógica e trilha de auditoria)
RNF07	Confiabilidade das respostas pela restrição ao contexto recuperado

3.2 Design e Desenvolvimento

A terceira fase materializou os requisitos em um artefato de software. Para garantir a manutenibilidade e a escalabilidade, a engenharia do GeFi foi segmentada em três pilares principais: arquitetura do backend, persistência de dados e pipeline de recuperação de informações.

3.2.1 Arquitetura do Backend e Modularização

O backend do GeFi foi implementado como uma API REST organizada em arquitetura em camadas [Fowler 2002], utilizando o framework NestJS sobre a linguagem TypeScript. A aplicação adota uma organização modular, na qual cada módulo encapsula um domínio coeso de negócio, garantindo alta coesão interna e baixo acoplamento estrutural.

A separação em camadas, conforme ilustrado na Figura 1, estrutura cada módulo em três níveis de responsabilidade. A camada de apresentação, composta pelos controladores, recebe as requisições HTTP, aplica os guardas de autenticação e delega o processamento. A camada de aplicação, composta pelos serviços, concentra a lógica de negócio e a orquestração das operações. A camada de infraestrutura, representada pelo serviço de acesso a dados e pelas integrações externas, encapsula a persistência e a comunicação com os modelos de linguagem. A comunicação flui da camada de apresentação para a de aplicação e desta para a de infraestrutura, sempre por injeção de dependências, o que preserva a substituibilidade das implementações e favorece a testabilidade.

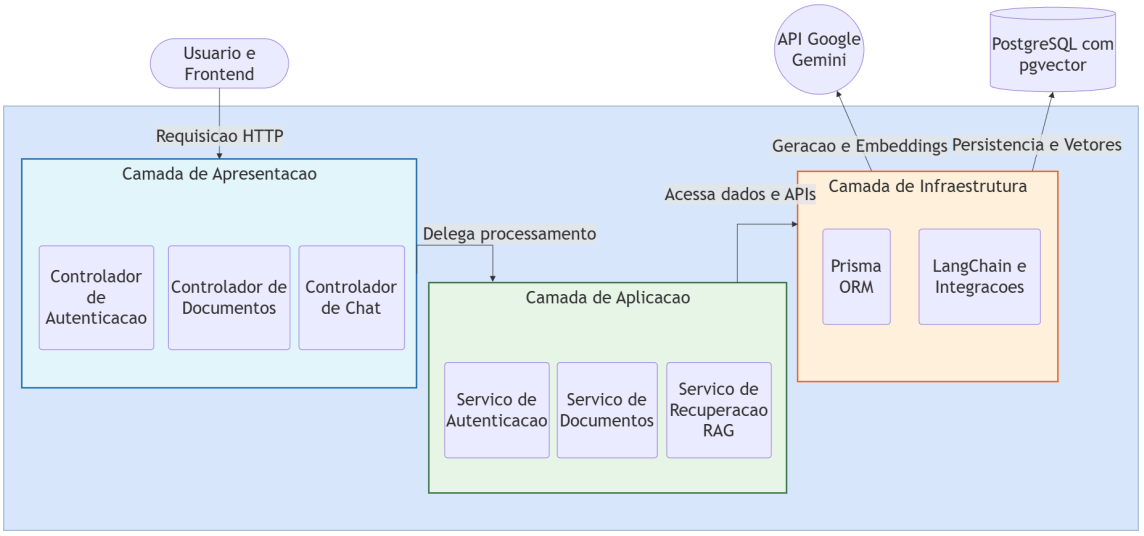


Figura 1. Arquitetura em camadas do GeFi, evidenciando a camada de apresentação, a camada de aplicação e a camada de infraestrutura, além das integrações com o banco de dados e com os modelos de linguagem.

A modularização do backend reflete os domínios identificados na fase de objetivos. O módulo de autenticação implementa o cadastro, o login e a emissão de tokens JWT, com hashing de senhas por bcrypt e controle de acesso baseado em papéis, distinguindo os perfis de administrador e de fiscal. O módulo de usuários gerencia os dados cadastrais e a busca de informações dos servidores. O módulo de documentos gerencia o ciclo de vida do acervo normativo. O módulo de conversas administra o histórico de interações e as perguntas dos fiscais. O módulo de recuperação concentra o pipeline RAG. Por fim,

o módulo de persistência encapsula a integração e a comunicação direta com o banco de dados PostgreSQL por meio do ORM Prisma.

3.2.2. Modelagem de Dados e Persistência

A persistência dos dados foi implementada com o ORM Prisma sobre o PostgreSQL. O Prisma adota uma abordagem declarativa, na qual o esquema de dados é descrito em um arquivo único de definição, a partir do qual é gerado um cliente de acesso fortemente tipado [Prisma Labs 2021]. Diferentemente de mapeadores que resolvem o mapeamento objeto relacional em tempo de execução, o Prisma gera código em tempo de compilação, o que favorece segurança de tipos em todas as operações e elimina uma classe inteira de erros que, de outro modo, somente seriam detectados em execução. O serviço de acesso a dados estende o cliente gerado e estabelece a conexão no ciclo de inicialização do módulo, sendo compartilhado por toda a aplicação como instância única.

O modelo de dados, definido no esquema do Prisma e versionado por migrações controladas, compreende cinco entidades principais: usuários, conversas, histórico de interações, fontes de resposta e documentos normativos. A entidade de fontes de resposta atua como tabela de junção entre o histórico de interações e os documentos normativos, implementando um relacionamento muitos para muitos que viabiliza a rastreabilidade. A Figura 2 ilustra a estrutura visual desse modelo e os seus respectivos relacionamentos. As migrações registradas evidenciam a evolução controlada do esquema desde o início do desenvolvimento.

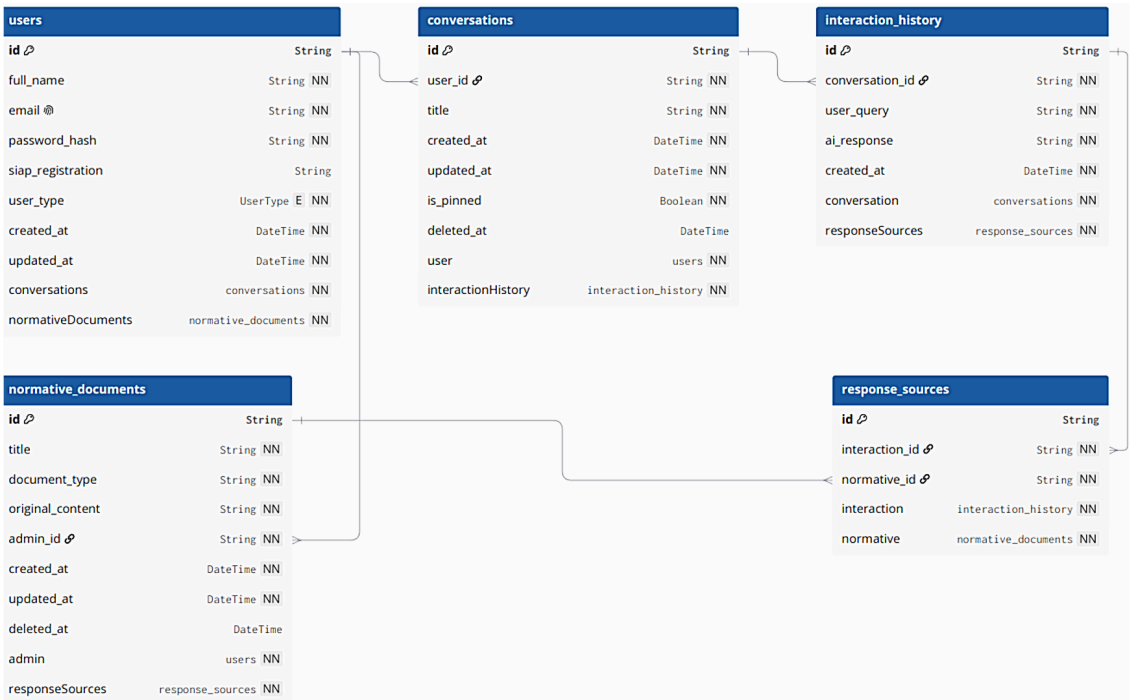


Figura 2. Modelo entidade relacionamento do GeFi, evidenciando as cinco entidades principais e seus relacionamentos.

3.2.3. Pipeline de Ingestão e Recuperação (RAG)

O pipeline de ingestão de documentos integra a camada de aplicação à infraestrutura de Inteligência Artificial, orquestrada pela biblioteca LangChain. Quando o administrador envia um documento normativo, o controlador aciona, de forma assíncrona, o carregador de PDF para extrair o conteúdo textual, evitando bloqueio da interface. O texto é então segmentado em fragmentos de 1000 caracteres, com sobreposição de 200 caracteres entre eles, preservando a continuidade semântica e reduzindo o corte de tabelas e cláusulas complexas.

Cada fragmento é convertido em uma representação vetorial densa pelo modelo de embeddings Gemini Embedding 2, responsável por capturar o significado semântico do trecho em um espaço de alta dimensionalidade. Os vetores são armazenados no PostgreSQL com a extensão pgvector [Kane 2021], que adiciona suporte a vetores e operadores de similaridade, incluindo a similaridade de cosseno utilizada na busca. Optou-se pelo pgvector, em vez de um índice vetorial dedicado como o FAISS [Souza et al. 2025], por dispensar um segundo sistema gerenciador e permitir filtragem relacional e busca por similaridade na mesma consulta. Essa abordagem unifica as bases relacional e vetorial em uma única infraestrutura, reduzindo a complexidade operacional.

Integrada a essa base, a consulta normativa é exposta por um ponto de acesso da API que recebe a pergunta do fiscal e o identificador da conversa. O controlador verifica o token de autenticação e a propriedade da conversa, enquanto o serviço de recuperação vetoriza a pergunta e executa a busca semântica, retornando os cinco fragmentos mais relevantes segundo a similaridade de cosseno. Esses fragmentos compõem o contexto enviado ao modelo Gemini 2.5 Flash, configurado com temperatura 0.3 para favorecer respostas mais determinísticas. O prompt de sistema restringe a geração ao conteúdo recuperado, mitigando o risco de alucinações jurídicas. Ao final, a interação é persistida juntamente com as fontes utilizadas, promovendo a rastreabilidade das respostas. A Figura 3 ilustra o fluxo de dados da etapa de consulta normativa e a geração da resposta.

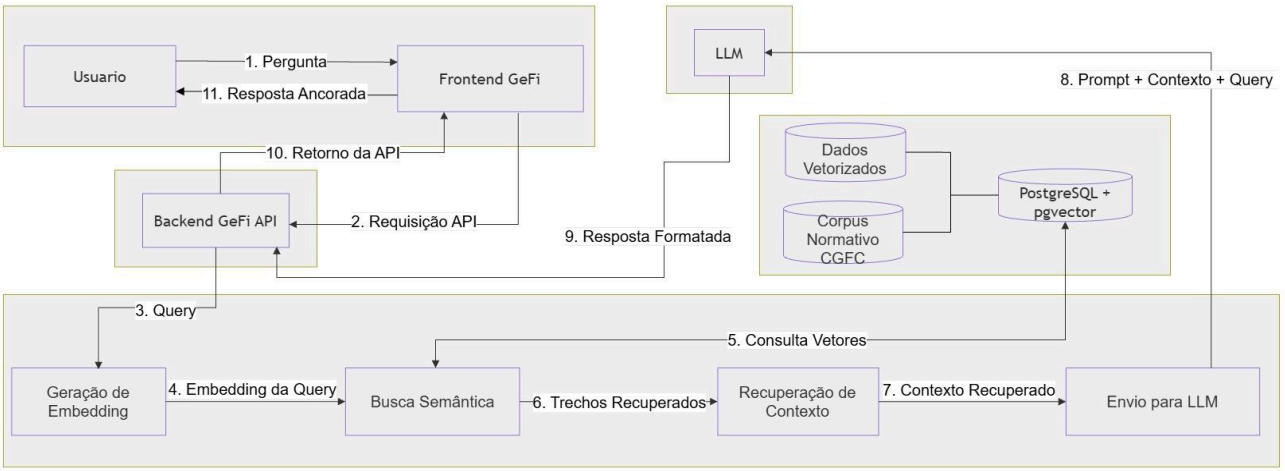


Figura 3. Pipeline RAG do GeFi ilustrando a etapa de consulta, desde a vetorização da pergunta até a montagem do contexto e geração da resposta.

4. Demonstração e Avaliação

Esta seção contempla as fases 4 e 5 da metodologia DSRM, focando na aplicação prática do GeFi e na análise de seus resultados preliminares. Para aferir a eficácia do sistema e mapear seus limites arquiteturais, foram conduzidos testes de uso controlados, cujos resultados empíricos estão sumarizados na Tabela 4.

Tabela 4. Resumo da avaliação empírica do pipeline RAG do GeFi.

Parâmetro de Avaliação	Cenário 1 (Controle)	Cenário 2 (Contexto Real)
Documento Testado	Documento normativo de teste	Instrução Normativa nº 05/2017
Extensão	02 páginas	Extensa, 57 paginas (Dezenas de artigos e anexos complexos)
Pergunta feita	Qual é o valor máximo mensal para que um serviço seja considerado de pequeno vulto?	Quais são as atribuições do fiscal administrativo de acordo com a IN 05/2017?
Resposta esperada	R$ 5.000,00 (cinco mil reais).	Acompanhar obrigações previdenciárias e trabalhistas (Art. 40), recebimento provisório (Art. 50), atuar na rescisão (Art. 64) e executar rotinas do Anexo VIII-B.
Resposta obtida	"O valor mensal de faturamento não pode ultrapassar o limite de R$ 5.000,00 (cinco mil reais)."	Resposta excludente e incompleta. O sistema citou apenas a atribuição na rescisão (Art. 64) e afirmou erroneamente que o Art. 40 não detalhava outras funções.
Limitações observada	Nenhuma	Falha na recuperação de informações dispersas e incapacidade de correlacionar os anexos com o corpo principal do texto.

4.1 Demonstração

A quarta fase da DSRM se destina a demonstrar o uso do artefato na resolução de uma instância do problema. Conforme ilustrado no primeiro cenário da Tabela 4, o sistema foi submetido a um documento de extensão moderada e estrutura textual homogênea. Após o envio pelo perfil administrador e a vetorização da pergunta pelo fiscal, o sistema recuperou os fragmentos semanticamente pertinentes e gerou uma resposta direta, acompanhada da indicação exata da fonte.

A interação foi corretamente registrada no histórico com o respectivo vínculo normativo. Esse resultado demonstra que o artefato atende, na prática, às especificações projetadas para cenários de complexidade moderada, garantindo a rastreabilidade das informações e a confiabilidade das respostas, uma vez que o modelo se manteve estritamente restrito ao acervo indexado.

4.2 Avaliação

A quinta fase examinou o comportamento do artefato no contexto real de grandes contratos (segundo cenário da Tabela 4), no qual os editais frequentemente ultrapassam dezenas de páginas e apresentam estrutura heterogênea. A avaliação revelou uma limitação associada à parametrização atual do pipeline RAG. Em documentos extensos, a estratégia de segmentação de tamanho fixo tende a separar dispositivos relacionados e romper a continuidade lógica das cláusulas (como tabelas e listas de sanções), resultando em respostas incompletas.

Foi precisamente a etapa de avaliação empírica prevista pela DSRM que permitiu mapear essa limitação antes de uma implantação definitiva no setor de contratos. A análise indicou que a complexidade e o volume dos editais demandam estratégias de processamento mais avançadas, tais como a segmentação semântica, a segmentação hierárquica e a introdução de uma etapa de reordenação por relevância. Cabe ressaltar que tais limitações não invalidam o artefato: elas constituem requisitos de escalabilidade de dados a serem incorporados ao próximo ciclo da metodologia. A separação em camadas assegura que essas melhorias possam ser introduzidas de forma isolada no módulo de recuperação, evidenciando o benefício de manutenibilidade projetado na Fase 3.

5. Conclusão e Trabalhos Futuros

A sexta fase da DSRM é a comunicação do conhecimento produzido, concretizada por este artigo, que descreveu um primeiro ciclo completo da metodologia aplicado ao GeFi: a partir de um problema concreto do setor de contratos do IFAL, definiram-se os objetivos de solução, construiu-se um artefato funcional e realizaram-se as fases de demonstração e avaliação.

O artefato, estruturado em arquitetura em camadas e apoiado pelo ORM Prisma, integrou a API REST, o banco de dados relacional vetorial e os modelos de linguagem do Google Gemini. A demonstração confirmou a aderência do sistema aos requisitos projetados em cenários de complexidade moderada, evidenciando o potencial do GeFi para reduzir a carga cognitiva e o tempo de resposta dos fiscais na consulta normativa.

O primeiro ciclo da DSRM foi concluído com êxito, gerando um artefato operacional e requisitos claros de escalabilidade de dados. Os trabalhos futuros incluem a adoção de estratégias avançadas de segmentação, a introdução de reordenação por relevância e a realização de uma avaliação empírica com fiscais, para mensurar quantitativamente os ganhos de produtividade proporcionados pelo artefato.

6. Referências

Almeida, C. W. L. de. (2022) "Fiscalização contratual na Lei nº 14.133/2021: governança e resultado na execução de contratos". Revista do TCU, n. 150.

Barbosa, R. F. et al. (2024) "Fiscalização de contratos nas instituições federais de ensino superior: desafios e perspectivas na gestão contratual". Revista Foco, v. 17, n. 12, e7100.

Bomfim, P. H. A. (2026) "Chat institucional com inteligência artificial para acesso a informações oficiais do IFBA". TCC (Licenciatura em Computação), Instituto Federal da Bahia, Porto Seguro.

Brown, T. et al. Language Models are Few Shot Learners. In: Advances in Neural Information Processing Systems, v. 33, pp. 1877 a 1901, 2020.

Fowler, M. Patterns of Enterprise Application Architecture. Addison Wesley Professional, 2002.

Gao, Y. et al. Retrieval Augmented Generation for Large Language Models: A Survey. arXiv:2312.10997, 2023.

Hevner, A. R.; March, S. T.; Park, J.; Ram, S. Design Science in Information Systems Research. MIS Quarterly, v. 28, n. 1, pp. 75 a 105, 2004.

Ji, Z. et al. Survey of Hallucination in Natural Language Generation. ACM Computing Surveys, v. 55, n. 12, pp. 1 a 38, 2023.

Kane, A. pgvector: Open source vector similarity search for Postgres. 2021. Disponível em: https://github.com/pgvector/pgvector.

Karpukhin, V. et al. (2020) "Dense passage retrieval for open-domain question answering". In: Proceedings of EMNLP, p. 6769 a 6781.

Lewis, P. et al. Retrieval Augmented Generation for Knowledge Intensive NLP Tasks. In: Advances in Neural Information Processing Systems, v. 33, pp. 9459 a 9474, 2020.

Matos, F. R. P.; Reis, D. L. (2023) "Gestão de contratos como instrumento de redução de custo: planejamento e fiscalização dos recursos públicos". Revista Científica Semana Acadêmica, v. 11, n. 232.

Peffers, K.; Tuunanen, T.; Rothenberger, M. A.; Chatterjee, S. A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems, v. 24, n. 3, pp. 45 a 77, 2007.

Prisma Labs. Prisma: Next generation ORM for Node.js and TypeScript. 2021. https://www.prisma.io/docs.

Ribeiro, M. M.; Segatto, C. I. (2025) "Inteligência artificial nas organizações públicas brasileiras: heterogeneidades e capacidades em TI". Revista de Administração Pública, v. 59, n. 1, e2024-0066.

Silva, E. C.; Moreira, A. A. A. P.; Nascimento, L. R.; Serafim, G. L.; Santos, H. S.; Mendes, M. C. (2024) “Uma revisão sistemática do uso de tecnologias digitais na modernização e agilização dos processos de licitação pública”. *Revista Foco*, v. 17, n. 2, e4491, p. 1–20. DOI: 10.54751/revistafoco.v17n2-106.

Souza, W. F. V. de et al. (2025) "Assistente virtual inteligente para instituições educacionais: aplicação da técnica RAG no IFBot". In: Anais do VI Workshop sobre as Implicações da Computação na Sociedade (WICS), p. 217 a 224. Porto Alegre: SBC.

STJ. 2023. Manual de Gestão e Fiscalização de Contratos. Superior Tribunal de Justiça.