Uso de Bases de Conhecimento em Soluções de IA

Metodologia, implementação e acompanhamento de projetos de IA

Junho 2026

MINISTÉRIO DA GESTÃO E DA INOVAÇÃO EM SERVIÇOS PÚBLICOS

SUMÁRIO

SUMÁRIO

SUMÁRIO

SUMÁRIO

SUMÁRIO

1. Resumo operacional para equipes técnicas

Esta subseção concentra as decisões e alertas que uma equipe técnica precisa consultar antes de iniciar qualquer projeto que envolva base de conhecimento. O restante do documento detalha e justifica cada ponto.

Quando usar base de conhecimento estruturada:

- Demanda recorrente do mesmo conhecimento por múltiplos usuários ou canais.
- Necessidade de rastreabilidade e auditoria por exigência legal, contratual ou regulatória.
- Integração prevista com pipeline RAG ou assistente conversacional.
- Risco de perda de conhecimento crítico por rotatividade de servidores.
- Inconsistência de versões entre áreas causando erros operacionais.

Quando não usar:

[1] Conhecimento predominantemente tácito, contextual e não redutível a conteúdo explícito sem perda de precisão relevante.

[2] Volume pequeno, estável e de uso restrito: o overhead de governança supera o benefício.

[3] Organização sem papéis mínimos de curadoria: a base degradará mais rápido do que um repositório simples gerenciado manualmente.

Decisões arquiteturais que devem ser tomadas antes da experimentação:

- Nível de acesso do conteúdo (Público / Interno / Restrito / Confidencial): distinguir entre (a) classificação legal de sigilo nos termos da LAI; (b) categoria de dado pessoal pela LGPD (dado pessoal comum, dado pessoal sensível); e (c) perfil operacional de acesso interno ao projeto. "Restrito" não equivale a "sigiloso" e controle de acesso não substitui base legal LGPD. Essa distinção determina onde a base pode ser hospedada e quais APIs externas podem ser usadas.
- Base legal aplicável conforme arts. 7º, 11 e 23 da LGPD: no setor público, avaliar prioritariamente obrigação legal, regulatória ou execução de política pública, evitando consentimento quando houver assimetria de poder ou prestação obrigatória de serviço público. A escolha determina os controles exigidos e a necessidade de RIPD.
- Integração com IA prevista: se sim, os requisitos de metadados, proveniência e controle de acesso pré-recuperação são obrigatórios desde o início.
- Modelo de hospedagem (*on-premises*, nuvem nacional, nuvem estrangeira, e existência de subprocessadores, localidade de backup, telemetria remota, modalidade de suporte e política do fornecedor quanto ao uso de dados para treinamento ou melhoria de modelos): impacta soberania, localização de dados e cláusulas contratuais exigidas.

Evidências que devem ser produzidas por fase:

- *Prospecção*: inventário de fontes e lacunas documentado.
- *Experimentação*: relatório de testes de busca e avaliação preliminar de privacidade.
- *Implementação*: matriz de metadados completa, taxonomia versionada, testes de controle de acesso.

- *Implantação*: plano de *rollback* e relatório de acessibilidade.
- *Sustentação*: relatório mensal de *zero-result*, incidentes, artigos expirados e acessos negados.
- *Desativação*: exportação validada, revogação de integrações, preservação de evidências.

Evidências mínimas para aprovação técnica (entrada em produção):

- Diagrama de arquitetura atualizado com fluxos de dados.
- Fluxo de dados com classificação de sensibilidade por categoria.
- Inventário de fontes com responsável, licença e periodicidade de atualização.
- Matriz de metadados completa para 100% dos artefatos publicados.
- Matriz de base legal por categoria de dado pessoal tratado.
- Plano de testes com cenários de acesso autorizado e negado.
- Relatório de acessibilidade (eMAG/WCAG) com resultado por critério.
- Relatório de segurança com resultados dos testes de controle de acesso.
- Plano de *rollback* documentado e com resultado de teste.
- Plano de resposta a incidentes de segurança e privacidade.
- Termo de aceite assinado pelo gestor responsável.

Riscos que bloqueiam entrada em produção:

- Artigos sem dono ativo identificado.
- Metadados obrigatórios incompletos.
- Controle de acesso não testado com usuário sem permissão.
- Base que alimenta IA sem filtro de permissão pré-recuperação testado.
- Dados pessoais na base sem base legal documentada.
- Nenhum mecanismo de revogação de acesso testado para desligamento de servidores, incluindo revogação de *tokens* de autenticação, chaves de API, contas de serviço, conectores de integração, índices vetoriais derivados, *embeddings* e caches associados.

1.1 Finalidade do anexo

Este documento integra o **GuIA — Guia unificado de Inteligência Artificial para o Setor Público Brasileiro** — e tem por finalidade orientar equipes técnicas, áreas demandantes, gestores, fiscais de contrato e equipes de sustentação quanto à concepção, implementação, governança e avaliação de soluções de inteligência artificial que utilizem bases de conhecimento como componente estruturante.

Uma base de conhecimento, neste contexto, não é um simples repositório de arquivos. É um sistema sociotécnico governado cujo propósito é converter conhecimento disperso e tácito em conteúdo explícito, encontrável, reutilizável, validado e auditável. Essa distinção importa

porque decisões de projeto que tratam a base como depósito passivo tendem a produzir acervos redundantes, desatualizados e sem impacto mensurável, independentemente da plataforma escolhida. A ISO 30401, referência normativa internacional para sistemas de gestão do conhecimento, estabelece exatamente esse entendimento: a gestão do conhecimento é uma capacidade organizacional contínua, com requisitos de estabelecimento, implementação, manutenção, revisão e melhoria, não um projeto pontual de digitalização de conteúdo.

O presente anexo não prescreve plataformas nem impõe uma única arquitetura. Estabelece os requisitos mínimos, os padrões técnicos aplicáveis, os critérios de governança e os indicadores de qualidade que qualquer solução baseada em bases de conhecimento deve satisfazer para operar de forma segura, auditável e compatível com o ordenamento jurídico e normativo brasileiro.

1.2 Relação com o GuIA e outros anexos

O **GuIA** define o marco geral de governança, os princípios éticos, as categorias de risco, os requisitos transversais de privacidade e acessibilidade, e os critérios de ciclo de vida aplicáveis a qualquer sistema de inteligência artificial no setor público. Este anexo não reproduz esse marco: pressupõe que o leitor o conhece e concentra-se nos aspectos específicos de bases de conhecimento que o documento principal não detalha.

A relação entre este anexo e o **GuIA** é de especificação técnica progressiva. O **GuIA** fixa o *o quê* — requisitos de transparência, rastreabilidade, responsabilização, não discriminação, privacidade, acessibilidade. Este anexo desdobra o *como* para a classe de soluções que dependem de bases de conhecimento estruturadas, sejam elas utilizadas em sistemas de Geração Aumentada por Recuperação (RAG), em serviços de autoatendimento, em assistentes conversacionais, em portais de informação pública ou em plataformas internas de gestão do conhecimento institucional.

Outros anexos técnicos do **GuIA**, como os relativos a Modelos de Linguagem de Grande Escala (LLM) e a sistemas de Geração Aumentada por Recuperação (RAG), dependem diretamente das definições e requisitos estabelecidos neste documento. Toda solução RAG pressupõe uma base de conhecimento; toda solução LLM que opere com documentos institucionais como contexto depende de uma base de conhecimento adequadamente estruturada, governada e auditável. Este anexo é, portanto, condição necessária para a correta aplicação daqueles.

Delimitação em relação a RAG e LLM: este anexo trata da camada de conteúdo, curadoria, metadados, proveniência, controle de acesso e governança da base de conhecimento. O anexo de RAG deve tratar da arquitetura de recuperação e orquestração do contexto, enquanto o anexo de LLM deve tratar do comportamento, avaliação e governança do modelo. A fusão desses anexos não é recomendada como regra, pois reduziria a clareza das responsabilidades técnicas; a integração entre eles deve ocorrer por referências cruzadas e critérios comuns de risco, métricas e aceite.

1.3 Público-alvo

Este documento se dirige a seis perfis com responsabilidades distintas no ciclo de vida de soluções que envolvem bases de conhecimento:

- **Equipes técnicas de desenvolvimento e arquitetura**: responsáveis por decisões de representação do conhecimento, escolha de plataformas, estruturação de metadados, integração com mecanismos de busca e pipelines de IA. Devem consultar este anexo para delimitar requisitos técnicos mínimos e padrões de interoperabilidade.
- **Gestores e equipes de produto**: responsáveis por priorização, roadmap e métricas de valor. Devem consultar este anexo para definir indicadores de desempenho adequados e evitar a armadilha de medir atividade no lugar de impacto.
- **Áreas demandantes e gestores de conhecimento**: responsáveis pela curadoria de conteúdo, pela definição de taxonomias e vocabulários, e pela governança editorial. Devem consultar este anexo para estabelecer papéis, fluxos e políticas de ciclo de vida.
- **Fiscais de contrato e equipes de compras**: responsáveis por especificar, avaliar e aceitar soluções contratadas. Devem consultar a 11 deste anexo, que consolida os requisitos mínimos de contratação e os critérios de aceite.
- **Equipes de segurança, privacidade e conformidade**: responsáveis por garantir aderência à LGPD, aos guias da ANPD, às diretrizes de acessibilidade (WCAG e eMAG) e à Lei 15.263/2025. Devem consultar as seções de governança de dados e riscos.
- **Equipes de sustentação**: responsáveis pela operação continuada, detecção de degradação de qualidade e gestão de incidentes. Devem consultar as seções de métricas, modos de falha e padrões operacionais.

1.3.1 Como usar este anexo por perfil

Leitura orientada por perfil: o anexo não precisa ser consumido linearmente por todos os leitores. Gestores de política pública e produto devem priorizar finalidade, aplicabilidade, ciclo de vida, métricas e critérios de aceite; desenvolvedores e arquitetos devem priorizar técnicas, arquitetura, metadados e integração; segurança, privacidade e conformidade devem priorizar governança de dados, riscos, controles e requisitos mínimos. A tabela a seguir funciona como trilha de leitura para reduzir sobrecarga informacional e direcionar cada perfil às seções indispensáveis.

Perfil	Decisões principais	Artefatos esperados	Seções obrigatórias
Arquiteto de solução	Modelo de hospedagem, arquitetura de índices, integração RAG	Diagrama de arquitetura, fluxo de dados, matriz de metadados	Seções 3, 4, 5, 8
Desenvolvedor	Estrutura de tópicos, metadados, APIs, controle de acesso	Código de ingestão, testes de acesso, logs de auditoria	Seções 4, 5, 8, 9
Segurança e privacidade	Controles de acesso, RAG *security*, transferência internacional	Relatório de segurança, testes de *prompt injection*, inventário de subprocessadores	Seções 8, 9, 11

Perfil	Decisões principais	Artefatos esperados	Seções obrigatórias
Encarregado / LGPD	Base legal, RIPD, incidentes, transferência internacional	Matriz de base legal, RIPD (quando aplicável), registro de incidentes	Seções 2, 9, 11
Fiscal de contrato	Critérios de aceite, obrigações de sustentação, cláusulas de privacidade	Termo de aceite, matriz de critérios, checklist de entregáveis	Seção 12 (requisitos)
Curador de conteúdo	Taxonomia, ciclo de vida editorial, qualidade	Inventário de fontes, taxonomia versionada, relatório de expiração	Seções 4, 5, 6, 7
Gestor de serviço	Métricas de impacto, indicadores de saúde, roadmap	Relatório mensal de métricas, plano de melhoria	Seção 7 (métricas)
Operador de sustentação	Monitoramento, incidentes, alertas de degradação	Relatório de zero-result, log de incidentes, relatório de artigos expirados	Seções 7, 9, 10

1.4 Como este anexo deve ser usado em projetos e contratações

Em projetos de desenvolvimento interno, este anexo deve ser referenciado nas fases de estruturação e experimentação, quando as decisões arquiteturais determinam se a base de conhecimento será estruturada o suficiente para sustentar as fases posteriores. Erros nessa etapa têm custo de correção exponencialmente maior do que na produção, em especial quando a base alimenta sistemas de IA que amplificam a qualidade — e os defeitos — do conteúdo subjacente.

Em processos de contratação, este anexo deve compor o referencial técnico do Estudo Técnico Preliminar (ETP) e do Termo de Referência (TR) sempre que o objeto envolver a implantação, migração, expansão ou sustentação de bases de conhecimento. A seção 0 consolida os requisitos mínimos funcionais, não funcionais, de segurança, de documentação e de aceite que devem constar nesses instrumentos.

Em revisões e avaliações de soluções existentes, este anexo fornece os critérios de diagnóstico de maturidade, os indicadores de saúde da base e os limites recomendados por criticidade do sistema.

1.5 O que este anexo complementa e o que não substitui

Este anexo complementa:

→ o documento principal do **GuIA**, quanto aos princípios, requisitos transversais e critérios de ciclo de vida;

→ os Anexos de LLM e RAG, fornecendo os fundamentos da camada de conteúdo que ambos pressupõem;

→ guias operacionais de gestão do conhecimento publicados pela ENAP e pelo TCU, adicionando a perspectiva técnica de arquitetura, busca e conformidade normativa.

Este anexo não substitui:

→ o *template* de caso de uso do GuIA, que deve ser preenchido para cada projeto ou iniciativa específica;

→ os artefatos formais de contratação pública (ETP, TR, contratos), dos quais este documento é apenas referência técnica;

→ o Relatório de Impacto à Proteção de Dados Pessoais (RIPD), quando recomendado ou exigido nos termos da LGPD e das orientações da ANPD, especialmente em tratamentos que possam gerar alto risco aos titulares;

→ normas e legislação setoriais que possam impor requisitos adicionais em domínios como saúde, justiça, educação ou defesa.

2. Escopo e aplicabilidade

2.1 Contexto institucional, regulatório e normativo

Bases de conhecimento operam em um ambiente normativo que, no Brasil, articula pelo menos cinco camadas regulatórias com implicações diretas sobre arquitetura, governança e operação.

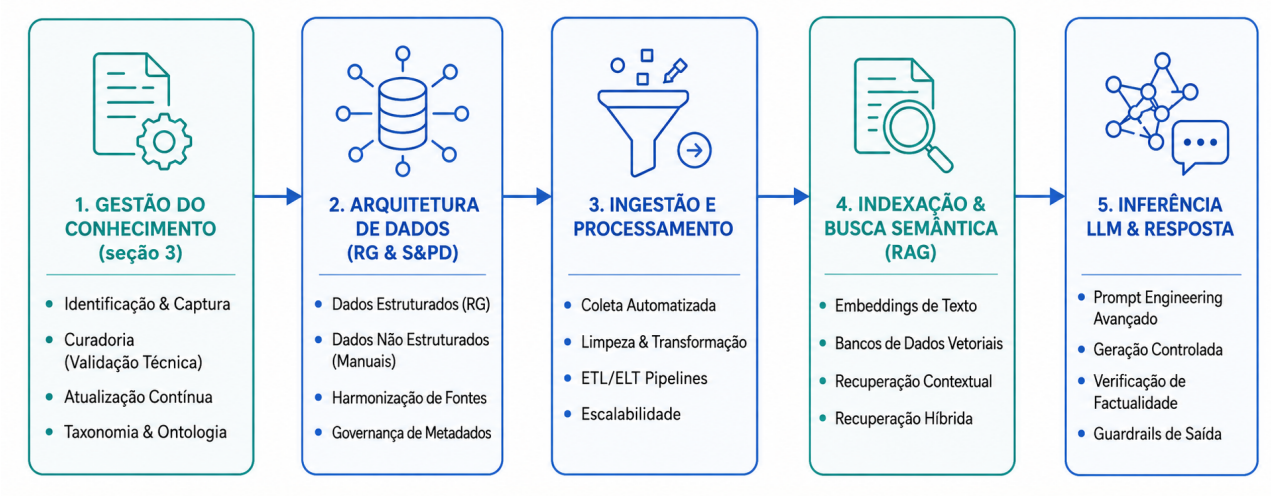

Figura 1 – Camadas regulatórias

A primeira camada é a gestão do conhecimento como sistema de gestão. A ISO 30401 estabelece que a gestão do conhecimento requer requisitos de estabelecimento, implementação, manutenção, revisão e melhoria contínua. Ela desloca a base de conhecimento de artefato editorial para capacidade organizacional sujeita a governança formal, papéis definidos, ciclos de revisão e evidências auditáveis. A ENAP e o TCU reforçam essa perspectiva no contexto do setor público brasileiro, exigindo diagnóstico prévio, identificação de conhecimentos críticos e estruturas de incentivo e institucionalização.

A segunda camada é a proteção de dados pessoais. A Lei Geral de Proteção de Dados Pessoais (LGPD, Lei 13.709/2018) incide sobre qualquer base de conhecimento que armazene ou processe dados pessoais, incluindo registros de atendimento, históricos de uso e conteúdo que identifique ou torne identificável uma pessoa natural. Os princípios de finalidade, necessidade, transparência e responsabilização exigem que o controlador mantenha registro das operações de tratamento, classifique dados sensíveis com controles reforçados e implemente critérios claros de publicação e acesso. A Autoridade Nacional de Proteção de Dados (ANPD) orienta, adicionalmente, que tratamentos baseados em legítimo interesse sejam documentados em suas três fases: finalidade, necessidade e balanceamento com salvaguardas.

Alerta operacional sobre legítimo interesse: No setor público, o legítimo interesse não é a base legal adequada quando o tratamento decorre de prerrogativa estatal típica. Para órgãos públicos, priorizar as bases dos arts. 7º, 11 e 23 da LGPD. Consentimento tende a ser inadequado quando há assimetria de poder ou prestação obrigatória de serviço público. Legítimo interesse não se aplica a dados pessoais sensíveis (art. 11 da LGPD). Quando excepcionalmente aplicável, o legítimo interesse deve passar por teste de: (1) finalidade legítima, (2) necessidade, (3) balanceamento com os direitos dos titulares e (4) salvaguardas implementadas. Verificar primeiro as bases legais específicas antes de recorrer a legítimo interesse.

A terceira camada é a acessibilidade digital. O Modelo de Acessibilidade em Governo Eletrônico (eMAG) adapta ao contexto brasileiro as diretrizes do *Web Content Accessibility Guidelines* (WCAG, W3C), estabelecendo padrões que qualquer portal ou sistema público deve atender. Bases de conhecimento de acesso público ou de uso por servidores com deficiência estão sujeitas a esses requisitos desde a concepção, não como adaptação posterior.

A quarta camada é a linguagem simples. A Lei 15.263/2025 instituiu a Política Nacional de Linguagem Simples, tornando obrigatório para a comunicação pública o uso de técnicas que permitam ao cidadão encontrar, compreender e usar a informação. Para bases de conhecimento públicas, isso implica revisão de vocabulário, estrutura de frases, organização por necessidade do leitor e eliminação de jargão técnico desnecessário como requisito de qualidade editorial, não como preferência estilística. Para bases de conhecimento de acesso público, exige-se evidência de revisão de linguagem simples como condição de aceite, não apenas conformidade declarada.

A quinta camada é a prontidão para inteligência artificial. O *NIST AI Risk Management Framework (AI RMF)* e o *Generative AI Profile* recomendam, como boa prática de gestão de risco aplicável a sistemas de IA, que aqueles que consomem bases de conhecimento tenham documentação de proveniência do conteúdo, rastreabilidade de linhagem, mecanismos de feedback humano e controles de exposição de dados pessoais. Quando a base de conhecimento alimenta sistemas de IA, os requisitos de governança deixam de ser somente editoriais e passam a integrar a cadeia de responsabilização do sistema de IA como um todo.

2.2 Definição operacional

Para os fins deste anexo, uma base de conhecimento é um sistema sociotécnico governado composto de conteúdo, metadados, vocabulários controlados, regras editoriais e fluxos operacionais, cujo propósito é converter conhecimento organizacional explícito em ativo encontrável, reutilizável, validado e auditável por pessoas e sistemas.

Essa definição distingue a base de conhecimento de infraestruturas adjacentes que frequentemente são confundidas com ela:

→ **Banco de dados relacional ou NoSQL**: projetado para armazenar dados transacionais estruturados com esquemas predefinidos, otimizado para leitura e escrita de alta frequência. Não é projetado para compreensibilidade humana nem para reutilização contextual de conteúdo.

→ **Data lake**: repositório de dados brutos e não estruturados armazenados a baixo custo, sem esquema prévio. Serve análise de dados, não gestão de conhecimento.

→ **Sistema de gerenciamento de conteúdo (CMS)**: focado em publicação, roteamento de URLs e apresentação visual para a web. A base de conhecimento pode ser implementada sobre um CMS, mas o CMS por si só não garante estruturação, governança nem encontrabilidade do conhecimento.

→ **Wiki**: ambiente colaborativo de edição distribuída. Viabiliza captura rápida, mas sem curadoria e taxonomia tende a produzir deriva terminológica, duplicação e degradação de consistência.

→ **FAQ**: subconjunto reativo de uma base de conhecimento, limitado a pares de perguntas e respostas. Não abrange o domínio cognitivo completo de uma organização.

→ **Sistema de gestão do conhecimento (KMS)**: ecossistema mais amplo que engloba a base de conhecimento, mas vai além, incorporando processos de criação, colaboração, análise de métricas e cultura de compartilhamento. A base de conhecimento é a camada de dados e conteúdo do KMS, não o KMS inteiro.

O conhecimento armazenado em uma base de conhecimento classifica-se em quatro tipos com implicações distintas para arquitetura e governança:

- **Declarativo**: fatos, conceitos, definições e políticas que formam a fundação teórica e normativa da organização.
- **Procedimental**: instruções, roteiros e processos que orientam a execução de tarefas, com estrutura sequencial e imperativa.
- **Episódico**: histórico de casos, lições aprendidas e registros de decisões anteriores, essencial para aprendizado organizacional.
- **Estratégico**: entendimento de contexto, valores e tomada de decisão sob incerteza, frequentemente tácito e de difícil externalização.

2.3 Tipos de sistema cobertos por este anexo

Este anexo cobre bases de conhecimento nos seguintes perfis de implantação:

- **Bases estruturadas**: conteúdo com esquemas fortemente tipados, metadados obrigatórios e recuperação determinística. Permitem recuperação rigorosa por máquinas, mas exigem esforço editorial maior.

- **Bases semiestruturadas**: texto livre com metadados ricos, marcação XML/JSON e categorização controlada. Equilíbrio entre flexibilidade editorial e precisão de recuperação; perfil dominante em organizações públicas de médio e grande porte.
- **Bases não estruturadas com camada de governança**: documentos brutos com metadados mínimos e vocabulário de classificação sobreposto. Exigem motores de busca semântica e controles compensatórios de qualidade.
- **Bases federadas**: conhecimento distribuído por domínios de negócio, unido por interface, taxonomia ou catálogo de busca central. Modelo recomendado para organizações com múltiplas áreas autônomas.
- **Bases que alimentam sistemas de IA**: qualquer configuração acima quando o conteúdo é consumido por pipelines de Geração Aumentada por Recuperação (RAG), assistentes conversacionais ou outros mecanismos de IA. Nesses casos, os requisitos de metadados, proveniência e permissão são mais estritos.

2.4 Tipos de sistema explicitamente fora do escopo

Este anexo não cobre:

- Bases de dados puramente transacionais sem componente de conhecimento organizacional (sistemas ERP, CRM, SGBD operacionais).
- Repositórios documentais com finalidade exclusiva de guarda e evidência, sem requisito de encontrabilidade ou reutilização contextual.
- Sistemas de informação geográfica, bancos de imagens ou acervos multimídia nos quais a unidade de conteúdo não é texto estruturado.
- Sistemas de *business intelligence* e *data warehouses*, cujo propósito é análise de dados, não recuperação e reutilização de conhecimento.
- Portais de transparência com publicação de dados abertos, regidos por normativa específica de dados governamentais.

Quando um projeto combinar uma base de conhecimento com qualquer infraestrutura acima, este anexo se aplica à camada de base de conhecimento e os requisitos da camada adjacente devem ser tratados nos documentos normativos pertinentes.

2.5 Casos de uso típicos

Os casos de uso a seguir representam os perfis de aplicação mais frequentes no setor público brasileiro e servem como referência para estruturação de projetos.

Contexto	Propósito	KPI prioritário	Risco típico
Atendimento e ITSM	Capturar soluções no fluxo de resolução de chamados e reduzir retrabalho	Tempo de resolução, reuso de artigos	Correções desatualizadas e duplicação
Políticas e processos internos	Centralizar normativos, manuais e procedimentos com controle de versão	Sucesso na busca, tempo para executar tarefa	Versões conflitantes, ambiguidade
Portal público de serviços	Orientar o cidadão com linguagem simples e navegação por tarefa	Conclusão de tarefa, redução de atendimento	Linguagem burocrática, inacessibilidade
Ambiente regulado (saúde, justiça)	Manter trilha de auditoria, versões imutáveis e permissões finas	Não conformidades, tempo de revisão	Vazamento, publicação indevida
Transferência de conhecimento crítico	Externalizar conhecimento tácito de especialistas antes de desligamento	Tempo para competência	Dependência de pessoas-chave
Assistentes e automação com IA	Fornecer camada autoritativa com proveniência para consumo por IA	Taxa de fundamentação, incidentes	Respostas incorretas ou sem permissão

2.6 Critérios de adoção

A adoção de uma base de conhecimento estruturada é recomendada quando o projeto ou a organização apresentar pelo menos uma das condições a seguir:

- Demanda recorrente pelo mesmo conhecimento por múltiplos usuários ou canais, indicando que a externalização e a reutilização geram ganho operacional mensurável.
- Necessidade de rastreabilidade e auditoria do conteúdo por exigência legal, contratual ou regulatória.
- Integração prevista com sistemas de IA, em especial pipelines RAG ou assistentes conversacionais, que exigem conteúdo com metadados e proveniência.
- Risco de perda de conhecimento crítico por rotatividade, aposentadoria ou desligamento de servidores especialistas.
- Inconsistência de versões e informações conflitantes entre áreas ou canais de atendimento, causando erros operacionais ou prejuízo ao cidadão.
- Exigência de conformidade com LGPD, WCAG/eMAG ou Lei 15.263/2025 que demande revisão estruturada de conteúdo, metadados e permissões.

2.7 Critérios de não adoção

A implantação de uma base de conhecimento estruturada não é recomendada quando:

- O conhecimento a ser registrado é altamente tácito, contextual e dependente de julgamento situacional, não sendo redutível a conteúdo explícito reutilizável sem perda significativa de precisão.
- O volume de conteúdo é pequeno, estável e de uso restrito, tornando o overhead de governança desproporcional ao benefício.
- A organização não dispõe de papéis mínimos de curadoria, governança e revisão; nesse caso, uma base de conhecimento sem curadoria degradará mais rapidamente do que um repositório simples gerenciado manualmente.
- O objetivo é exclusivamente preservação documental com finalidade de evidência, sem requisito de encontrabilidade ou reutilização.

2.8 Comparativo com alternativas tecnológicas

A tabela a seguir posiciona a base de conhecimento estruturada em relação às principais alternativas tecnológicas segundo critérios relevantes para o setor público. O objetivo não é indicar uma solução superior em termos absolutos, mas orientar a decisão arquitetural com base nas necessidades concretas do projeto.

Critério	Base de conhecimento estruturada	Repositório documental	Wiki colaborativa	Data lake	Knowledge graph
Rastreabilidade de conteúdo	Alta	Alta	Média	Baixa	Alta
Encontrabilidade pelo usuário	Alta	Baixa	Média	Baixa	Alta
Atualização contínua	Alta	Média	Alta	Baixa	Média
Consistência terminológica	Alta	Baixa	Baixa	Baixa	Alta
Custo de implantação	Médio	Baixo	Baixo	Baixo	Alto
Custo de manutenção editorial	Médio	Baixo	Médio	Baixo	Alto
Governança de dados (LGPD)	Alta	Média	Baixa	Média	Alta
Acessibilidade (WCAG/eMAG)	Alta	Baixa	Média	N/A	Média
Prontidão para IA (RAG/LLM)	Alta	Média	Baixa	Média	Alta

Critério	Base de conhecimento estruturada	Repositório documental	Wiki colaborativa	Data lake	Knowledge graph
Adequação ao setor público	Alta	Alta	Média	Baixa	Média
Complexidade de implantação	Média	Baixa	Baixa	Média	Alta
Inferência e relações semânticas	Baixa	Baixa	Baixa	Baixa	Alta

Legenda: Alta / Média / Baixa / N/A referem-se à adequação do critério para o perfil típico de cada tecnologia, não ao desempenho absoluto de produtos específicos. Configurações híbridas podem combinar características de mais de uma coluna.

A análise da tabela sugere três decisões arquiteturais recorrentes. Primeiro: quando o requisito dominante é rastreabilidade e evidência documental sem necessidade de encontrabilidade semântica, o repositório documental é suficiente e mais econômico. Segundo: quando o requisito dominante é captura rápida e colaboração distribuída sem exigência de consistência terminológica, a wiki é mais ágil, desde que acompanhada de curadoria mínima. Terceiro: quando o projeto prevê integração com IA generativa ou inferência sobre relações entre entidades, a base de conhecimento estruturada com camada semântica ou *knowledge graph* é a arquitetura adequada, a despeito do custo de modelagem maior.

A maioria das organizações públicas de médio e grande porte beneficia-se de uma arquitetura híbrida: base semiestruturada com taxonomia controlada, metadados mínimos obrigatórios, proveniência registrada e camada de busca semântica. Essa configuração equilibra custo de manutenção, governança e prontidão para IA sem exigir a complexidade de um *knowledge graph* completo.

3. Ciclo de vida sugerido para a solução

3.1 Impactos da tecnologia no ciclo de vida

A adoção de uma base de conhecimento afeta todas as etapas do ciclo de vida de soluções de IA. As decisões tomadas nas etapas iniciais — em especial sobre arquitetura de informação, vocabulário controlado e metadados — têm impacto exponencialmente maior do que as revisões feitas em produção. Uma base estruturada inadequadamente que alimenta um sistema de IA distribui erros em escala.

Os impactos mais significativos ocorrem em três pontos: na experimentação, onde a qualidade da base determina a qualidade dos protótipos de IA e condiciona decisões de arquitetura que serão caras de reverter; na implantação, onde a ausência de governança produz degradação imediata de qualidade sem mecanismos de detecção; e na sustentação, onde a ausência de revisão cadenciada acumula conteúdo obsoleto que corrói a confiança do usuário e a precisão dos sistemas de IA que dependem da base.

Figura 2 - Ciclo de vida sugerido da solução

A imagem traduz a lógica do ciclo de vida da solução em sete etapas, da prospecção à desativação segura. Ela evidencia as saídas principais de cada fase, como inventário de lacunas, escopo validado, base estruturada, produção monitorada e acessos revogados. O diagrama reforça que decisões iniciais sobre arquitetura, vocabulário e metadados afetam todo o ciclo.

3.2 Atividades típicas por etapa

Compatibilidade com o GuIA: para cada etapa do ciclo de vida, os artefatos do projeto devem manter a mesma lógica de organização do GuIA, explicitando entregáveis, ponto de decisão, matriz RACI e recomendações, políticas e boas práticas. As subseções seguintes detalham as atividades específicas da tecnologia, mas devem ser usadas em conjunto com essa estrutura comum.

3.2.1 Prospecção de desafios

Entradas: catálogo de serviços, mapeamento de demandas recorrentes, entrevistas com equipes operacionais.

Atividades: identificar domínios de conhecimento crítico, mapear demandas recorrentes não documentadas, avaliar maturidade de gestão do conhecimento existente, identificar riscos de

perda de conhecimento tácito.

Saídas: inventário de lacunas de conhecimento, mapa de criticidade por domínio, avaliação de maturidade (nível 1–4 APQC).

Evidências: relatório de diagnóstico, atas de entrevista, inventário de artefatos existentes.

3.2.2 Estruturação do desafio de IA no serviço público

Entradas: inventário de lacunas, caso de uso de IA proposto.

Atividades: definir se a base de conhecimento é o componente adequado para o caso de uso (cf. critérios de adoção na Seção 2), delimitar escopo temático, identificar regulação aplicável (LGPD, setorial, WCAG, linguagem simples), definir nível de acesso do conteúdo.

Saídas: escopo validado, mapeamento normativo, decisão de arquitetura.

Evidências: documento de escopo, parecer de conformidade LGPD, decisão de arquitetura documentada.

3.2.3 Experimentação

Entradas: escopo validado, amostra de conteúdo para prova de conceito.

Atividades: definir taxonomia inicial com *card sorting e tree testing*, estruturar amostra de conteúdo nos tipos de tópico adequados, configurar motor de busca com parâmetros iniciais, testar recuperação com usuários reais, avaliar qualidade de consumo por IA se aplicável.

Saídas: taxonomia v0, *templates* validados, configuração inicial do motor de busca, relatório de testes de recuperação.

Critério de passagem: taxa de sucesso de busca 70% na amostra de conteúdo, sem artigos com metadados obrigatórios ausentes.

Evidências: resultados de *card sorting e tree testing*, logs de busca da PoC, avaliações de usuário.

3.2.4 Implementação

Entradas: taxonomia e *templates* validados, plataforma selecionada.

Atividades: estruturar o acervo completo, aplicar taxonomia e metadados a todos os artigos, configurar fluxos editoriais e papéis, implementar controles de acesso, integrar com sistemas externos, configurar *analytics*, treinar equipe editorial, realizar auditoria inicial de qualidade.

Saídas: base estruturada com cobertura mínima do escopo definido, fluxos operacionais ativos, integrações testadas, *dashboard* de métricas ativo.

Critério de passagem: cobertura de pelo menos 80% das demandas recorrentes identificadas no diagnóstico, todos os artigos com metadados obrigatórios preenchidos, controles de acesso validados.

Evidências: auditoria de metadados, relatório de cobertura, testes de integração, registro de treinamento da equipe.

3.2.5 Implantação e *rollout*

Entradas: base validada, plano de comunicação.

Atividades: implantação gradual por domínio ou canal, monitoramento de métricas de uso e qualidade, ajuste de ranqueamento com base em *analytics* reais, coleta de feedback dos primeiros usuários, correção de lacunas emergentes.

Saídas: base em produção com métricas dentro dos limiares aceitáveis.

Evidências: relatório de métricas das primeiras semanas, registro de feedbacks atendidos.

3.2.6 Sustentação

Entradas: base em produção, *dashboard* de métricas, *backlog* editorial.

Atividades: revisão cadenciada de artigos por criticidade, análise mensal de logs de busca e zero resultados, auditoria semestral de qualidade, manutenção do vocabulário controlado, gestão de domínios e proprietários de conteúdo, atualização de integrações, relatórios periódicos de impacto.

Saídas: base com qualidade sustentada, relatórios de impacto, *backlog* priorizado.

Evidências: relatórios mensais de métricas, atas de revisão de artigos, relatório semestral de auditoria.

3.2.7 Desativação segura

Entradas: decisão de descontinuação do caso de uso ou da plataforma.

Atividades: exportar todo o conteúdo em formato aberto, preservar registros de proveniência e trilha de auditoria, redirecionar URLs históricas, documentar o inventário final, desativar integrações e acessos, comunicar usuários e sistemas dependentes. A desativação segura deve incluir: revogação de chaves de API e *tokens* de acesso; exclusão de *embeddings* e índices vetoriais derivados; limpeza de caches; invalidação de snapshots; exportação auditável do conteúdo; descarte conforme política de retenção; e comprovação formal de eliminação por fornecedores terceiros, quando aplicável.

Saídas: acervo exportado, registros preservados, integrações desativadas.

Evidências: arquivo de exportação, lista de redirecionamentos, registro de desativação de acessos.

4. Governança de dados, artefatos e conhecimento

4.1 Tipos de dados e artefatos tratados

Uma base de conhecimento trata três categorias de dados com requisitos de governança distintos:

- **Conteúdo de conhecimento**: artigos, tópicos, políticas, procedimentos, referências. Não contém dados pessoais de forma estruturada, mas pode mencioná-los incidentalmente em exemplos ou casos. Requisito: revisão editorial periódica, versionamento, proveniência.
- **Dados de autoria e gestão**: nome do autor, aprovador, dono, histórico de edições com identificação do usuário. Dados pessoais de servidores sujeitos à LGPD. Requisito: minimização, acesso restrito a gestores, retenção limitada ao período de responsabilização.
- **Dados de uso e analytics**: logs de busca, avaliações, cliques, sessões. Podem conter dados pessoais se associados a usuários identificáveis. Requisito: anonimização antes de armazenamento de longo prazo, base legal documentada, finalidade restrita à melhoria da base.

4.2 Fontes de dados e critérios de admissibilidade

Nem todo conhecimento disponível deve ser incorporado à base. Os critérios de admissibilidade abaixo devem ser aplicados antes da ingestão:

- **Autoridade**: o conteúdo é produzido ou validado por área com competência técnica ou legal para o tema?
- **Atualidade**: o conteúdo reflete o estado vigente da política, processo ou tecnologia?
- **Relevância**: o conteúdo atende a demanda real documentada (log de busca, chamado recorrente, lacuna editorial)?
- **Licenciamento e sigilo**: o conteúdo pode ser publicado no nível de acesso pretendido sem violar sigilo, propriedade intelectual ou LGPD?
- **Completude**: o conteúdo resolve o problema do usuário de forma autossuficiente ou requer contexto que não estará disponível no ponto de uso?

4.3 Curadoria, qualidade e atualização

A qualidade do conteúdo de uma base de conhecimento organiza-se em seis dimensões que devem ser avaliadas em conjunto:

- **Exatidão**: as afirmações, procedimentos e dados são verificáveis e aderentes à realidade operacional atual.
- **Completude**: o tópico abrange os casos de uso pretendidos, incluindo casos de exceção relevantes.
- **Consistência**: terminologia, estrutura e estilo seguem os padrões do guia editorial e do vocabulário controlado.
- **Pontualidade**: o conteúdo reflete a versão atual das políticas, sistemas e processos que descreve.
- **Clareza**: o texto é compreensível pelo público-alvo declarado, em conformidade com a Lei 15.263/2025.

- **Encontrabilidade**: o conteúdo é recuperável pelas consultas típicas do público-alvo, medida por taxa de sucesso de busca.

Figura 3 - Ciclo editorial de estruturação e publicação da base

Auditorias de conteúdo devem ser conduzidas periodicamente com quatro etapas: inventário quantitativo (todos os artigos com metadados), inspeção de desempenho por *analytics* (artigos de alta tração vs. artigos nunca acessados), identificação de redundâncias e conflitos semânticos, e manutenção operacional (links quebrados, imagens ausentes, *alt-text* faltante).

A imagem complementa a discussão sobre curadoria, qualidade e atualização ao mostrar o percurso editorial do conteúdo até a publicação. Ela destaca identificação da demanda, elicitação, estruturação, validação e publicação como controles antes da exposição ao usuário. O foco nas saídas evidencia que artigo, metadados e proveniência precisam estar completos antes da entrada em operação.

4.4 Versionamento de bases, modelos, *prompts* e configurações

Toda alteração significativa em qualquer artefato da base deve gerar uma nova versão com identificação cronológica (*snapshot versioning*). A versão anterior não é deletada; é arquivada com preservação da URL histórica e do registro de proveniência. Isso garante a trilha de auditoria e a retroatividade forense para conformidade legislativa.

O versionamento se aplica a: artigos e tópicos de conteúdo, esquemas de metadados, taxonomias e tesauros, configurações do motor de busca (pesos e políticas de ranqueamento), e — quando a base alimenta IA — aos *prompts* de sistema e aos parâmetros de recuperação.

4.5 Classificação da informação e controle de acesso

Todo conteúdo deve receber uma classificação de acesso no momento da criação:

- **Público**: acessível a qualquer usuário sem autenticação.
- **Interno**: acessível a servidores autenticados.
- **Restrito**: acessível a papel específico (área, cargo, projeto).
- **Confidencial**: acessível apenas a lista nominalmente definida, com registro de cada acesso.

A classificação deve ser revisada a cada ciclo de revisão do artigo. Conteúdo que contém dados pessoais, informações sigilosas ou dados que possam prejudicar terceiros se publicados deve ser classificado como restrito ou confidencial independentemente de sua natureza editorial.

Atenção — duas camadas distintas de classificação: A classificação operacional de acesso descrita acima (Público, Interno, Restrito, Confidencial) organiza o controle de acesso na plataforma, mas não se confunde com a classificação legal de sigilo da Lei de Acesso à Informação (LAI, Lei 12.527/2011). Para informações com sigilo formal, aplicam-se os graus legais da LAI: *reservado* (prazo máximo de 5 anos), *secreto* (até 15 anos) e *ultrassecreto* (até 25 anos), com autoridade competente e ato formal de classificação. Conteúdo operacionalmente "confidencial" na base de conhecimento não é equivalente a informação sigilosa nos termos da LAI; a confusão entre esses regimes pode gerar obrigações legais indevidas ou omissões de divulgação.

4.6 Retenção, descarte e minimização

Artigos arquivados devem ser retidos pelo prazo mínimo estabelecido pela política de gestão documental da organização, considerando exigências legais setoriais. Após o prazo, a decisão de descarte definitivo deve ser documentada e aprovada pelo responsável pela governança.

Dados de uso anonimizados podem ser retidos por período maior para fins de análise de tendência. Dados de uso identificáveis (logs com ID de usuário) devem ser anonimizados em prazo não superior a 90 dias.

4.7 Rastreabilidade entre entrada, processamento e saída

Cada artigo publicado deve ter um registro de proveniência que responda: *quem* produziu o conteúdo original, *a partir de qual fonte*, *quando*, *quem validou*, *quando foi publicado* e *quais alterações foram feitas em cada versão*. O padrão PROV-DM do W3C fornece o modelo de dados para esse registro de forma interoperável.

Quando o conteúdo é consumido por sistemas de IA, o registro de proveniência deve ser transmitido junto com o conteúdo via metadados da resposta da API, para que o sistema de IA possa citar a fonte e o gestor possa auditar qual versão do artigo fundamentou qual resposta gerada.

Sistemas generativos que consomem a base devem documentar como os dados de proveniência são rastreados ao longo do pipeline de recuperação e geração. Quando o conteúdo da base contiver dados pessoais identificáveis (PII) em exemplos, históricos de casos ou registros de atendimento, esses dados devem ser removidos ou anonimizados antes da ingestão no pipeline de IA, salvo quando a base legal e a finalidade do sistema justificarem o tratamento e houver controles de acesso correspondentes documentados.

4.8 Responsáveis técnicos e *data stewards*

Papel	Responsabilidades	Requisitos mínimos
Knowledge Manager (KM)	Arquitetura da base, taxonomia, padrões editoriais, métricas, relatórios de qualidade	Conhecimento de gestão do conhecimento e arquitetura de informação
Domain Owner	Propriedade do conteúdo do domínio, aprovação de publicações, revisão periódica	Competência técnica ou legal no domínio
Data Steward	Classificação de informação, conformidade LGPD, tratamento de dados pessoais na base	Conhecimento de LGPD e políticas de privacidade da organização
Editor / Redator	Captura, estruturação e manutenção de artigos	Domínio do guia editorial, vocabulário controlado e templates
Revisor Técnico	Validação de exatidão e completude	Competência técnica no domínio do artigo

5. Caracterização do sistema de base de conhecimento

5.1 Fundamentos conceituais

Uma base de conhecimento repousa sobre três tradições intelectuais que, na prática contemporânea, operam de forma integrada.

A primeira é a gestão do conhecimento organizacional. Alavi e Leidner estabeleceram que sistemas de gestão do conhecimento devem apoiar criação, transferência e aplicação do conhecimento, não apenas seu armazenamento. Nonaka formalizou o modelo SECI, que descreve a espiral de conversão entre conhecimento tácito e explícito em quatro dinâmicas: socialização (tácito para tácito, absorção por observação e prática compartilhada), externalização (tácito para explícito, a fase crítica de KM em que intuições e macetes se tornam artigos, políticas e procedimentos), combinação (explícito para explícito, síntese e estruturação de fragmentos documentados em bases maduras) e internalização (explícito para tácito, assimilação operacional por profissionais que executam tarefas a partir da documentação). Para uma base de conhecimento, o desafio central está na fase de externalização: converter o que os especialistas

sabem mas não articulam em conteúdo estruturado, encontrável e reutilizável.

A segunda tradição é a organização do conhecimento e a ciência da informação. Tesauros, taxonomias, vocabulários controlados e padrões de metadados como SKOS, ISO 25964 e Dublin Core provêm os mecanismos de indexação, navegação e recuperação sem os quais o conteúdo excelente permanece inacessível. A distinção entre conceito e termo — central no SKOS — resolve um problema recorrente: o mesmo conceito recebe nomes diferentes em áreas distintas da organização, fragmentando a busca e multiplicando duplicatas.

A terceira é a engenharia de conteúdo e a Web Semântica. DITA formaliza a unidade de conteúdo como tópico reutilizável e tipado. PROV-DM registra proveniência de forma interoperável. SHACL valida a integridade de grafos RDF. OWL e RDF viabilizam ontologias formais para domínios que exigem inferência. O resultado contemporâneo é um campo híbrido: editorial, semântico, organizacional e computacional simultaneamente.

5.2 Componentes arquiteturais fundamentais

Uma base de conhecimento madura é composta pelos seguintes componentes:

- **Camada de conteúdo**: o conjunto de artigos, tópicos, políticas, procedimentos e referências que constituem o acervo. A unidade mínima é o tópico atômico: autocontido, focado em um único propósito, reutilizável em múltiplos contextos. Este anexo adota quatro tipos DITA como perfil mínimo operacional, com implicações estruturais distintas:
    - *Concept*: arcabouço teórico, definições e contexto histórico (responde a "o que é?");
    - *Task*: instrução procedimental sequencial com pré-requisitos, passos e resultado esperado (responde a "como fazer? ");
    - *Reference*: dados factuais de consulta imediata, como tabelas de parâmetros e especificações técnicas (responde a "qual é o valor? ");
    - *Troubleshooting*: diagnóstico de falha com sintoma, causa e remédio (responde a "por que falhou e o que fazer? ").
- **Camada de metadados**: conjunto de atributos que descrevem, classificam e controlam o conteúdo. Os metadados obrigatórios mínimos para qualquer base de conhecimento no âmbito deste anexo são: título funcional, resumo, dono (pessoa ou área responsável), domínio temático, estado editorial (rascunho, em revisão, publicado, arquivado), data da última revisão, versão, público-alvo, permissão de acesso, produto ou processo associado e classificação taxonômica. Para conteúdo que alimenta sistemas de IA, acrescentam-se: origem do conteúdo, contexto de validade e registro de proveniência.
- **Vocabulário controlado**: conjunto de termos preferidos, sinônimos, hierarquias e mapeamentos que padronizam a linguagem de indexação e recuperação. Pode ser implementado como lista simples, taxonomia hierárquica, tesauro (com relações de equivalência, hierarquia e associação) ou ontologia formal. A escolha depende da complexidade do domínio e do custo de manutenção aceitável.
- **Motor de busca e recuperação**: componente que indexa o conteúdo e responde a con-

sultas. Motores como *Elasticsearch* e *Apache Solr* oferecem escalabilidade para grandes acervos e expressividade analítica; *Typesense e Meilisearch* priorizam latência baixa e usabilidade. O algoritmo BM25 (variante de TF-IDF) é o mecanismo padrão de relevância em texto completo; *field boosting* e *freshness boost* ajustam a relevância com base no campo pesquisado e na atualidade do conteúdo, respectivamente. Busca facetada e *autocomplete* completam a experiência de recuperação.

- **Camada de proveniência e auditoria**: registros que documentam quem produziu cada artefato, quando, a partir de qual fonte, com qual aprovação e com qual histórico de alterações. PROV-DM fornece o modelo de dados interoperável. Essa camada é condição necessária para uso seguro da base em sistemas de IA e para conformidade com exigências de auditoria do setor público.
- **Camada de governança**: políticas, fluxos de aprovação, papéis, cadências de revisão e métricas operacionais que sustentam a qualidade e a atualidade do acervo ao longo do tempo. Sem governança, os demais componentes degradam-se progressivamente independentemente da qualidade da plataforma.
- **Interface de acesso**: portal, API, *widget* de busca embutida ou conector com sistemas de atendimento, que expõe o conteúdo aos usuários e aos sistemas que o consomem.

5.3 Entradas, saídas e artefatos gerados

Categoria	Exemplos	Observações
Entradas primárias	Demandas de atendimento, incidentes resolvidos, normativos, políticas, manuais legados, lições aprendidas, entrevistas com especialistas	Devem ser triadas antes de incorporação
Entradas de retroalimentação	Logs de busca, avaliações de artigos, relatórios de falha de recuperação, feedbacks de usuários e sistemas de IA	Alimentam melhoria contínua e priorização editorial
Artefatos de conteúdo	Tópicos, artigos, políticas, procedimentos, FAQs, referências, guias de troubleshooting	Classificados por tipo DITA quando aplicável
Artefatos de estrutura	Taxonomia, tesauro, ontologia, mapa de tópicos, esquema de metadados	Versionados separadamente do conteúdo
Artefatos de governança	Registros de revisão, trilhas de auditoria, registros de proveniência, histórico de versões, relatórios de qualidade	Obrigatórios para ambientes regulados e para uso com IA
Saídas para usuários	Páginas de ajuda, respostas de busca, conteúdo embutido em portais, exportações em PDF ou HTML	Devem atender WCAG e Lei 15.263/2025
Saídas para sistemas	Respostas de API, contexto para pipelines RAG, *snippets* para assistentes conversacionais, *webhooks* de notificação	Devem incluir metadados de proveniência e permissão

5.4 Fluxo fundamental de funcionamento

O ciclo operacional de uma base de conhecimento organiza-se em fluxos distintos conforme o momento de execução e o tipo de ator envolvido.

Esquema operacional de referência: fontes autorizadas → curadoria e metadados → indexação e controle de acesso → recuperação por pessoas, APIs ou sistemas de IA → feedback, auditoria e revisão contínua.

5.4.1 Fluxo *offline*: ingestão e estruturação

O fluxo offline compreende todas as operações que ocorrem antes da publicação e da recuperação pelo usuário final. É nesse fluxo que a qualidade da base é determinada:

- **Identificação da demanda**: a necessidade de novo conteúdo é detectada por análise de logs de busca (termos sem resultado), volume de chamados recorrentes, lacunas identificadas em auditorias ou solicitações de áreas especializadas.
- **Elicitação e captura**: o conhecimento é externalizado por meio de entrevistas estruturadas, análise de protocolo (verbalização concorrente), método do incidente crítico ou captura direta no fluxo de resolução de demandas (modelo KCS). Documentos preexistentes são triados e reformatados conforme o tipo de tópico adequado.
- **Estruturação**: o conteúdo bruto é organizado segundo o modelo de conteúdo da base (tipos DITA, templates, esquema de metadados), com aplicação do vocabulário controlado e preenchimento dos metadados obrigatórios.
- **Validação**: revisão técnica (exatidão do conteúdo), revisão editorial (clareza, conformidade com linguagem simples e acessibilidade), revisão jurídica ou de compliance quando aplicável, e validação de metadados por checklist automatizado ou manual.
- **Publicação**: o conteúdo aprovado é publicado no canal adequado (portal interno, portal público, API), com indexação automática pelo motor de busca e registro de proveniência.

Figura 4 - Fluxo ***offline*** de ingestão e estruturação

A imagem sintetiza o fluxo *offline* descrito nesta subseção, desde a identificação da demanda até a publicação. Ela evidencia que a qualidade da base é definida antes do uso pelo usuário final, com validação técnica, editorial, de *compliance* e de metadados. Também explicita as saídas esperadas: artigo publicado, metadados completos e proveniência registrada.

5.4.2 Fluxo online: recuperação e uso

O fluxo online compreende as operações em tempo real disparadas pelo usuário ou por sistemas que consomem a base:

- **Formulação da consulta**: o usuário digita termos, navega pela taxonomia ou o sistema envia uma consulta via API. O vocabulário controlado expande sinônimos e corrige variações ortográficas.
- **Recuperação**: o motor de busca aplica algoritmo de relevância (BM25 com ajustes de campo e atualidade), retornando resultados ranqueados. Em sistemas de IA, o pipeline de recuperação pode incluir *embedding search* ou recuperação híbrida.
- **Apresentação**: os resultados são apresentados com título, resumo, metadados de contexto e, quando disponível, trecho relevante destacado. Sistemas de IA recebem o conteúdo como contexto para geração de resposta.
- **Retroalimentação**: cliques, avaliações, ausência de clique e sinais de abandono são registrados como dados de *analytics* para alimentar o fluxo *offline* de melhoria.

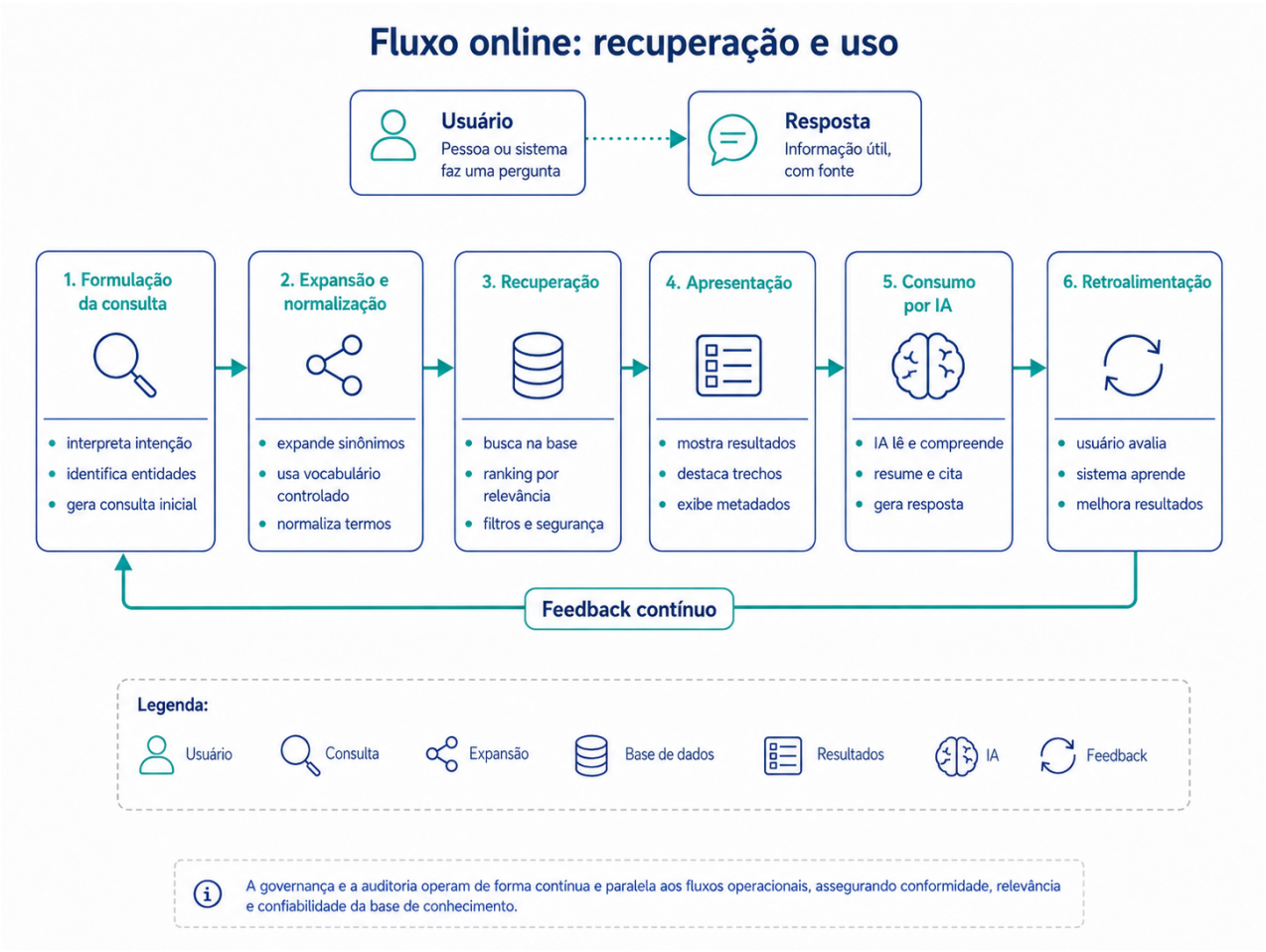

Figura 5 - Fluxo online - recuperação e uso

A imagem relaciona as etapas em tempo real do fluxo online: formulação da consulta, expansão, recuperação, apresentação, consumo por IA e retroalimentação. Ela reforça que a resposta ao usuário depende tanto da qualidade da recuperação quanto do registro de sinais de uso. Esses sinais retornam ao fluxo de melhoria contínua da base.

5.4.3 Fluxo de governança e auditoria

O fluxo de governança opera em paralelo aos fluxos *offline* e online. Para que funcione na prática: nomeie um responsável por artigo, defina prazo de revisão, bloqueie publicação sem metadados obrigatórios e realize auditoria mensal de artigos expirados. As etapas operacionais são:

- **Monitoramento de qualidade**: *dashboards* de métricas (taxa de sucesso de busca, taxa de zero resultados, artigos expirados, *backlog* de revisão) são acompanhados pelo núcleo central de governança.
- **Revisão cadenciada**: cada artigo possui prazo de revisão definido por criticidade. Artigos não revisados dentro do prazo geram notificação automática ao dono e sinalização de estado desatualizado para o usuário.
- **Arquivamento e descontinuação**: conteúdo obsoleto é arquivado com preservação da URL histórica e do registro de proveniência, nunca simplesmente deletado, para manter a trilha de auditoria.

- **Auditoria periódica**: inventário quantitativo, inspeção de desempenho por *analytics*, identificação de redundâncias e conflitos semânticos, verificação de conformidade com LGPD, WCAG e linguagem simples.

Figura 6 - Governança e auditoria contínuas da base de conhecimento

A imagem organiza visualmente o ciclo de governança descrito no texto, destacando monitoramento de qualidade, revisão cadenciada, arquivamento, auditoria e priorização de ajustes. O diagrama deixa claro que a sustentação da qualidade não é uma atividade pontual, mas um ciclo contínuo. Essa visão conecta métricas, responsáveis e evidências de auditoria.

5.5 Arquitetura de referência

As camadas arquiteturais são independentes entre si, permitindo substituição de componentes sem reescrita do modelo de governança.

5.5.1 Fronteiras entre camadas

A operação correta da base de conhecimento depende de fronteiras claras entre as camadas que a compõem e os sistemas que a consomem:

- **Aplicação e conteúdo**: a aplicação (portal, API, conector) não deve conter lógica de negócio embutida no conteúdo. Regras condicionais, exceções e variações de contexto pertencem ao conteúdo e aos metadados, não ao código da aplicação.
- **Conteúdo e modelo de IA**: quando a base alimenta um modelo de linguagem, o conteúdo é a fonte de verdade. O modelo não deve ser tratado como substituto da base nem como árbitro de conflitos entre artigos. Contradições no conteúdo que o modelo tenta conciliar produzem alucinações, não sínteses.
- **Dados e segurança**: permissões de acesso residem na camada de metadados e são aplicadas pela camada de interface e recuperação, não pelo conteúdo em si. Artigo marcado como restrito não deve ser retornado a usuários sem as permissões correspondentes, independentemente da consulta.
- **Observabilidade**: logs de busca, métricas de uso e registros de auditoria são produzidos por todas as camadas e consumidos exclusivamente pela camada de governança. Dados de uso não devem retroalimentar automaticamente o conteúdo sem validação humana.

5.6 Políticas de decisão do componente de recuperação

O componente de recuperação toma decisões que afetam diretamente o que o usuário ou sistema recebe como resposta. As políticas a seguir devem ser documentadas e auditáveis:

- **Política de relevância**: critérios de ranqueamento (pesos de campo, decaimento por idade, *boost* por tipo de tópico) devem ser versionados e revisados periodicamente, não tratados como parâmetro invisível de plataforma.
- **Política de resultado vazio**: quando a busca retorna zero resultados, o sistema deve registrar o termo, sugerir reformulação e escalar o termo para a fila de lacunas editoriais. Nunca deve sugerir conteúdo não relacionado como substituto.
- **Política de conteúdo expirado**: artigos fora do prazo de revisão devem ser sinalizados ao usuário e ao sistema de IA como "pendente de atualização", não simplesmente suprimidos nem apresentados sem ressalva.
- **Política de permissão**: artigos restritos devem ser filtrados antes do ranqueamento, não após. A ausência de resultado por restrição de permissão não deve ser indistinguível da ausência de conteúdo.

5.7 Dependências externas, soberania tecnológica e exposição de dados

Bases de conhecimento no setor público frequentemente dependem de infraestrutura de terceiros: motores de busca como serviço, plataformas de KM em nuvem, modelos de linguagem para geração ou indexação semântica. Essas dependências introduzem riscos que devem ser gerenciados explicitamente:

- **Soberania sobre o conteúdo**: o contrato de serviço deve garantir que o conteúdo da base pertence à organização e pode ser exportado em formato aberto a qualquer momento,

sem aprisionamento tecnológico (*vendor lock-in*).

- **Residência dos dados**: conteúdo classificado como sigiloso ou sensível sob a LGPD não deve ser enviado a APIs externas de processamento sem base legal adequada, contrato de processamento de dados e avaliação de risco.
- **Disponibilidade**: a base de conhecimento não pode ter disponibilidade dependente de um único provedor externo para serviços críticos de atendimento. Estratégias de degradação segura devem prever funcionamento local mínimo em caso de indisponibilidade de serviços externos.
- **Modelos de linguagem externos**: quando a base alimenta um modelo de linguagem hospedado por terceiros, cada consulta transmite fragmentos de conteúdo potencialmente sensível ao provedor do modelo. A política de uso desse provedor deve ser avaliada antes da integração.

5.8 Limitações técnicas inerentes

As limitações a seguir são inerentes à categoria de solução e devem ser consideradas no planejamento, não tratadas como defeitos de implementação:

- **Conhecimento tácito**: bases de conhecimento capturam apenas o que pode ser externalizado em linguagem. Conhecimento profundamente tácito, como julgamento situacional e intuição especializada, não é redutível a artigos sem perda de precisão relevante.
- **Degradação temporal**: todo conteúdo envelhece. Sem governança ativa de revisão, a base acumula artigos desatualizados que degradam a confiança do usuário e a qualidade das respostas de sistemas de IA.
- **Cobertura de cauda longa**: demandas raras e altamente específicas frequentemente não têm cobertura na base. Sistemas de autoatendimento ou IA que dependem exclusivamente da base falharão nesse segmento.
- **Qualidade de recuperação sem metadados**: motores de busca em texto completo sem vocabulário controlado e metadados adequados produzem resultados inconsistentes e dependentes da variação linguística da consulta.
- **Amplificação de erros por IA**: quando a base alimenta um sistema de IA generativa, erros, contradições e desatualizações do conteúdo são amplificados nas respostas geradas. A qualidade da saída do modelo é limitada pela qualidade da base.

6 Técnicas, arquiteturas e métodos

6.1 Representação do conhecimento

A escolha do modelo de representação do conhecimento determina o que a base pode expressar, como o conteúdo é recuperado e quão bem sistemas de IA o consomem. Os modelos

formam um espectro de complexidade crescente:

Modelo	Definição	Quando usar	Quando evitar
Folksonomia	Lista de *tags* geradas pelos usuários sem hierarquia formal	Captura inicial, exploração terminológica	Sempre como mecanismo principal de recuperação
Taxonomia	Vocabulário controlado em hierarquia pai-filho	Navegação em menus, arquivamento por categorias	Domínios com relações cruzadas complexas
Tesauro	Taxonomia mais relações de equivalência (sinônimos/termos preferidos) e associação	Indexação de grandes acervos, domínios com variação terminológica alta	Quando custo de manutenção do vocabulário supera o ganho de recuperação
Ontologia (OWL/RDF)	Modelo formal com restrições lógicas e inferência	Domínios regulados, integração entre sistemas, inferência automatizada	Ambientes sem equipe de modelagem semântica
Knowledge graph	Implementação física de ontologia em banco de grafos	Catálogos complexos, IA, rastreabilidade de relações entre entidades	Quando taxonomia hierárquica simples resolve
SKOS	Padrão W3C ponte entre taxonomias, tesauros e linked data	Publicação de vocabulários, mapeamento entre bases, interoperabilidade	Como único mecanismo de governança editorial

Para a maioria das organizações públicas, o ponto de equilíbrio está em tesauro gerenciado com SKOS, complementado por metadados Dublin Core e proveniência PROV-DM. Knowledge graphs completos devem ser considerados apenas quando o volume de relações entre entidades justifica o custo de modelagem.

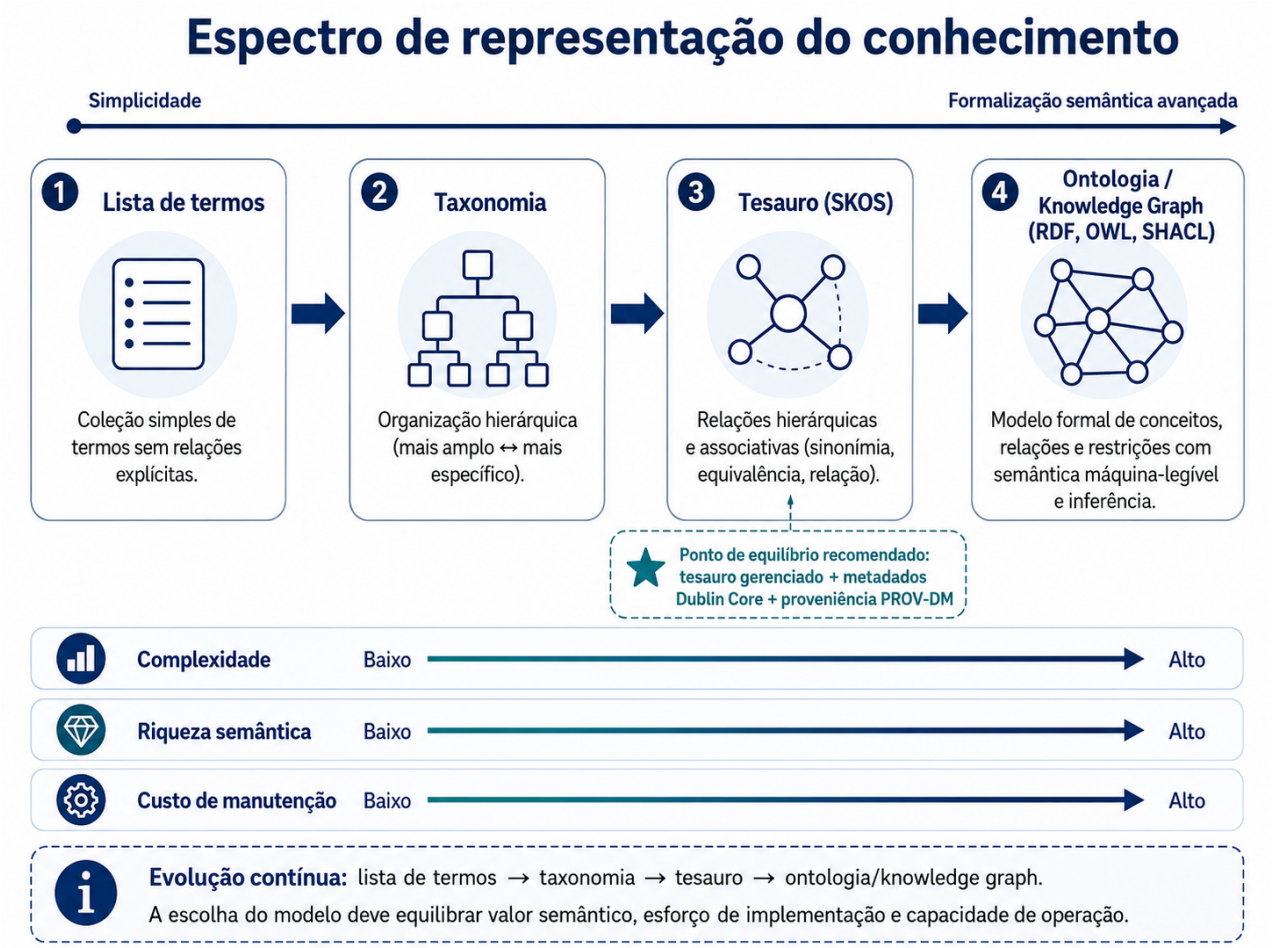

Figura 7 - Espectro de representação do conhecimento

A imagem apresenta a progressão entre lista de termos, taxonomia, tesauro e ontologia/*knowledge graph.* Ela contextualiza a recomendação do texto de equilibrar riqueza semântica, custo de manutenção e capacidade operacional. O destaque para tesauro gerenciado com metadados e proveniência reforça o ponto de equilíbrio sugerido para a maior parte das organizações públicas.

6.2 Estruturação de conteúdo

6.2.1 Tipos de tópico DITA

Este anexo adota quatro tipos DITA como perfil mínimo operacional: *Concept, Task, Reference e Troubleshooting*. A especificação DITA 1.3 contém outros tipos, como *general task, machinery task, glossary entry e glossary group.* A tipificação garante que máquinas consigam distinguir semanticamente uma instrução operacional de uma premissa teórica, o que é condição necessária para recuperação precisa e consumo por IA:

- ***Concept***: contexto, definições, "o que é" e "por que existe". Proibido conter passos de execução ou dados de referência. Parâmetros mínimos: título declarativo, resumo em uma frase, corpo com hierarquia de cabeçalhos.
- ***Task***: instrução sequencial com pré-requisitos, passos numerados e resultado esperado. Cada passo deve conter uma única ação verificável. Risco associado: passos vagos ou com múltiplas ações acopladas, que impedem automação e aumentam erros de execução.

- ***Reference***: dados factuais de consulta imediata, como tabelas de parâmetros, especificações técnicas, valores de configuração e listas de códigos. Não deve conter narrativa explicativa. Parâmetros mínimos: nome do campo, tipo, valor padrão, faixa válida e descrição curta.
- ***Troubleshooting***: diagnóstico com sintoma (condição observável), causa raiz e remédio (passos de mitigação ou resolução). Estrutura condicional: "se X, então Y". Controle mínimo: cada sintoma deve mapear para uma causa distinta.

Figura 8 - Tipos de tópico DITA e seus usos

A imagem resume os quatro tipos DITA adotados como perfil mínimo operacional: *Concept, Task, Reference e Troubleshooting*. Ela relaciona cada tipo à pergunta que responde e à estrutura mínima exigida. Essa tipificação apoia a recuperação precisa e o consumo por IA porque separa definições, procedimentos, dados de referência e diagnóstico de falhas.

6.2.2 Granularidade e *interlinking*

A unidade de conteúdo deve ser o menor tópico autocontido que resolve um único propósito. Tópicos longos e enciclopédicos reduzem a precisão de recuperação e aumentam o risco de janela de contexto truncada em sistemas de IA. Referências cruzadas entre tópicos devem ser implementadas via metadados e IDs absolutos, nunca por posição textual relativa ("consulte o item anterior"), pois essas referências colapsam quando o conteúdo é consumido isoladamente via busca ou API.

6.2.3 Conteúdo para consumo por IA

Bases que alimentam sistemas de IA generativa requerem cuidados adicionais de estruturação:

- **Eliminação de ambiguidade referencial**: pronomes e referências implícitas ("isso", "ele", "o sistema") devem ser substituídos pela entidade explícita, pois modelos de linguagem frequentemente perdem a vinculação referencial quando o trecho é consumido fora de contexto.
- **Resolução de contradições**: dois artigos com diretrizes mutuamente exclusivas produzem alucinações, não sínteses. Conflitos semânticos devem ser resolvidos com condicionais explícitas ("se contexto A, prazo X; se contexto B, prazo Y") documentadas em um único nó canônico.
- **Contexto explícito**: premissas, siglas e jargões organizacionais que qualquer humano infere pelo contexto cultural devem ser definidos textualmente, pois modelos os interpretam literalmente.
- **Metadados como controle**: campos como validade_expiracao, jurisdicao e publico_alvo devem ser preenchidos e processados pelo pipeline de recuperação para filtrar conteúdo inaplicável antes de enviar ao modelo.
- **Exemplos concretos**: bases com exemplos empíricos práticos produzem melhor fundamentação em modelos do que bases com apenas definições axiomáticas abstratas.

6.3 Esquemas de metadados

Esquema	Características	Uso recomendado
Dublin Core (DCMI)	Elementos descritivos, suportados por RDF, interoperabilidade máxima com *Linked Data*, padrão ISO e NISO	Metadados mínimos portáveis, coleções digitais, interoperabilidade entre bases
Schema.org	Focado em web semântica e SEO, metadados embutidos em HTML via JSON-LD	Portais públicos com necessidade de indexação por mecanismos de busca externos
Metadados customizados	Extensões específicas do domínio (estado de aprovação, ID de ticket, grau de sigilo, validade)	Rastreabilidade além do domínio semântico padrão, controle operacional

6.4 Arquitetura de busca e recuperação

6.4.1 Algoritmos de relevância

O algoritmo BM25 (*Best Matching 25*) é o mecanismo padrão de relevância em recuperação de texto completo. Ele equaliza a frequência do termo no documento com sua raridade no corpus, compensando documentos longos por um fator de normalização. Para bases de conhecimento, três ajustes são recomendados:

- ***Field boosting***: ocorrências do termo no título e no resumo recebem peso maior do que

ocorrências no corpo. Implementado como fator multiplicativo por campo.

- ***Freshness boost***: artigos mais recentes ou recém-revisados recebem *boost* decrescente por função de decaimento temporal. Parâmetro de meia-vida deve ser calibrado por criticidade do domínio.
- ***Type boost***: tópicos do tipo *Task* podem receber peso maior em consultas que contêm verbos de ação; tópicos *Troubleshooting* em consultas que contêm termos de erro.

6.4.2 Classificação de motores de busca

Motor	Força	Limitação
Elasticsearch / Solr	Escalabilidade, agregações complexas, expressividade analítica profunda	Latência de configuração inicial alta, curva de UX pobre sem customização
Typesense / Meilisearch	Baixa latência em cenários interativos (a validar por teste de carga no ambiente-alvo), tolerância a erros tipográficos nativa, UX instantânea	Limitações em agregações profundas e volumes de petabytes
Busca semântica (vetorial)	Recuperação por similaridade semântica, robusta a variações lexicais	Exige embeddings, custo computacional maior, menos explicável

6.4.3 Busca facetada e *analytics*

Busca facetada permite ao usuário filtrar resultados por múltiplas dimensões ortogonais (tipo de tópico, domínio, produto, data) simultaneamente. Valores dentro de uma faceta devem ser mutuamente exclusivos. O design de facetas deve se limitar a três a sete dimensões primárias para evitar sobrecarga cognitiva.

Search analytics é o mecanismo mais confiável de diagnóstico de saúde da base. A taxa de zero resultados é o indicador mais direto de lacunas de cobertura ou desalinhamento entre o vocabulário da base e o vocabulário dos usuários. Termos sem resultado devem alimentar automaticamente a fila editorial de lacunas.

6.5 Padrões arquiteturais recomendados

O padrão recomendado para a maioria das organizações públicas é a arquitetura híbrida de referência: base semiestruturada com taxonomia controlada, metadados mínimos obrigatórios, motor de busca com BM25 ajustado e proveniência registrada. Os três padrões alternativos válidos são:

- **Wiki com curadoria ativa**: adequada para equipes pequenas com conhecimento emergente. Exige curador dedicado e revisão taxonômica trimestral. Inapropriada para bases que alimentam IA sem camada de validação posterior.
- ***Docs-as-code* com versionamento Git**: adequada para documentação técnica de sistemas de software. Permite rastreabilidade total via histórico de *commits*, mas exige proficiência técnica dos autores e é inacessível para equipes não técnicas.

- ***Knowledge graph* com ontologia de domínio**: adequada para domínios com relações complexas entre entidades (legislação, saúde, catálogos de serviços). Exige equipe de modelagem semântica e manutenção ontológica contínua.

6.6 Padrões arquiteturais avançados

- **SSOT (*Single Source of Truth*)**: cada fato ou instrução reside em um único local canônico. Outros pontos de apresentação referenciam ou transcludem o conteúdo original sem duplicá-lo. Evita inconsistências quando uma política se altera.
- **Recuperação híbrida**: combinação de BM25 (busca léxica) com busca vetorial semântica, com *reranking* por relevância cruzada. Recomendado para bases que alimentam pipelines RAG com acervos mistos.
- **Modelo federado com catálogo central**: domínios mantêm bases autônomas; um catálogo central agrega metadados e expõe busca unificada. A taxonomia e o esquema de metadados são de responsabilidade central; o conteúdo é de responsabilidade dos domínios.

6.7 Integração com sistemas externos

Bases de conhecimento no setor público integram-se tipicamente com:

- Sistemas de ITSM (*ServiceNow*, *Jira Service Management*, *TOPdesk)* para captura no fluxo KCS;
- Portais de atendimento ao cidadão, com exposição de conteúdo via *widget* de busca ou API REST;
- Pipelines RAG e LLM, com exposição de conteúdo estruturado, metadados e proveniência via API;
- Sistemas de autenticação e autorização (SSO, LDAP) para controle de acesso baseado em papel;
- Ferramentas de *analytics* e observabilidade para coleta de métricas de uso.

Toda integração deve ser documentada com: sistema integrado, tipo de dado trafegado, direção do fluxo, autenticação, frequência de sincronização e classificação LGPD do conteúdo.

6.8 Segurança desde a concepção e controles preventivos

- Permissões de acesso devem ser definidas no nível do artigo e herdadas da taxonomia, não apenas no nível do portal.
- Conteúdo restrito nunca deve aparecer em resultados de busca de usuários sem a permissão correspondente, mesmo que parcialmente.
- Campos de metadados com dados pessoais (histórico de atendimento, nome do servidor como autor) devem ser classificados e minimizados.

- Logs de busca e de uso devem ser anonimizados antes do armazenamento de longo prazo.
- Integrações com modelos de linguagem externos devem filtrar conteúdo restrito antes do envio, nunca após.

6.9 *Trade-offs* técnicos e critérios de escolha

Decisão	Opção A	Opção B	Critério de escolha
Folksonomia vs vocabulário controlado	Velocidade de captura, terminologia emergente	Consistência de recuperação, governança	Volume e criticidade; híbrido é o padrão
Taxonomia vs *knowledge graph*	Custo baixo, suficiente para navegação	Inferência, relações complexas	Existência de relações entre entidades que taxonomia não expressa
Wiki vs DITA	Adoção rápida, custo baixo	Modularidade, reuso, prontidão para IA	Volume, audiência técnica, necessidade de multicanal
Elasticsearch vs Typesense	Análise profunda, grande escala	Latência baixa, UX imediata	Volume de acervo e perfil de uso (exploratório vs dirigido)
Publicar rápido vs validar fortemente	Atualidade, aderência ao uso	Consistência, confiabilidade	Criticidade do domínio (suporte vs saúde vs jurídico)

6.10 Faça e não faça — decisões técnicas críticas

Faça	Não faça
Classificar o nível de acesso antes de configurar a indexação	Indexar todo o conteúdo no mesmo índice independentemente do nível de acesso
Filtrar permissões antes do ranqueamento no pipeline de recuperação	Filtrar permissões após o ranqueamento ou confiar no LLM para não revelar conteúdo restrito
Testar a busca com usuários reais do público-alvo declarado	Declarar sucesso de busca com base apenas em testes internos da equipe técnica
Revisar conteúdo gerado por IA antes de publicar na base	Publicar automaticamente conteúdo gerado por IA sem revisão humana
Documentar a base legal para cada categoria de dado pessoal tratado	Usar consentimento como base legal padrão no setor público
Usar legítimo interesse como base legal apenas após análise de finalidade, necessidade e balanceamento	Usar legítimo interesse para dados pessoais sensíveis ou para tratamento que decorre de prerrogativa estatal típica
Enviar ao pipeline de IA apenas conteúdo com metadados de permissão e proveniência completos	Tratar vector database como repositório confiável sem governança de conteúdo
Aplicar filtro de conteúdo proibido antes de enviar ao modelo externo	Enviar documentos com nível de acesso Restrito ou Confidencial para API externa de LLM sem base legal, contrato de processamento e avaliação de risco

Faça	Não faça
Manter trilha de qual documento fundamentou qual resposta do sistema de IA	Aceitar resposta de IA sem citação de fonte verificável
Implementar mecanismo de feedback humano sobre respostas do sistema de IA	Deixar o LLM decidir o nível de acesso ao conteúdo
Documentar controles mínimos de segurança para ingestão, armazenamento, busca e APIs	Medir sucesso da base apenas pelo número de artigos publicados
Testar revogação de acesso após desligamento de servidor	Assumir que remoção do usuário no IdP revoga acesso automaticamente em todos os sistemas integrados

7. Avaliação e métricas

7.1 Estratégia geral de avaliação

A armadilha mais comum na avaliação de bases de conhecimento é medir atividade no lugar de impacto. Segundo pesquisas da APQC, a maioria das organizações rastreia volume de publicações e satisfação de usuários; uma minoria mede impacto operacional e calcula ROI. Programas que medem impacto têm significativamente mais chance de ser percebidos como eficazes pela liderança e de obter aprovação orçamentária. Isso não é preferência metodológica: é condição de sobrevivência institucional da iniciativa.

A estratégia de avaliação deve ser definida antes da implantação, não após. As métricas devem ser coletadas automaticamente por *analytics* e não depender de esforço manual de coleta. O painel mínimo deve cobrir as sete dimensões a seguir.

7.2 Métricas da capacidade principal: encontrabilidade e busca

Taxa de sucesso de busca (*Findability Rate*): proporção de sessões de busca em que o usuário clica em um resultado e não retorna imediatamente para reformular a consulta. Interpretação: abaixo de 60% indica problema crítico de cobertura ou ranqueamento. Fórmula: sessões com clique e sem reformulação imediata / total de sessões de busca.

Taxa de zero resultados: proporção de consultas que não retornam nenhum resultado. Interpretação: acima de 10% indica lacuna de cobertura ou desalinhamento de vocabulário. Cada termo sem resultado deve ser registrado para triagem editorial.

Taxa de abandono: proporção de sessões em que o usuário acessa a base e sai sem clicar em nenhum resultado. Interpretação: abandono alto, com taxa de zero resultados - baixa, indica problema de relevância ou apresentação dos resultados.

7.3 Métricas de qualidade da saída

Pontuação de confiança do artigo (*Article Confidence Score*): índice composto que combina recência da última revisão, volume de visualizações no período, taxa de avaliação positiva e

taxa de abandono imediato. Artigos com pontuação abaixo do limiar crítico devem entrar automaticamente na fila de revisão.

Taxa de artigos expirados: proporção de artigos que ultrapassaram o prazo de revisão definido por sua criticidade. Interpretação: acima de 15% indica falha no fluxo de governança editorial.

Taxa de aprovação editorial: proporção de novos artigos aprovados sem retorno para correção. Interpretação: abaixo de 70% indica necessidade de melhoria nos *templates* ou no treinamento da equipe.

7.4 Métricas de impacto operacional

Taxa de deflexão (*Deflection Rate*): proporção de sessões na base que resultam em resolução sem abertura de chamado de suporte. Calculada por:

$$Deflection\ Rate = \frac{sessões\ resolvidas\ sem\ chamado}{sessões\ resolvidas\ sem\ chamado + chamados\ abertos} \times 100$$

Interpretação: cada ponto percentual de aumento representa redução direta de volume de atendimento. Bases maduras atingem taxas de 30% a 50% em portais de autoatendimento.

Taxa de autoatendimento (Self-Service Rate): proporção do total de demandas resolvidas pela base sem intervenção humana. Métrica mais abrangente que a deflexão, pois inclui demandas que nunca chegariam ao canal de suporte.

Tempo de resolução: para demandas que passam pela base antes de abrir chamado, tempo médio de resolução comparado à linha de base sem a base. Redução de tempo de resolução é um dos indicadores de ROI mais citados em benchmarks APQC.

7.5 Métricas de conformidade e prontidão para IA

Cobertura de metadados obrigatórios: proporção de artigos com todos os campos obrigatórios preenchidos. Deve ser 100%. Qualquer valor abaixo indica falha de processo ou de validação.

Cobertura de proveniência: proporção de artigos com registro de origem, dono e histórico de aprovação completo. Requisito mínimo para bases que alimentam sistemas de IA.

Taxa de incidentes de privacidade: número de ocorrências de conteúdo publicado indevidamente com dados pessoais ou com nível de acesso incorreto. Deve ser zero em condições normais de operação.

Taxa de fundamentação em IA (*Grounding Rate*): quando a base alimenta um sistema de IA, proporção de respostas geradas que citam ao menos um artigo da base como fonte. Baixa fundamentação indica lacuna de cobertura ou problema de recuperação.

7.6 Limiares recomendados por criticidade

Métrica	Criticidade baixa	Criticidade média	Criticidade alta
Taxa de sucesso de busca	≥ 60%	≥ 70%	≥ 80%
Taxa de zero resultados	≤ 15%	≤ 10%	≤ 5%
Artigos expirados	≤ 20%	≤ 15%	≤ 5%
Cobertura de metadados	≥ 90%	≥ 95%	%
Cobertura de proveniência	≥ 80%	≥ 90%	%
Incidentes de privacidade	0	0	0
Taxa de deflexão (se aplicável)	≥ 20%	≥ 30%	≥ 40%

8. Padrões operacionais recomendados

8.1 Padrões por tipo de sistema

Tipo	Arquitetura recomendada	Controles essenciais	Métricas prioritárias	Critério de produção
Atendimento / ITSM	KB semiestruturada + fluxo KCS + motor de busca	Captura no fluxo, revisão de artigos usados, metadados de produto	Deflexão, reuso, tempo de resolução	Deflexão 20%, zero artigos expirados críticos
Políticas internas	KB estruturada + taxonomia corporativa + versionamento	Revisão jurídica, controle de versão, SSOT	Sucesso de busca, cobertura, conformidade	Metadados 100%, sem contradições documentadas
Portal público	KB semiestruturada + linguagem simples + acessibilidade	WCAG, Lei 15.263/2025, triagem de dados pessoais	Conclusão de tarefa, deflexão de atendimento	Conformidade WCAG AA, linguagem simples validada
Ambiente regulado	KB estruturada + trilha de auditoria + permissões finas	Aprovação formal, imutabilidade de versões, LGPD	Não conformidades, tempo de revisão	Trilha de auditoria completa, zero exposição de dados
Base para IA (RAG)	KB estruturada + proveniência + metadados ricos	Qualidade mínima verificada, filtragem de permissão pré-recuperação, feedback humano	*Grounding* rate, incidentes, cobertura	Qualidade limiares da Tab. 10, proveniência 100%

Regra obrigatória para uso de LLM externo, nuvem estrangeira ou subprocessador internacional: Antes da contratação ou integração, verificar: (a) ocorrência de transferência internacional de dados pessoais; (b) identidade e localidade dos subprocessadores; (c) localidade de armazenamento e backup; (d) política de retenção e descarte pelo fornecedor; (e) vedação contratual expressa de uso dos dados do órgão para treinamento, ajuste fino ou melhoria de modelos; (f) adoção de cláusulas contratuais compatíveis com a Resolução CD/ANPD nº 19/2024, que regulamentou os mecanismos de transferência internacional de dados pessoais.

8.2 Controles mínimos de segurança por componente

Componente	Controles obrigatórios
Ingestão de conteúdo	Validação de fonte e autoridade; verificação de *malware* para arquivos binários; assinatura ou hash do artefato ingerido; revisão humana antes da publicação; quarentena para fontes não confiáveis.
Armazenamento	Criptografia em repouso para conteúdo Restrito e Confidencial (TLS em trânsito mínimo 1.2); segregação de índice por domínio e nível de acesso; controle de acesso baseado em papel (RBAC); backup diário com teste trimestral de restauração; trilha de auditoria imutável.
Busca e recuperação	Filtro de permissão aplicado antes do ranqueamento em todos os componentes do pipeline (léxico e vetorial); resultado negado por restrição de acesso distinguível de resultado vazio por ausência de conteúdo; logs de acesso a conteúdo restrito retidos conforme política de auditoria.
Pipeline LLM / RAG	Filtro de conteúdo proibido antes do envio ao modelo; logging sem dados excessivos (minimização de PII nos logs); red team antes da implantação em produção; bloqueio de conteúdo fora do escopo de permissão do usuário; mecanismo de feedback humano operacional antes de go-live.
APIs externas	Autenticação forte (OAuth2 / chave de API com escopo mínimo); rate limiting; gestão de segredos em cofre (nunca em código ou variável de ambiente sem controle); monitoramento de chamadas e alertas de anomalia; cláusula contratual proibindo uso de dados para treinamento do modelo.
Desligamento de acesso	Processo formal de revogação de acesso para servidores desligados ou remanejados; teste periódico de que a revogação no IdP se propaga aos sistemas integrados; relatório mensal de contas sem dono ativo.

8.2.1 Controles de segurança para bases usadas em RAG

Bases de conhecimento que alimentam pipelines de Geração Aumentada por Recuperação (RAG) requerem controles adicionais específicos:

- **Filtro de permissão pré-recuperação**: o filtro deve ocorrer antes do ranqueamento e antes da montagem do contexto enviado ao LLM. Aplicar filtro apenas após a geração constitui falha crítica.
- **Isolamento de *tenant***: índices vetoriais devem isolar conteúdo por *tenant* ou perfil de acesso, impedindo recuperação cruzada não autorizada.

- **Rastreabilidade de *chunks***: registrar os IDs dos trechos recuperados, versão da base, política de recuperação aplicada e vínculo com a resposta gerada.
- **Bloqueio de conteúdo expirado**: conteúdo expirado não deve alimentar resposta automatizada sem sinalização explícita de validade e regra de *fallback* definida.
- **Testes de *prompt injection***: verificar se entradas do usuário conseguem alterar o comportamento do sistema ou extrair conteúdo além do autorizado.
- **Testes de vazamento de *system prompt***: verificar se o *prompt* de sistema ou instruções internas são expostos ao usuário final.
- **Testes de recuperação indevida**: confirmar que conteúdo restrito não aparece em respostas para usuários sem permissão via busca lexical, semântica ou assistente RAG.
- **Detecção de envenenamento da base**: monitorar inserções anômalas de conteúdo que possam manipular respostas do sistema de IA.

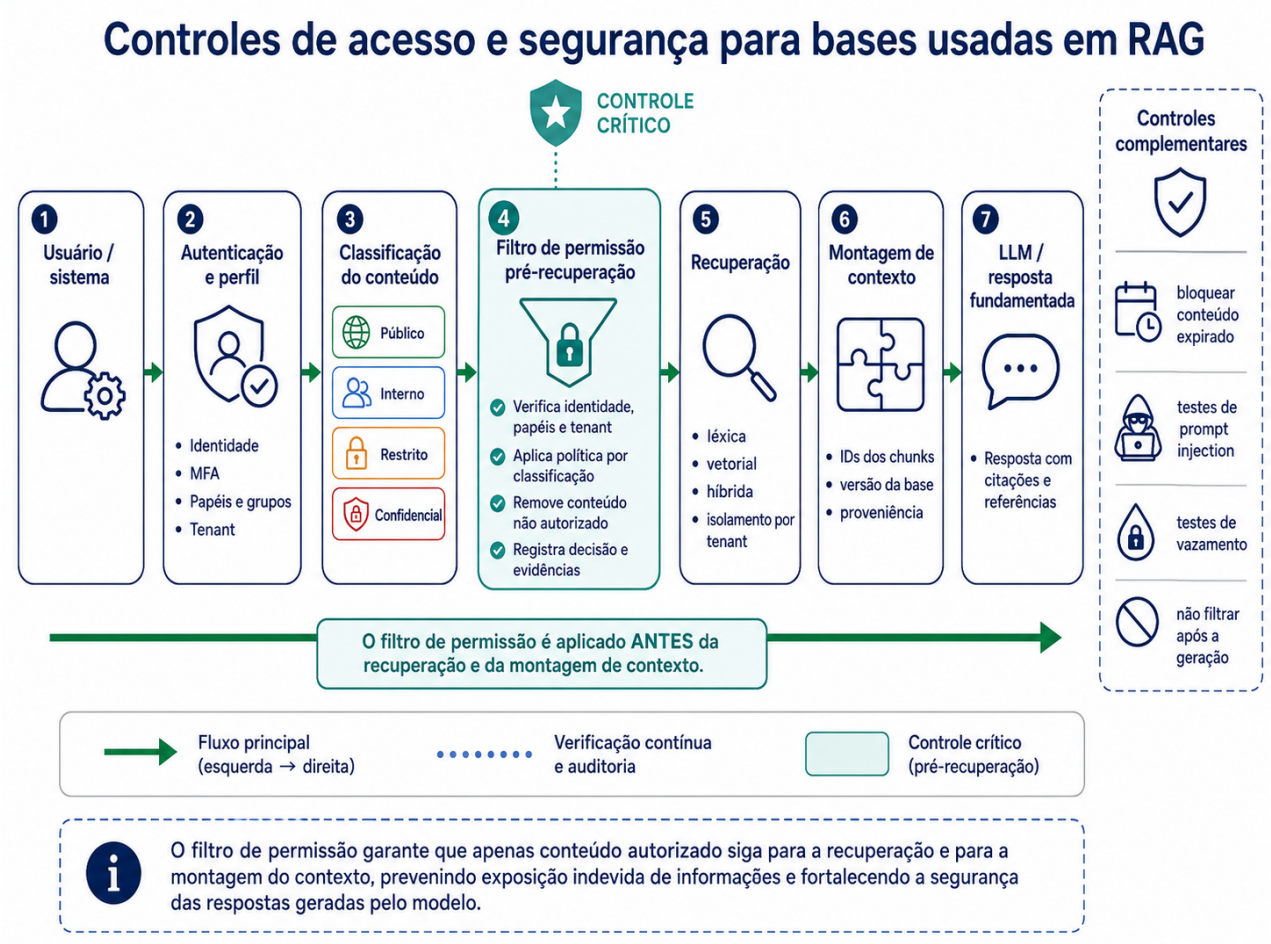

Figura 9 - Controles de acesso e segurança para bases usadas em RAG

A imagem consolida os controles mínimos descritos nesta subseção para bases que alimentam pipelines RAG. Ela combina autenticação, classificação de conteúdo, filtro pré-recuperação, isolamento por *tenant*, rastreabilidade de *chunks* e resposta fundamentada. Os controles complementares reforçam testes contra *prompt injection*, vazamento e uso indevido de conteúdo expirado.

8.3 *Checklist* mínimo de produção

O *checklist* a seguir deve ser verificado antes de qualquer base de conhecimento entrar em operação, independentemente do tipo de sistema:

- Todos os artigos possuem os metadados obrigatórios preenchidos.
- Todos os artigos possuem dono ativo identificado.
- A taxonomia está documentada e versionada.
- Os fluxos de aprovação estão configurados e testados.
- Os controles de acesso estão implementados e validados.
- O *dashboard* de métricas está ativo e acessível à equipe de governança.
- O processo de revisão cadenciada está configurado com prazos por criticidade.
- A política de zero resultados está configurada (registro de termos, escalada editorial).
- A conformidade com WCAG AA foi testada (para portais públicos e sistemas acessíveis a servidores com deficiência).
- O conteúdo foi revisado para conformidade com a Lei 15.263/2025 (para portais públicos).
- A análise de dados pessoais foi realizada e documentada conforme LGPD.
- O plano de exportação e migração foi testado (backup de conteúdo em formato aberto).
- Se a base alimenta IA: os limiares de qualidade da Tabela do item 7.3 estão sendo atingidos; a filtragem de permissão pré-recuperação foi testada; o mecanismo de feedback humano está operacional.

8.4 *Checklist* de privacidade antes da publicação

Este *checklist* deve ser verificado por todo editor antes de publicar qualquer artigo, independentemente do tipo ou do nível de acesso pretendido:

- O artigo contém dados pessoais (nome, matrícula, e-mail, CPF, cargo + setor, foto, protocolos identificáveis)?
- Os exemplos, estudos de caso ou *prints* de tela incluem dados pessoais ou de pessoas identificáveis?
- O artigo contém logs, registros de atendimento ou histórico de uso que identifique usuários?
- A base legal para o tratamento dos dados pessoais identificados está documentada?
- O nível de acesso atribuído ao artigo é compatível com a sensibilidade dos dados presentes?
- Os dados pessoais identificados são necessários para o propósito do artigo ou podem ser removidos sem prejudicar a utilidade do conteúdo?
- Existe versão anonimizada ou pseudonimizada possível?
- O conteúdo é público, interno, restrito ou sigiloso nos termos da LAI?

- Há aviso de privacidade ou aviso de sigilo aplicável ao conteúdo?

Se qualquer resposta nos itens 1 a 3 for positiva, o artigo deve ser revisado pelo Data Steward antes da publicação. Se houver dados pessoais sensíveis, a publicação deve ser precedida de análise pelo encarregado pelo tratamento de dados pessoais (DPO).

8.5 Checklist de linguagem simples e acessibilidade editorial

Para portais públicos e bases de acesso por servidores com diferentes perfis de formação, verificar antes da publicação:

Linguagem simples (requisito da Lei 15.263, de 14 de novembro de 2025):

- O título expressa a tarefa do usuário, não o nome interno do processo?
- A primeira frase responde diretamente ao que o usuário precisa saber ou fazer?
- As frases têm em média até 25 palavras?
- A voz ativa predomina ("o servidor envia o formulário", não "o formulário deve ser enviado pelo servidor")?
- As siglas são explicadas na primeira ocorrência?
- Os passos são numerados e cada passo contém uma única ação?
- O texto foi testado com ao menos um representante do público-alvo?
- O juridiquês desnecessário foi eliminado?

Acessibilidade (WCAG 2.2 AA e eMAG):

- Todas as imagens têm texto alternativo descritivo?
- A estrutura de cabeçalhos segue hierarquia semântica (H1, H2, H3)?
- O contraste de texto atende ao mínimo de 4.5:1?
- O conteúdo é navegável por teclado sem armadilhas de foco?
- Links descrevem o destino (não "clique aqui" ou "leia mais")?
- Tabelas têm cabeçalhos identificados corretamente?

8.6 Gestão de incidentes

Fluxo obrigatório de resposta a incidentes com dados pessoais:

- Identificar e conter o incidente; preservar evidências.
- Avaliar se há risco ou dano relevante a titulares de dados pessoais.
- Se houver risco ou dano relevante: comunicar à ANPD e aos titulares afetados no prazo de **três dias úteis** a partir da ciência do incidente.
- Registrar a decisão de comunicar ou não comunicar, com justificativa documentada.

- Manter todos os registros do incidente por **pelo menos cinco anos**.

Incidentes em bases de conhecimento classificam-se em três severidades:

- **Severidade 1 (crítica)**: exposição de dados pessoais, conteúdo incorreto em base que alimenta IA em produção, indisponibilidade total. Tempo de resposta: imediato. Ação: remoção imediata do conteúdo problemático, notificação ao encarregado, *rollback* para versão anterior, comunicação aos usuários afetados. **Obrigação legal:** incidentes que envolvam dados pessoais e que possam acarretar risco ou dano relevante aos titulares devem ser comunicados pelo controlador à ANPD e aos titulares afetados em até três dias úteis a partir da ciência do incidente, conforme Resolução CD/ANPD no 15, de 24 de abril de 2024, publicada no Diário Oficial da União em 26 de abril de 2024.
- **Severidade 2 (alta)**: degradação significativa de qualidade de busca, conteúdo contraditório publicado, artigo crítico desatualizado. Tempo de resposta: 24 horas. Ação: marcação do artigo como em revisão, notificação ao dono do domínio.
- **Severidade 3 (moderada)**: artigo expirado não crítico, metadado incorreto, link quebrado. Tempo de resposta: próximo ciclo de revisão. Ação: inclusão no *backlog* editorial prioritário.

8.7 Observabilidade, logs e auditoria

Para pipelines RAG, registrar: IDs dos *chunks* recuperados, versão da base, política de recuperação aplicada, perfil de usuário ou papel, decisão de bloqueio por permissão, e identificador ou *hash* da resposta gerada. Evitar guardar *prompts* completos quando contiverem dados pessoais sem necessidade justificada. Logs devem ser suficientes para reconstituir a decisão técnica, mas minimizados para não constituir novo repositório de dados pessoais desnecessários.

Os prazos de retenção de logs devem ser definidos por matriz vinculada à finalidade, à base legal, à política de gestão documental da organização e aos requisitos de segurança e auditoria. Os prazos abaixo são orientativos e devem ser validados pela equipe jurídica e pelo encarregado da organização:

- **Logs de busca** (termos, resultados, cliques): finalidade de melhoria da base; base legal deve ser documentada; anonimizar antes do armazenamento de longo prazo. **Atenção:** termos de busca podem conter dados pessoais ou sensíveis digitados pelo usuário; aplicar minimização e mascaramento antes da retenção; controlar o acesso aos logs de busca.
- **Logs de aprovação editorial** (quem aprovou, quando, qual versão): pelo prazo de vigência do artigo acrescido do prazo de prescrição aplicável, conforme tabela de temporalidade da organização.
- **Logs de acesso a conteúdo restrito**: pelo prazo da política de segurança da informação e de auditoria da organização; manter ao menos pelo período exigido para responsabilização civil ou administrativa.
- **Registros de proveniência** (origem, versão, aprovação): pelo prazo de existência da base e pelos requisitos de auditoria associados; não descartar sem deliberação formal.
- **Logs de integração com sistemas de IA** (qual artigo foi enviado, para qual consulta, com

qual resultado): pelo prazo de auditoria do sistema de IA integrado, com retenção mínima vinculada ao prazo de responsabilização por incidentes.

- **Registros de incidentes de segurança com dados pessoais**: mínimo de 5 anos a partir da ciência do incidente, conforme orientações da ANPD.

8.8 Mínimo produto seguro para bases que alimentam IA

Antes de conectar uma base de conhecimento a um pipeline LLM ou RAG em produção, os seguintes requisitos devem ser verificados e documentados:

- 100% dos conteúdos utilizados pelo pipeline têm metadados obrigatórios preenchidos, incluindo nível de acesso.
- Proveniência completa registrada para todos os artigos consumíveis pelo pipeline.
- Filtro de permissão pré-recuperação testado e documentado (não confiar no modelo para não revelar conteúdo restrito).
- Avaliação de *grounding* concluída: proporção de respostas de teste com citação verificável dentro do limiar aceitável.
- Avaliação de conteúdo proibido: teste de que o pipeline não retorna conteúdo fora do escopo, ofensivo ou factualmente incorreto.
- *Red team* realizado: tentativas de manipulação por *prompt injection* direta e indireta documentadas e mitigadas.
- Plano de *rollback* documentado e testado: como desconectar o pipeline sem afetar a base e sem perda de dados.
- Trilha de qual documento fundamentou qual resposta operacional.
- Mecanismo de *feedback* humano sobre respostas incorretas ou problemáticas operacional e integrado à fila editorial.
- Cláusula contratual com o provedor do modelo proibindo uso dos dados para treinamento.

9. Riscos específicos, ameaças e controles

9.1 Metodologia de classificação de risco

Cada risco é classificado por probabilidade (alta, média, baixa) e impacto (alto, médio, baixo), resultando em classificação geral (crítico, relevante, moderado, baixo). O molde de cada risco segue o padrão: nome, descrição, cenário de ameaça, ativos afetados, probabilidade, impacto, classificação, mecanismos de detecção, controles obrigatórios, controles recomendados, risco residual e evidências para auditoria.

9.2 Riscos sobre dados e bases de conhecimento

9.2.1.1 R01 — Desatualização silenciosa

Artigos que descrevem processos, políticas ou sistemas que se alteraram sem que o conteúdo seja revisado. Em bases que alimentam IA, o sistema replica informação desatualizada em escala. Probabilidade: alta. Impacto: alto em domínios operacionais e regulados.

Controles obrigatórios: revisão cadenciada com prazo por criticidade, sinalização automática de artigos expirados, métrica de artigos expirados no *dashboard*.

Controles recomendados: detecção de desatualização por alinhamento de tópico com mudanças em sistemas integrados.

9.2.1.2 R02 — Conteúdo contraditório

Dois ou mais artigos com diretrizes mutuamente exclusivas sobre o mesmo tópico, produzidos por áreas distintas sem coordenação. Em sistemas de IA, produz alucinações. Probabilidade: média. Impacto: alto.

Controles obrigatórios: identificação de redundâncias em auditorias periódicas, política de SSOT, resolução de conflito com condicionais explícitas.

Controles recomendados: análise semântica automatizada para detecção de sobreposições.

9.2.1.3 R03 — Envenenamento da base (data poisoning)

Ingestão de conteúdo incorreto, tendencioso ou malicioso, seja por erro de curadoria ou por ação deliberada interna. Probabilidade: baixa a média. Impacto: alto em bases que alimentam IA e portais públicos.

Controles obrigatórios: fluxo de aprovação com revisão técnica antes da publicação, trilha de auditoria de quem aprovou cada artigo, log imutável de alterações.

9.2.1.4 R04 — Exposição de dados pessoais

Publicação acidental de conteúdo com dados pessoais (nome de servidores, dados de atendimento) em nível de acesso mais aberto do que o adequado. Probabilidade: média. Impacto: alto (violação de LGPD).

Controles obrigatórios: *checklist* editorial de verificação de dados pessoais antes da publicação, validação de nível de acesso no fluxo de aprovação, auditoria amostral mensal de conteúdo publicado.

9.3 Riscos sobre o processo editorial

9.3.1.1 R05 — Deriva taxonômica

O vocabulário controlado deixa de ser atualizado e o conteúdo novo é classificado com termos

informais ou inconsistentes, degradando a recuperação. Probabilidade: alta sem governança ativa. Impacto: médio (degradação gradual de qualidade).

Controles obrigatórios: processo formal de atualização do vocabulário com validação pelo KM, treinamento periódico da equipe editorial.

9.3.1.2 R06 — Perda de propriedade do conteúdo

Artigos sem dono definido, especialmente após desligamento de servidores especialistas. Conteúdo órfão não é revisado e acumula desatualização. Probabilidade: alta em organizações com rotatividade. Impacto: médio a alto.

Controles obrigatórios: metadado de dono obrigatório, processo de transferência de propriedade em desligamentos, relatório mensal de artigos sem dono ativo.

9.4 Riscos sobre integração com sistemas de IA

9.4.1.1 R07 — Amplificação de erros por IA

Erros, contradições e desatualizações na base são amplificados em escala pelo sistema de IA que a consome. Um único artigo incorreto pode fundamentar milhares de respostas erradas. Probabilidade: alta em bases com qualidade inconsistente. Impacto: alto. **Controles obrigatórios**: limiares mínimos de qualidade da base antes de integração com IA (cf. Tabela 10), mecanismo de feedback humano sobre respostas do sistema de IA conectado à fila editorial, rastreabilidade de qual artigo fundamentou qual resposta.

9.4.1.1 R08 — Vazamento de conteúdo restrito por IA

O pipeline de recuperação envia conteúdo restrito ao modelo de linguagem, que o inclui em respostas a usuários sem permissão para acessá-lo. Probabilidade: média se o controle de acesso não for aplicado antes da recuperação. Impacto: alto.

Controles obrigatórios: filtragem por permissão antes do ranqueamento, nunca após; teste de penetração do pipeline de recuperação antes da implantação.

Figura 10 - Filtro de permissão pré-recuperação no fluxo RAG

A imagem torna explícito o controle preventivo exigido para evitar vazamento de conteúdo restrito em sistemas RAG. O filtro de permissão aparece antes da recuperação e antes da montagem do contexto, que é exatamente a salvaguarda indicada no texto. Essa posição reduz o risco de que conteúdos não autorizados cheguem ao modelo de linguagem.

9.5 Riscos operacionais e de disponibilidade

9.5.1.1 R09 — Dependência de fornecedor único

Aprisionamento em plataforma proprietária sem exportação de dados em formato aberto, inviabilizando migração. Probabilidade: média. Impacto: alto no longo prazo.

Controles obrigatórios: cláusula contratual de exportação em formato aberto, teste periódico de exportação e reimportação.

9.5.1.2 R10 — Indisponibilidade de serviços externos

Motores de busca como serviço ou modelos de linguagem externos ficam indisponíveis, afetando a base inteira. Probabilidade: baixa a média. Impacto: alto se não houver *fallback*. **Controles obrigatórios**: estratégia de degradação segura com busca local mínima, SLA contratual com penalidade, monitoramento de disponibilidade.

9.6 Riscos específicos de RAG, vetores e sistemas generativos

9.6.1.1 R11 — Injeção de prompt indireta por conteúdo ingerido

Documento incorporado à base contém instruções ocultas ou texto projetado para manipular o comportamento do modelo de linguagem que consome o conteúdo via pipeline RAG. O conteúdo atua como vetor de ataque contra o sistema de IA sem que o usuário final seja o atacante. Probabilidade: baixa a média em bases abertas à ingestão de fontes externas. Impacto: alto — pode produzir respostas manipuladas, vazamento de informações do *prompt* de sistema ou desvio de comportamento.

Controles obrigatórios: validação de fonte antes da ingestão; revisão humana de conteúdo externo antes da publicação; quarentena de documentos de fontes não confiáveis; monitoramento de outputs do sistema de IA para detecção de padrões anômalos.

9.6.1.2 R12 — Fraquezas em vetores e embeddings

Bases vetoriais utilizadas em pipelines RAG podem ser alvo de envenenamento semântico: documentos com conteúdo construído para ficarem próximos semanticamente de consultas legítimas, deslocando conteúdo correto do ranking. Adicionalmente, *embeddings* podem vazar informações sobre o conteúdo da base por reconstrução aproximada a partir de representações vetoriais. Probabilidade: baixa. Impacto: alto em ambientes com múltiplos locatários ou com dados sensíveis.

Controles obrigatórios: segregação de índices vetoriais por domínio e nível de acesso; não tratar o vector *store* como repositório confiável sem governança de conteúdo; filtro de permissão aplicado antes da recuperação semântica, nunca após; auditoria de conteúdo incorporado no índice vetorial. **Referência:** OWASP Top 10 for LLM *Applications* 2025.

9.6.1.3 R13 — Bypass de controle de acesso na recuperação semântica

Em pipelines de recuperação híbrida (léxica + vetorial), o filtro de permissão pode não ser aplicado ao componente vetorial, permitindo que conteúdo restrito seja recuperado semanticamente e entregue ao modelo, mesmo que o usuário não tenha acesso ao artigo original. Probabilidade: média se a arquitetura não for auditada. Impacto: alto.

Controles obrigatórios: garantir que o filtro de permissão seja aplicado antes da recuperação em todos os componentes do pipeline, inclusive no componente vetorial; testar o *bypass* de ACL como parte do teste de aceitação.

9.6.1.4 R14 — Dependência de modelo externo e uso de dados para treinamento

Quando a base alimenta um modelo hospedado por terceiros, o conteúdo transmitido nas consultas pode ser utilizado pelo fornecedor para treinamento ou aprimoramento do modelo, salvo cláusula contratual expressa em contrário. Probabilidade: média. Impacto: alto para conteúdo sigiloso ou sujeito à LGPD.

Controles obrigatórios: cláusula contratual explícita proibindo uso dos dados transmitidos para treinamento sem autorização; avaliação de risco antes da integração; localização ou residência de dados quando exigido por política institucional ou regulação setorial.

9.7 Matriz de riscos

ID	Risco	Probabilidade	Impacto	Classif.	Controle principal
R01	Desatualização silenciosa	Alta	Alto	Crítico	Revisão cadenciada
R02	Conteúdo contraditório	Média	Alto	Crítico	SSOT + auditoria
R03	Envenenamento da base	Baixa	Alto	Relevante	Fluxo de aprovação
R04	Exposição de dados pessoais	Média	Alto	Crítico	Checklist LGPD
R05	Deriva taxonômica	Alta	Médio	Relevante	Governança de vocabulário
R06	Perda de propriedade	Alta	Médio	Relevante	Metadado dono obrigatório
R07	Amplificação de erros por IA	Alta	Alto	Crítico	Limiares de qualidade
R08	Vazamento por IA	Média	Alto	Crítico	Filtragem pré-ranqueamento
R09	Aprisionamento de fornecedor	Média	Alto	Relevante	Cláusula de exportação
R10	Indisponibilidade externa	Baixa	Alto	Moderado	Fallback local
R11	Injeção de prompt indireta	Baixa	Alto	Relevante	Validação de fonte + revisão humana
R12	Fraquezas em vetores/embeddings	Baixa	Alto	Relevante	Segregação de índice + filtro pré-recuperação
R13	Bypass de ACL na recuperação semântica	Média	Alto	Crítico	Filtro de permissão em todos os componentes
R14	Uso de dados por modelo externo	Média	Alto	Relevante	Cláusula contratual expressa

10. Modos de falha por componente

Componente	Modo de falha	Sintoma	Mitigação	Responsável
Motor de busca	Ranqueamento incorreto ou desatualizado	Usuários não encontram artigos existentes; abandono alto	Reconfigurar pesos; reindexar acervo	Equipe técnica
Motor de busca	Índice desatualizado	Artigos publicados não aparecem nos resultados	Forçar reindexação; verificar pipeline de ingestão	Equipe técnica

Componente	Modo de falha	Sintoma	Mitigação	Responsável
Camada de conteúdo	Artigos desatualizados	Informações incorretas ou obsoletas entregues ao usuário	Revisão de urgência; marcação de estado; notificação ao dono	KM + Domain Owner
Camada de conteúdo	Conteúdo contraditório	Respostas inconsistentes entre artigos; alucinações em IA	Auditoria imediata; consolidação em SSOT	KM
Vocabulário controlado	Deriva terminológica	Artigos novos não são encontrados por consultas padrão	Atualização do tesauro; reindexação	KM
Metadados	Metadados ausentes ou incorretos	Artigos não recuperados por filtros; falha em controle de acesso	Auditoria de metadados; correção em lote	Editor + KM
Fluxo editorial	Backlog de aprovação acumulado	Conteúdo relevante não publicado; lacunas reportadas por usuários	Revisão do processo; aumento de capacidade de revisores	KM
Controle de acesso	Conteúdo restrito acessível a usuários sem permissão	Reclamação de usuário; incidente de privacidade	Remoção imediata; auditoria de permissões; notificação ao encarregado pelo tratamento de dados pessoais (DPO)	Equipe técnica + encarregado
Integração com ITSM	Falha na captura KCS	Resoluções de chamados não geram artigos; acervo para de crescer	Verificar webhook; reprocessar fila de captura	Equipe técnica
Integração com IA	Vazamento de conteúdo restrito	Resposta de IA contém informação que o usuário não deveria acessar	Desligar integração; auditoria de filtragem; corrigir antes de reativar	Equipe técnica + encarregado
Analytics	Métricas incorretas	Decisões editoriais baseadas em dados errados	Validar coleta de eventos; reconciliar com logs de servidor	Equipe técnica
Plataforma	Indisponibilidade	Base inacessível; sistemas de IA sem contexto	Ativar fallback; comunicar usuários; acionar SLA	Equipe técnica

11. Requisitos mínimos para contratação e aceite

11.1 Requisitos funcionais mínimos

- O sistema deve permitir criação, edição, versionamento e arquivamento de artigos com histórico completo de alterações.
- O sistema deve suportar ao menos quatro tipos de tópico distintos (equivalentes a *concept, task, reference e troubleshooting*) ou permitir a configuração de *templates* equivalentes.

- O sistema deve dispor de mecanismo de busca em texto completo com suporte a sinônimos, tolerância a erros tipográficos e busca facetada por ao menos três dimensões.
- O sistema deve permitir a definição e gestão de vocabulário controlado (taxonomia ou tesauro) integrado ao mecanismo de indexação.
- O sistema deve suportar fluxo de aprovação configurável com ao menos dois estágios (revisão e publicação) e registro de quem aprovou cada versão.
- O sistema deve permitir definição de prazo de revisão por artigo ou por categoria, com notificação automática ao dono quando o prazo é atingido.
- O sistema deve disponibilizar *dashboard* de métricas operacionais com ao menos: volume de artigos por estado, taxa de zero resultados, artigos expirados e avaliações de usuário.
- O sistema deve expor API REST para integração com sistemas externos, com autenticação e controle de acesso por *token*.

11.2 Requisitos não funcionais mínimos

- Disponibilidade: 99,5% em horário comercial para sistemas operacionais; 99,9% para portais públicos de alta demanda.
- Latência de busca: 500ms para o percentil 95 das consultas em condições normais de carga.
- Capacidade: suporte ao volume projetado de artigos com degradação de desempenho inferior a 20% até o dobro do volume inicial.
- Escalabilidade: capacidade de expansão horizontal sem redesenho de arquitetura.
- Backup: cópia de segurança diária com retenção mínima de 30 dias e teste de restauração trimestral documentado.

11.3 Requisitos de segurança e privacidade

- O sistema deve implementar controle de acesso baseado em papel (RBAC) com no mínimo quatro níveis (público, interno, restrito, confidencial).
- O sistema deve garantir que conteúdo restrito não seja retornado em buscas de usuários sem a permissão correspondente, com filtragem aplicada antes do ranqueamento.
- O sistema deve registrar acesso a conteúdo confidencial com identificação do usuário, data, hora e artigo acessado.
- O sistema deve suportar exportação completa do acervo em formato aberto (JSON, XML ou equivalente documentado) sem dependência de licença adicional.
- O sistema deve aplicar criptografia em trânsito (TLS 1.2 mínimo) e em repouso para conteúdo classificado como restrito ou confidencial.
- O fornecedor deve assinar Acordo de Processamento de Dados conforme LGPD quando o sistema processar dados pessoais.

11.4 Teste obrigatório de controle de acesso

Teste obrigatório de controle de acesso: simular solicitação de informação restrita por usuário sem permissão via (a) busca lexical, (b) busca semântica e (c) assistente RAG. O sistema deve negar a recuperação, registrar a tentativa e não expor nenhum trecho do conteúdo restrito em nenhuma das modalidades.

11.5 Requisitos de rastreabilidade e auditoria

- Cada artigo deve ter registro imutável de: autor, aprovador, datas de criação e aprovação de cada versão, e histórico completo de alterações.
- O sistema deve manter versões arquivadas de cada artigo pelo prazo mínimo de 5 anos após o arquivamento.
- O sistema deve gerar relatório de auditoria exportável com todos os eventos de criação, edição, aprovação e arquivamento em um período definido.

11.6 Requisitos de documentação técnica

- O fornecedor deve entregar documentação de arquitetura do sistema com diagrama de componentes, fluxos de dados e integrações.
- O fornecedor deve entregar guia de administração, guia editorial e guia de integração via API, todos em português.
- O fornecedor deve entregar modelo lógico de dados e esquema de metadados suportados pelo sistema.

11.7 Matriz de aceite para contratação

O aceite técnico da solução está condicionado à apresentação das evidências listadas a seguir. A ausência de qualquer evidência obrigatória constitui critério de reprovação.

Requisito	Evidência exigida	Teste de aceite	Responsável	Critério de reprovação
Metadados obrigatórios	Relatório de auditoria de metadados	Verificação automatizada de cobertura de campos	Fornecedor	Cobertura 100%
Controle de acesso pré-recuperação	Log de tentativas negadas	Teste com usuário sem permissão em cada nível de acesso	Fiscal de contrato	Conteúdo restrito retornado a usuário sem permissão

Requisito	Evidência exigida	Teste de aceite	Responsável	Critério de reprovação
Exportação em formato aberto	Arquivo exportado e reimportado em ambiente alternativo	Importação completa e verificada em ambiente isolado	Fiscal de contrato	Falha na importação ou perda de conteúdo
Acessibilidade	Laudo WCAG 2.2 AA (para portais públicos e sistemas acessíveis a servidores com deficiência)	Teste automático + teste manual de navegação por teclado	Equipe de acessibilidade	Não conformidade em critérios WCAG 2.2 AA
Privacidade	Inventário de dados pessoais e análise de necessidade; RIPD quando recomendado pela ANPD	Revisão pelo encarregado pelo tratamento de dados pessoais (DPO)	Encarregado / Fiscal	Dado pessoal sem base legal documentada
Busca	Relatório de testes de recuperação com usuários reais	Taxa de sucesso dentro do limiar por criticidade (Tab. Item 7.6)	Fiscal de contrato	Taxa de sucesso abaixo do limiar
Proveniência (se integrar IA)	Demonstração de rastreabilidade artigo-resposta	Auditoria de qual artigo fundamentou qual resposta	Equipe técnica	Resposta de IA sem citação verificável de fonte
Filtro pré-recuperação (se integrar IA)	Relatório de teste de *bypass* de ACL no pipeline	Tentativa de recuperação *cross-tenant ou cross-*permissão	Equipe técnica	Conteúdo restrito recuperado por usuário sem permissão
Rollback	Plano de *rollback* documentado e testado	Execução do procedimento em ambiente de homologação	Equipe técnica	Procedimento não testado ou não documentado
Plano de sustentação	*Dashboard* de métricas ativo; processo de revisão cadenciada configurado	Verificação de alertas de artigos expirados	Fiscal	*Dashboard* inativo ou sem alertas de expiração

11.8 Requisitos contratuais para bases externas e modelos de terceiros

É vedado ao fornecedor usar conteúdo, *prompts*, logs, documentos, *embeddings* ou interações do órgão para treinamento, ajuste fino ou melhoria de modelos de terceiros, sem autorização expressa do órgão, base legal LGPD documentada e avaliação de risco formalizada.

Quando a base de conhecimento for hospedada por terceiros ou integrada a modelos de linguagem externos, o contrato deve obrigatoriamente incluir:

- Proibição expressa de uso dos dados transmitidos para treinamento ou aprimoramento do modelo sem autorização formal do contratante.

- Localização ou residência de dados no território nacional quando exigido por política institucional, regulação setorial ou classificação do conteúdo.
- Identificação de suboperadores e dos países para os quais os dados podem ser transferidos.
- Garantia de exportação dos dados em formato aberto a qualquer tempo, sem custo adicional e sem dependência de funcionalidade proprietária.
- Reversibilidade: procedimento documentado de migração para outra plataforma com exportação de conteúdo, metadados, taxonomia e trilhas de auditoria.
- SLA de disponibilidade com penalidades proporcionais ao impacto do serviço.
- Obrigação de notificação de incidente de segurança em prazo compatível com o prazo legal de comunicação à ANPD (até 3 dias úteis), com identificação dos dados afetados.
- Acesso a logs de auditoria das operações realizadas pelo fornecedor sobre os dados do contratante.
- Segregação de *tenant*: garantia de que dados do contratante não são acessíveis por outros clientes do fornecedor.
- Exclusão segura e comprovada de todos os dados ao término do contrato, incluindo cópias de segurança e logs residuais.
- Acordo de processamento de dados nos termos do artigo 39 da LGPD quando houver tratamento de dados pessoais.

11.9 Obrigações de sustentação e atualização

O contrato deve prever:

- Suporte técnico com SLA de atendimento por severidade (mínimo: S1 em 4h, S2 em 24h, S3 em 5 dias úteis).
- Atualizações de segurança aplicadas em prazo máximo de 72 horas para vulnerabilidades críticas.
- Notificação antecipada de pelo menos 90 dias para descontinuação de funcionalidades ou versões.
- Transferência de conhecimento à equipe gestora ao final do contrato, incluindo documentação de configurações e customizações.

12. Erros comuns em implantações de bases de conhecimento no setor público

Esta seção descreve os erros mais frequentes observados em implantações públicas de bases de conhecimento. O objetivo é tornar os riscos do anexo anterior tangíveis para equipes técnicas.

- **Ingestão de PDF legado sem curadoria.** Digitalização em massa de documentos sem revisão de estrutura, metadados, dados pessoais ou nível de acesso. O resultado é um acervo volumoso com qualidade inconsistente que degrada a busca e pode expor conteúdo indevido.
- **Indexação de conteúdo sigiloso por engano.** Ausência de verificação de nível de acesso antes da indexação. Documentos restritos ou sigilosos acabam acessíveis a usuários sem permissão.
- **Planilhas soltas como base canônica.** Conhecimento crítico mantido em planilhas compartilhadas em drives sem versionamento, sem dono formal e sem trilha de auditoria. Quando o servidor responsável sai, o conhecimento se perde ou fica desatualizado sem que ninguém perceba.
- **Duplicação de normativos em múltiplos artigos.** A mesma portaria ou instrução normativa é descrita em dezenas de artigos de domínios distintos, gerando versões conflitantes quando o normativo é atualizado.
- **Artigos sem dono.** Conteúdo publicado sem responsável formal de revisão. Após desligamento ou remanejamento do servidor original, o artigo jamais é atualizado.
- **Artigos sem data de validade.** Procedimentos, prazos e contatos que mudam periodicamente publicados sem prazo de revisão. A base acumula informações incorretas sem sinalização ao usuário.
- **Consentimento como base legal padrão no setor público.** O setor público raramente tem base legal em consentimento para tratamento de dados de servidores ou cidadãos em serviços obrigatórios. O consentimento exige liberdade real de recusa, o que geralmente não existe quando o serviço é compulsório.
- **Permitir que o LLM decida o nível de acesso.** Assumir que o modelo de linguagem não revelará conteúdo restrito por instrução no *prompt*, sem filtro técnico pré-recuperação. Modelos podem vazar informações por *jailbreak, prompt injection* ou limitações de instrução.
- **Aceitar resposta de IA sem citação verificável.** Implantar assistentes que respondem sem citar o artigo que fundamenta a resposta. O usuário não pode verificar a fonte e a equipe não pode auditar erros.
- **Medir apenas número de artigos publicados.** Tratar volume de publicação como indicador de sucesso. Bases com milhares de artigos obsoletos têm desempenho inferior a bases menores, bem curadas e atualizadas.
- **Não testar com o servidor que executa a tarefa real.** Validar a busca apenas com a equipe técnica. O vocabulário da equipe técnica raramente coincide com o vocabulário do usuário operacional.
- **Ausência de plano de *rollback* ao conectar IA.** Implantar pipeline RAG em produção sem procedimento testado de desconexão em caso de incidente. Quando o sistema produz respostas incorretas em escala, a equipe não sabe como reverter sem afetar a base.

13. Glossário técnico

13.1 Termos técnicos

Termo	Definição
Article Confidence Score	Índice composto que combina recência, volume de visualizações, avaliação positiva e taxa de abandono para estimar a confiabilidade de um artigo.
BM25	Algoritmo de ranqueamento de relevância em recuperação de texto completo, variante ponderada de TF-IDF que normaliza por comprimento do documento.
Card sorting	Técnica de design de taxonomia em que usuários agrupam itens conceituais para revelar modelos mentais de classificação.
Conteúdo canônico	Versão oficial e autoritativa de um tópico, da qual outras referências derivam sem duplicar. Implementa o princípio SSOT.
Coreference resolution	Processo de resolução de referências pronominais e anafóricas em texto, necessário para que modelos de linguagem mantenham o contexto ao consumir trechos isolados.
Data poisoning	Ingestão deliberada ou acidental de conteúdo incorreto ou tendencioso que degrada a qualidade da base e dos sistemas que a consomem.
Deflection rate	Taxa de sessões na base de conhecimento que resultam em resolução sem abertura de chamado de suporte.
DITA	*Darwin Information Typing Architecture*. Arquitetura de conteúdo estruturado que organiza informação em tópicos reutilizáveis e tipados (*concept, task, reference, troubleshooting*).
Domain Owner	Responsável pela propriedade e qualidade do conteúdo de um domínio temático específico da base.
Dublin Core	Conjunto de 15 elementos de metadados interoperáveis mantido pelo DCMI, base para descrição de recursos digitais com suporte a *Linked* Data.
Findability rate	Taxa de sessões de busca em que o usuário localiza e seleciona um resultado pertinente sem reformular a consulta.
Folksonomia	Sistema de classificação por *tags* geradas livremente pelos usuários, sem hierarquia formal ou controle de vocabulário.
Freshness boost	Ajuste de ranqueamento de busca que favorece artigos mais recentes ou recém-revisados por função de decaimento temporal.
Grounding rate	Proporção de respostas geradas por um sistema de IA que citam ao menos um artigo da base de conhecimento como fonte de fundamentação.
KCS	*Knowledge-Centered Service*. Metodologia operacional que integra captura e manutenção de conhecimento ao fluxo de resolução de demandas de atendimento.
Knowledge graph	Implementação em banco de grafos de uma ontologia formal, representando entidades como nós e relações como arestas rotuladas.

Termo	Definição
KMS	*Knowledge Management System*. Ecossistema de processos, práticas e plataformas que suporta criação, compartilhamento, uso e gestão do conhecimento organizacional. A base de conhecimento é a camada de dados do KMS.
Ontologia	Modelo formal que descreve conceitos, atributos e relações de um domínio com restrições lógicas, permitindo inferência automatizada. Implementada tipicamente em OWL sobre RDF.
OWL	*Web Ontology Language*. Linguagem de ontologia do W3C baseada em lógica de descrição, que permite definir restrições e inferência formal sobre grafos RDF.
PROV-DM	*Provenance Data Model*. Padrão W3C para representação interoperável de proveniência: quem produziu o quê, quando e em que contexto.
RAG	*Retrieval-Augmented Generation*. Arquitetura de IA que recupera trechos relevantes de uma base de conhecimento para fundamentar a geração de respostas por um modelo de linguagem.
RDF	*Resource Description Framework*. Modelo de dados do W3C para representar afirmações como triplas sujeito-predicado-objeto.
RBAC	Role-Based Access Control. Modelo de controle de acesso em que permissões são atribuídas a papéis, e papéis são atribuídos a usuários.
SECI	Modelo de conversão do conhecimento de Nonaka e Takeuchi: Socialização, Externalização, Combinação e Internalização.
SHACL	*Shapes Constraint Language*. Padrão W3C para validação de grafos RDF contra restrições estruturais definidas.
SKOS	*Simple Knowledge Organization System*. Padrão W3C para publicação e mapeamento de vocabulários controlados, taxonomias e tesauros como Linked Data.
SSOT	*Single Source of Truth*. Princípio arquitetural em que cada fato ou instrução reside em um único local canônico, eliminando inconsistências por duplicação.
Snapshot versioning	Estratégia de versionamento que congela e arquiva cada versão de um artefato com identificador cronológico, preservando o histórico sem deletar versões anteriores.
Tesauro	Vocabulário controlado que vai além da taxonomia hierárquica ao mapear relações de equivalência (sinônimos e termos preferidos), hierarquia e associação entre conceitos. Padrão ISO 25964.
Topic Maps	Padrão ISO/IEC 13250 que separa a camada de conhecimento lógico da camada de recursos físicos, permitindo sobrepor uma rede semântica a repositórios não estruturados sem modificar os dados brutos.
Tree testing	Técnica de validação de taxonomia em que usuários navegam por uma estrutura hierárquica para localizar itens específicos, revelando falhas de organização sem o viés visual do design da interface.
Zero-result rate	Taxa de consultas de busca que não retornam nenhum resultado. Indicador primário de lacunas de cobertura ou desalinhamento de vocabulário.

13.2 Siglas

Sigla	Significado
ANPD	Autoridade Nacional de Proteção de Dados
APQC	*American Productivity & Quality Center*
BM25	*Best Matching 25*
CMS	*Content Management System*
DCMI	*Dublin Core Metadata Initiative*
RIPD	Relatório de Impacto à Proteção de Dados Pessoais
DPIA	*Data Protection Impact Assessment (equivalente internacional; não é o termo legal brasileiro)*
DITA	*Darwin Information Typing Architecture*
DMS	*Document Management System*
DPO	*Data Protection Officer*
ECM	*Enterprise Content Management*
eMAG	Modelo de Acessibilidade em Governo Eletrônico
ENAP	Escola Nacional de Administração Pública
ETP	Estudo Técnico Preliminar
GuIA	Guia de Inteligência Artificial para o Setor Público Brasileiro
ISO	*International Organization for Standardization*
ITSM	*Information Technology Service Management*
KB	*Knowledge Base (Base de Conhecimento)*
KCS	*Knowledge-Centered Service*
KM	*Knowledge Management* (Gestão do Conhecimento)
KMS	Knowledge Management System
LAI	Lei de Acesso à Informação (Lei 12.527/2011)
LGPD	Lei Geral de Proteção de Dados Pessoais (Lei 13.709/2018)
LLM	*Large Language Model*
NIST	*National Institute of Standards and Technology*
OASIS	*Organization for the Advancement of Structured Information Standards*
OWASP	*Open Worldwide Application Security Project*
ePING	Padrões de Interoperabilidade de Governo Eletrônico
OWL	*Web Ontology Language*
PII	*Personally Identifiable Information*
PPSI	*Política de Privacidade e Segurança da Informação*
PROV	*Provenance (W3C PROV-DM)*
RAG	*Retrieval-Augmented Generation*

Sigla	Significado
RBAC	*Role-Based Access Control*
RDF	*Resource Description Framework*
SECI	Socialização, Externalização, Combinação, Internalização
SHACL	*Shapes Constraint Language*
SKOS	*Simple Knowledge Organization System*
SLA	*Service Level Agreement*
SME	*Subject Matter Expert*
SSO	*Single Sign-On*
SSOT	*Single Source of Truth*
TCU	Tribunal de Contas da União
TF-IDF	*Term Frequency–Inverse Document Frequency*
TR	Termo de Referência
W3C	*World Wide Web Consortium*
WCAG	*Web Content Accessibility Guidelines*

14. Referências bibliográficas

14.1 Normas e legislação

[1] BRASIL. **Lei no 12.527, de 18 de novembro de 2011**. Lei de Acesso à Informação (LAI). Brasília: Diário Oficial da União, 2011.

[2] BRASIL. **Lei no 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília: Diário Oficial da União, 2018.

[3] BRASIL. **Lei no 15.263, de 14 de novembro de 2025**. Institui a Política Nacional de Linguagem Simples. Brasília: Diário Oficial da União, 2025.

[4] AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Resolução CD/ANPD no 15, de 24 de abril de 2024, publicada no Diário Oficial da União em 26 de abril de 2024**. Regulamento de Comunicação de Incidente de Segurança com Dados Pessoais. Brasília: ANPD, 2024.

[5] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 30401:2018** – Knowledge management systems – Requirements. Geneva: ISO, 2018.

[6] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 25964-1:2011** – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval. Geneva: ISO, 2011.

[7] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION; INTERNATIONAL ELECTROTECHNICAL COMMISSION. **ISO/IEC 27001:2022** – Information security, cybersecurity and privacy protection – Information security management systems – Requirements. Geneva: ISO/IEC, 2022.

[8] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION; INTERNATIONAL ELECTROTECHNICAL COMMISSION. **ISO/IEC 42001:2023** – Artificial intelligence – Management system. Geneva: ISO/IEC, 2023.

[9] ISO/IEC. **ISO/IEC 13250:2003** – Topic Maps. Geneva: ISO, 2003.

14.2 Guias governamentais

[10] AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Guia Orientativo sobre Tratamento de Dados Pessoais pelo Poder Público**. Brasília: ANPD, 2022.

[11] AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Guia Orientativo: Legítimo Interesse**. Brasília: ANPD, 2022.

[12] AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Guia Orientativo: Relatório de Impacto à Proteção de Dados Pessoais (RIPD)**. Brasília: ANPD, 2021.

[13] ESCOLA NACIONAL DE ADMINISTRAÇÃO PÚBLICA. **Gestão do Conhecimento: Guia para Implementação**. Brasília: ENAP, 2023.

[14] GOVERNO FEDERAL. **Guia de Requisitos Mínimos de Privacidade e Segurança para Aplicações Web**. Brasília: SGD/ME, 2022.

[15] GOVERNO FEDERAL. **Manual de Linguagem Simples**. Brasília, 2021.

[16] GOVERNO FEDERAL. **Padrões de Interoperabilidade de Governo Eletrônico (ePING)**. Brasília: SGD/ME, versão vigente.

[17] GOVERNO FEDERAL. **Instrução Normativa SGD/ME no 94, de 23 de dezembro de 2022**. Dispõe sobre o processo de contratação de soluções de Tecnologia da Informação e Comunicação pelos órgãos integrantes do SISP. Brasília: Diário Oficial da União, 2022.

[18] GOVERNO FEDERAL. **Política de Privacidade e Segurança da Informação 2.0 instituída pela Portaria SGD/MGI nº 9.511/2025, com vigência a partir de 1º de janeiro de 2026 (Guia do Framework PPSI 2.0)**. Brasília: SGD/MGI, 2025.

[19] GOVERNO FEDERAL. **Modelo de Acessibilidade em Governo Eletrônico (eMAG)**. Brasília: SGD/ME, versão vigente.

[20] TRIBUNAL DE CONTAS DA UNIÃO. **Referencial de Gestão do Conhecimento para o Setor Público**. Brasília: TCU, 2022.

14.3 Frameworks de risco e segurança

[21] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. **AI Risk Management Framework (AI RMF 1.0)**. Gaithersburg: NIST, 2023. NIST AI 100-1.

[22] NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY. **Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile**. Gaithersburg: NIST, 2024. NIST AI 600-1.

[23] OWASP. **OWASP Top 10 for Large Language Model Applications 2025**. OWASP Foundation, 2025. Disponível em: https://owasp.org/www-project-top-10-for-large-language-model-applications/.

14.4 Padrões técnicos W3C e OASIS

[24] W3C. **SKOS Simple Knowledge Organization System Reference**. World Wide Web Consortium, 2009.

[25] W3C. **PROV-DM: The PROV Data Model**. World Wide Web Consortium, 2013.

[26] W3C. **Shapes Constraint Language (SHACL)**. World Wide Web Consortium, 2017.

[27] W3C. **Web Content Accessibility Guidelines (WCAG) 2.2**. World Wide Web Consortium, 2023.

[28] W3C. **PROV Overview**. World Wide Web Consortium, 2013. (*Inclui PROV-DM e PROV-N.*)

[29] W3C. **RDF 1.1 Concepts and Abstract Syntax**. World Wide Web Consortium, 2014.

[30] W3C. **OWL 2 Web Ontology Language Primer**. World Wide Web Consortium, 2012.

[31] OASIS. **Darwin Information Typing Architecture (DITA) 1.3**. OASIS Standard, 2015.

[32] DUBLIN CORE METADATA INITIATIVE. **DCMI Metadata Terms**. DCMI, 2020.

14.5 Referências técnicas e acadêmicas

[33] ALAVI, M.; LEIDNER, D. E. Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. **MIS Quarterly**, v. 25, n. 1, p. 107–136, 2001.

[34] APQC. **Knowledge Management Priorities and Trends Survey**. Houston: APQC, 2026.

[35] APQC. **Knowledge Management Measurement Framework**. Houston: APQC, 2023.

[36] CONSORTIUM FOR SERVICE INNOVATION. **Knowledge-Centered Service v6**. Lafayette: CSI, 2019.

[37] HEDDEN, H. Taxonomies and controlled vocabularies: Best practices for metadata. **Journal of Digital Asset Management**, v. 6, n. 5, p. 279–284, 2010.

[38] KRUSCHWITZ, U.; HULL, C. **Searching the Enterprise**. Synthesis Lectures on Information Concepts, Retrieval, and Services. San Rafael: Morgan & Claypool, 2017.

[39] NONAKA, I.; TAKEUCHI, H. **The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation**. New York: Oxford University Press, 1995.

[40] *ZENG*, M. L.; DEXTRE CLARKE, S. G. The evolution of thesauri and information systems. **Cataloging & Classification Quarterly**, v. 56, n. 5–6, p. 374–395, 2018.

14.6 Referências bibliográficas - metodologia

[41] CPQD. Guia unificado de Inteligência Artificial: Projeto INSPIRE – Meta 3.2. Metodologia de desenvolvimento de soluções de IA. FINEP. 2026.

[42] CPQD. Anexo: glossário unificado. Projeto INSPIRE – Meta 3.2: Metodologia de desenvolvimento de soluções de IA. FINEP.2026.