

# Projeto ATHENAS: Proposta de Projeto para Inteligência de Acesso ao Conhecimento

**Nome do projeto:** ATHENAS

- **Autores / Função:** Amanda Zanata (Tech Solutions Engineer); Washington Gama (Tech Solutions Engineer)
- **Versão:** v1.0

## 1. Problema & Oportunidade

**Contexto atual:** Em toda a LCGBR, assim como em muitas organizações globais, os colaboradores do conhecimento (“knowledge workers”) enfrentam um desafio diário oneroso: a busca por informações internas. Dados críticos, documentos e insights estão fragmentados em um ecossistema digital complexo e crescente, incluindo SharePoint, Confluence, drives de rede, wikis internas e outras plataformas. O processo atual de acesso a essas informações é manual, ineficiente e depende de buscas por palavras-chave sem compreensão de contexto. Em suma, é difícil encontrar a informação certa na hora certa. Estudos de mercado pintam um quadro alarmante desta realidade: um relatório da McKinsey revela que funcionários gastam, em média, 1.8 a 3.6 horas por dia, cerca de 9.3 horas por semana, procurando e compilando informações. Isso representa 25 a 30% de um dia de trabalho; em outras palavras, para cada quatro funcionários focados em suas tarefas principais, um quinto colaborador está efetivamente ocupado em uma “caça ao tesouro” informacional, sem gerar valor direto ao negócio. Este cenário não é um mero inconveniente operacional, mas sim um gargalo sistêmico que freia a produtividade e a agilidade da organização.

**Impacto negativo:** A ineficiência no acesso ao conhecimento traz consequências profundas e multifacetadas para a LCGBR:

- **Custo Financeiro Direto:** O tempo perdido em buscas infrutíferas traduz-se em prejuízo financeiro substancial. A empresa de pesquisa IDC quantificou esse impacto, estimando que a perda de produtividade devido a problemas na gestão de documentos custa às empresas cerca de \$19.732 USD por trabalhador do conhecimento por ano. Para uma organização com 1.000 funcionários nessa categoria, essa ineficiência equivale a quase \$2,5 milhões USD por ano em valor desperdiçado (cálculo direto). Trata-se de um custo oculto, porém real, que corrói a margem de lucro.
- **Dreno de Produtividade e Inovação:** Tempo gasto procurando informações é tempo não investido em tarefas de alto valor, pensamento estratégico ou inovação. Essa fricção constante leva muitas vezes à duplicação de esforços (funcionários recriam documentos ou análises porque não sabem que já

existem em algum silo) e sufoca a criatividade. Em suma, a LCGBR perde agilidade de resposta ao mercado e oportunidades de inovação ficam comprometidas.

- **Moral e Retenção de Talentos:** A frustração diária de não encontrar as informações necessárias gera estresse, ansiedade e burnout. Pesquisas da APQC indicam que funcionários em ambientes assim têm satisfação significativamente menor e são 44% mais propensos a deixar a empresa. Num mercado de trabalho competitivo, a falta de ferramentas adequadas se torna um grave risco de perda de talentos.
- **Risco Organizacional (Perda de Conhecimento):** Quando funcionários experientes saem, seu conhecimento tácito, o “como” e o “porquê” por trás de decisões e processos, muitas vezes vai embora com eles. Essa perda de “memória organizacional” é um risco estratégico, pois torna o onboarding de novos colaboradores mais lento e aumenta a chance de erros repetidos. Sem mecanismos eficazes de captura e disseminação do conhecimento, a empresa pode involuntariamente “reinventar a roda” com frequência.

É importante notar que a natureza desse problema é auto-perpetuante. A cada novo sistema ou ferramenta introduzida (CRM, plataforma de projeto, etc.), cria-se um novo silo de dados, multiplicando a complexidade de se encontrar informação.

**Oportunidade:** O projeto ATHENAS propõe quebrar esse ciclo, atuando como uma “camada de inteligência unificadora” que se torna mais valiosa conforme o ecossistema tecnológico cresce. Ou seja, à medida que adicionamos mais sistemas e fontes de dados, a ATHENAS amplia seu conhecimento e utilidade, posicionando-se como um investimento estratégico de longo prazo para o futuro da LCGBR.

## 2. Objetivo de Negócio

Para guiar o sucesso do projeto ATHENAS, foi definida uma métrica principal que mede diretamente o impacto no negócio e facilita a tradução em retorno sobre investimento (ROI):

- **Métrica Selecionada:** Redução do Tempo Médio Gasto na Recuperação de Informações ou Tempo de Busca (TIR, *Time to Information Retrieval*).
- **Baseline Atual:** Estudos indicam ~9,3 horas por semana gastas buscando dados por trabalhador do conhecimento. Tomaremos esse valor (aproximadamente 9,3 h/semana) como referência inicial, refinando-o com uma pesquisa interna antes do lançamento do MVP para estabelecer a linha de base específica da LCGBR.
- **Meta (90 dias de MVP):** Reduzir em 20% o TIR do grupo piloto (Vendas & Marketing). Isto equivale a baixar de ~9,3h para ~7,4 horas/semana,

recuperando cerca de 2 horas de tempo produtivo por colaborador por semana.

Essa meta de 20% não é apenas um ganho de eficiência abstrato; é uma economia de custo concreta e calculável. Multiplicando as horas economizadas pelo custo médio da hora de trabalho, o valor gerado pelo projeto torna-se claro para a liderança financeira. Em outras palavras, cada ponto percentual de redução no TIR traduz-se em reais poupados, o que torna o caso de negócio da ATHENAS extremamente sólido. A escolha desta *North Star Metric* foi deliberada justamente para conectar o projeto de IA a resultados de negócio irrefutáveis, ao apresentarmos que “a ATHENAS devolve X horas para a empresa, equivalentes a Y reais”, a proposta deixa de ser apenas tecnologia e passa a ser um investimento com ROI tangível.

### 3. Proposta de Solução

**Conceito em 1 linha:** A ATHENAS é um assistente IA conversacional de nível empresarial que fornece respostas instantâneas, sintetizadas e com fontes verificáveis para perguntas dos colaboradores, utilizando todo o acervo de conhecimento interno da LCGBR. Em vez de fazer o usuário garimpar documentos e páginas, a ATHENAS traz a informação pronta, de forma confiável, em linguagem natural.

**Como a IA é aplicada:** A solução ATHENAS é construída sobre a arquitetura de ponta conhecida como RAG (*Retrieval-Augmented Generation*), que combina o poder de *Large Language Models* (LLMs) com a precisão dos dados internos da empresa. O funcionamento pode ser resumido em três etapas principais:

- **Recuperação (Retrieval):** Quando o usuário faz uma pergunta na interface de chat (por ex: “*Por onde começar a investigar vazamentos de audiência no Adobe Target?*”), o sistema primeiro traduz a pergunta para uma representação vetorial (*embedding*) e realiza uma busca semântica num banco de dados vetorial contendo o conhecimento interno (documentos do SharePoint, páginas do Confluence, etc.). Essa busca retorna os trechos (*chunks*) de informação mais relevantes para a pergunta, em vez de documentos inteiros. (*Em termos simples, a ATHENAS “lê” milhares de páginas em segundos e pinça apenas os pedaços relevantes.*)
- **Aumento de Contexto (Augmentation):** Os trechos recuperados são então combinados com a pergunta original para formar um *prompt* contextualizado. Esse *prompt* enriquecido, contendo a pergunta do usuário + conteúdo relevante encontrado, é enviado a um LLM de última geração (no MVP, o GPT-4 da OpenAI). Assim, garantimos que o modelo tenha toda a base factual necessária antes de gerar qualquer resposta.
- **Geração da Resposta (Generation):** O LLM, instruído a basear-se exclusivamente no contexto fornecido, gera uma resposta concisa em

linguagem natural. Crucialmente, a ATHENAS também fornece citações e links diretos para as fontes originais de onde extraiu a informação. Esse mecanismo de *grounding* (fundamentação) reduz drasticamente o risco de alucinações ou erros, um ponto crítico para adoção de IA em ambiente corporativo, onde a confiança na informação é primordial.

## 4. Diferencial / Originalidade

A ATHENAS representa um salto qualitativo em relação às ferramentas de busca e gestão de conhecimento tradicionais, oferecendo diversos diferenciais únicos:

- **Acesso Unificado vs. Buscas Isoladas:** Diferente de buscas *siloe*d (em que o SharePoint busca apenas no SharePoint, Notion apenas no Notion, o Confluence apenas no Confluence, etc.), a ATHENAS atua como uma camada unificada de inteligência sobre todas as fontes de dados. Uma única interface conversacional consulta tudo que a empresa sabe, independente de onde a informação reside. Isso elimina as barreiras entre repositórios e acaba com a necessidade de o colaborador repetir a mesma busca em múltiplos sistemas.
- **Respostas Sintetizadas vs. Listas de Links:** Enquanto uma busca tradicional devolve uma lista de documentos, transferindo para o usuário o trabalho de abrir cada item e achar manualmente a resposta, a ATHENAS entrega uma resposta direta já sintetizada a partir das fontes relevantes. Ou seja, ela não é apenas um mecanismo de busca, mas sim um mecanismo de respostas e descoberta de conhecimento. Isso economiza tempo e reduz a carga cognitiva do usuário.
- **Confiança e Verificabilidade:** Em ambiente empresarial, confiar na informação é fundamental. Cada resposta da ATHENAS vem acompanhada de citações das fontes originais, com um clique direto para o documento fonte. Isso permite aos usuários verificar instantaneamente a precisão e contexto da resposta, fomentando confiança no sistema e garantindo *accountability*. Essa transparência embutida é essencial para adoção em larga escala de qualquer solução de IA corporativa.
- **Proatividade e Alertas Inteligentes:** Ao contrário de ferramentas passivas de busca, a ATHENAS poderá atuar de forma proativa. Ela não apenas responde perguntas, mas também poderá monitorar padrões de uso e dados para detectar anomalias ou tendências relevantes. Por exemplo, se houver um volume anormal de consultas sobre um certo produto ou política (indicando possivelmente um problema emergente ou informação desatualizada), a ATHENAS pode gerar alertas inteligentes para as equipes responsáveis. Da mesma forma, ao integrar dados de negócios, a ATHENAS poderia notificar gerentes sobre indicadores fora do padrão (ex.: "*Investigações no GA4 - anomalia detectada*"), antecipando discussões e ações corretivas antes

mesmo que alguém pergunte. Essa inteligência proativa transforma a ATHENAS em um radar estratégico para o negócio.

- **Geração Automática de Conteúdo:** Além de encontrar e compilar informações, a ATHENAS terá a capacidade de gerar documentos e apresentações automaticamente. Com base em solicitações do usuário, o assistente poderá, por exemplo, montar um resumo executivo em Word ou preparar slides PowerPoint (ou um documento Loop) com os principais insights sobre um determinado assunto. Essa funcionalidade, possível graças aos LLMs, agiliza a criação de materiais de suporte (propostas comerciais, relatórios, etc.) mantendo consistência e qualidade. Em vez de gastar horas coletando dados e formatando uma apresentação, o colaborador pode pedir à ATHENAS: *“Faça uma apresentação de 5 slides com os destaques do projeto do cliente X”*, e obter um *draft* inicial pronto para refinamento.
- **Análise Multimodal (Texto + Imagem):** A ATHENAS não se limitará a texto, ela será capaz de “ler” e interpretar imagens e gráficos também. Por exemplo, poderá extrair dados de um gráfico em um PDF, entender um diagrama de arquitetura ou organograma, ou resumir o conteúdo de um infográfico. Isso será viabilizado integrando tecnologias de OCR (reconhecimento ótico de caracteres) e modelos multimodais de IA. Com isso, informações atualmente presas em imagens ou documentos escaneados passarão a ser pesquisáveis e utilizáveis pelo assistente, ampliando significativamente o alcance do conhecimento acessível.
- **Integração com Ferramentas de Marketing (Adobe):** Uma grande originalidade da ATHENAS será conectar-se a fontes vivas de dados, não apenas documentos estáticos. Planejamos integrar a ATHENAS com plataformas como Adobe Analytics e Adobe Target, utilizadas pelo Marketing, para trazer insights de campanhas e testes A/B direto para a conversa. Por exemplo, um BUD poderá perguntar: *“ATHENAS, como está a taxa de conversão da campanha Y esta semana e como se compara ao teste A/B anterior?”*, e o assistente recuperará os números do Adobe Analytics em tempo real, fornecendo a resposta com contexto. Essa integração de dados de analytics permite que decisões de marketing sejam tomadas mais rapidamente, sem precisar navegar por painéis separados, a ATHENAS torna-se um console unificado de inteligência.
- **Funcionalidade Agentic (Execução de Ações):** Mais que um FAQ avançado, a ATHENAS evoluirá para um agente que executa tarefas em nome do usuário. Isso significa que, integrada a sistemas corporativos, ela poderá acionar comandos além de responder. Por exemplo, um vendedor poderá dizer: *“ATHENAS, crie um rascunho de e-mail para o cliente XX destacando nossos 3 principais argumentos de venda do produto X”*, e a ATHENAS gerará o e-mail no Outlook com aquelas informações (pronto para revisão e envio). Ou um

analista de suporte pode pedir: “*Abra um ticket no ClickUp com a descrição desse problema*”, e a ATHENAS preencherá os campos no ClickUp automaticamente. Até mesmo gerar documentos completos entra aqui - “*ATHENAS, reúna os aprendizados do último projeto e crie um documento de lições aprendidas*”. Essa capacidade *agentic* posiciona a ATHENAS não só como um assistente passivo de consulta, mas como um parceiro ativo de produtividade, economizando tempo em tarefas manuais.

- **Perfis Inteligentes por Tipo de Usuário:** A ATHENAS poderá adaptar seu comportamento conforme o perfil do usuário. Diferentes funções na empresa têm necessidades distintas, o que um gerente financeiro espera de uma resposta pode diferir de um engenheiro de software. Planejamos incorporar personalização por perfil, onde a ATHENAS leva em conta o departamento, cargo ou preferências do usuário para priorizar informações mais relevantes para ele. Por exemplo, ao perguntar “*Como foi o desempenho do produto X?*”, um usuário de Vendas pode receber a resposta focada em métricas de vendas e feedback de clientes, enquanto um usuário de Tech receberia detalhes mais técnicos ou de qualidade. Essa inteligência de contexto por usuário aumenta a relevância das respostas e a satisfação, pois o assistente “entende” quem está perguntando. (T tecnicamente, isso pode envolver *fine-tuning* adicional ou filtros por persona.)
- **Governança e Explicabilidade:** Para adoção corporativa, não basta que a IA seja útil, ela precisa ser governável e explicável. A ATHENAS incluirá um painel de governança que dá visibilidade e controle à administração de TI e *compliance*. Nele, será possível monitorar em tempo real o uso do sistema (quais departamentos mais perguntam o quê), o desempenho (tempos de resposta, *uptime*), os custos de API e métricas de qualidade. Além disso, teremos recursos de explicabilidade, por exemplo: logs detalhados de cada interação, quais fontes foram usadas em cada resposta, e até possibilidade de inspecionar por que o modelo chegou a certa resposta (via visualização dos *embeddings* relevantes, etc.). Isso garante que a ATHENAS opere de forma transparente, segura e alinhada às políticas da empresa. Qualquer comportamento anômalo pode ser detectado e auditado. Essa governança robusta é um diferencial importante em comparação a soluções *black-box* no mercado.
- **Gamificação & Engajamento Interno:** Para impulsionar a adoção da ATHENAS dentro da empresa, pretendemos implementar elementos de gamificação. A ideia é tornar o uso da ATHENAS algo recompensador e até competitivo de forma saudável. Por exemplo, poderemos ter um dashboard interno de engajamento mostrando quais equipes ou usuários mais interagiram com o assistente (quantidade de perguntas, feedbacks positivos, etc.), instituir pequenos reconhecimentos (badges virtuais, menções em comunicações internas) para *power users*, ou até competições mensais do

tipo “Qual time extraiu mais valor da ATHENAS?”. Essa abordagem lúdica incentiva os colaboradores a incorporarem a ATHENAS no dia a dia, ao mesmo tempo em que fornece à gestão uma métrica visível de adoção por área. Em resumo, ao “jogar o jogo do conhecimento” dentro da empresa, cultivamos uma cultura de busca ativa de informações e melhoria contínua.

**User Flow resumido (opcional):** O fluxo de interação do usuário com o ATHENAS segue as etapas da arquitetura RAG descritas acima, garantindo uma experiência intuitiva e eficiente. Em alto nível, o processo ponta-a-ponta é:

1. O usuário faz sua pergunta através de uma interface de chat simples (web ou aplicativo).
2. A ATHENAS processa essa pergunta, realiza a busca dos conteúdos relevantes no repositório de conhecimento interno, e envia a pergunta + contexto para o LLM.
3. O LLM retorna a resposta em linguagem natural, que a ATHENAS exibe para o usuário com referências.
4. O usuário pode então fornecer feedback (por exemplo, marcando se a resposta foi útil ou não), permitindo aprimoramentos contínuos do sistema.

*(Um diagrama visual poderia ilustrar: Usuário → [Interface Chat] → [Módulo de Busca] → [LLM] → Resposta → Usuário. O feedback do usuário fecha o ciclo, sinalizando ao módulo de melhora contínua do ATHENAS.)*

## 5. Fontes de Dados & Privacidade

A gestão de dados e privacidade são pilares fundamentais do projeto ATHENAS. Para o MVP, focaremos em fontes de dados de alta relevância para o grupo piloto, com políticas rígidas para garantir conformidade à LGPD e às normas internas da LCGBR.

**Tabela de Fontes de Dados (MVP):**

Fonte	Tipo de Conteúdo	Status de Acesso	Privacidade (LGPD)
SharePoint (Sites de Vendas & Marketing)	Docs corporativos (PDF, DOCX, PPTX)	OK para uso	Contém PII (dados pessoais identificáveis)
Confluence (Espaços de Projetos)	Páginas wiki, anexos	A validar	Potencialmente contém PII
Drive de Rede Compartilhado (Documentos legados)	Vários formatos (DOC, XLS, etc.)	A validar	Contém PII

Como mostra a tabela, é inevitável que nas fontes designadas haja dados pessoais e sensíveis dispersos. Assim, um ponto de atenção crítico é o envio de conteúdo para serviços externos de IA (como a API do LLM). Dados pessoais como nomes de

clientes, informações de funcionários, números identificadores etc., não podem vazarem para fora do ambiente seguro da empresa. Qualquer solução de IA que não aborde explicitamente essa questão estaria incompleta e inviável em nosso contexto corporativo.

**Política de envio para LLMs: Mascaramento de PII (anonimização):** Para mitigar riscos, implementaremos um passo obrigatório de pré-processamento em toda informação antes de enviá-la a qualquer API externa. Em específico, usaremos a biblioteca *Microsoft Presidio*, uma ferramenta *open-source* de referência para detecção e anonimização de PII. O Presidio combina técnicas de reconhecimento de entidades (NER), padrões (regex) e regras para identificar com alta precisão elementos como nomes, emails, telefones, CPFs, etc., substituindo-os por placeholders genéricos ou tokens anonimizados. O processo será assim:

- **Detecção:** Antes de gerar embeddings de um documento, ele passa pelo motor do Presidio para marcar todas as ocorrências de PII.
- **Anonimização:** Cada dado sensível identificado é substituído por um marcador (exemplo: [NOME], [EMAIL], [TEL]). Apenas o texto anonimizado segue adiante.
- **Uso Seguro:** Os embeddings e as consultas enviadas ao LLM serão baseados nesse texto anonimizado. Assim, nenhuma informação pessoal sai do ambiente da LCGBR, mesmo que o LLM processe o conteúdo, ele verá apenas tokens genéricos.
- **Mapeamento Interno (opcional):** Manteremos internamente um mapeamento criptografado dos placeholders para os dados reais, possibilitando que usuários autorizados vejam a informação original se necessário. Dessa forma equilibramos privacidade e utilidade: a ATHENAS pode responder, e se a resposta contiver um [NOME], um gerente autorizado saberá a quem se refere por meio do mapa seguro (apenas se for preciso).

Essa abordagem garante conformidade total com LGPD e políticas internas, sem sacrificar a funcionalidade. É uma medida não-negociável: somente com anonimização robusta podemos usar um LLM terceiro de forma segura. Além de proteger a empresa, isso demonstra maturidade técnica e compromisso com segurança de dados, aumentando a confiança dos stakeholders no projeto.

*(Obs: Para fontes futuras como bases de RH ou CRM, políticas semelhantes ou ainda mais rigorosas serão aplicadas, possivelmente excluindo certas informações de uso na ATHENAS.)*

## 6. Arquitetura Técnica (MVP)

A arquitetura do MVP do ATHENAS foi desenhada para ser robusta, escalável e, acima de tudo, viabilizar um desenvolvimento rápido com validação efetiva da



proposta de valor. Ela segue estritamente o fluxo RAG, utilizando componentes modernos e comprovados no ecossistema de IA:

- **Ingestão de Dados:** Será orquestrada por scripts Python usando *Document Loaders* do framework LangChain. Em particular, utilizaremos o *SharePointLoader* do LangChain para conectar nativamente aos repositórios de Vendas e Marketing. Durante a ingestão, cada documento passará pelo pipeline de anonimização (Presidio) já descrito, garantindo que só dados seguros sejam indexados.
- **Orquestração (LangChain):** O LangChain funcionará como espinha dorsal da solução, gerenciando toda a pipeline RAG. Ele cuidará de operações como:
  - *Chunking*: Dividir documentos em partes menores e semanticamente consistentes (usaremos *RecursiveCharacterTextSplitter* para manter contexto).
  - *Gestão de Prompt*: Montar dinamicamente o *prompt* a ser enviado ao LLM, combinando a pergunta do usuário com os conteúdos recuperados.
  - *Chamadas de Componentes*: Orquestrar as chamadas para o modelo de embedding, para o banco vetorial e para o LLM conforme o fluxo.
- **Modelo de Embeddings:** Para transformar textos em vetores semânticos usaremos, no MVP, o modelo *text-embedding-ada-002* da OpenAI. Esse modelo oferece excelente equilíbrio entre performance, custo e maturidade, sendo uma escolha padrão da indústria para iniciar projetos RAG. (No futuro, podemos migrar para um modelo proprietário ou *fine-tunado*, mas para começar, o Ada-002 garante velocidade e qualidade suficientes).
- **Banco de Dados Vetorial:** O repositório de embeddings e seus metadados será o *ChromaDB* (open-source). O Chroma foi escolhido por ser facilmente *deployável* (em memória ou disco) sem configurações complexas, ideal para prototipação rápida. Ele elimina a necessidade de infraestrutura pesada de banco na fase inicial, podemos rodá-lo em container junto com a aplicação, o que nos permite focar na lógica principal em vez de manutenção de BD.
- **LLM (Large Language Model):** O motor gerador de respostas será o GPT-4 (versão otimizada empresarial). As capacidades avançadas do GPT-4 em compreensão de contexto, raciocínio e seguimento de instruções são críticas para produzir respostas de alta qualidade e fiéis às fontes. Embora existam modelos *open-source*, a opção pela OpenAI se deve à sua comprovada eficácia e apoio técnico, importantes para o sucesso do MVP. (Obs: Avaliaremos constantemente custo-benefício; se o uso crescer muito, poderemos considerar modelos dedicados.)

- **Frontend (Interface do Usuário):** Para o MVP adotaremos uma interface web simples de chat construída com *Streamlit*, uma biblioteca Python que permite criar UIs de forma ágil, sem exigir desenvolvimento frontend complexo. O *Streamlit* nos possibilita entregar rapidamente uma experiência interativa (campo de chat, botão de envio, histórico de conversa, etc.) e iterar conforme o feedback do usuário. Embora não tenha o acabamento de um app web profissional, ele é suficiente para validar o conceito. Em fases posteriores, poderemos migrar para um *framework* web completo (React, Angular) conforme necessidade de customização.
- **Backend (API):** O cérebro da ATHENAS rodará em um backend Python. Usaremos FastAPI para expor endpoints (ex: um endpoint /chat que recebe pergunta e retorna resposta). O LangChain operará nesse backend. A separação backend/frontend traz flexibilidade, poderíamos no futuro integrar o backend da ATHENAS a outras interfaces (por ex, Microsoft Teams, chat do SharePoint, assistentes de voz, etc.) apenas desenvolvendo conectores, sem reescrever a lógica de Q&A.
- **Deployment:** Toda a aplicação (backend FastAPI + LangChain, banco vetorial ChromaDB, e UI Streamlit) será containerizada via Docker. Para o MVP, esses containers serão implantados em uma Máquina Virtual (VM) no Azure, dentro da subscrição da LCGBR. Essa abordagem proporciona um ambiente controlado, seguro e escalável. Podemos escolher uma VM de médio porte (e.g. Standard B2ms, 8GB RAM) para começar. A containerização facilita futura migração para orquestração (Kubernetes) caso o projeto escale significativamente.
- **Repositório de Código:** Todo o código-fonte será armazenado em repositório Git interno. Planejamos criar (após aprovação) um repositório no GitHub Enterprise da LCGBR, por exemplo: *github.com/LCGBR/athenas-knowledge-assistant*, assegurando versionamento e colaboração. (Políticas de acesso serão definidas, inicialmente somente equipe do projeto.)

*(Diagrama de arquitetura poderia ser apresentado aqui, ilustrando os componentes e fluxo de dados.)*

## 7. Escopo do MVP (90 dias)

O escopo do MVP (90 dias) está focado em entregar uma ferramenta funcional e de alto valor para o grupo piloto (Vendas & Marketing), de forma rápida e iterativa. O trabalho será organizado em 6 sprints quinzenais (2 semanas cada). Abaixo, destacamos as épicas e principais funcionalidades previstas, com priorização e alocação por sprint:

- **Épica 1: Ingestão & Setup**
- Configuração do ambiente Azure e repositório Git (Pri 1, Sprint S1)

- Desenvolvimento de script conector para SharePoint (Pri 1, Sprint S1–S2)
- Implementação da anonimização de PII (Presidio) (Pri 1, Sprint S2)
- Implementação da pipeline de *chunking* e *embeddings* (Pri 1, Sprint S2)
- **Épica 2: Core API**
  - Construção da cadeia RAG usando LangChain (Pri 1, Sprint S3)
  - Desenvolvimento do endpoint /chat (FastAPI) (Pri 1, Sprint S3–S4)
  - Integração com ChromaDB e API da OpenAI (Pri 1, Sprint S4)
- **Épica 3: Frontend**
  - Desenvolvimento da UI de chat (Streamlit) (Pri 2, Sprint S4–S5)
  - Implementação do mecanismo de feedback (👍 / 👎) (Pri 2, Sprint S5)
  - Exibição das citações de origem nas respostas (Pri 2, Sprint S5)
- **Épica 4: Métricas & Testes**
  - Criação do "Golden Dataset" (50 pares de Q&A) (Pri 1, Sprint S6)
  - *Logging* básico de perguntas e respostas (Pri 2, Sprint S6)
  - Dashboard inicial de uso e feedback (PowerBI) (Pri 2, Sprint S6)

*Legenda:* **Pri** = Prioridade (1 = essencial, 2 = desejável); **Sprint** S1–S6 corresponde às quinzenas durante os ~90 dias.

Como visto acima, as funcionalidades essenciais (prioridade 1) concentram-se nos primeiros sprints, garantindo que até a metade do ciclo já teremos um núcleo utilizável (ingestão de dados completada, backend RAG funcional, e uma interface mínima). As funcionalidades complementares (prioridade 2), como melhorias de frontend e métricas, entram na segunda metade conforme o produto se estabiliza.

Essa cadência permite demonstrar valor incremental rapidamente: por exemplo, ao final do Sprint 3 já poderemos mostrar internamente um protótipo funcional de Q&A para alguns usuários-chave e coletar feedback precoce, mesmo que sem todas as comodidades de UI.

*(Um quadro Kanban detalhado será mantido no Jira para acompanhamento das tarefas dentro de cada sprint.)*

## 8. Métricas de Sucesso / KPIs

O sucesso do ATHENAS será avaliado através de uma combinação de métricas de impacto no negócio, satisfação do usuário e performance técnica do sistema. Essa visão multidimensional garante que prestaremos contas a todos os stakeholders do projeto, desde executivos até a equipe técnica. As principais KPIs planejadas são:

Métrica	Baseline	Meta (90 dias)	Como medir
Tempo Médio de Busca (TIR)	~9.3 h/semana	7.4 h/semana (redução de 20%)	Inquérito quinzenal com grupo piloto
Satisfação do Usuário (CSAT)	N/A (novo)	> 4 de 5	Feedback na UI (👍 / 👎) + pesquisa pós-uso
Taxa de Adoção (uso ativo)	0% (antes do MVP)	≥ 75% do piloto usando ≥3x/semana	Monitoramento de logins e uso (backend)
Relevância das Respostas (Retrieval)	N/A	> 85% relevantes	Avaliação humana/LLM em conjunto de teste ( <i>Golden Dataset</i> )
Fidelidade das Respostas (Generation)	N/A	> 90% fiéis às fontes	Avaliação humana/LLM (métrica de <i>faithfulness</i> )

#### Observações:

- **TIR (Tempo de Busca):** Métrica principal de produtividade. Será medida via enquete rápida junto aos usuários piloto a cada duas semanas, perguntando quanto tempo estimam ter gasto buscando informação naquela quinzena, comparado a antes do ATHENAS. A redução de ~2h/semana por pessoa gera um *business case* de ROI forte (como mencionado).
- **CSAT:** Mediremos satisfação tanto implicitamente (pelos cliques de 👍 ou 👎 nas respostas fornecidas pelo ATHENAS na interface) quanto explicitamente (através de uma breve pesquisa de satisfação ao final do piloto). Meta de >4/5 significa os usuários considerarem a experiência global como “Muito Boa” em média.
- **Adoção:** 75% de adoção no grupo piloto significa 3 em cada 4 pessoas usando o ATHENAS regularmente. A definição de “uso regular” adotada é 3 ou mais vezes por semana. Vamos instrumentar o backend para logar usuários ativos e suas frequências de uso. Essa métrica mostra se o ATHENAS está realmente se integrando ao fluxo de trabalho ou não.
- **Relevância (Retrieval):** Mede se o módulo de busca está trazendo os conteúdos certos para o LLM. Para avaliar, criaremos um conjunto de perguntas de teste (*Golden Dataset*) com gabarito de quais documentos/trechos seriam relevantes, e verificaremos a porcentagem de vezes que o ATHENAS recupera aqueles contextos corretos. Almejamos >85% de *hits*. Possivelmente utilizaremos uma abordagem de *LLM-as-a-judge* combinada com revisão humana para validar relevância.

- **Fidelidade (Generation):** Verifica se as respostas geradas estão de fato apoiadas nas fontes fornecidas, sem “inventar” detalhes não suportados. Novamente usando o *Golden Dataset*, compararemos as respostas do ATHENAS com as respostas ideais esperadas, medindo a aderência (*faithfulness*). Meta >90% significa praticamente nenhuma alucinação ou informação desconexa.

Importante notar que separamos KPIs de negócio (como TIR, CSAT, Adoção) e KPIs técnicos (Relevância, Fidelidade). A liderança de negócio se importa mais com produtividade e satisfação, enquanto o time técnico foca na qualidade e precisão do sistema. Apresentar ambos os lados demonstra uma compreensão abrangente do que constitui sucesso e permite ajustar tanto aspectos de uso quanto do algoritmo conforme necessário.

Para suportar a medição rigorosa dos KPIs técnicos, será criado um “Golden Dataset” de referência. Ele consistirá em dezenas de perguntas reais do domínio com respostas aprovadas por especialistas, junto com as fontes esperadas. Esse conjunto permite rodar avaliações automáticas periódicas (ex: a cada mudança de modelo ou parâmetro, rodamos o dataset e medimos quantas respostas continuam corretas e referenciadas). Essa prática, utilizada na indústria, garante controle de qualidade contínuo e evita regressões no desempenho do ATHENAS conforme ele evolui.

## 9. Riscos & Mitigações

Como em qualquer projeto inovador, há riscos a serem gerenciados. Abaixo identificamos os principais riscos do MVP do ATHENAS e nossos planos para mitigá-los:

- **Qualidade dos Dados (“dados sujos”)** (*Probabilidade: Média; Impacto: Alto*)  
*Mitigação:* Implementar scripts de pré-processamento e limpeza de texto para remover ruído e padronizar documentos antes da ingestão. Focar inicialmente em repositórios curados pelo negócio, garantindo fontes de boa qualidade e relevância. (Se necessário, envolver donos do dado para sanear conteúdo crítico.)
- **“Alucinações” do LLM (respostas incorretas)** (*Probabilidade: Baixa; Impacto: Alto*)  
*Mitigação:* Realizar engenharia de *prompt* rigorosa, instruindo explicitamente o LLM a se basear apenas no contexto fornecido e não usar conhecimento externo/preexistente. Ajustar parâmetros (e.g. *temperature* próximo de 0) para reduzir a criatividade aleatória. Além disso, as citações de fonte em cada resposta permitem verificação humana imediata, caso algo pareça estranho, o usuário confere o link e identifica qualquer deslize, mantendo a confiança no sistema.

- **Vazamento de PII (dados sensíveis)** (*Probabilidade: Baixa; Impacto: Alto*)  
*Mitigação:* Anonimização mandatória de dados via Presidio (conforme seção de Privacidade). Nenhum dado bruto contendo PII será enviado ao LLM sem passar por mascaramento. Esse controle é inegociável. Adicionalmente, revisaremos periodicamente amostras de log para assegurar que nada sensível escapou.
- **Baixa Adoção pelos Usuários** (*Probabilidade: Média; Impacto: Média*)  
*Mitigação:* Envolver o grupo piloto desde o início do desenvolvimento, fazer *demos* a cada entrega importante, coletar feedback e ajustar. Garantir que estamos resolvendo uma “dor real” e comunicando os benefícios claramente. Faremos sessões de treinamento e divulgação para os usuários piloto, ensinando melhores práticas de uso da ATHENAS. Implementaremos um ciclo rápido de melhoria contínua: feedback dos usuários será capturado (via UI e reuniões quinzenais) e priorizado para ação, aumentando a chance de eles abraçarem a ferramenta.
- **Limites/Custos da API OpenAI** (*Probabilidade: Média; Impacto: Baixa*)  
*Mitigação:* Implementar cache local de respostas para perguntas repetidas frequentes, evitando chamadas redundantes ao LLM. Otimizar tamanho dos *chunks* e do *prompt* para enviar apenas o necessário (reduz consumo de tokens). Monitorar de perto o uso e os custos em um dashboard, se os gastos ameaçarem estourar o orçamento, podemos ajustar a taxa de uso (por exemplo, limitando consultas muito extensas ou implementando um sistema de cotas). Em último caso, avaliar modelos *open-source* rodando localmente se o custo do OpenAI se mostrar proibitivo. (No horizonte de MVP, entretanto, o custo está projetado como baixo, ver seção de Budget).
- **Resultados Irrelevantes na Busca** (*Probabilidade: Média; Impacto: Alto*)  
*Mitigação:* Iniciaremos com uma estratégia robusta de *chunking* (recursiva) e *embeddings* gerais. Se durante os testes iniciais percebermos muita irrelevância nos trechos recuperados, experimentaremos ajustes: por exemplo, *chunking* semântico (dividir documentos por tópicos latentes) ou afinamento do modelo de embedding. No *roadmap* futuro, está prevista até substituição do modelo de embedding por um especializado na linguagem da LCGBR. Em paralelo, iteraremos constantemente o *prompt* de busca e utilizaremos técnicas como filtros por palavra-chave para complementar a busca vetorial (*Hybrid Search*) caso necessário.

(*Prob.* = Probabilidade de ocorrência; *Impacto* = impacto caso ocorra. *Mitigações* compõem tanto prevenção quanto planos alternativos.)

Em suma, abordamos os riscos com proatividade: muitos mitigadores (limpeza de dados, anonimização, *prompt tuning*, envolvimento do usuário) já estão incorporados

no plano do MVP. Isso aumenta a resiliência do projeto e demonstra para a liderança que temos estratégias para navegar obstáculos comuns em projetos de IA.

## 10. Recursos Necessários

Para entregar o MVP em 90 dias, estimamos a necessidade dos seguintes recursos:

### Pessoas (alocação por 90 dias):

- **1.0 FTE Tech Solutions Engineer (Líder Técnico):** Responsável pela arquitetura, desenvolvimento *backend* e integração de todos os componentes. (Perfil: você mesmo, ou alguém com background em engenharia de dados/IA).
- **0.5 FTE Product Owner (Patrocinador de Negócio / Analista):** Ponto de contato do negócio, responsável por definição de requisitos, priorização das funcionalidades com maior valor, e coleta de feedback contínuo do grupo piloto. (Ex.: alguém da equipe de Vendas/Marketing com tempo parcial dedicado ao projeto).
- **0.2 FTE UI/UX Designer (consultoria pontual):** Especialista em design de interface, envolvimento breve no início para orientar a criação de uma UI intuitiva no Streamlit e alinhada à identidade visual da empresa. (Não é full-time; atuação talvez de 1 dia por sprint nas fases de design/testes de usabilidade).
- **0.3 FTE Engenheiro de Dados / QA:** Apoio em dados e qualidade, auxiliando na criação e curadoria do Golden Dataset, preparação de dados de treino/teste, e definição/execução de casos de teste para garantir que o sistema responda corretamente. (Perfil: alguém de Data Analytics ou QA Automation com conhecimento de NLP seria ideal).

### Infraestrutura:

- **Azure VM:** 1 máquina virtual em Azure (Linux) para hospedar a aplicação containerizada. Recomenda-se uma VM de uso geral (por ex. Standard B2ms com 2 vCPU, 8 GB RAM) para início. Ela rodará o backend + ChromaDB + UI. Pode-se escalar verticalmente caso necessário mais poder de processamento durante testes.
- **Azure Blob Storage:** 1 container de Blob Storage (100 GB) para armazenar artefatos como logs, dados intermediários (por exemplo, uma cópia dos embeddings gerados, backups do ChromaDB) e possivelmente arquivos de configuração. Isso assegura durabilidade e conformidade (backups automáticos, etc.).

### Budget - Estimativa de Custo (OpenAI API) para 90 dias de MVP:

Embora o ATHENAS utilize recursos *cloud* e API pagos, os custos projetados são relativamente modestos. Abaixo detalhamos a estimativa de gastos com a API

OpenAI (principal componente variável), considerando premissas realistas de uso durante o piloto:

- **Ingestão (Embeddings iniciais):** ~10.000 documentos indexados, totalizando ~34 milhões de tokens processados pelo modelo Ada-002 (preço: \$0.10 por 1.000.000 tokens) → ~R\$ 20.
- **Geração de Respostas (Consultas):** ~22.500 perguntas dos usuários ao longo do piloto (50 usuários \* 5 perguntas/dia \* ~90 dias), totalizando ~54 milhões de tokens gerados no GPT-4 (soma de entrada+saída). Preço GPT-4 estimado: \$0.03 por 1.000 tokens (considerando \$0.06/1k *output* + \$0.03/1k *input*). Custo → ~R\$ 1.000.
- **Buffer de Segurança:** +100% do valor acima para cobrir testes, iterações e folga (caso haja reprocessamentos, ajustes ou uso acima do esperado) → ~R\$ 1.000.
- **Total estimado (90 dias):** ~R\$ 2.020.

(Nota: Câmbio considerado 1 USD = ~5 R\$, valores arredondados.)

O custo de ingestão inicial é notavelmente baixo, indexar todo o conhecimento relevante custa apenas alguns reais. O principal custo recorrente é o das consultas ao LLM, mas mesmo com uma folga generosa, estamos estimando em torno de R\$ 2 mil para três meses de operação. Para um projeto de IA com potencial de impacto tão grande, R\$ 2k é um investimento muito baixo em comparação aos ganhos esperados, representando claramente um caso de baixo risco e alto retorno. Em termos de orçamento global da empresa, é um valor quase desprezível, o que facilita a justificativa financeira do MVP.

(Obs: Os custos de VM Azure e storage também são baixos, uma VM B2ms custa em torno de R\$ 300/mês, total ~R\$ 900 no trimestre; Storage 100 GB < R\$ 50. Portanto, mesmo somando tudo, o MVP opera dentro de ~R\$ 3 mil diretos, sem contar horas-pessoa.)

## 11. Roadmap Próximos Passos

Este *roadmap* delinea a visão de evolução do ATHENAS, desde o MVP focado em um caso de uso específico até se tornar uma plataforma crítica de inteligência empresarial. As fases propostas estão alinhadas com a curva de maturidade de soluções RAG observada na indústria e as necessidades crescentes da LCGBR.

### MVP (0-3 Meses): Go-Live com Vendas & Marketing

**Objetivo:** Provar rapidamente o valor fundamental da solução em ambiente controlado.

- <Data + 2 semanas>: Finalizar a ingestão e processamento dos documentos prioritários dos depts. de Vendas e Marketing (garantir que o *corpus* mais importante está disponível na ATHENAS).

- <Data + 4 semanas>: Homologar toda a pipeline RAG e a API de chat em ambiente



de teste. Garantir que o núcleo técnico (ingestão → busca → geração) está funcional e confiável.

- <Data + 8 semanas>: Lançamento de um Beta interno para ~10 usuários-chave (*early adopters*) dos departamentos piloto. Foco em coletar feedback intensivo: usabilidade da interface, relevância das respostas, eventuais ajustes de fontes ou performance.

- <Data + 12 semanas>: Lançamento completo do MVP para todo o grupo piloto de Vendas & Marketing (aprox. 50 usuários). Nesta data marcamos também uma apresentação formal dos resultados do MVP para a liderança executiva, mostrando métricas de sucesso atingidas, depoimentos de usuários e casos de uso. Será o momento de decidir o *Go/No-Go* para expansão do projeto.

*(Indicador de sucesso da fase MVP: liderança concordar em prosseguir com investimento para próximas fases, baseado nos ganhos comprovados.)*

## **Fase 2 (3-9 Meses): Expansão e Otimização (Production RAG)**

**Objetivo:** Com o valor comprovado no piloto, o foco muda para robustez, escalabilidade e primeiros incrementos de funcionalidades.

- **Integração de Novas Fontes:** Expandir significativamente a base de conhecimento da ATHENAS, conectando-a a outras fontes de dados críticas. Prioridades prováveis: Confluence (onde muitas documentações de projetos vivem) e eventualmente bases de dados estruturados (via técnicas de *text-to-SQL*). Também planeja-se integrar ferramentas de marketing como Adobe Analytics/Target para permitir consultas sobre dados de campanhas (trazendo automaticamente dados de engajamento, conversões, etc. nas respostas da ATHENAS). Isso ampliará o alcance e utilidade do assistente, que passará a responder perguntas combinando múltiplas fontes.

- **Busca e Recuperação Avançadas:** Implementar técnicas mais sofisticadas para melhorar a precisão da recuperação de contexto. Avaliaremos adoção de *Hybrid Search*, combinando busca vetorial semântica com busca tradicional por palavras-chave (ex.: BM25) para casos onde detalhes específicos (como números ou nomes exatos) importam. Introduzir um componente *Re-ranker*: um modelo auxiliar leve que reordena os resultados inicialmente retornados, garantindo máxima relevância nos trechos usados pelo LLM. Essas melhorias visam elevar ainda mais a qualidade das respostas, reduzindo chances de algum documento relevante ser ignorado.

- **Melhoria de UI/UX:** Com base no extenso feedback dos usuários do MVP, aprimorar a interface da ATHENAS. Provavelmente isso envolve migrar de Streamlit para um *frontend* web customizado (React ou similar), permitindo uma experiência mais fluida e integrada visualmente ao portal da empresa. Funcionalidades adicionais na UI podem incluir: histórico de consultas, filtros por fonte ou data, botão de exportar resposta (para Word/Excel), etc., conforme as demandas observadas.

- **Avaliação Contínua & Aprendizado:** Automatizar por completo a pipeline de avaliação com o *Golden Dataset*. Cada nova versão do modelo ou mudança significativa no sistema será testada automaticamente contra o dataset e resultados reportados no *dashboard*. Além disso, expandir o *Golden Dataset* para cobrir novos domínios e casos de uso conforme a ATHENAS se expande para outras áreas,

garantindo que mantemos alta qualidade uniformemente. Poderemos também introduzir aprendizado incremental: por exemplo, incorporar no dataset perguntas reais mal respondidas para treinar as próximas versões (*failures* -> *fixes*).

*(Ao final de 9 meses, esperamos ter a ATHENAS operando em produção para grande parte dos funcionários de conhecimento, com performance aprimorada e UI profissional. As bases estarão prontas para dar o próximo salto de capacidades.)*

### **Fase 3 (9-18 Meses): Rumo à Inteligência Agentic (Escala Enterprise)**

**Objetivo:** Evoluir a ATHENAS de uma ferramenta reativa de Q&A para um assistente proativo e agente operacional, integrado profundamente aos fluxos de trabalho da empresa.

- **Capacidades de Agente:** Tornar a ATHENAS capaz de realizar ações em nome dos usuários nos sistemas corporativos conectados. Como mencionado, ao invés de apenas responder perguntas, a ATHENAS passará a executar comandos complexos via linguagem natural. Exemplos: *"Crie um rascunho de e-mail com os 3 principais argumentos de venda para o produto X e salve nos meus rascunhos do Outlook"; "Registre um ticket de suporte no Jira com esta descrição de problema"; "Atualize o status do documento Y no SharePoint para 'Aprovado' e notifique a equipe", ou "Gerar uma apresentação PPT resumindo os resultados da campanha Z"*, e a ATHENAS executará essas tarefas integrando-se via APIs (Outlook, Jira, SharePoint, PowerPoint/Office 365). Essa evolução transformará a ATHENAS em um verdadeiro copiloto digital para cada funcionário, automatizando passos e agilizando processos.
- **Personalização e Fine-Tuning Específico:** Realizar *fine-tuning* dos modelos de embedding e possivelmente do LLM nos dados específicos da LCGBR. Isso permitirá que a ATHENAS compreenda profundamente o jargão interno, acrônimos, nomes de produtos/projetos e nuances do negócio, aumentando ainda mais a relevância e precisão das respostas. Além disso, implementar perfis inteligentes de usuário (conforme mencionado nos diferenciais), ou seja, a ATHENAS aprende preferências de cada usuário ou função. Ela saberá, por exemplo, que um diretor prefere respostas mais resumidas e visuais, enquanto um analista quer detalhes e planilhas anexas, adaptando a forma de responder conforme quem pergunta.
- **Multimodalidade Avançada:** Ampliar as capacidades de ingestão e geração da ATHENAS para além de texto puro. Nesta fase, a ATHENAS poderá indexar e pesquisar conteúdo em imagens e áudio/vídeo. Por exemplo, *"ATHENAS, encontre aquele diagrama de arquitetura do projeto ABC"*, ela buscará não por texto, mas por similaridade visual em slides e imagens armazenadas. Ou então, suportaremos busca em transcrições de reuniões e vídeos: reuniões gravadas no Teams poderão ser transcritas e integradas ao conhecimento da ATHENAS, permitindo perguntas como *"Quais decisões foram tomadas na última reunião da diretoria?"* com respostas extraídas das atas/transcrições. Essa multimodalidade tornará a ATHENAS um ponto central de todo conhecimento corporativo, independente do formato.
- **Governança, Explicabilidade e Otimização de Custos:** Expandir os *dashboards* de monitoramento para nível enterprise. Incluir métricas de latência (tempos de resposta da ATHENAS), uso por departamento, custos segregados por centro de

custo, etc. Implementar também uma camada de explicabilidade ao usuário final: por exemplo, funcionalidade “*Por que a ATHENAS me deu essa resposta?*” que mostre quais fatores ela considerou, ou um resumo das fontes utilizadas. Isso aumenta ainda mais a confiança e ajuda na adoção entre perfis menos técnicos. Em paralelo, otimizar custos, se o uso explodir, avaliar *deploy* de modelos proprietários internos ou *caching* avançado das consultas mais comuns. A governança também abrangerá aspectos de segurança: controle de permissões (a ATHENAS respeitará as permissões dos documentos subjacentes, ex: se um usuário não tem acesso a certa pasta no SharePoint, a ATHENAS não revelará conteúdo dela). Tudo isso garante que ao operar em larga escala, a ATHENAS se mantenha eficiente, segura e transparente.

- **Engajamento e Cultura de Conhecimento:** A essa altura, com a ATHENAS disseminada na empresa, reforçaremos os mecanismos de gamificação e engajamento interno. Podemos lançar programas de incentivo para uso criativo da ATHENAS (prêmios simbólicos para melhores cases de uso do assistente), integrar a ATHENAS nas metas de produtividade de equipes, e continuar aprimorando a comunicação interna destacando os ganhos. A ideia é consolidar a ATHENAS não apenas como uma ferramenta, mas como parte da cultura da LCGBR, ser uma empresa que valoriza e facilita o acesso ao conhecimento em todos os níveis.

*(Após 18 meses, vislumbramos a ATHENAS como um elemento onipresente no dia a dia da organização, aumentando produtividade, acelerando processos e até mudando a forma como colaboradores interagem com a informação e entre si. A partir daqui, entraremos em um ciclo de melhorias contínuas e manutenção, buscando novos horizontes, como integração com IA generativa para clientes etc., mas isso já extrapola o escopo inicial.)*

## 12. Deliverables & Links (pós evento)

*(Esta seção será preenchida após a conclusão do MVP e do evento de apresentação, contendo materiais entregues e links úteis para consulta.)*

**Deliverables previstos:** - Repositório de Código (GitHub): Código-fonte completo do ATHENAS MVP, incluindo README de deploy, ex: [github.com/LCGBR/athenas-knowledge-assistant](https://github.com/LCGBR/athenas-knowledge-assistant). - Documentação de API e Arquitetura: Documentos detalhando endpoints da API, configuração de ambiente, credenciais necessárias, etc. - Apresentação Final (Slides): PPT ou Loop utilizado na banca executiva final, resumindo resultados obtidos, lições aprendidas e próximos passos. - Vídeo Demonstrativo: Gravação de uma demonstração do ATHENAS em ação, mostrando casos de uso reais durante o piloto. - Relatório de Resultados do MVP: Documento pós-implementação apresentando as métricas atingidas versus metas (TIR reduzido, adoção, CSAT), feedback qualitativo dos usuários pilotos, e recomendações para escala.

*(Links serão adicionados conforme os deliverables forem gerados e armazenados no SharePoint ou repositório apropriado.)*

### 13. Anexos / Evidências (pós evento)

*(Também a serem preenchidos após execução do MVP, fornecendo evidências adicionais do sucesso do projeto.)*

Possíveis anexos/evidências incluirão: - Dashboard de Uso do ATHENAS: Capturas de tela ou link para o dashboard PowerBI mostrando estatísticas de uso, principais consultas, times mais engajados, etc., durante o piloto. - Exemplos de Q&A do ATHENAS: Comparativo de perguntas feitas e respostas fornecidas pela ATHENAS (incluindo as citações), demonstrando na prática a qualidade e utilidade das respostas. - Depoimentos de Usuários: Comentários ou *quotes* curtos de participantes do piloto sobre como o ATHENAS ajudou em seu trabalho. (Ex: "Nunca foi tão fácil achar informações de produto, ganhei horas da minha semana graças à ATHENAS!" - [Nome, Gerente de Vendas]) - Referências Bibliográficas: Lista de fontes e pesquisas citadas no projeto (ex: relatório McKinsey, estudo IDC, press release APQC etc.), reforçando a base conceitual e dados de mercado mencionados ao longo do documento.

*(Estes itens serão compilados e anexados conforme disponibilidade pós-MVP.)*

---