# Guia Unificado 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

# Introdução

Compreender as dores e os desafios enfrentados pelos órgãos e pela sociedade é essencial para o desenvolvimento de soluções relevantes. O Núcleo de IA (NIA) colabora no mapeamento detalhado das necessidades que podem ser endereçadas com o uso da Inteligência Artificial (IA).

Além disso, para garantir que as soluções propostas estejam alinhadas com o estado da arte da IA e para aprender com experiências bem-sucedidas, o NIA acompanha de perto as tendências tecnológicas, os casos de sucesso e as lições aprendidas, tanto no setor público quanto no setor privado.

Este documento tem como objetivo ser um guia para projetos de IA para serviços públicos, com adoção de um caráter genérico, podendo ser aplicado às diversas **abordagens técnicas** existentes no campo da IA, como aprendizado de máquina supervisionado e não supervisionado, deep learning, IA generativa e LLMs, reconhecimento e síntese de voz, sistemas de recomendação, visão computacional, entre outros.

Este guia é parte integrante da Metodologia de desenvolvimento de Inteligência Artificial (MIA) e orienta a prospecção e o desenvolvimento ou aquisição de soluções digitais, cujo núcleo funcional seja baseado em algoritmos ou modelos de Inteligência Artificial (IA), distinguindo claramente:

- o desenvolvimento da IA (modelo, dados, experimentação) e
- o desenvolvimento da solução aplicacional (“casca”), como chatbots, assistentes virtuais ou sistemas inteligentes.

## Como usar este guia

Este guia pode ser para apoio a:

- Planejamento ou execução de um projeto de IA aplicada;
- Necessidade de diferenciação entre etapas de pesquisa e desenvolvimento de IA de engenharia de software;
- Servidores que atuam como gestor, técnico, jurídico ou órgão de controle.

Desta forma, diversos perfis da administração pública encontram apoio, conforme visão demandada.

- **Alta Administração:** Visão executiva · classificação de risco · decisões-chave
- **Gestores de Projeto:** Fases do Guia · riscos · indicadores
- **Equipes Técnicas:** Pipeline de IA · arquitetura da solução · testes
- **Órgãos de Controle:** Governança · conformidade · evidências

Este guia está dividido em etapas representando o ciclo de vida de uma solução de IA, desde sua ideação até sua descontinuidade, representada na figura a seguir.


**Figura 1 - Ciclo de vida de soluções de IA**

Cada etapa, por sua vez, está também dividida em fases que representam macro atividades. Cada fase possui um descritivo resumido, seus resultados esperados e recomendações de boas práticas, além de uma matriz RACI[1] para as macroatividades da fase. O marco de passagem de uma fase para a seguinte sempre envolve uma aprovação dos resultados e resultados, por parte dos decisores (Accountable - responsáveis por aprovar) indicados na matriz RACI.

Neste documento, os papéis das partes interessadas envolvidas no ciclo de vida de soluções de IA incluem:

- **Alta Gestão** (órgão)**:** nível superior de liderança e decisão na administração pública, responsável por definir diretrizes estratégicas, deliberar sobre prioridades institucionais e representar o órgão.
- **CISC gov.br (Centro Integrado de Segurança Cibernética**): Instância central responsável pela prevenção, detecção, tratamento e resposta a incidentes cibernéticos no âmbito do Governo Digital. Deve ser notificada imediatamente pelos órgãos em caso de incidentes (ex.: vazamento de dados decorrente do uso de IA). Compete ainda ao CISC: realizar testes de intrusão (*penetration tests*), testes de segurança estáticos e dinâmicos; conduzir análises contínuas e atividades de inteligência cibernética; monitorar padrões maliciosos; e emitir alertas e determinações, com prazos definidos, para correção de vulnerabilidades críticas.
- **Comitê de IA / Ética**: Comitê interno ao órgão responsável por orientar, avaliar e supervisionar o uso de inteligência artificial, assegurando conformidade legal, ética, técnica e institucional. Nem todos os órgãos possuem este tipo de comitê.
- **Curador de dados (*business data steward*).** Agente público responsável pela gestão de ativos de dados, internos ou externos ao órgão ou entidade, designado por liderança na estrutura organizacional. Responsável por garantir a qualidade, confiabilidade, conformidade e uso adequado dos dados do órgão, ao longo de todo o seu ciclo de vida. Atua como um guardião dos dados, assegurando que sejam corretos, atualizados, bem documentados, acessíveis a quem precisa e utilizados de forma ética e conforme normas legais. Suas principais responsabilidades são: manter atualizada a catalogação de dados e metadados sob sua custódia nas unidades negociais; classificar os dados quanto ao nível de acesso definidos conforme legislações vigentes; assegurar a proteção dos dados pessoais, observando as orientações do encarregado e conforme o disposto pela lei 13.709, de 14 de agosto de 2018 [9].

1 Sigla para (R) *Responsable*, (A) *Accountable*, (C) *Consulted*, (I) *Informed*, é uma técnica que descreve as responsabilidades de várias partes interessadas na conclusão de tarefas ou entregas.

- **Demandante**: representante de uma unidade de um órgão ou órgão demandante para solução de um desafio.
- **Encarregado pelo tratamento de dados pessoais,** doravante denominado ETDP (DPO - *Data Protection Officer*) - pessoa indicada pelo controlador como braço técnico consultivo para atuar como canal de comunicação (intermediador), orientador sobre boas práticas, fiscalizador de normas e relator sobre riscos e necessidades de melhorias na proteção de dados pessoais, conforme a Lei Geral de Proteção de Dados (LGPD – Lei nº 13.709/2018). Não toma decisões sobre as finalidades do tratamento de dados.**Executor** (Parceiro executor): entidade, órgão ou empresa responsável por desenvolver uma solução de IA (denominado Produtor de IA na ABNT NBR ISO/IEC 22989:2023, 5.19). Este papel pode ser desempenhado por um parceiro ou pelo próprio órgão.
- **Executivo de Dados (*data owner*):** é um servidor público ocupante de cargo efetivo, empregado público ou militar de carreira, responsável pela implementação e manutenção do Programa de Governança de Dados no âmbito do órgão ou entidade, atuando no nível estratégico, desvinculado da área de tecnologia da informação, e como ponto focal de comunicação tanto internamente quanto para os órgãos de monitoramento desta política e os demais atores do ecossistema de dados. [12]
- **Facilitador**: pessoa jurídica, servidor público do órgão ou parceiro com capacidade de executar a fase de descoberta de hipóteses e de solução na etapa de estruturação do desafio.
- **Gestor de Segurança da Informação**: Responsável por orientar, conduzir diagnósticos, planejar e monitorar todas as medidas de segurança da informação do órgão. Neste guia, ele atua como a principal autoridade técnica que aprova o mapeamento de riscos cibernéticos, valida os testes de segurança antes da implantação e garante a supervisão ininterrupta do sistema durante a operação. [10]
- **Gestor de TIC (tecnologia da Informação e Comunicação)**: responsável por planejar, executar e monitorar projetos tecnológicos, alinhando-os à Estratégia de Governo Digital (EGD) e ao Plano Diretor de TIC (PDTIC) do órgão.
- **Responsável Setorial pela Gestão de Integridade** – **RSGI**: Responsável por coordenar e gerir riscos para a integridade institucional relacionados ao projeto. Ele atua como um "guardião ético e moral", atuando nas fases iniciais de avaliação de risco para garantir que o algoritmo desenvolvido não resulte em vieses discriminatórios, falhas de *compliance*, corrupção algorítmica ou qualquer impacto que manche a reputação e a integridade do serviço público.
- **Sustentador** (Parceiro de sustentação): entidade, órgão ou empresa responsável por sustentar a solução de IA, após sua implantação. Este papel pode ser desempenhado por um parceiro ou pelo próprio órgão.

Para os casos de **aquisição**, por não envolver **desenvolvimento de uma nova solução**, na etapa de experimentação deve-se considerar provas de conceito das diversas opções de soluções propostas para aquisição. A etapa de Implementação, passa a ser utilizada para configurações e estruturação de ambiente de homologação necessário para aprovação final da solução e plano de implantação.

# 1. Etapa - Prospecção de desafios e priorização de IA no serviço público

Fundamental para alinhar a aplicação de IA no serviço público, a etapa de prospecção busca maximizar o valor e o impacto das soluções. Nesta etapa, o Plano Brasileiro de Inteligência Artificial (PBIA) é o direcionador das tratativas, permitindo a identificação e a seleção criteriosa dos desafios e necessidades que possuam o maior potencial para gerar valor e impacto positivo nos serviços públicos, garantindo que recursos e os esforços de IA sejam aplicados onde são mais necessários.


Figura 2 - Fases da Prospecção

## 1.1 Fase - Descoberta de desafio

A identificação de desafios no serviço público com potencial de resolução por meio de soluções de IA pode ocorrer de forma orgânica - a partir de uma demanda de um servidor - ou de forma estruturada pelo órgão. Uma vez identificados os desafios, o órgão realiza uma apuração inicial, avaliando cada um deles quanto ao alinhamento estratégico e à disponibilidade sobre os dados envolvidos, para serem classificados por seu impacto potencial, compondo uma lista de desafios priorizados, filtrados por um funil de critérios.

De forma alternativa à identificação interna de desafios, a descoberta também pode ocorrer a partir de rodadas periódicas de prospecção, aplicadas pelo NIA aos diversos órgãos. O principal objetivo é assegurar que o funil seja alimentado com demandas minimamente estruturadas, o que é essencial para facilitar a subsequente triagem e priorização.

A priorização interna de projetos do órgão deve ocorrer de maneira estruturada e para tal o órgão tem à sua disposição um questionário padronizado, contemplando questões que direcionam a construção de um ONEPAGE, indicação se a proposta destina-se realmente a uma solução de Inteligência Artificial e, em caso afirmativo, cria uma pontuação para que o órgão possa considerar para priorização, comparando-a a outros projetos avaliados pelo questionário. As orientações e esclarecimentos do questionário constam do [6] ANEXO - Critérios de priorização de desenvolvimento de soluções de IA.

### 1.1.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **ONEPAGE**[2] **-** Formulário deDesafio Preliminar | Documento preenchido pelo demandante, contendo:<br>• O desafio;<br>• Como resolve os desafios atualmente;<br>• Como entende que IA poderia resolver o desafio;<br>• Benefícios esperados com uso da IA;<br>• Panorama de dados. |
| **Desafios Priorizados e Alinhados** | Conjunto de desafios aprovados, alinhados à Alta Gestão e com impacto potencial analisado. |

### 1.1.2 O que verificar antes de avançar

**Analisar o impacto potencial com o uso da solução** (por exemplo, operacional no órgão ou em relação a resultados de política pública). A priorização no nível estratégico interno é realizada por cada entidade da APF (órgão, secretaria ou ministério) que participou da etapa de identificação dos desafios e das necessidades de suas unidades.

**A aprovação exige a resposta afirmativa para todos os itens da lista a seguir.** Caso haja reprovação, será avaliada a possibilidade de ajuste para posterior aprovação, ou a descontinuação do desafio por inviabilidade de resolução.

- O desafio só pode ser resolvido por meio de um sistema de IA?
- No caso de envolvimento de dados pessoais ou sensíveis, foi verificado se existe necessidade de adequação à LGPD (Lei Geral de Proteção de Dados Pessoais)?
- Os dados estão disponíveis?
- A acessibilidade dos dados foi confirmada?
- Trata-se de um desafio estratégico para políticas públicas?

**Decisão:**

- **Aprovação.** Seguir para próxima fase. O desafio é estratégico, possui dados disponíveis, acessíveis e demanda IA para resolução.
  - **Ação.** Seguir para 2.2 Fase - Aprofundamento de desafios.
- **Reprovação.** Dados inexistentes ou de difícil acesso.
  - **Ação.** Retornar para 2.1 Fase - Descoberta de desafio.
- **Reprovação.** Desafio solucionável por automação tradicional.
  - **Ação.** Cancelar desafio. Finalizar processo justificando o cancelamento.

2 ONEPAGE - documento de uma única página que reúne todas as informações básicas do desafio.

### 1.1.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC |
|---|---|---|---|---|
| **Identificação interna de desafios** | R | C | A | C |
| **Preenchimento do ONEPAGE sobre desafio** | R | C | A | C |
| **Análise de disponibilidade dos dados** | A | R | C | C |
| **Estabelecimento de alinhamento estratégico** | R | C | A | C |
| **Análise de impacto e priorização interna** | R | C | A | C |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor); - (Não atua).

### 1.1.4 Recomendações, políticas e boas práticas

- A definição e identificação de curadores de dados é uma importante ação de governança de dados que agiliza a implementação da solução no futuro.

## 1.2 Fase - Aprofundamento de desafios

A fase de aprofundamento dos desafios visa rever e detalhar o contexto e o (s) desafio (s) validado (s) /priorizado (s) pelo órgão para a produção do *briefing* (detalhamento do desafio). Isso é feito por meio da **geração de hipóteses** de solução, que incluem a **definição da arquitetura de alto Nível** (ex: para *LLM*, Regressão ou Visão Computacional).

A fase também exige a **avaliação preliminar da viabilidade de dados**, identificando o responsável/curador, a acessibilidade, a centralização, o tipo de dado necessário e os trâmites legais para acesso.

Por fim, é realizada uma **avaliação preliminar de risco ético,** utilizando o **Questionário de Avaliação de Risco** do framework de Autoavaliação de Impacto Ético (AIE)[3] e a **identificação de riscos gerais e fatores críticos** (barreiras). A AIE deve considerar, no mínimo:

- Contexto de uso e partes afetadas;
- Potenciais impactos éticos e sociais
- Riscos de discriminação, exclusão, opacidade e uso indevido;
- Proporcionalidade do uso de IA;
- Medidas de mitigação e supervisão humana.

A AIE complementa, mas não substitui, a avaliação de qualidade, segurança e conformidade

3 Disponível em: www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados/inteligencia-artificial-1/publicacoes/framework-para-a-autoavaliacao-de-impacto-etico-em-inteligencia-artificial-no-setor-publico-federal.

dos dados. O principal resultado desta fase é o **Estudo Técnico Preliminar (ETP) oficial** e uma **versão parcial do mapa de riscos** de aplicação de IA. É oportuno notar que risco ético se refere às consequências sociais, morais e humanas do uso da tecnologia. Mesmo um sistema 100% seguro (sem *hacks*) pode causar um impacto ético negativo se seu *design* ou aplicação forem prejudiciais.

### 1.2.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **AIE - Avaliação de Impacto Ético** | Avaliação preliminar de risco ético realizada (utilizar parte 1 do framework AIE). |
| **Mapas de riscos** | Riscos em geral e fatores críticos e barreiras identificados. |
| **Briefing de desafios (visão geral dos desafios)** | Documento resumindo os desafios identificados e uma visão geral dos riscos associados (éticos e outros). |
| **Estudo Técnico Preliminar (ETP)** | Documento para fundamentar tecnicamente a decisão de desenvolver, adquirir ou contratar uma solução de IA, demonstrando necessidade, viabilidade, riscos, alternativas e aderência normativa (neste ponto para responder a pergunta "devemos ou não adotar uma solução de IA, e de que forma?" |

### 1.2.2 O que verificar antes de avançar

Este O que verificar antes de avançar considera a aprovação ou não dos resultados.

**Decisão:**

- **Aprovação.** Seguir para próxima fase, caso todos resultados tenham sido analisados e aprovados.
  - **Ação.** Seguir para 2.3 Fase - Validação de desafios pelo NIA (opcional).
- **Reprovação.**
  - **Ação.** Rever premissas da proposta de sistema de IA, seja para reduzir riscos identificados ou mesmo para reduzir custos estimados. Na impossibilidade de ajustes e inviabilidade pela reprovação, cancelar a demanda e registrar justificativa para aprendizados futuros.

### 1.2.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de segurança da informação | NIA |
|---|---|---|---|---|---|---|---|
| **Elaboração de ETP - Estudo Técnico Preliminar** | R | C | A | R | C | C | - |
| **Avaliação de dados** | A | C | - | R | C | - | - |
| **AIE - Avaliação de Impacto Ético** | A | C | I | R | C | - | - |
| **Identificação de riscos e barreiras (mapas de riscos)** | A | C | C | R | C | C | - |
| **Decisão de envolvimento do NIA** | R | C | A | C | - | C | - |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor); - (Não atua).

### 1.2.4 Recomendações, políticas e boas práticas

O uso de *frameworks* de referência e normas técnicas aplicáveis a esta fase.

**Gestão de dados e qualidade:**

- **DAMA-DMBOK (*Data Management Body of Knowledge*) [3]:** *Framework* abrangente para gestão de dados, cobrindo governança de dados, qualidade de dados, arquitetura de dados, segurança, integração, metadados e dados mestres. Fundamental para garantir que os dados utilizados em sistemas de IA atendam aos padrões de qualidade necessários.
- **ISO/IEC 23894**: Tecnologia da informação – Inteligência artificial – Orientações sobre gestão de riscos. Processos de gestão de riscos (identificação, análise e mitigação) em projetos de IA.
- **ISO/IEC TR 24028**: Propriedades técnicas (explicabilidade, robustez e ética) que tornam o sistema confiável. Examina tópicos relacionados à confiabilidade em sistemas de IA, incluindo: abordagens para estabelecer confiança em sistemas de IA por meio de transparência, explicabilidade, controlabilidade, etc.; armadilhas de engenharia e ameaças e riscos típicos associados a sistemas de IA, juntamente com possíveis técnicas e métodos de mitigação; e abordagens para avaliar e alcançar disponibilidade, resiliência, confiabilidade, precisão, segurança e privacidade de sistemas de IA. A especificação de níveis de confiabilidade para sistemas de IA não faz parte do escopo deste TR.. **Privacidade e Proteção de Dados:**
- **ABNT NBR ISO/IEC 27001:** Sistema de Gestão de Segurança da Informação (SGSI).
- **ABNT NBR ISO/IEC 27701:** Extensão da ISO 27001 focada em Sistema de Gestão de Informações de Privacidade (SGIP), especialmente relevante para conformidade com LGPD.

- **Diretrizes da ANPD:** Consultar orientações da Autoridade Nacional de Proteção de Dados (ANPD) sobre tratamento de dados pessoais em sistemas de IA incluindo minimização de dados, limitação de finalidade e garantia de direitos dos titulares.
- ***CIS Controls*** (Center for Internet Security) e seu *Privacy Companion Guide*: apoia os objetivos dos Controles CIS, alinhando os princípios de privacidade e destacando possíveis preocupações com a privacidade que podem surgir com o uso dos Controles CIS. São apresentadas considerações para que as equipes de TI, jurídicas e quaisquer outros funcionários com responsabilidades em privacidade possam identificar oportunidades para integrar as considerações de privacidade aos controles de segurança de dados.
- ***Secure Controls Framework* (SCF)**: trata de forma nativa e integrada temáticas de privacidade e segurança da informação.
- **Autoavaliação de Impacto Ético da Inteligência Artificial no Setor Público** (AIE)[4].

## 1.3 Fase - Validação de desafios pelo NIA (opcional)

**Se houver um processo seletivo de soluções,** a fase de validação pelo NIA tem como objetivo principal apoiar a validação e a priorização dos desafios sob a perspectiva da Administração Pública Federal.  Esta fase somente realizada nos casos em que houver um processo seletivo de soluções no qual o NIA deva apoiar.

### 1.3.1 Atividades e critérios de priorização

O NIA realiza uma análise detalhada do conteúdo da *One Page* e do *briefing* para executar as seguintes avaliações:

- **Confirmar o alinhamento** do desafio com a estratégia da Administração Pública Federal.
- **Avaliar o potencial de adoção** de IA para solucionar o desafio.
- **Validar preliminarmente os dados** necessários, verificando sua disponibilidade e validade.
- **Avaliar o potencial de impacto** do uso da solução, considerando aspectos tanto positivos quanto negativos no trabalho, na economia, na sociedade, de ética, entre outros.

**Priorização final:**

Os projetos que já foram priorizados por cada instituição são avaliados e passam por uma **priorização final** pelo NIA. Essa decisão é baseada em critérios, tais como:

- Viabilidade técnica;
- Potencial de impacto;
- Alinhamento estratégico federal;
- Potencial da adoção de IA para solucionar o desafio.

4 Disponível em: www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados/inteligencia-artificial-1/publicacoes/framework-para-a-autoavaliacao-de-impacto-etico-em-inteligencia-artificial-no-setor-publico-federal.

O NIA acompanha tendências tecnológicas, casos de sucesso e lições aprendidas nos setores público e privado para garantir que as soluções propostas estejam alinhadas com o estado da arte da IA.

### 1.3.2 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Validação de desafios** | Lista de desafios avaliados e orientados quanto a necessidade de parceiros. |

### 1.3.3 O que verificar antes de avançar

As análises realizadas serão a base para tomada de decisão quanto à priorização e à necessidade de parcerias. Para que a viabilidade seja aprovada, a lista de checagem a seguir deve ser verificada:

- O potencial de adoção de IA para solucionar o desafio foi avaliado como viável?
- O desafio atende aos critérios de viabilidade técnica (avaliação do NIA)?
- A avaliação preliminar da disponibilidade e validade dos dados necessários foi satisfatória?
- O potencial de impacto positivo do uso da solução é relevante e justificado?

**Decisão:**

- **Aprovação. A aprovação exige a resposta afirmativa para todos os itens da lista.** O desafio é estratégico, possui dados disponíveis, acessíveis e demanda IA para resolução.
  - **Ação.** Seguir para 2.4 Fase - Definição de facilitador.
- **Reprovação**.
  - **Ação.** Avaliar a possibilidade de ajuste para posterior aprovação e seguir para 2.2 Fase - Aprofundamento de desafios.
  - Caso não seja possível ajuste, cancelar desafio. Finalizar processo justificando o cancelamento.

### 1.3.4 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | NIA |
|---|---|---|---|---|---|
| **Validação do desafio** | R | C | I | C | A |

Legenda: A (Autoridade responsável); I (Informado); R (Responsável Executor).

### 1.3.5 Recomendações, políticas e boas práticas

- **Transparência e Explicabilidade:** Avaliar a necessidade de explicabilidade do sistema de IA de acordo com o contexto de uso e o impacto potencial. Sistemas de alto risco que afetam direitos fundamentais devem priorizar modelos interpretáveis ou implementar mecanismos de explicação *pós-hoc*.
- **Avaliação de impacto socioeconômico:** Considerar não apenas os benefícios técnicos e econômicos, mas também os impactos sociais, incluindo potenciais efeitos sobre diferentes grupos populacionais, acessibilidade, inclusão digital e equidade.
- **Conformidade com PBIA:** Verificar alinhamento com os eixos estratégicos do **Plano Brasileiro de Inteligência Artificial**, incluindo capacitação, inovação, aplicação no setor produtivo e no poder público, e desenvolvimento socioeconômico responsável.

## 1.4 Fase - Definição de facilitador

A seleção de um especialista em facilitação deve priorizar, primeiramente, a fluência metodológica e a capacidade de adaptação. Mais do que o domínio rígido de frameworks como *Design Thinking* ou *Lean Inception*, o especialista precisa demonstrar capacidade para ajustar as dinâmicas à cultura e à maturidade do órgão.

Fundamentalmente, o facilitador deve possuir uma visão sistêmica que integre negócio/política pública, tecnologia e usuário. É crucial que ele atue como um tradutor competente, capaz de validar a viabilidade técnica junto à engenharia e arquitetura, enquanto alinha os impactos e resultados com os demandantes e a alta gestão. Essa competência impede a criação de escopos inalcançáveis e assegura que a solução desenhada seja não apenas desejável para o demandante, mas tecnicamente factível e financeiramente sustentável.

Por fim, o facilitador deve dominar a arte de fatiar o escopo, transformando ideias complexas em um MVP enxuto e um *backlog* priorizado pronto para o desenvolvimento.

O especialista em facilitação pode ser uma pessoa jurídica ou um servidor público capacitado para tal atividade.

### 1.4.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Definição de facilitação para estruturação da solução** | Escolha de pessoa jurídica, servidor capacitado no próprio órgão ou parceiro, com habilidades comprovadas para facilitação de desafios. |

### 1.4.2 O que verificar antes de avançar

Decisão sobre quem seria o facilitador para a Estruturação.

### 1.4.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | NIA | Facilitador |
|---|---|---|---|---|---|---|
| **Definição do facilitador** | R | I | A | I | C* | - |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor).

* Consultado, se necessário para apoio.

### 1.4.4 Recomendações, políticas e boas práticas

- Capacitação de servidores ou agentes - recomenda-se que cada órgão tenha em seu quadro servidores ou colaboradores com capacitação em métodos diversos de facilitação.
- Nesta fase, órgãos e entidades como os laboratórios de inovação do governo, o LIIA (Laboratório de Inovação em Inteligência Artificial) da ENAP (Escola Nacional de Administração Pública) são possibilidades para acorde de cooperação, além de parcerias com universidades.

## 2. Etapa - Estruturação do desafio de IA no serviço público

O objetivo principal desta etapa é **garantir a compreensão profunda do desafio** para que seja formulada uma proposta de solução de IA viável, realizando uma avaliação detalhada dos riscos e estabelecendo métricas claras de sucesso. Para isso, o facilitador define a Metodologia de Estruturação (ex: *Design Thinking* ou *Design Sprint*) e a aplica para estruturar o detalhamento dos desafios e necessidades da equipe técnica.

Nesta etapa, é fornecido acesso preliminar a dados para que seja realizada a análise da viabilidade da experimentação, detalhando a quem pertencem, formatos e sistemas envolvidos.

Nesta etapa a área de negócio ou política pública estabelece métricas claras e quantificáveis de sucesso, critérios mínimos de aceitação e cenários de teste. Um mapeamento de riscos e mitigação é realizado, abrangendo riscos de dados (sigilo, privacidade) e riscos associados ao comportamento do sistema (vieses, questões éticas), com possível colaboração do NIA. Com base nesta análise, o executor elabora a Proposta de Solução e Tecnologia, ajustando-a à infraestrutura de TI disponível. Por fim, é produzido o Plano de Experimentação, com detalhamento da abordagem de IA, a arquitetura de alto nível e as hipóteses a testar, culminando no Relatório de Estruturação, contendo o sumário, o cronograma e a recomendação fundamentada sobre o prosseguimento.

A estratégia definitiva de contratação ou desenvolvimento ("***Make or Buy***") é definida a partir dos resultados do MVP.

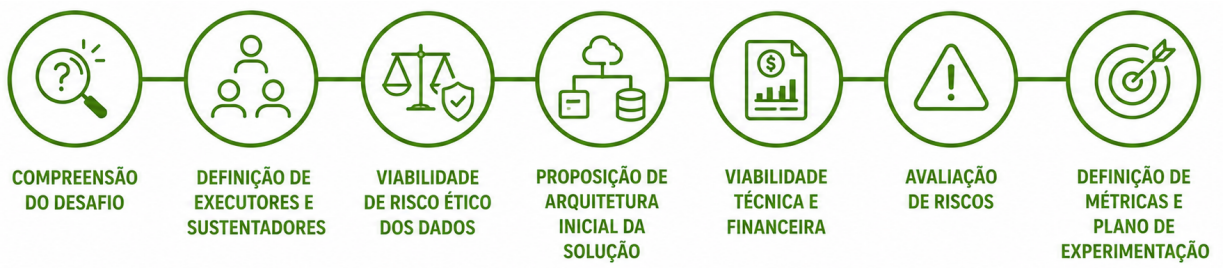


Figura 3 - Fases de Estruturação

## 2.1 Fase - Compreensão do desafio

A fase de compreensão do desafio tem como objetivo detalhar e estruturar a necessidade antes de propor a solução. Esta fase deve ser conduzida pelo facilitador selecionado, em colaboração com as áreas de negócio do órgão público. Independente da abordagem de facilitação, alguns passos são cruciais para o alcance dos objetivos desta fase:

- **Pesquisar a anterioridade,** levantando informações disponíveis sobre o desafio (como pesquisas anteriores, documentos e processos) para verificar se o problema exige uma abordagem específica de IA.
- **Criar um mapa ou processo atual do desafio,** identificando os *stakeholders*, validando o público-alvo ou a *persona* e detalhando suas dores e necessidades.
- **Selecionar a abordagem de estruturação** mais adequada (como *Design Sprint*, *Lean Inception*, *Canvas*, *Service Blueprint* ou BPM), identificando os participantes e providenciando a logística para rodar a metodologia. **Aplicar a abordagem de *design* selecionada para identificar e consolidar a proposta de solução**. A depender da metodologia, esta atividade envolve também a definição de hipóteses e a realização de testes/validação funcional com usuários reais para sua confirmação.
- **Realizar a validação da proposta de MVP** junto aos *stakeholders*.

Nesta fase deve ser definido se a solução será adquirida ou desenvolvida para que, na próxima fase, se necessário, definir a estratégia de parcerias de desenvolvimento e sustentação.

### 2.1.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Proposta de MVP** | Proposta de MVP elaborada e validada. |

### 2.1.2 O que verificar antes de avançar

A proposta de MVP deve ser validada junto aos *stakeholders*. Investir na implementação total sem antes testar as hipóteses estabelecidas para a proposta de MVP pode levar à criação de sistemas ociosos, que consomem orçamento e infraestrutura sem entregar o resultado esperado.

**Decisão**:

- **Aprovação. A aprovação demanda a resposta afirmativa dos envolvidos.**
  - **Ação.** Seguir para 3.2 Fase - Definição de executores e sustentadores.
- **Reprovação.**
  - **Ação.** Avaliar a possibilidade de iteração, executando novamente a fase 3.1Fase - Compreensão do desafio novamente.
  - Caso seja detectado, a partir dos resultados do MVP, a inviabilidade, é possível cancelar desafio. Finalizar processo justificando o cancelamento, sendo o motivo registrado para histórico.

### 2.1.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | NIA | Facilitador |
|---|---|---|---|---|---|---|
| **Definição de necessidade ou não de parcerias (Make or buy)** | A | C | I | R | C* | - |
| **Construção de proposta de MVP** | A | C | C | C | I** | R |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor); - (Não atua).

* O NIA atuará apenas para apoio, quando houver processo seletivo de soluções.

** Nesta, assim como nas demais matrizes RACI, o NIA pode ser "informado" por meio do sistema de organização e registro de requisição de IA (SORRIA), caso haja um processo seletivo de soluções que demande seu apoio.

### 2.1.4 Recomendações, políticas e boas práticas

Participação das áreas de negócio ou políticas públicas do órgão nas próximas etapas do processo, com o objetivo de entender a solução proposta e apontar riscos e restrições sobre desenvolvimento e/ou sustentação que estarão sob sua responsabilidade.

*Design* **Centrado no Usuário.** Aplicar princípios de *design* centrado no usuário, garantindo que a solução seja desenvolvida com foco nas necessidades reais dos cidadãos ou servidores, considerando acessibilidade (WCAG 2.1), usabilidade e linguagem simples.

**Validação Incremental.** Adotar abordagem iterativa com validações frequentes junto aos *stakeholders*, permitindo ajustes rápidos e reduzindo riscos de desenvolvimento de soluções inadequadas.

A **decisão "*Make or Buy*"** analisada na elaboração do MVP, deve considerar a interoperabilidade (nos padrões ePING[5]) e o ciclo de vida da solução, conforme a norma ISO/IEC 5338 - *AI*

5 Disponível em: eping.governoeletronico.gov.br.

*system life cycle processes*[6].

## 2.2 Fase - Definição de executores e sustentadores

A estratégia de aquisição ou parceria (definida em "*Make or Buy*") para a execução da experimentação e para a sustentação da solução é elaborada somente caso tenha sido aprovada em fases anteriores. Nesta fase, será chamado de executor o ator responsável pela execução da experimentação, seja ele o órgão, o parceiro ou o terceiro contratado. A busca pode ser realizada em uma lista ou cadastro de parceiros já homologados.

Neste ponto, tanto o executor de desenvolvimento da experimentação quanto o sustentador devem ser definidos para que participem das validações tecnológicas. A definição do sustentador nesta etapa não é um detalhe administrativo: projetos de IA que chegam ao fim do desenvolvimento sem um sustentador identificado e envolvido tendem a ser descontinuados — não por falha técnica, mas porque quem vai operar a solução não acompanhou sua construção e não se apropria dela, principalmente por incopatiblidade de padrões e componentes tecnológicos não aprovados.

A escolha de parceiros em projetos de IA deve ir além de critérios comerciais ou de capacidade de entrega imediata, considerando de forma sistemática a maturidade técnica, a governança de dados e o alinhamento ético do fornecedor. Em ambientes organizacionais complexos, parceiros frequentemente têm acesso a dados sensíveis, influenciam decisões técnicas críticas e podem introduzir riscos que se propagam por todo o ciclo de vida do sistema de IA.

### 2.2.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Execução da estratégia de aquisição ou parceria** | Estratégia elaborada para as demais etapas, incluindo aquisições ou parcerias necessárias para execução e sustentação. No caso de internalização da solução pelo próprio órgão, a área de TI torna-se o sustentador da solução. |
| **Parceiro de execução e sustentação contratados** | Os parceiros definidos para execução e sustentação da solução são contratados. |

### 2.2.2 O que verificar antes de avançar

As análises e a estratégia devem indicar as alternativas e decisões quanto à aquisição ou parcerias necessárias.

6 Disponível em: www.iso.org/standard/81118.html.

### 2.2.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | NIA | Executor | Sustentador |
|---|---|---|---|---|---|---|---|
| **Definição da parceria** | A | I | I | R | C* | - | - |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor).

* O NIA pode ser consultado ou informado, caso tenha sido acionado pelo demandante para colaborar.

### 2.2.4 Recomendações, políticas e boas práticas

**Participação ativa.** No caso da opção de parceiros, representantes designados por estes devem participar ativamente das próximas etapas do processo, com o objetivo de entender a solução proposta e apontar riscos e restrições sobre desenvolvimento e/ou sustentação que estarão sob sua responsabilidade.

**Evitar *Vendor Lock-in*.** Ao selecionar parceiros, priorizar soluções baseadas em padrões abertos e que garantam portabilidade de dados e modelos. Estabelecer cláusulas contratuais que preservem a autonomia do órgão na gestão da solução. A solução deve:

- **Mitigar aprisionamento tecnológico (*Vendor Lock-in*)**, o que inclui priorizar padrões abertos e evitar dependência exclusiva de um único fornecedor ou nuvem. O objetivo é garantir que o sistema possa ser migrado ou adaptado sem custos proibitivos, protegendo o órgão caso o fornecedor aumente preços ou descontinue ferramentas essenciais.
- **Garantir a sustentabilidade da manutenção**, criando um projeto estruturado para ter continuidade, independentemente da equipe ou da empresa contratada. Isso envolve código auditável, documentação completa e uso de tecnologias consolidadas no mercado, evitando sistemas fechados ("caixas-pretas") que impeçam atualizações futuras ou correções pelo próprio órgão.

***Due Diligence* de Segurança.** Realizar avaliação de segurança dos parceiros potenciais, verificando certificações relevantes **(ISO 27001, SOC 2)**, práticas de desenvolvimento seguro e histórico de incidentes de segurança.

**Geração de SBOM (*Software Bill of Materials*) [5].** Considerando-se que apenas *Due Diligence* de Segurança manual é ineficaz contra os ataques sofisticados de cadeia de suprimentos, são requisitos mandatórios a geração de SBOM e assinaturas digitais. Sem um **SBOM automatizado** (*CycloneDX*), o órgão não saberá se uma biblioteca usada na Fase de Experimentação possui uma vulnerabilidade crítica (CVE) recém-descoberta. Sem assinatura digital (*Sigstore/Cosign*), um modelo aprovado na homologação pode ser substituído por um modelo malicioso (envenenado) antes de entrar em produção sem ser detectado. Sendo assim, recomenda-se que nenhum modelo, biblioteca ou componente externo, seja integrado implantado ou colocado em produção sem a entrega de um SBOM no padrão CycloneDX (**padrão aberto** de SBOM, mantido pela OWASP, amplamente adotado em práticas de ***DevSecOps, supply chain security***

***e compliance*). Implementar a exigência de SBOM (CycloneDX) significa **garantir visibilidade, rastreabilidade e controle de segurança** sobre **todos os componentes externos**, incluindo **modelos de IA**, antes que eles sejam usados em sistemas produtivos.

**Transferência de conhecimento.** Garantir que acordos de parceria incluam provisões para transferência de conhecimento e capacitação das equipes internas, visando a autonomia futura do órgão na operação e evolução da solução.

## 2.3 Fase - Viabilidade e risco ético dos dados

Esta fase, focada na **viabilidade e risco ético dos dados**, tem como objetivo detalhar os ativos de dados necessários para a solução e os riscos associados. O processo inicia-se com o detalhamento dos dados disponíveis, identificando seus responsáveis (guardiões ou curadores), formato, volume mínimo e mecanismo de extração.

É realizada a **estimativa de custo de nuvem/inferência**, aspecto crucial onde muitos projetos encontram barreiras. Recomenda-se utilizar o *FinOps Framework*[7] e calculadoras oficiais de provedores para fundamentar a previsão orçamentária do ETP, baseada em métricas de mercado (*tokens* ou horas de GPU).

Tão importante quanto a estimativa é a atenção necessária aos **dados sensíveis e restritos**, no caso de utilização de soluções em nuvem, pois neste caso , **como regra geral**, devem estar restritos à **nuvem de governo**[8]. Esses dados devem ser hospedados nesse tipo de nuvem para garantir seu armazenamento físico em solo brasileiro, garantindo a soberania operacional sobre eles.

Prossegue-se com a **verificação do nível de segurança**, que envolve a classificação dos dados quanto ao sigilo e à sensibilidade, a determinação da necessidade de anonimização e a validação da conformidade com a LGPD e com outros marcos regulatórios relevantes. Após essa análise, deve-se **formalizar a logística de acesso** para garantir que os insumos estejam disponíveis para a fase de Experimentação. Por fim, a etapa de Estruturação é concluída com a consolidação do relatório de análise de dados e a realização da **AIE completa - classificação de risco preliminar (baixo, médio, alto, extremo)**. É essencial definir papéis e responsabilidades claras para avaliação ética e de dados, incluindo responsáveis técnicos, jurídicos e de negócio.

7 FinOps Framework é uma metodologia de gestão que une as equipes financeira, de operações e de engenharia. O objetivo é dar visibilidade e controle sobre os gastos variáveis da nuvem.

8 Esse conceito foi estabelecido a partir da publicação da Portaria SGD nº 5950/2023.

### 2.3.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Relatório de análise de dados** | Documento consolidado que detalha:<br>• **Viabilidade técnica dos dados.** Contendo conter informações sobre os responsáveis (guardião/curador), o formato, o volume mínimo e o mecanismo de extração.<br>• **Classificação de segurança e conformidade.** Registro que classifica os dados quanto ao sigilo e à sensibilidade. Inclui a determinação da necessidade de anonimização e a verificação de conformidade com a LGPD e com o marco regulatório do órgão.<br>• **Formalização de acesso.** Documentação que estabelece a logística de acesso aos dados para a fase posterior de experimentação. |
| **ETP - Estudo Técnico Preliminar** | Documento que demonstra a viabilidade técnica e econômica da solução proposta (neste caso, a solução de IA), justificando a necessidade da contratação e definindo os requisitos básicos do projeto. Ele serve de base para o Termo de Referência ou Projeto Básico subsequente. Neste ponto o ETP é usado para responder como será contratado o que já foi decidido contratar, por exemplo. |
| **Autoavaliação de Impacto Ético (AIE) - classificação de risco** | Realização da Autoavaliação de Impacto Ético, utilizando a segunda parte do *framework* específico. |

### 2.3.2 O que verificar antes de avançar

Para que o projeto avance para a fase de Proposta de Solução, todos os itens abaixo devem ser avaliados positivamente:

- **Detalhamento técnico:** o formato, volume mínimo e mecanismo de extração dos dados foram definidos?
- **Custódia e responsabilidade:** o responsável (guardião/curador) pelos dados foi devidamente identificado?
- **Classificação de segurança:** os dados foram classificados quanto ao sigilo e à sensibilidade?
- **Conformidade legal:** o uso proposto está em conformidade com a LGPD (quando aplicável) e com o marco regulatório do órgão?
- **Necessidade de anonimização:** a necessidade de anonimização foi determinada e o plano para executá-la está definido?
- **Garantia de acesso:** a logística de acesso para a fase de Experimentação foi formalizada?
- **Impacto ético:** a AIE completa (parte 2) foi realizada e os riscos são aceitáveis e/ou mitigáveis?

**Decisão:**

- **Aprovação.**
  - Ação. Seguir para 3.4 Fase - Proposição de arquitetura inicial da solução.

- **Reprovação.** Seja por risco ético extremo, não conformidade legal, impossibilidade de anonimização ou custos acima do orçamento.
    - **Ação**. Deve ser avaliada a possibilidade de ajuste para posterior aprovação 2.2 Fase - Aprofundamento de desafios, ou a descontinuação do projeto por inviabilidade de resolução.

### 2.3.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor deTIC | ETDP | Gestor de Segurança da Informação | NIA | Executor | Sustentador |
|---|---|---|---|---|---|---|---|---|---|
| **Detalhar dados (formato, volume, extração)** | I | R | I | A | I | - | - | I | I |
| **Verificar nível de segurança necessário (sigilo, LGPD e anonimização)** | C | R | I | C | A | A | - | I | I |
| **Formalizar acesso (logística para experimentação)** | A | R | I | C | C | C | - | I | I |
| **Consolidar relatório de análise** | A | R | I | C | C | C | - | I | I |
| **Realizar AIE complementar (Parte 2)** | C | R | A | C | I | I* | I* | I | I |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor).

* Informado via sistema de organização e registro de soluções de IA.

### 2.3.4 Recomendações, políticas e boas práticas

**Avaliar riscos de segurança.** Identificar onde os dados sensíveis estão armazenados e quais proteções são necessárias. A arquitetura de segurança deve ser planejada para evitar vulnerabilidades como o abuso de privilégios e a vulnerabilidade à injeção de SQL.

**Incorporar requisitos de segurança.** As necessidades de segurança devem ser capturadas na fase de análise do projeto (considerando o SDLC - *Software Development Life Cycle*) para evitar a necessidade de adaptações posteriores ("*retrofit*").

**Realizar avaliação inicial de qualidade dos dados**. Focar em um conjunto de dados pequeno e crítico para quantificar as não-conformidades, identificar riscos e formular hipóteses sobre as causas-raiz dos problemas.

**Avaliar risco por contexto de uso.** Os riscos devem ser avaliados considerando o contexto de uso do modelo, o público afetado e o grau de autonomia da decisão (informativo, recomendação ou decisão automatizada), pois o mesmo dado pode ser aceitável em um projeto e crítico

em outro. Por exemplo, um modelo usado apenas para apoio interno pode ter uma avaliação diferente do modelo que afeta diretamente a experiência do cidadão.

**Escolher fontes de dados.** Identificar e avaliar fontes de dados (internas e externas) necessárias para preencher lacunas analíticas, priorizando fontes confiáveis e autorizadas.

**Adquirir e ingerir fontes.** Planejar a ingestão dos dados (*load* e *transform*) e garantir a captura do metadado crítico (origem, tamanho, formato, segurança) para catalogar o *data lake* e evitar que se torne um "pântano de dados".

**Governança ética.** Considerar o impacto ético dos modelos e das análises para evitar vieses não intencionais e garantir que as análises não violem a privacidade.

**Conformidade com LGPD.** Verificar a conformidade com a Lei Geral de Proteção de Dados (Lei 13.709/2018), garantindo:

- base legal adequada para tratamento de dados (Art. 7º e 11 da LGPD);
- implementação de princípios de finalidade, adequação, necessidade e minimização;
- garantia de direitos dos titulares (acesso, correção, anonimização, eliminação);
    - avaliação da necessidade de Relatório de Impacto à Proteção de Dados Pessoais (RIPD), conforme orientações da ANPD.

**Classificação de Dados.** Adotar esquema de classificação de sensibilidade, alinhado com:

- Decreto nº 7.845/2012 (informações sigilosas no âmbito da Administração Pública Federal);
- Política de segurança da informação do órgão;
    - Categorias especiais de dados da LGPD (Art. 5º, II e dados sensíveis Art. 11).

***Frameworks* e normas aplicáveis**

- **Normativos do GSI/PR.** O Gabinete de Segurança Institucional da Presidência da República (GSI/PR) emite diretrizes relacionadas à segurança da informação e cibernética que se aplicam ao uso de tecnologias emergentes, incluindo a IA, na administração pública federal.
- **NIST AI RMF (*AI Risk Management Framework*).** *Framework* voluntário estruturado em quatro funções principais - *Govern* (Governar), *Map* (Mapear), *Measure* (Medir) e *Manage* (Gerenciar) - para incorporar considerações de confiabilidade no *design*, desenvolvimento, uso e avaliação de sistemas de IA.
- **ABNT NBR ISO/IEC 42001:2024.** Norma brasileira publicada em abril de 2024 que especifica requisitos para estabelecer, implementar, manter e melhorar continuamente um Sistema de Gestão de Inteligência Artificial (SGIA). Segue a Estrutura de Alto Nível (HLS), facilitando integração com ISO 27001 e ISO 9001. Aborda desafios de ética, transparência, aprendizado contínuo e gestão de riscos específicos de IA.
- **ISO/IEC 5338.** Norma internacional que define processos de ciclo de vida para sistemas de IA cobrindo planejamento, desenvolvimento, implantação, operação, manutenção e descontinuação.
- **ISO/IEC 23894**: Tecnologia da informação – Inteligência artificial – Orientações sobre gestão de riscos. Processos de gestão de riscos (identificação, análise e mitigação) em projetos de IA. Orienta as organizações sobre como identificar, avaliar e tratar os riscos que surgem ao longo de todo o ciclo de vida de um sistema de IA.

**Recomendações técnicas**

Nos casos em que são usados LLMs, recomenda-se verificar vulnerabilidades listadas no *OWASP Top 10 for LLM Applications [1]*. Ao decidir entre desenvolver uma solução de IA internamente (*make*) ou utilizar componentes externos (*buy*), deve-se ter como pontos de atenção:

- **Escolha do Modelo.** É preciso garantir que a estratégia de desenvolvimento (treinar um novo modelo, usar um modelo existente com ou sem ajuste fino ou acessar via API externa) seja **adequada às necessidades** (requisitos e restrições) do projeto.
- **Avaliação de fornecedores:** Ao optar por um fornecedor de modelos externo, deve-se realizar uma **avaliação de *due diligence* da postura de segurança** desse fornecedor.
- **Bibliotecas e componentes externos.** Ao utilizar bibliotecas externas, é obrigatória a realização de *due diligence* para garantir que a biblioteca possua controles que **impeçam o carregamento de modelos não confiáveis** e evitem a execução imediata de código arbitrário.
- **Importação de modelos.** Modelos de terceiros ou pesos serializados devem ser tratados como **código de terceiros não confiável**. É necessário implementar **varredura e isolamento (*sandbox*)** ao importá-los, pois podem permitir a execução remota de código.
- **Uso de API externa.** Caso a opção seja uma API externa, devem ser aplicados **controles apropriados sobre os dados enviados** a serviços fora do controle da organização. Isso inclui exigir que os usuários façam *login* e confirmem antes de enviar informações potencialmente confidenciais.
- **Integridade dos dados.** É fundamental aplicar **verificações e higienização adequadas aos dados e entradas**, incluindo o *feedback* do usuário ou os dados de aprendizado contínuo, reconhecendo que os dados de treinamento definem o comportamento final do sistema.
- **Segurança em IA.** Consultar o ***OWASP Top 10 for LLM Applications*** para identificar vulnerabilidades comuns em aplicações baseadas em modelos de linguagem, incluindo: injeção de *prompt*, envenenamento de dados de treinamento, negação de serviço do modelo e vazamento de informações sensíveis.
- **Ameaças adversariais.** Utilizar o **MITRE ATLAS (*Adversarial Threat Landscape for AI Systems*)** [2] como referência para identificar táticas e técnicas de ataques específicos a sistemas de IA incluindo roubo de modelo, envenenamento de dados e ataques de evasão. O ATLAS documenta 14 táticas e 82 técnicas baseadas em ataques reais.
- **Interoperabilidade.** Garantir aderência aos **Padrões de Interoperabilidade de Governo Eletrônico (ePING)**, que definem premissas, políticas e especificações técnicas para o uso de TIC na administração pública federal, cobrindo cinco segmentos: interconexão, segurança, meios de acesso, organização e intercâmbio de informações e áreas de integração.
- **Ciclo de vida de sistemas de IA.** Adotar a **ISO/IEC 5338** para os processos de ciclo de vida de sistemas de IA garantindo o planejamento adequado desde a concepção até a descontinuação, de modo a evitar *vendor lock-in* e facilitar a manutenção.
- **Otimização de custos de nuvem.** Aplicar princípios do ***FinOps Framework*** para gestão financeira de nuvem, incluindo: visibilidade em tempo real de custos, otimização contínua

de recursos e colaboração entre equipes de negócio, tecnologia e finanças. Para estimativas iniciais recomenda-se acessar Portaria que descreve a modalidade de remuneração baseada em Unidade de Serviços de Nuvem (USN), utilizando Multicatálogo de serviços de Computação em Nuvem Padronizados[9].

## 2.4 Fase - Proposição de arquitetura inicial da solução

O objetivo principal desta fase é o desenvolvimento do **desenho conceitual da solução de IA** estabelecendo a abordagem técnica candidata que será testada na fase de experimentação. O processo envolve a elaboração de uma proposta que identifica as **tecnologias mais adequadas** (como IA Generativa/LLM ou Modelos de Classificação) e verifica a existência de **componentes para reuso** dentro da plataforma do NIA. Esta definição técnica é o que guiará o detalhamento dos recursos e custos na fase subsequente.

A proposta arquitetural da solução gerada nesta fase consolida a visão técnica da solução de IA proposta.

### 2.4.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Proposta inicial de solução de IA** | O documento deve conter:<br>• O desenho conceitual e a abordagem técnica;<br>• **Identificação de tecnologias** e ferramentas específicas selecionadas para o desafio;<br>• **Identificação de infraestrutur**a estimada para experimentação e escalabilidade para produção;<br>• **Análise de reuso de componente**s, indicando quais ativos da plataforma do NIA podem ser aproveitados no projeto. |

### 2.4.2 O que verificar antes de avançar

Para validar os resultados dessa fase e aprovar o avanço no processo as questões a seguir devem ser consideradas.

- **Sobre o Desenho Conceitual e Abordagem Tecnológica:**
  - **Separação de Camadas:** O desenho conceitual separa claramente a arquitetura nas três frentes obrigatórias: **Camada de IA** (modelos e dados), **Camada aplicacional** (interface e lógica) e **Camada de governança** (auditoria e monitoramento)?

9 Portaria SGD/MGI n° 5.950, de 2023

- **Adequação da Tecnologia:** A lista de tecnologias selecionadas (ex: IA Generativa ou IA Clássica) justifica tecnicamente como resolverá o desafio proposto?
- **Padrões de Projeto:** O desenho adotou padrões arquiteturais adequados para a escalabilidade, como o encapsulamento em *Model-as-a-Service*?

- **Sobre Otimização e Reuso:**
  - **Validação de Reuso (NIA):** O relatório de análise verificou o ecossistema atual e listou explicitamente quais ativos e componentes da plataforma do Núcleo de IA (NIA) serão aproveitados para evitar desenvolvimento redundante?
- **Sobre Segurança e Privacidade (*Security by Design*):**
  - **Soberania Digital e Residência de Dados:** O desenho garante que dados sensíveis ou sigilosos serão processados e armazenados priorizando infraestruturas nacionais (nuvem de governo)?
  - **Isolamento de Ambientes e Criptografia:** A arquitetura prevê criptografia (em trânsito e repouso) e o isolamento claro entre os ambientes de desenvolvimento, homologação e produção?
  - **Controles de Acesso:** O princípio do menor privilégio e mecanismos robustos de autenticação foram incorporados ao *design*?
- **Sobre Observabilidade e Governança Contínua:**
  - **Detecção de *Drift*:** A arquitetura técnica contempla ferramentas para monitoramento de derivação de dados (*data drift*) e de conceito (*concept drift*)?
  - **Auditoria e Desempenho:** Há previsão de *logs* estruturados para rastreamento de predições, *feedback* e coleta de métricas de performance (latência, *throughput*)?

**Decisão:**

- **Aprovação.** Todas as questões foram respondidas afirmativamente. A visão técnica está sólida o suficiente para orientar o **detalhamento de recursos de TI, equipe e custos de nuvem (se houver)** na fase subsequente.
  - **Ação.** Seguir para a 3.5 **Fase - Viabilidade técnica e financeira.**
  - **Reprovação.** Uma ou mais questões foi respondida negativamente.
  - **Ação.** O ítem avaliado negativamente deve ser reavaliado e uma solução arqutitural deve ser proposta para nova avalidação.

### 2.4.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de segurança da Informação | NIA | Executor | Sustentador |
|---|---|---|---|---|---|---|---|---|---|
| **Desenho conceitual da solução de IA e abordagem técnica** | I | C | - | A | C | C | - | R | C |

Legenda: C (Consultado); I (Informado); R (Responsável Executor); - (Não atua).

### 2.4.4 Recomendações, políticas e boas práticas

**Arquitetura em Camadas.** Separar claramente as camadas da solução em:

- Camada de IA (*core*): Modelos, dados, treinamento, inferência;
- Camada aplicacional: Interface, orquestração, lógica de negócio;
- Camada de governança: Auditoria, logs, monitoramento, explicabilidade.

**Padrões de Projeto para IA.** Considerar padrões arquiteturais apropriados:

- *Model-as-a-Service*: Encapsular modelos como serviços independentes.

**Observabilidade e monitoramento.** Incluir no desenho arquitetural:

- Monitoramento de performance do modelo (acurácia, latência, *throughput*);
- Detecção de derivação de dados e conceito (*data/concept drift*);
- Rastreamento de predições e *feedback* para aprendizado contínuo;
- *Logs* estruturados para auditoria e depuração;

**Segurança desde o projeto (*by Design*).** Incorporar controles de segurança desde a concepção:

- Isolamento de ambientes (desenvolvimento, homologação, produção);
- Princípio do menor privilégio para acesso a dados e modelos;
- Criptografia de dados em trânsito e em repouso;
- Mecanismos de autenticação e autorização robustos;
- Definição da residência dos dados (Soberania Digital), priorizando infraestruturas nacionais para dados sensíveis ou sigilosos.

**Infraestrutura estimada**. Definir pelo menos:

- O tipo de modelo (generativo, preditivo, visão computacional etc.);
- Onde vai processar (nuvem, on-premise, híbrido);
- O volume estimado de dados e requisições;

- Se há necessidade de GPU/TPU ou compute convencional.

**Model Card e Transparência do Modelo**

Como parte das boas práticas de desenvolvimento responsável de sistemas de IA, recomenda-se a elaboração de um *Model Card* — documento de transparência que acompanha o modelo e descreve suas características essenciais para todas as partes interessadas (desenvolvedores, gestores, auditores e usuários finais).

O *Model Card* deve contemplar, no mínimo:

- **Descrição e finalidade do modelo** Definição clara do propósito para o qual o modelo foi desenvolvido, seus casos de uso pretendidos e, explicitamente, os casos de uso para os quais ele **não** deve ser utilizado.
- **Dados de treinamento e distribuição entre grupos** Descrição da origem e composição dos dados utilizados no treinamento, incluindo a distribuição entre grupos relevantes — como região geográfica, faixa etária, gênero, idioma ou qualquer outra dimensão pertinente ao contexto da aplicação. A ausência ou sub-representação de determinados grupos nos dados deve ser explicitamente declarada, pois pode resultar em desempenho desigual ou discriminatório do modelo.
- **Vieses identificados e limitações conhecidas** Registro transparente de eventuais vieses detectados durante o desenvolvimento, avaliação ou operação do modelo, acompanhado das medidas adotadas para mitigação. Limitações técnicas ou contextuais do modelo também devem ser documentadas, evitando expectativas inadequadas sobre seu desempenho.
- **Métricas de desempenho desagregadas** Apresentação dos resultados de avaliação do modelo segmentados por grupos ou subpopulações relevantes, de modo a evidenciar possíveis disparidades de desempenho que não seriam visíveis em métricas agregadas.
- **Considerações éticas e de equidade** Análise dos riscos de impacto diferenciado sobre grupos ou populações afetadas pelo modelo, com indicação das salvaguardas adotadas para garantir tratamento justo e não discriminatório.
- **Responsáveis e ciclo de revisão** Identificação dos responsáveis pela elaboração, validação e atualização do *Model Card*, bem como a periodicidade prevista para sua revisão, especialmente diante de mudanças no modelo, nos dados ou no contexto de uso.

A elaboração do *Model Card* deve ser tratada como requisito do processo de desenvolvimento, e não como etapa opcional, sendo atualizada a cada nova versão do modelo disponibilizada em ambiente produtivo.

Esta recomendação está ancorada em um conjunto consolidado de referenciais normativos nacionais e internacionais, descritos a seguir:

**NIST AI Risk Management Framework (AI RMF) — NIST AI 100-1 (2023)** Framework do Instituto Nacional de Padrões e Tecnologia dos Estados Unidos que estrutura a gestão de riscos de IA em quatro funções: *Governar*, *Mapear*, *Medir* e *Gerenciar*. Orienta a documentação de características do modelo, rastreabilidade de dados e monitoramento contínuo de desempenho e vieses como elementos centrais de um sistema de IA confiável.

**ISO/IEC 42001:2023 — Sistemas de Gestão de Inteligência Artificial** Primeira norma internacional de sistema de gestão voltada especificamente para IA, publicada pela Organização Internacional de Normalização. Estabelece requisitos para que organizações desenvolvam, implantem e operem sistemas de IA de forma responsável, incluindo controles sobre transparência, gestão de riscos, impacto social e governança do ciclo de vida do modelo.

**ISO/IEC 23894:2023 — Gestão de Riscos em IA** Norma complementar à ISO/IEC 42001 que fornece diretrizes específicas para identificação, análise e tratamento de riscos associados a sistemas de IA, com atenção a riscos de viés, falhas de desempenho e impactos sobre direitos fundamentais.

**ISO/IEC 25059:2023 — Qualidade de Sistemas de IA (série SQuaRE)** Extensão da série de normas de qualidade de software (ISO/IEC 25000) ao contexto de IA, definindo características de qualidade específicas como explicabilidade, robustez, equidade e rastreabilidade, que devem ser mensuradas e documentadas ao longo do ciclo de vida do modelo.

**ISO/IEC 38507:2022 — Governança de TI e Implicações do Uso de IA** Norma que aborda as implicações para a governança organizacional decorrentes do uso de sistemas de IA, orientando conselhos e alta direção sobre responsabilidades, supervisão e prestação de contas no emprego dessas tecnologias. Referenciada pelo próprio TCU em suas diretrizes internas de governança de IA.

**Estratégia Brasileira de Inteligência Artificial — EBIA (2024–2028)** Documento orientador da Administração Pública Federal brasileiro que estabelece diretrizes para o desenvolvimento ético, seguro e inclusivo da IA no país, com ênfase em soberania digital, não discriminação, transparência algorítmica e proteção de dados. Alinha-se diretamente aos princípios da Lei Geral de Proteção de Dados (LGPD — Lei nº 13.709/2018), que impõe obrigações de transparência sobre decisões automatizadas que afetem titulares de dados. O acompanhamento da implementação da EBIA está sob monitoramento do TCU, por força do Acórdão nº 1.139/2022 – Plenário.

**Regulamento Europeu de Inteligência Artificial — EU AI Act (Regulamento UE 2024/1689)** Primeiro marco regulatório vinculante de IA no mundo, em vigor desde agosto de 2024. Classifica sistemas de IA por nível de risco e impõe, para sistemas de alto risco, obrigações expressas de documentação técnica, transparência, rastreabilidade de dados de treinamento, monitoramento pós-implantação e registro em base de dados europeia. Embora de aplicação direta na União Europeia, constitui referência regulatória relevante para organizações brasileiras com operações internacionais ou que adotem padrões globais de governança.

**CGI.br — Comitê Gestor da Internet no Brasil: Resolução CGI.br/RES/2009/003/P e atuação em IA** O CGI.br, instância multissetorial responsável por estabelecer diretrizes estratégicas para o uso e desenvolvimento da Internet no Brasil, tem atuação crescente na governança de IA. Seus dez Princípios para a Governança e Uso da Internet (Resolução CGI.br/RES/2009/003/P, 2009) — que incluem liberdade, privacidade, direitos humanos, universalidade, diversidade, neutralidade e transparência — constituem base normativa nacional diretamente aplicável ao desenvolvimento responsável de sistemas de IA. Por meio de sua Resolução CGI.br/RES/2024/049, o Comitê incluiu a inteligência artificial entre os temas centrais de sua agenda estratégica até 2027, reforçando a necessidade de análise prévia de risco, proteção de direitos fundamentais e abordagem regulatória assimétrica e baseada em riscos para sistemas de IA. O Observatório Brasileiro de Inteligência Artificial (OBIA), coordenado pelo NIC.br e CGI.br, atua

como repositório de documentos orientadores e como ponto de convergência de conhecimento sobre adoção de IA no Brasil.

**TCU — Tribunal de Contas da União: Diretrizes para IA na Administração Pública** O TCU consolidou-se como referência nacional na fiscalização e orientação do uso de IA na Administração Pública Federal. Seu conjunto de deliberações e publicações institui um corpo de diretrizes de observância recomendada para organizações públicas e, por extensão, para seus fornecedores de soluções tecnológicas:

- **Acórdão nº 1.139/2022 – Plenário (TC 006.662/2021-8):** Determinou ao TCU que avalie periodicamente o nível de maturidade dos órgãos federais no uso de IA, que desenvolva referencial metodológico próprio para auditoria de algoritmos e aplicações de IA, e que elabore guia de diretrizes, parâmetros e riscos para orientar líderes e gestores públicos na implementação ou contratação de soluções de IA.
- **Guia de Uso de IA Generativa no TCU (2023):** Estabelece orientações para uso responsável e ético da IA generativa, cobrindo proteção de dados, confidencialidade, revisão humana obrigatória de conteúdos gerados e restrições ao uso de ferramentas externas não homologadas. Constitui referência de boas práticas para toda a Administração Pública Federal, tendo sido apresentado a mais de 80 instituições federais.
- **Acórdãos nº 1.868/2024, 2.198/2024, 2.493/2024 e 2.591/2024 – Plenário:** Consolidam a jurisprudência do TCU incentivando e orientando o uso estratégico, ético e governado da IA em diferentes setores da gestão pública, com ênfase em eficiência, transparência, mitigação de erros e equidade na prestação de serviços.

**Recomendação da UNESCO sobre Ética da IA (2021)** Primeiro instrumento normativo global sobre ética em IA, adotado por 193 países membros, incluindo o Brasil. Estabelece princípios de transparência, responsabilização, equidade, privacidade e supervisão humana como fundamentos para o desenvolvimento e uso responsável de sistemas de IA.

**Mitchell et al. — Model Cards for Model Reporting (2019)** Artigo seminal publicado na *ACM Conference on Fairness, Accountability, and Transparency (FAccT)* que propôs formalmente o conceito de *Model Card* como instrumento de comunicação transparente sobre modelos de aprendizado de máquina. Referência metodológica base para esta recomendação. DOI: https://doi.org/10.1145/3287560.3287596

## 2.5 Fase - Viabilidade técnica e financeira

O objetivo desta fase é consolidar todos os insumos necessários para avaliar a viabilidade do projeto, tendo como base a proposta de solução e a análise de reuso de componentes existentes.

O processo começa com o levantamento detalhado dos recursos tecnológicos e humanos necessários para o desenvolvimento, a implementação e a sustentação futura da solução — sejam eles do próprio órgão ou de parceiros. É neste momento que a arquitetura inicial se traduz em necessidades concretas de infraestrutura: tipo de processamento, ambientes (desenvolvimento, homologação e produção), volumes estimados de dados e requisições, e eventuais dependências de GPU ou recursos de nuvem.

A responsabilidade por esse levantamento é compartilhada e bem delimitada. À equipe técnica de IA cabe identificar as tecnologias necessárias e estimar volumes de uso — não dimensionar a infraestrutura. O dimensionamento e a verificação de viabilidade técnica são responsabilidade da área de TI do órgão ou do parceiro executor e/ou sustentador, que conhece o ambiente onde a solução vai operar. Ao órgão cabe identificar e confirmar o orçamento disponível para arcar com esses recursos.

Com essas informações consolidadas, apuram-se os custos totais do projeto — incluindo custos variáveis de nuvem e *tokens* — e realiza-se uma análise orçamentária que exige, obrigatoriamente, a participação do representante responsável pela sustentação. Esse alinhamento é o que garante que a aprovação financeira reflita a realidade do órgão e não apenas uma estimativa técnica desconectada da capacidade de execução.

### 2.5.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Levantamento de recursos e custos** | **Levantamento de recursos e custos.** Detalhamento sobre a infraestrutura de TI, a equipe técnica necessária e a estimativa de custos totais da solução. Para identificação de custos, deve ser especificada qual infraestrutura será fisicamente/virtualmente necessária, quanto ela custará e se o órgão tem os recursos e o orçamento para sustentá-la. |
| **Relatório de análise orçamentária** | Parecer sobre a disponibilidade de orçamento e a viabilidade financeira do projeto. |

### 2.5.2 O que verificar antes de avançar

Para que a viabilidade seja aprovada, a lista de checagem a seguir deve ser verificada:

- **Recursos tecnológicos.** Os recursos e infraestrutura de TI necessária para a sustentação foi totalmente mapeada?
- **Recursos humanos.** Os técnicos (próprios ou parceiros) que manterão a solução estão identificados?
- **Custo estimado da solução.** O custo total estimado (incluindo nuvem, inferência e pessoal) foi apurado?
- **Disponibilidade financeira.** Há orçamento confirmado para a execução e sustentação futura?
- **Participação da Sustentação.** O responsável pela sustentação futura validou os custos e recursos?

**Decisão:**

- **Aprovação.** Os recursos foram identificados para execução e sustentação.
  - **Ação.**Seguir para a próxima fase.
- **Reprovação.** Devido à insuficiência de recursos
  - **Ação**. Deve-se verificar a possibilidade de ajuste. Se não houver meios de sustentação técnica, ele deve ser encerrado.

**Atenção.** Nesta fase, se a definição do sustentador estiver vinculada a um processo licitatório, não seria possivel consultá-lo para fins de apuração de custos e recursos, o que torna-se um **risco a ser gerenciado no projeto**.

### 2.5.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Executor | Sustentador |
|---|---|---|---|---|---|---|---|---|---|
| **Apurar Recursos e Custos** | A | C | C | C | C | C | - | R | C |
| **Análise Orçamentária** | R | I | A | C | - | - | I* | C | C |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor).

* Informado via sistema de organização e registro de soluções de IA.

### 2.5.4 Recomendações, políticas e boas práticas

Definir o **método de avaliação do ativo de dados.** A Governança de Dados deve **definir e padronizar o método de avaliação** para quantificar consistentemente o valor dos ativos de dados. Isso é fundamental para justificar o investimento.

- O custo do risco legal ou regulatório, como **penalidades potenciais, custos de remediação e despesas com litígios**, deve ser avaliado e ser compensado pelo custo da intervenção operacional para proteger os dados.
- O orçamento de viabilidade deve incluir a **quantificação do custo da baixa qualidade de dados**, que se manifesta em desperdício, retrabalho, baixa produtividade e multas por não conformidade.

Definir **metas de melhoria de Qualidade de Dados (DQ) ligadas a valor.** As metas de melhoria de DQ devem ser **conectadas à quantificação do valor** para o negócio (ex.: redução de custos, melhor atendimento ao cidadão, potencializar resultado de política pública), justificando o investimento.

- A avaliação de viabilidade deve ir além dos custos de obtenção e armazenamento, quantificando o **valor** que o dado gera (que é a diferença entre o custo e o benefício derivado). Para o governo, o valor deve incluir a **capacidade de orientar as atividades operacionais, táticas e estratégicas** do órgão ou da agência.

## 2.6 Fase - Avaliação de riscos

O objetivo central desta fase é identificar os riscos críticos e construir um **plano de mitigação robusto** para garantir a segurança e a viabilidade do projeto. O processo envolve o mapeamento detalhado de riscos relacionados a dados, segurança cibernética (*Adversarial ML*), tecnologia, regulação e infraestrutura.  É recomendável a execução da AIE novamente, visto que o processo prevê iterações para ajustes e validações e, nesse cenário, podem ocorrer novos eventos que necessitem de verificação.

Um componente essencial é a criação de um plano de mitigação que inclua a logística formal para obtenção de dados, definindo caminhos de liberação, termos de requisição (finalidade, tipo, volume) e o ambiente de armazenamento seguro acessível apenas pela equipe de IA.

### 2.6.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Matriz de riscos composta** | Documento que consolida os riscos de dados, regulatórios, éticos, de segurança da informação e de infraestrutura. |
| **Autoavaliação de Impacto Ético (AIE) Total** | Relatório final da avaliação ética do projeto. |
| **Plano de mitigação** | Estratégia detalhada para minimizar os riscos identificados. |
| **Termo de requisição de dados** | Documento formal especificando finalidade, tipo, volume e período dos dados necessários. |

### 2.6.2 O que verificar antes de avançar

Para que o plano de riscos seja aprovado e o projeto avance, os seguintes itens devem estar em conformidade:

- **Mapeamento estimado de riscos.** Riscos críticos (relacionados a dados, infraestrutura, regulação e ética) foram identificados?
- **AIE finalizada.** A AIE total foi executada?
- **Logística de dados.** O caminho formal para liberação e acesso aos dados foi estabelecido com os proprietários?
- **Ambiente seguro.** O ambiente de trabalho para armazenamento dos dados preliminares foi definido?
- **Plano de mitigação.** As ações propostas são suficientes para reduzir os riscos a níveis aceitáveis?

**A aprovação exige a resposta afirmativa para todos os itens da lista** para que o plano de mitigação seja considerado sólido e os riscos residuais sejam aceitáveis. Caso existam lacunas na mitigação ou na logística de dados, o plano deve ser revisado. Se os riscos (éticos, legais ou técnicos) forem críticos e sem mitigação possível, o processo é encerrado.

### 2.6.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Executor | Sustentador | RSGI | Comitê de IA - Ética |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Mapear riscos** | A | C | - | C | C | A | I* | R | C | C | C |
| **Avaliar risco ético total** | I | I | I | I | I | I | - | R | I | C | A |
| **Criar plano de mitigação** | C | C | - | C | C | A | - | R | C | C | A |
| **Formalizar pedido de dados** | R | A | - | I | I | I | - | C | C | - | I |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor); - (Não atua).

* Informado via sistema de organização e registro de soluções de IA.

### 2.6.4 Recomendações, políticas e boas práticas

**Gerenciar o risco ético e reputacional.** O risco ético deve ser ativamente mitigado, uma vez que a ausência de práticas íntegras pode resultar em perda de reputação — fator crítico para a manutenção da confiança pública e para a legitimidade das operações governamentais.

## 2.7 Fase - Definição de métricas e proposta de experimentação

O objetivo principal desta fase é estabelecer as bases quantitativas e operacionais para a fase de testes. O processo consiste em refinar o *Baseline* (estado atual) e definir métricas claras e quantificáveis, tanto técnicas quanto de impacto, que servirão de critério para avaliar o sucesso da solução. Com base nisso, elabora-se uma **proposta de experimentação**, consolidando os objetivos, as hipóteses a serem testadas, o conjunto de dados *Ground Truth* para validação dos dados necessários e o cronograma. A fase encerra-se com a geração do **relatório de estruturação**, que fundamenta a viabilidade de prosseguimento para a execução do piloto.

### 2.7.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Proposta de experimentação** | Consolidação dos objetivos, hipóteses, dados necessários, estratégia de avaliação (automática ou humana), cronograma com macro atividades (não detalhado) e definição do ambiente para o piloto.<br>**Definição de métricas de negócio e *Thresholds*.** Indicadores de desempenho e os critérios mínimos de aceite (limiares) para a solução.<br>Sumário executivo final da etapa de estruturação com recomendação fundamentada.<br>• **Métricas de negócio:** define KPIs reais, como redução de custos, otimização de tempo ou aumento da satisfação do usuário. |

### 2.7.2 O que verificar antes de avançar

Para concluir a **estruturação** e autorizar a **experimentação**, os seguintes itens devem estar validados:

- **Métricas definidas.** As métricas de impacto de negócio são claras e quantificáveis?
- **Critérios de aceite.** Foram estabelecidos os *thresholds* (limites mínimos) de negócio para considerar o experimento um sucesso?
- **Hipóteses claras.** As hipóteses a serem testadas no piloto estão bem descritas na proposta?
- **Recursos de dados.** Os dados necessários para o experimento foram confirmados e mapeados?
- **Cronograma de macroatividades.** O cronograma macro de experimentação é exequível para as partes envolvidas?
- **Recursos orçamentários.** O orçamento para a experimentação está aprovado?
- **Riscos**. Os riscos estão dentro do patamar aceitável pela organização?

A aprovação da proposta está condicionada à conformidade integral de todos os itens de verificação, exigindo consenso formal entre demandante e executor. Caso as métricas sejam estimadas como irreais ou a proposta careça de dados suficientes para os testes, a proposta deverá ser revisada obrigatoriamente antes do início da fase de experimentação

### 2.7.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Executor | Sustentador | CISC | RSGI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Definir métricas de sucesso** | R | - | A | - | - | I | - | R | C | - | - |
| **Elaborar plano de experimentação** | A | I | I | C | I | C | - | R | I | I | I |
| **Consolidar relatório final** | A | I | - | C | - | I | I* | R | C | - | I |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor); - (Não atua).

* Informado via sistema de organização e registro de soluções de IA.

### 2.7.4 Recomendações, políticas e boas práticas

**Priorização por risco técnico.** A experimentação deve priorizar a validação das hipóteses mais críticas para a viabilidade do projeto, especialmente aquelas que, se não atendidas, inviabilizam sua continuidade. Por exemplo, questões relacionadas à disponibilidade, qualidade ou adequação dos dados devem ser testadas antes de investimentos em arquiteturas complexas ou ajustes finos de modelos. Essa abordagem reduz desperdício de recursos e permite decisões rápidas e informadas.

**Gestão de tempo da pesquisa e da experimentação.** Projetos de ciência de dados e IA devem operar dentro de janelas temporais claramente definidas para a fase de pesquisa e experimentação. A ausência de limites tende a prolongar indefinidamente a busca por soluções ideais, sem entrega de valor concreto. O estabelecimento de **prazos rígidos** para testes e comparações de alternativas favorece decisões tempestivas, mesmo que baseadas em soluções imperfeitas, mas operacionalmente viáveis.

## 3. Etapa - Experimentação

Esta etapa tem como objetivo principal desenvolver protótipos e realizar testes técnicos rigorosos para validar o desempenho, a segurança e a viabilidade econômica da solução de IA proposta. É aplicável tanto para desenvolvimento interno (*make*) quanto à avaliação de soluções de mercado (*buy*).


Figura 4 - Fases da Experimentação

## 3.1 Fase - Planejamento e detalhamento do uso da IA

O objetivo principal desta fase é elaborar o **Plano de Experimentação**, baseado na proposta acordada na fase anterior, definir as histórias de usuário, o comportamento esperado da IA e as métricas técnicas específicas para cada abordagem.

Na elaboração do cronograma detalhado, a equipe executora pode obter recomendações gerais e atividades específicas recomendadas, acessando os anexos técnicos deste guia, conforme a abordagem de IA foco da experimentação (por exemplo, RAG, LLM).

### 3.1.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Matriz de métricas de sucesso da solução de IA** | • **Matriz de métricas de sucesso da solução de IA.** Destacar métricas de negócio que devem ser atingidas e que foram acordadas na proposta de experimentação.<br>• **Finalidade:** refinar o baseline (estado atual) e estabelecer métricas claras que permitirão medir o impacto e a eficácia da solução.<br>• **Componentes técnicos:** inclui métricas específicas por tipo de tecnologia, como acurácia, precisão, F1-Score e recall para modelos clássicos, ou fluência e coerência para IA Generativa.<br>• **Critérios de aceite (*Thresholds*):** estabelece os valores mínimos (limiares) que a solução deve atingir para ser considerada viável (ex: performance superior a 80% do limiar crítico).<br>**Matriz de premissas segurança e privacidade.** Detalhar as restrições, riscos e diretrizes de conformidade que a solução deve respeitar para garantir a integridade dos dados e a segurança jurídica.<br>• **Segurança e conformidade:** avalia a aderência à LGPD, garantindo que o tratamento de dados pessoais siga o marco regulatório do órgão.<br>• **Riscos técnicos específicos:** lista itens de verificação como toxicidade, alucinações (em LLMs), prompt injection, vazamento de dados via contexto e vieses discriminatórios.<br>• **Privacidade desde o projeto (by *design*):** define regras de minimização de dados, anonimização, retenção e descarte, além de prever trilhas de auditoria.<br>• **Controles de acesso:** mapeia a estratégia de autenticação, autorização (RBAC) e gestão de segredos (como *tokens* e senhas).<br>**Cronograma detalhado.** Detalhar as macroatividades apresentadas na proposta da fase anterior.<br>**Estrutura de comunicação.** Responsáveis e como os *stakeholders* e a equipe se comunicam. |
| **Casos de uso / histórias** | Detalhamento de jornadas/histórias e comportamento esperado da solução. |

### 3.1.2 O que verificar antes de avançar

Para dar prosseguimento à **experimentação**, os seguintes itens devem estar validados. Caso contrário, os resultados devem ser revistos e sofrer ajustes para atender às premissas.

- O caso de uso está claramente definido?
- Todas as hipóteses apontadas na proposta estão contempladas nos casos de uso?
- Os recursos de infraestrutura e acesso aos dados (amostra) estão liberados?
- As hipóteses a serem testadas no piloto estão bem descritas na proposta?
- Os dados necessários para o experimento foram disponibilizados?
- O cronograma detalhado para experimentação é exequível para as partes envolvidas?

### 3.1.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Executor | Sustentador | CISC | RSGI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Especificar casos de uso da solução** | A | C | - | - | I | I | - | R | I | I | I |
| **Definir métricas da solução** | - | C | - | - | I | I | - | R | A | - | - |
| **Definir premissas de segurança e privacidade** | I | C | I | I | A | A | - | R | C | C | C |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor); - (Não atua).

### 3.1.4 Recomendações, políticas e boas práticas

**Métricas técnicas por tipo de abordagem:**

- **Para modelos de IA Generativa e *LLMs*.** Consultar as diretrizes de avaliação do ***OWASP Top 10 for LLM Applications*** para métricas de segurança (toxicidade, alucinações, vazamento de contexto) e o **NIST AI RMF - Função *Measure* (Medir)** para métricas de confiabilidade, robustez e viés. Métricas incluem: fluência, coerência, relevância, fundamentação (*groundedness*) e taxa de alucinação.
- **Para modelos de *ML* Clássico (classificação/regressão).** Utilizar métricas padrão documentadas em literatura técnica e *frameworks* como *scikit-learn metrics* (como acurácia, precisão, *recall*, *F1-Score*, *AUC-ROC* para classificação; *RMSE*, *MAE*, $R^2$ para regressão). No contexto governamental, deve-se priorizar a interpretabilidade das métricas para atuantes não técnicos. Métricas de performance do modelo: FPS (*Frames Per Second*): importante para modelos de visão computacional em tempo real (ex.: monitoramento de tráfego ou segurança). Define o requisito mínimo de hardware e a latência aceitável para a tomada de decisão. Latência de Inferência: O tempo que o sistema leva para processar uma imagem individualmente, garantindo que o serviço público seja ágil.
- **Para modelos de visão computacional.** Métricas de detecção de objetos (*mAP*, *IoU*), segmentação (*Dice coefficient, IoU*) e classificação de imagens. Atentar para vieses em dados de treinamento que possam comprometer a equidade no serviço público.

***Frameworks* de referência para medição e avaliação:**

- **NIST AI RMF - Função *Measure* (Medir).** Fornece orientações sobre como medir e avaliar as características de confiabilidade de sistemas de IA incluindo: validade e confiabilidade de métricas, avaliação de viés e equidade, robustez adversarial e rastreabilidade.
- **ISO/IEC 42001:2023 - Cláusulas de teste e validação.** Estabelece requisitos para pro-

cessos de verificação e validação de sistemas de IA incluindo: definição de critérios de aceitação, documentação de resultados de teste e rastreabilidade entre requisitos e testes.

- **ABNT NBR ISO/IEC 25010 (Qualidade de *software*).** Modelo de qualidade de produto de *software* aplicável também a sistemas de IA abordando características como adequação funcional, confiabilidade, usabilidade, eficiência de desempenho, manutenibilidade e portabilidade.

**Qualidade de dados para experimentação:**

- Aplicar as dimensões de qualidade de dados do **DAMA-DMBOK** [3] para validar os conjuntos de dados de treinamento e teste: acurácia, completude, consistência, atualidade e validade.
- Documentar métricas de qualidade do conjunto de dados (*dataset quality metrics*): balanceamento de classes, cobertura de cenários, representatividade de grupos populacionais (para análise de equidade).

**Listas de checagem auxiliares:**

- **Lista de checagem de métricas técnicas:** para cada possibilidade de abordagem técnica, elaborar matriz relacionando: métrica técnica, fórmula/definição, *threshold* mínimo, justificativa do *threshold* e método de medição.
- **Lista de checagem de métricas de negócio:** relacionar métricas técnicas com Indicadores-Chave de Desempenho (*KPIs*) de impacto (ex: "redução de 30% no tempo de atendimento" vinculado a "latência de resposta menor que 2 segundos").
- **Lista de checagem de conformidade:** verificar aderência a requisitos de segurança (*OWASP*), privacidade (LGPD), equidade (análise de viés por grupo) e explicabilidade (capacidade de justificar decisões do modelo).

## 3.2 Fase - Experimentação de modelo de IA

Os objetivos principais desta fase são: garantir a infraestrutura computacional necessária para o problema, realizar análise exploratória de dados para bases numéricas ou estruturação e padronização para bases categóricas, e executar o treinamento ou a iteração de *prompts* conforme a tecnologia escolhida, observando estatisticamente a qualidade dos resultados obtidos.

Mesmo em experimentações que utilizam modelos prontos, sem geração ou treinamento de modelo próprio, a análise exploratória de dados é uma atividade necessária e não deve ser suprimida.

Nesse contexto, o objetivo da análise exploratória não é selecionar variáveis ou definir arquitetura de modelo, mas verificar se os dados de entrada do órgão são adequados para o uso pretendido. Trata-se de uma validação de aderência: os dados precisam estar no formato esperado pelo modelo, com qualidade suficiente para produzir resultados confiáveis e sem conteúdo sensível que exija tratamento antes do envio.

### 3.2.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Ambiente** | Ambiente de experimentação configurado. |
| **Modelos ou algoritmos** | Modelos selecionados e otimizados. |
| **Relatório de análise exploratória** | Em caso de aprofundamento e/ou limpeza dos dados. |

### 3.2.2 O que verificar antes de avançar

Neste ponto, **dependendo da abordagem técnica de IA**, deve-se verificar se os dados estão prontos ou se necessitam de limpeza e *feature engineering.* Caso seja necessário, atividades específicas para seleção, criação e tratamento de novos dados deverão ser executadas para dar prosseguimento à **experimentação**.

**Decisão em caso de abordagem de IA (modelo) que necessite treino:**

- **Aprovação.** A análise exploratória confirma que os dados são suficientes para o treino.
  - **Ação.** Executar o treinamento do modelo. Prosseguir a experimentação para encaminhamento para validação técnica.
- **Reprovação.** Dados enviesados, incompletos ou sem correlação com o alvo.
  - **Ação.** Retornar à 3.3Fase - Viabilidade e risco ético dos dados, para buscar novas fontes ou realizar nova limpeza e *feature engineering*.

**Decisão em caso de abordagem de IA use modelos pré-treinados:**

- **Aprovação.** Os dados estão no formato e na estrutura compatíveis com o modelo escolhido, dados sensíveis foram tratados adequadamente  e resultados dentro da margem esperada. Corpus com documentos atualizados, bem extraídos e sem lacunas temáticas. Documentos não se contradizem.
  - **Ação.** Prosseguir para a próxima fase para validação técnica.
- **Reprovação.** Dados enviesados, incompletos ou sem correlação com o alvo.
  - **Ação.** Retornar à 3.3Fase - Viabilidade e risco ético dos dados, para buscar novas fontes, mais atualizadas e coerentes.

### 3.2.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Executor | Sustentador | CISC | RSGI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Estruturação de ambiente** | I | - | - | A | C | I | I* | R | C | I | I |
| **Seleção de modelos** | - | I | - | - | I | A | - | R | I | - | - |
| **Análise exploratória** | I | A | I | I | I | - | - | R | I | - | - |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor); - (Não atua).

* Informado via sistema de organização e registro de soluções de IA.

### 3.2.4 Recomendações, políticas e boas práticas

**Qualidade de dados.** Definir quais dados (incluindo *PII - Personally Identifiable Information* e *PHI - Protected Health Information*) são críticos para a missão do órgão. Estabelecer regras de negócio e/ou de qualidade de dados mensuráveis e alinhadas aos requisitos de conformidade.

Realizar **avaliação inicial de qualidade de dados,** com ênfase em subconjuntos de dados críticos para quantificar não-conformidades e identificar **causas-raiz** que impactam a confiabilidade dos registros públicos.

**Planejar a ingestão dos dados**, assegurando a captura de **metadados críticos** (origem, segurança, formato). O processo de ingestão e classificação deve ser rigoroso em qualquer volume de dados para proteger a informação e garantir a confiança na solução de IA.

**Normalizar e estruturar dados.** Normalizar e estruturar dados, aplicar transformação e padronização, e garantir, onde for aplicável, a consistência dos dados mestres e de referência (MDM/RDM) e a correta identificação do Sistema de Registro.

**Enriquecer dados (*data augmentation*)** para aumentar a qualidade e utilidade (ex: geocodificação de endereços de cidadãos) e documentar o processo, garantindo que não viole a confidencialidade ou segurança.

**Carregar dados (ingestão).** Coletar Metadados sobre o processo (origem, tamanho, segurança, qualidade) e classificar a confidencialidade/restrição da informação, aplicando *tags* para controle de acesso.

**Validar ingestão / qualidade de dados**. Realizar profiling de dados e implementar monitoramento contínuo de DQ para identificar e isolar dados imprecisos ou inconsistentes antes que comprometam as análises e decisões públicas.

**Preparar coleção para experimentação.** Aplicar *Data Masking/Obfuscation* em informações sensíveis (PII, PHI) ao criar conjuntos de dados em ambientes não produtivos para mitigar riscos de exposição e violação regulatória.

**Desenvolver algoritmos.** Aplicar algoritmos estatísticos e de *ML* e garantir que o desenvolvimento suporte a rastreabilidade (linhagem) das variáveis e o rastreamento de experimentos (*Experiment Tracking*), crucial para auditorias em contextos governamentais.

**Bloqueio de formatos de serialização inseguros**. Ao desenvolver algoritmos, para evitar a execução remota de códigos (RCE), deve-se aplicar uma medida técnica de segurança para bloquear serialização insegura (processo de salvamento ou transmissão de um objeto, seja modelo, dados, estruturas em memória) e posterior carregamento (desserialização).

**Realizar *feature engineering*.** Garantir a governabilidade ética e a justiça na seleção de atributos (*features*), evitando vieses não intencionais que possam resultar em discriminação ou tratamento desigual de cidadãos.

**Desenvolver e simular algoritmos**. Treinar o modelo e verificar hipóteses, assegurando que os conjuntos de dados de treinamento sejam representativos e que o processo seja documentado (Metadados).

**Validar simulações.** Usar conjuntos de validação independentes para estimar o erro de generalização antes da implantação, garantindo a robustez do modelo em um contexto de serviço público.

**Testar qualidade**. Implementar controles de monitoramento e auditoria de dados para verificar se as regras de qualidade e a conformidade regulatória são mantidas durante a experimentação.

**Validar modelo.** Assegurar que o modelo não viole os princípios éticos de Respeito às Pessoas e Beneficência e que o uso de dados anônimos ou agregados não permita a recombinação de informações sensíveis do cidadão.

**Qualidade de dados.** Aplicar dimensões de qualidade conforme DAMA-DMBOK [3]:

- Acurácia: os dados representam corretamente a realidade?
- Completude: os dados necessários estão presentes?
- Consistência: os dados estão em conformidade com as regras e definições estabelecidas, incluindo coerência entre diferentes sistemas (se existirem)?
- Atualidade: os dados são atuais e disponíveis quando necessários?
- Validade: os dados atendem às regras de negócio e aos formatos esperados?
- Além disso, realizar *profiling* de dados para quantificar não-conformidades antes da experimentação.

**Anonimização e pseudonimização.** Quando aplicável, utilizar técnicas adequadas de anonimização (k-anonimato, l-diversidade, t-*closeness*) ou pseudonimização, conforme requisitos de proteção. Atentar para riscos de nova identificação em modelos de IA.

**Linhagem de dados (*data lineage*).** Documentar a origem, as transformações e o destino dos dados utilizados, garantindo rastreabilidade e auditabilidade. Esta etapa é fundamental para a explicabilidade e a conformidade regulatória.

**Categorias de tecnologias aplicáveis** (lista não exaustiva).

- IA generativa: modelos de linguagem de grande escala (LLMs) de código aberto ou com acesso via API comercial, conforme política de uso de dados e requisitos de soberania informacional.
- Frameworks de orquestração: frameworks de encadeamento e orquestração de pipelines de inferência de LLM.
- ML clássico: algoritmos de classificação e regressão supervisionada ou de agrupamento não supervisionado, conforme a natureza do problema.
- Infraestrutura de busca vetorial (RAG): bancos de dados vetoriais dedicados ou extensões vetoriais de bancos de dados relacionais.

**Rastreabilidade de experimentos**. ***Model registry***. Utilizar catalogação e versionamento de modelos desde o início da experimentação para garantir a reprodutibilidade dos testes**.**

## 3.3 Fase - Validação técnica, segurança e viabilidade

Esta fase visa avaliar se o modelo atingiu os limiares de performance, validar a robustez contra ataques (ex: *prompt injection*) e confirmar se os custos de operação são sustentáveis.

### 3.3.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Relatório de validação técnica** | Ambiente de experimentação configurado. |
| **Parecer de segurança e compliance (LGPD)** | Parecer da área responsável por segurança e *compliance*, com relação aos requisitos necessários conforme solução proposta. |
| **Análise de ROI** | Relatório de análise de Retorno de Investimento |
| **Registro de lições aprendidas** | Relatório detalhando as lições aprendidas durante a experimentação |

### 3.3.2 O que verificar antes de avançar

Para dar prosseguimento à **experimentação**, os seguintes itens devem estar validados. Caso contrário, os resultados devem sofrer ajustes para atender as premissas.

- A performance é superior ao limiar definido?
- A latência é viável para a experiência do usuário (UX)?
- A conformidade com a LGPD foi verificada?
- Vulnerabilidades críticas foram zeradas?

**Decisão.**

- **Aprovação.** Performance acima do *threshold* (acurácia/F1-Score), latência viável para UX e ausência de vulnerabilidades críticas.
  - **Ação.** Seguir para Etapa - Implementação.
- **Reprovação.** Modelo apresenta alucinações inaceitáveis, toxicidade ou performance insuficiente para o negócio.
  - **Ação.** Retornar à Fase - Planejamento e detalhamento do uso da IA (4.1), para ajustar as métricas ou à Fase - Proposição de arquitetura inicial da solução (3.4), para testar um algoritmo/modelo diferente.

### 3.3.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | ETDP | Gestor de Segurança da Informação | NIA | Executor | Sustentador | CISC | RSGI |
|---|---|---|---|---|---|---|---|---|---|---|
| **Definir métricas e estórias** | A | C | - | I | - | - | R | I | - | - |
| **Executar experimentos / Treino** | I | C | - | - | - | - | R/A | C | - | - |
| **Validar segurança e compliance** | C | C | - | A | A | - | R | I | C | C |
| **Analisar viabilidade econômica** | A | C | I | I | I | I* | R | C | - | - |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor); - (Não atua).

* Informado via sistema de organização e registro de soluções de IA.

### 3.3.4 Recomendações, políticas e boas práticas

**Recomendações de ferramentas:** *Apache JMeter*, *Gatling*, *K6* ou *Locust* para testes de carga.

**Requisito de teste de carga**: ferramentas de teste de carga capazes de simular múltiplas requisições concorrentes e medir latência, taxa de erros e *throughput* sob carga sustentada.

**Recomendações para implementação**:

- Detalhamento da arquitetura testada e proposta para implementação. Orientações sobre ajustes necessários na solução, próximos passos, requisitos técnicos e de dados para a etapa de Implementação.
- Recomenda-se a conteinerização baseada em padrões abertos (OCI - *Open Container Initiative*) e a preparação para orquestração do ambiente, garantindo portabilidade, padroni-

zação e facilidade de replicação da solução entre diferentes ambientes (desenvolvimento, homologação e produção). O uso de imagens de contêineres imutáveis permite que todas as dependências sejam configuradas, reduzindo riscos de inconsistências.

- Sugere-se a definição de um repositório central de imagens seguro (*Trusted Registry*) com varredura automática de vulnerabilidades, pipeline de CI/CD para build e deploy automatizados, bem como a criação de manifestos de Infraestrutura como Código (IaC, ex: *Helm Charts ou Kubernetes Manifests*). Devem ser documentadas claramente as variáveis de ambiente (injetadas preferencialmente via soluções de Gestão de Segredos/*Vault*), protocolos de rede, volumes e políticas de escalabilidade *(Horizontal Pod Autoscaling*) para facilitar a manutenção e evolução do sistema.
- Relatório de Lições Aprendidas: Conhecimentos adquiridos durante a experimentação (técnicos, processuais, de negócio).

## 4. Etapa - Implementação

Esta etapa tem por objetivos integrar, configurar e implantar a solução de IA de forma robusta e escalável, garantindo que a arquitetura seja segura (security by *design*) e que existam mecanismos de monitoramento e contingência antes do lançamento oficial.


Figura 5 – Etapa de implementação da solução

### 4.1 Fase - Arquitetura e *design* da solução

O objetivo principal desta fase é definir as camadas da arquitetura (dados, processamento, exposição e interfaces) e os componentes de segurança, como criptografia, controles de acesso (RBAC - Role-Based Access Control) e gestão de segredos.

### 4.1.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Arquitetura e *design* de solução** | Documentação da arquitetura segura, estratégia de identidade e acesso mapeada. |
| **Plano de privacidade** | Plano de privacidade (com base no *RIPD* - Relatório de Impacto à Proteção de Dados) iniciado. |

### 4.1.2 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Executor | Sustentador | CISC | RSGI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Definir Arquitetura segura** | I | C | - | C | A | A | - | R | C | C | C |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor); - (Não atua).

### 4.1.3 Recomendações, políticas e boas práticas

A documentação da arquitetura deve detalhar pontos relativos à identidade e acesso. Ela deve prever a **segregação entre as camadas** (*front-end, backend*, orquestrador de *prompts e LLM*) para evitar que um acesso indevido em uma interface comprometa todo o sistema ou os dados sensíveis de treinamento.

- **Princípio do menor privilégio**. Garantia de que as permissões mínimas necessárias sejam aplicadas em todos os níveis da solução.
- **Identidade verificada**. Análise da necessidade de integração com soluções de identidade digital.
- **Engenharia de atributos.** Avaliar a implementação de um repositório centralizado de *features* para garantir consistência entre treinamento e inferência e promover reuso.
- **Observabilidade e monitoramento detalhado.**
  - Definir ferramentas para detecção de *Data/Concept Drift* (desvio de dados).
  - Implementar monitoramento de performance técnica (latência, *throughput*) e de negócio.
  - Estabelecer logs estruturados para auditoria e depuração avançada.
- **Estratégias de *deploy*.**
  - **Testes A/B e *Canary*.** Preparar a arquitetura para suportar experimentação controlada em produção.

- **Segurança aplicada.**
  - Implementação técnica de criptografia de dados em trânsito (TLS) e em repouso.
  - Configuração de mecanismos robustos de autenticação e autorização (ex: OAuth2, mTLS).

**Documentação de requisitos.** Documentar requisitos funcionais e não funcionais de forma clara, incluindo requisitos específicos de IA (acurácia mínima, latência máxima, interpretabilidade), de segurança e de conformidade regulatória.

Em relação a Minimização e Segregação de Dados.

- A arquitetura deve garantir a segregação entre as camadas de *front-end, backend*, orquestrador e o modelo de IA.
- A estratégia de acesso mapeia quem pode ver quais dados, aplicando o princípio de minimização, onde o sistema acessa apenas o estritamente necessário para a finalidade proposta

Em relação ao **Controle de Acesso Baseado em Funções** (RBAC). O mapeamento define perfis de acesso (RBAC), garantindo que apenas usuários autorizados interajam com dados sensíveis ou funções administrativas. Isso é fundamental para o RIPD, pois comprova que existem controles técnicos para evitar acessos indevidos ou vazamentos.

Em relação à **gestão de segredos e identidade**. A política de gestão de segredos (uso de *Vaults* ou *Secrets Managers*) deve ser definida para que chaves e *tokens* de acesso não fiquem expostos na arquitetura. Em casos específicos, analisa-se a integração com identidades verificadas para aumentar o nível de confiança no acesso.

Em relação a **trilhas de auditoria e retenção.** A estratégia de acesso deve prever o registro de **trilhas de auditoria**, permitindo rastrear quem acessou o quê e quando. Esses logs são evidências essenciais para a finalização do RIPD e para a conformidade com a LGPD.

**Assinatura digital de artefatos.** O orquestrador só deve aceitar containers/modelos que possuam assinatura criptográfica válida da fase de homologação. Ou seja, somente containers e modelos de IA oficialmente aprovados e assinados criptograficamente durante a homologação podem ser executados em produção, impedindo a execução de artefatos não autorizados ou adulterados.

**Requisitos de infraestrutura e orquestração.** Adotar tecnologia de conteinerização para portabilidade e isolamento de dependências entre ambientes. Frameworks de orquestração de pipelines de LLM devem ser selecionados conforme requisitos de rastreabilidade e auditabilidade. O provisionamento de infraestrutura deve ser realizado via ferramentas de Infraestrutura como Código (IaC), garantindo reprodutibilidade e trilha de auditoria das configurações.

## 4.2 Fase - Construção e infraestrutura

Em paralelo à arquitetura e *design* da solução, esta fase visa a preparação da infraestrutura de dados e integração, configurar o pipeline de retreino automatizado e construir painéis de controle para monitoramento em tempo real.

Nesta fase, devem ser implementados testes automatizados de robustez adversarial (ART) no pipeline, a fim de validar resistência contra evasão e inferência [5].

### 4.2.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Pipeline de retreino** | Pipeline de retreino automatizado. (Com etapa de validação humana - Human-in-the-loop). |
| **Interface de usuário** | Interface de usuário concluída ou versão estável *(release candidate).* |
| ***Dashboards*** | *Dashboards* de monitoramento configurados. |

### 4.2.2 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Executor | Sustentador | CISC | RSGI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Preparar infra e *pipelines*** | I | I | - | A | C | C | - | R | C | I | I |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor); - (Não atua)

### 4.2.3 Recomendações, políticas e boas práticas

**Recomendações de ferramentas.** *Dashboards* para métricas de qualidade, performance (latência) e segurança (detecção de ataques). Implementação de "*Model Registry*" para versionamento estrito dos artefatos gerados.

## 4.3 Fase - Testes e validação

Esta fase visa a execução de testes integrados, testes de carga e testes de segurança (*Penetration Test*, *Red Team*), além de validar o desempenho final e formalizar o processo de **gestão de mudança (*GMUD*)** para o deploy seguro. Esta fase inclui os itens a seguir.

- Testes do modelo isolado: desconectados de qualquer interface ou sistema, com a intenção de avaliar performance algorítmica através de métricas como Acurácia, *Recall* e *F1-Score*, usando um conjunto de dados de teste.
- Testes da aplicação: verifica se o sistema tem bom funcionamento para todas as respostas possíveis do modelo de IA avaliando recebimento da requisição, conexão ao modelo, tempo de resposta e formato correto.
- Testes fim a fim: testa o fluxo completo da experiência do usuário, do clique inicial à resposta final (passando por login, pergunta e resposta), sempre com foco em verificar se a solução completa resolve o problema do usuário.

- Testes de abuso, erro e uso indevido: testa a integridade, segurança e os limites éticos da solução, realizando tentativas de *jailbreak*, *prompt injection*, viés preconceituoso, extração de dados sensíveis, entre outras possibilidades.
- Testes de segurança: os testes de segurança de Robustez Adversarial (ART) devem ser executados automaticamente para verificação. Isto evitará erros humanos por checagem manual. Por exemplo, um desenvolvedor pode, por pressa, usar uma imagem de container *Docker* com privilégios de *root* ou expor segredos no código, e isso só ser descoberto (ou não) na auditoria final.

### 4.3.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Testes de carga** | Relatório de testes de carga realizados. |
| **Plano de *rollback*** | Plano de *rollback* construído e testado. |
| **Manuais operacionais** | Manuais operacionais elaborados e divulgados *(Runbooks)*. |
| **Termo de aprovação** | Termo de aprovação final *(Go-Live)*. |

### 4.3.2 O que verificar antes de avançar

Para dar prosseguimento à Implantação (entrada em Produção), os seguintes itens devem estar validados. Caso contrário, os resultados devem sofrer ajustes para atender as premissas.

- O tempo de resposta (latência) permanece nos níveis aceitáveis e informado nas premissas da área de negócio?
- O plano de *rollback* foi devidamente testado?
- As vulnerabilidades críticas e a conformidade com a LGPD foram totalmente verificadas?
- O *pipeline* de retreino funcional e monitoramento foi configurado?

**Decisão.**

- **Aprovação.** Testes fim a fim aprovados, plano de *rollback* testado e conformidade LGPD verificada.
  - Ação. Seguir para Etapa - Implantação (*rollout*).
- **Reprovação.** Falhas de integração, latência excessiva em ambiente de homologação ou erros de interface.
  - **Ação.** Retornar à Fase - Construção e infraestrutura, para ajustes técnicos.

### 4.3.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Executor | Sustentador | CISC | RSGI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Testes de segurança e carga** | I | I | - | A | I | A | - | R | C | R | I |
| **Elaboração do plano de rollout (*GMUD*)** | C | I | I | A | C | A | I* | C | R | I | I |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor); - (Não atua).

* Informado via sistema de organização e registro de soluções de IA.

### 4.3.4 Recomendações, políticas e boas práticas

**Recomendações de tecnologias.** Estratégias de *Safe Deployment* como *Shadow Mode* (execução em paralelo) ou *Canary Release* (liberação gradual).

## 5 Etapa - Implantação (*rollout*)

Esta etapa possui uma única fase de Implantação (*rollout*) .


Figura 6 – Etapa Implantação

### 5.1 Fase - Implantação

Esta fase visa a execução do plano de *rollout*, a fim de validar o processo de gestão de mudança (*GMUD*) para o *deploy* seguro e formalizá-lo, transferindo a solução para a equipe de sustentação.

### 5.1.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Termo de aceite** | Termo de aceite da equipe de sustentação, validando que receberam códigos, acessos e documentação. |
| **Plano de comunicação** | Plano de comunicação executado. |

### 5.1.2 O que verificar antes de avançar

Para dar prosseguimento à sustentação, o parceiro (ou responsável) de sustentação deve decidir sobre o aceite da solução em produção. Caso o parceiro ou responsável pela sustentação rejeite, deve proceder à comunicação dos envolvidos e adiar a entrada em produção até que o ambiente e solução estejam conforme homologado.

### 5.1.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Executor | Sustentador | CISC | RSGI |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Execução do rollout *(GMUD)*** | A | I | I | A | C | A | I* | R | R | I | I |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor).

* Informado via sistema de organização e registro de soluções de IA.

## 6. Etapa - Sustentação

Esta etapa tem como objetivo principal garantir que a solução continue gerando valor, mantendo a estabilidade técnica, a segurança e a precisão do modelo através de monitoramento contínuo e ciclos de melhoria (*MLOps e FinOps*).


Figura 7 - Etapa Sustentação

## 6.1 Fase - Monitoramento contínuo

O objetivo desta fase é observar continuamente o sistema em tempo real para detectar falhas de infraestrutura, degradação de performance, desvios orçamentários e riscos éticos ou de segurança.

Para um monitoramento de ataques de inversão e extração adequado, devem ser utilizadas ferramentas de *Adversarial Robustness* [5], a fim de testar e fortalecer modelos de IA contra possíveis ataques e manipulações intencionais, garantindo que o sistema permaneça seguro, confiável e conforme, mesmo sob uso malicioso. Deve-se integrar alertas se o padrão de consultas (API *requests*) indicar tentativa de reconstrução de dados de treino. Esse tipo de monitoramento é imprescindível, pois um modelo pode manter alta acurácia e baixo ***drift*** enquanto está sendo vítima de um **ataque de inversão de modelo** (extração de dados sensíveis de treinamento), algo que um monitoramento tradicional não detectaria.

### 6.1.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Painéis de observabilidade *(dashboards)*** | Interfaces gráficas que exibem métricas técnicas (latência, erro), métricas de modelo (acurácia, *recall*) e métricas de negócio (ROI). |
| **Alertas de *drift* e anomalias** | Notificações automatizadas disparadas quando há mudança na distribuição dos dados de entrada (*Data Drift*) ou na relação entre variáveis (*Concept Drift*). |
| **Relatórios de segurança contínua** | Registros de varreduras de vulnerabilidades, tentativas de *jailbreak* e conformidade periódica com a LGPD. |

### 6.1.2 Alertas de monitoramento contínuo

Para que a solução continue ativa, os seguintes itens devem ser validados periodicamente:

- Desempenho: a acurácia/precisão permanece acima do limite mínimo (*threshold*) estabelecido?
- Segurança: não foram detectadas vulnerabilidades críticas ou alucinações inaceitáveis?
- Sustentabilidade: os custos de infraestrutura e *tokens* estão dentro do orçamento previsto o e as cotas de consumo não foram violadas?

### 6.1.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | ETDP | Gestor de Segurança da Informação | Gestor de TIC | NIA | Sustentador | CISC | RSGI |
|---|---|---|---|---|---|---|---|---|---|---|
| ***Monitoramento contínuo*** | I | I | I | I | I | A | I* | R | I | I |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor).

* Informado via sistema de organização e registro de soluções de IA.

### 6.1.4 Recomendações, políticas e boas práticas

**Requisitos de monitoramento.** Ferramentas de detecção de desvio de dados e de modelo (*data drift e model drift*) devem produzir relatórios comparativos entre distribuições de referência e produção. A infraestrutura deve dispor de coleta de métricas e visualização em tempo real (latência, taxa de erro, *throughput*). Para gestão de custos de inferência (FinOps), adotar solução com granularidade por serviço e alertas de cota.

## 6.2 Fase - Gestão, retreino e evolução

O objetivo desta fase é tratar incidentes detectados e atualizar o modelo para que ele aprenda com novos dados do mundo real.

### 6.2.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| ***Logs* de retreinamento e versão** | Registro completo da linhagem do novo modelo, incluindo ID único, *dataset* utilizado, hiperparâmetros e métricas de validação. |
| **Plano de contingência (*failover*)** | Documentação de regras de negócio determinísticas ou sistemas legados que assumem o serviço caso o modelo de IA falhe ou se degrade criticamente. |
| **Relatório de lições aprendidas e auditoria** | Documento para governança contendo justificativas técnicas de mudanças, aprovações de comitês e aprendizados com falhas. |

### 6.2.2 Alertas de monitoramento de evolução

Para que a solução continue ativa, os seguintes itens devem ser validados periodicamente:

- Desempenho: a acurácia/precisão permanece acima do limite mínimo (*threshold*) estabelecido?
- Segurança: não foram detectadas vulnerabilidades críticas ou alucinações inaceitáveis?

- Conformidade: o uso ético e a aderência à LGPD foram revalidados pelo ETDP/Comitê?
- Sustentabilidade: os custos de infraestrutura e *tokens* estão dentro do orçamento previsto o e as cotas de consumo não foram violadas?

**A aprovação exige a resposta afirmativa para todos os itens da lista.** Caso existam respostas negativas às questões e os riscos não forem mitigados, a equipe de sustentação deve comunicar os responsáveis para possível ação de desativação (não permanência em operação).

### 6.2.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | ETDP | Gestor de Segurança da Informação | Gestor de TIC | NIA | Sustentador | CISC | RSGI |
|---|---|---|---|---|---|---|---|---|---|---|
| **Decisão sobre permanência** | A | C | I | C | A | A | I* | R | A | C |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor).

* Informado via sistema de organização e registro de soluções de IA.

### 6.2.4 Recomendações, políticas e boas práticas

Nos casos de soluções de IA que envolvam retreino por motivo de inclusão de novos atributos (*features*), as atividades citadas a seguir devem ser previstas nesta fase.

- Executar a Autoavaliação de Impacto Ético para identificar e mitigar riscos introduzidos pelos novos atributos — especialmente quanto à viés, privacidade, equidade, segurança, transparência — com critérios de aprovação definidos antes do retreino.
- Executar o RIPD para mapear os riscos à privacidade introduzidos pelos novos atributos, verificar a base legal do tratamento e documentar as medidas de mitigação adotadas, em conformidade com o Art. 38 da LGPD, com participação do Encarregado de Dados (ETDP).
- Atualizar *o System Card* para refletir as mudanças de comportamento, capacidades e riscos introduzidos pelos novos atributos, identificando claramente as alterações em relação à versão anterior e comunicando limitações relevantes aos usuários e partes interessadas.

A relevância destas recomendações reside no fato de que cada documento responde a uma dimensão distinta de responsabilidade: a Autoavaliação de Impacto Ético orienta decisões internas de mitigação; o RIPD documenta conformidade legal perante a LGPD; e o *System Card* é o único cuja falha é visível externamente. Um RIPD desatualizado é um risco legal; um *System Card* desatualizado é um risco de confiança pública — ele comunica algo sobre o modelo que já não é verdade.

# 7. Etapa - Desativação


Figura 8 - Fases de desativação

Esta etapa tem como objetivo principal garantir que um modelo ou algoritmo desatualizado seja removido de forma controlada, sem impactos negativos nos sistemas dependentes, preservando a rastreabilidade e a conformidade legal.

## 7.1 Fase - Identificação e planejamento

O objetivo desta fase é reconhecer a necessidade de descontinuação (devido a desempenho degradado, mudança de negócio ou requisitos regulatórios) e planejar a transição.

### 7.1.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Ticket de *retirement* (justificativa técnica)** | Documento formalizando o motivo da desativação, como queda de acurácia, obsolescência de dados ou novos requisitos da LGPD ou corte de orçamento. |
| **Mapa de dependências e janela de transição** | Levantamento de todos os sistemas e APIs que utilizam o modelo e definição do cronograma para o desligamento. |

### 7.1.2 O que verificar antes de encerrar

Para que a solução seja desativada, os seguintes itens devem ser avaliados positivamente:

- Os motivos da desativação (viés, erro elevado, mudança regulatória) estão documentados?
- Os *stakeholders* (operação, TI, compliance) foram formalmente notificados?
- Não há dependências críticas de negócio que ficarão descobertas sem o modelo?

**A aprovação exige a resposta afirmativa para todos os itens da lista.** Caso existam respostas negativas às questões, a equipe de sustentação deve comunicar os responsáveis para possível ajuste para a desativação correta.

### 7.1.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Comitê de IA | Sustentador | CISC |
|---|---|---|---|---|---|---|---|---|---|---|
| **Identificar necessidade de desativação / substituição** | A | - | - | I | I | I | - | C | R | I |
| **Elaboração de plano de desativação** | A | C | A | A | C | A | I* | C | R | C |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor).

* Informado via sistema de organização e registro de soluções de IA.

### 7.1.4 Recomendações, políticas e boas práticas

Devem ser previstos os seguintes itens para controle:

- Critérios para retreinamento
- Gestão de versões
- Critérios de desligamento
- Preservação de evidências

## 7.2 Fase - Execução e substituição

O objetivo desta fase é realizar a migração para um algoritmo sucessor ou *fallback* e desativar tecnicamente os componentes do modelo antigo.

### 7.2.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Plano de *cutover* e *shadow mode*** | Relatório comparativo validando o modelo sucessor em paralelo antes da substituição definitiva. |
| **Certificado de desativação segura** | Registro técnico da desconexão de endpoints, arquivamento de pesos/artefatos e desalocação de recursos de *hardware* (GPU/*Storage*) com evidência de parada de faturamento. |

### 7.2.2 O que verificar antes de encerrar

Para que a solução seja desativada, os seguintes itens devem ser avaliados positivamente:

- No caso de substituição, o modelo sucessor atingiu métricas de performance e *fairness* superiores?
- Dados sensíveis foram removidos ou arquivados em conformidade com a LGPD?
- Os recursos de nuvem/hardware foram liberados e a cobrança encerrada?

### 7.2.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Comitê de IA | Sustentador | CISC |
|---|---|---|---|---|---|---|---|---|---|---|
| **Executar desligamento técnico** | A | - | - | I | I | I | - | C | R | I |
| **Aprovação ou rejeição do relatório de desativação** | A | I | I | A | C | A | I* | C | R | A |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); R (Responsável Executor).

* Informado via sistema de organização e registro de soluções de IA.

## 7.3 Fase - Rastreabilidade e auditoria (governança)

O objetivo desta fase é preservar a memória técnica do projeto e garantir a prestação de contas (*accountability*) após o desligamento.

### 7.3.1 Resultados esperados da fase

| Resultado | Descrição |
|---|---|
| **Dossiê de arquivamento (*model registry*)** | Repositório contendo a versão final do código, dados de treino, logs de decisão e razões do *retirement* para futuras auditorias. |
| **Relatório de descontinuação e aprendizados** | Documento final alimentando o inventário de modelos do NIA com o histórico de performance e impacto ético observado. |

### 7.3.2 O que verificar antes de encerrar

Para que o processo de desativação seja considerado concluído, os seguintes itens devem ser avaliados positivamente:

- Todos os *endpoints* foram desconectados e os recursos financeiros/nuvem foram liberados?
- O inventário do NIA foi atualizado e o status do modelo alterado para "Desativado"?
- O Comitê de IA e a Auditoria validaram a rastreabilidade das decisões tomadas pelo modelo durante sua vida útil?

### 7.3.3 Quem faz o quê - papéis e responsabilidades (matriz RACI)

| Atividade | Demandante | Curador de dados | Alta Gestão | Gestor de TIC | ETDP | Gestor de Segurança da Informação | NIA | Comitê de IA | Sustentador | CISC |
|---|---|---|---|---|---|---|---|---|---|---|
| **Auditar e arquivar dados** | I | I | - | I | I | A | - | R | I | I |
| **Atualizar inventário do órgão** | I | I | - | I | I | I | I* | A | R | I |

Legenda: A (Autoridade responsável); C (Consultado); I (Informado); E (Executor); - (Não atua).

* Informado via sistema de organização e registro de soluções de IA.

# 8. Referências bibliográficas


*[1]* OWASP Foundation. Threat Modeling. OWASP Community – The Open Web Application Security Project. Disponível em: https://owasp.org/www-community/Threat_Modeling. Acesso em: 10 dez. 2025.

*[2] MITRE. AI Security 101. ATLAS – Adversarial Threat Landscape for Artificial-Intelligence Systems.* Disponível em: https://atlas.mitre.org/resources/ai-security-101. Acesso em: 10 dez. 2025.

[3] DAMA International. *DAMA-DMBOK: Data Management Body of Knowledge*. 2. ed. Bradley Beach, NJ: Technics Publications, 2017.

[4] National Cyber Security Centre (NCSC). Guidelines for Secure AI System Development: Secure *Design*. NCSC – National Cyber Security Centre, 2023. Disponível em: https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development/guidelines/secure--design. Acesso em: 11 dez. 2025.

[5] CPQD. MLSecOps: Estrutura Técnica e Controles de Segurança para Operações Confiáveis em IA/ML. 2025. Projeto INSPIRE – Meta 4. FINEP.

[6] CPQD. *Metodologia de Desenvolvimento de Soluções de IA*. Projeto INSPIRE. Frente M3.2 – Subfrente M3.2.1. Meta 3. Convênio nº 01.25.0728.00. MGI/Finep. Campinas-SP, 2026.

[7] CPQD. *Metodologia de Desenvolvimento de Soluções de IA: Critérios de identificação e priorização de desenvolvimento de soluções de IA*: Projeto INSPIRE – Meta 3. Frente M3.2 – Subfrente M3.2.1. FINEP. 2026.

[8] CPQD. *Metodologia de Desenvolvimento de Soluções de IA: Anexo técnico – Geração Aumentada por Recuperação (RAG)*. Projeto INSPIRE. Frente M3.2 – Subfrente M3.2.1. Meta 3. Convênio nº 01.25.0728.00. MGI/Finep. Campinas-SP, 2026.

[9] CPQD. *Metodologia de Desenvolvimento de Soluções de IA: Glossário Unificado*. Projeto INSPIRE. Frente M3.2 – Subfrente M3.2.1. Meta 3. Convênio nº 01.25.0728.00. MGI/Finep. Campinas-SP, 2026.

[10] BRASIL. Presidência da República/Gabinete de Segurança Institucional. Instrução Normativa GSI/PR nº 9, de 8 de janeiro de 2026. Altera a Instrução Normativa GSI/PR nº 1, de 27 de maio de 2020. Diário Oficial da União, Brasília, DF, 8 jan. 2026.

[11] BRASIL. Ministério da Gestão e da Inovação em Serviços Públicos/Secretaria de Governo Digital. Portaria SGD/MGI nº 9.511, de 28 de outubro de 2025. Institui o Programa de Privacidade e Segurança da Informação no âmbito dos órgãos e das entidades da administração pública federal direta, autárquica e fundacional, que possuem unidades que integram o Sistema de Administração dos Recursos de Tecnologia da Informação - SISP do Poder Executivo federal. Diário Oficial da União, Brasília, DF, edição 208, seção 1, p. 91, 31 out. 2025.

[12] BRASIL. Ministério da Gestão e da Inovação em Serviços Públicos/Secretaria de Governo Digital. Cartilha do Catálogo Nacional de Dados. Disponível em: https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados/governancadedados/arquivos/CartilhaGovDadosvol3.pdf.