



TOMADA DE SUBSÍDIO	
Edital Nº <b>0692/2025</b>	Objeto <b>IA Soberana</b>
ANEXO I – OPORTUNIDADE E CONTRIBUIÇÕES ESPERADAS	

## 1. OPORTUNIDADE DE NEGÓCIO

1.1 A presente Tomada de Subsídio tem como finalidade reunir contribuições do mercado, incluindo empresas nacionais, Instituições Científicas, Tecnológicas e de Inovação (ICTs) e universidades, para subsidiar a modelagem de uma parceria estratégica com o SERPRO voltada ao desenvolvimento conjunto de uma **solução tecnológica de Inteligência Artificial Soberana**.

1.2 A oportunidade de negócio decorre da identificação de uma lacuna estratégica no setor público de dependência de soluções de IA providas por grandes corporações estrangeiras, sem controle nacional sobre dados, algoritmos e infraestrutura. Ao lado de um parceiro tecnológico, o SERPRO busca o desenvolvimento de uma **plataforma nacional de IA baseada em LLMs (Large Language Models)**, treinada majoritariamente em língua portuguesa, com dados da realidade brasileira, com capacidade de geração e refinamento contínuo de modelos, incluindo mecanismos de explicabilidade, governança e controle ético, bem como ações de transferência de conhecimento e capacitação.

1.3 Entre os benefícios esperados, destacam-se:

- **Soberania de Dados e Conformidade Regulatória**, assegurando que os dados tratados e os modelos gerados estejam sob domínio nacional;
- **Redução de Custos e Menor TCO**, ao oferecer uma alternativa pública à dependência de soluções estrangeiras;
- **Raciocínio e Respostas em Português Brasileiro**, com aderência à realidade local, incluindo expressões regionais, normas jurídicas e contexto sociocultural;
- **Controle Algorítmico Ético e Auditável**, com mitigação de vieses conforme princípios legais e éticos do país;
- **Segurança Estratégica**, com proteção de setores críticos e dados sensíveis do Estado;
- **Interoperabilidade com o Ecossistema Público**, com integração facilitada com sistemas como gov.br, SEI, entre outros;
- **Fomento à Inovação Nacional**, com incentivo à pesquisa aplicada, desenvolvimento tecnológico e capacitação de mão de obra especializada.

1.4 A solução proposta aproveitará a infraestrutura computacional soberana e segura do SERPRO, garantindo sua operação em conformidade com os princípios da **Lei Geral de Proteção de Dados (LGPD)**, do **Plano Brasileiro de Inteligência Artificial (PBIA)** e da **Estratégia Nacional de Segurança Cibernética**.

1.5 Estruturada no modelo **Software como Serviço (SaaS)**, a solução visa garantir escalabilidade, continuidade evolutiva, redução de custos operacionais e aderência às normas legais nacionais, permitindo também o uso colaborativo e a retroalimentação por múltiplas instituições públicas.

1.6 A solução deverá ser ofertada a órgãos da Administração Pública nas esferas federal, estadual e municipal, operando em **infraestrutura computacional segura, sob controle estatal**, em conformidade com as diretrizes de soberania de dados, interoperabilidade e segurança cibernética, visando posicionar o Brasil como protagonista no desenvolvimento de soluções de IA que respeitem sua soberania, promovam a autonomia tecnológica do Estado e ampliem o uso de tecnologias emergentes de forma ética, segura e sustentável.

## 2. CONTRIBUIÇÕES ESPERADAS

2.1 Para a construção do conhecimento que visa auxiliar na modelagem da parceria em oportunidade de negócio, espera-se o recebimento de contribuições relacionadas às questões a seguir:

### 2.2 Aspectos Comerciais e Mercadológicos

2.2.1 Quais são as principais necessidades e problemas enfrentados pela Administração Pública para adoção de soluções baseadas em IA Soberana?

2.2.2 Quais são os desafios regulatórios (ex: LGPD, LAI, Marco Civil) ou normativos internos (portarias, diretrizes setoriais) que afetam a adoção desse tipo de solução?

2.2.3 Quais são os entraves enfrentados para contratação dessas soluções tecnológicas? (orçamento, licitação, equipe técnica, suporte etc.)

2.2.4 Qual é o público-alvo específico no setor público (federal, estadual ou municipal)? Existem setores prioritários?

2.2.5 Quais atividades ou processos da Administração Pública mais demandam soluções de IA de linguagem?

2.2.6 Qual o potencial de crescimento deste mercado considerando as necessidades da Administração Pública?

2.2.7 Quais modelos de comercialização são mais adequados para esse tipo de solução (ex: por transação, assinatura, usuários, módulos)?

2.2.8 É comum a adoção de práticas comerciais adaptadas a diferentes portes de clientes públicos? (planos escalonados, franquias, uso progressivo etc.)

2.2.9 É comum a adoção da prática de descontos progressivos conforme a escala contratada?

2.2.10 Qual o preço médio praticado para esse tipo de solução no mercado público e/ou privado?

2.2.11 É comum a formação de consórcios entre empresas, ICTs ou universidades para entregar esse tipo de solução?

---

2.2.12 A solução poderá ser implementada de forma nacional ou precisa de adaptações regionais?

2.2.13 A subcontratação é comum nesse mercado? Se sim, quais etapas costumam ser subcontratadas?

2.2.14 Quais critérios e práticas costumam ser utilizados para garantir conformidade, auditoria e transparência?

### **2.3 Aspectos Técnico-Operacionais**

2.3.1 Existem soluções similares no mercado nacional ou internacional?

2.3.2 A solução costuma ser implementada de forma centralizada ou segmentada?

2.3.3 Quais funcionalidades ou módulos são considerados estratégicos? (dashboards, NLP, automação de atendimento, etc.)

2.3.4 Quais painéis interativos são comumente oferecidos? (ex: KPIs, volume de requisições, uso de tokens, tempo de resposta)

2.3.5 Quais são os principais requisitos técnicos da solução? (infraestrutura, integração com sistemas, escalabilidade, segurança, multiusuário etc.)

2.3.6 A customização da solução por órgão contratante é comum? Se sim, quais são os limites e possibilidades?

2.3.7 Esse tipo de solução poderá ser oferecida em nuvem pública, nuvem governamental ou operar em infraestrutura do SERPRO?

2.3.8 Esse tipo de solução é compatível com arquitetura multiusuário (*multi-tenancy*) com segurança e isolamento entre instâncias?

2.3.9 Nesse tipo de solução, os processos são automatizados por padrão ou exigem personalização?

2.3.10 Nesse tipo de solução, é comum haver interoperabilidade com sistemas públicos como Gov.br, SEI, Sigepe, e-SIC, etc.?

2.3.11 Nesse tipo de solução, quais mecanismos de segurança são normalmente utilizados? (criptografia, MFA, segregação, logs de auditoria, perfis etc.)

2.3.12 Nesse tipo de solução, quais mecanismos de contingência, replicação ou alta disponibilidade são normalmente utilizados?

2.3.13 Nesse tipo de solução, é comum a adoção de acessibilidade digital (WCAG 2.1, eMAG)? Se sim, como costuma ser a implementação e validação?

2.3.14 Nesse tipo de solução, qual é o prazo médio de implementação e quais fatores costumam impactar o prazo?

2.3.15 Nesse tipo de solução, quantos recursos humanos e tecnológicos são normalmente exigidos para implantação e suporte?

---

2.3.16 Nesse tipo de solução, é comum a adoção de avaliação contínua e retreinamento de modelos? Se sim, em qual frequência?

## **2.4 Fontes e Tratamento dos Dados**

2.4.1 Nesse tipo de solução, quais fontes textuais em português podem ser utilizadas para treinamento? (sites, redes sociais, documentos públicos etc.)

2.4.2 Nesse tipo de solução, os dados costumam ser obtidos sob quais regimes legais? (públicos, licenciados, próprios, parcerias etc.)

2.4.3 Nesse tipo de solução, como costumam ser tratados a privacidade, a anonimização e o consentimento?

2.4.4 Nesse tipo de solução, a base de dados considera as variações regionais e culturais do Brasil, como sotaques e gírias? É comum usar *fine-tuning* para adaptar a solução à linguagem e realidade brasileira, mesmo com dados não estruturados?

2.4.5 Nesse tipo de solução, é comum o uso de metodologia para enriquecimento semântico, contextual ou pragmático dos dados?

2.4.6 Nesse tipo de solução, como costuma ser tratado o direito ao esquecimento?

2.4.7 Nesse tipo de solução, quais garantias costumam ser oferecidas para segurança dos dados em todo o ciclo de vida, incluindo garantias de que os dados não serão transferidos ou utilizados para fora do território nacional?

## **2.5 Soberania Nacional**

2.5.1 Nesse tipo de solução, qual tem sido o entendimento do conceito de Inteligência Artificial Soberana?

2.5.2 Nesse tipo de solução, é comum a adoção de compatibilidade com políticas e diretrizes de soberania digital estabelecidas pelo governo brasileiro?

## **2.6 Características Técnicas dos Modelos**

2.6.1 Nesse tipo de solução, quais tipos de modelos podem ser desenvolvidos ou adaptados? (LLM, SLM, QA, geração etc.)

2.6.2 Nesse tipo de solução, é comum a adoção de modelos multimodais (texto + imagem + áudio)?

2.6.3 Nesse tipo de solução, quais seriam as capacidades médias para treinamento e manutenção dos modelos (em número de parâmetros)?

2.6.4 Nesse tipo de solução, é comum a adoção de modelos de entrada e saída multimodal?

2.6.5 Nesse tipo de solução, é comum a adoção de compatibilidade com os padrões de arquitetura dos sistemas públicos?

---

2.6.6 Nesse tipo de solução, é comum a adoção de mecanismos de governança e controle sobre os parâmetros dos modelos?

## **2.7 Ética, Viés e Segurança**

2.7.1 Nesse tipo de solução, quais metodologias são normalmente utilizadas para detectar, auditar e mitigar vieses?

2.7.2 Nesse tipo de solução, os modelos costumam ser avaliados quanto a *fairness*, toxicidade, alinhamento e segurança?

2.7.3 Nesse tipo de solução, é comum a adoção de mecanismos de *Explainable AI*? (transparência das decisões)

2.7.4 Nesse tipo de solução, os modelos costumam ser auditáveis? Se sim, é comum haver histórico de auditorias internas ou externas?

2.7.5 Nesse tipo de solução, é possível garantir que os modelos pré-treinados foram desenvolvidos com dados públicos e éticos?

2.7.6 Nesse tipo de solução, é possível garantir a diversidade e equidade em modelos de imagem, linguagem e multimodais?

2.7.7 Nesse tipo de solução, como costuma ser feita a curadoria e anotação de dados de forma ética e representativa?

## **2.8 Repasse de Conhecimento**

2.8.1 Nesse tipo de solução, quais formas de capacitação são normalmente oferecidas? (online, presencial, tutoriais, videoaulas etc.)

## **2.9 Infraestrutura**

2.9.1 Nesse tipo de solução, quais infraestruturas de hardware costumam ser utilizadas para treinamento e operação?

2.9.2 É comum que esse tipo de solução opere em infraestrutura nacional, seja em nuvem ou *on-premise*? Em caso de operação *on-premise*, é comum a disponibilização de hardware próprio para essa finalidade?

2.9.3 A solução é escalável para atender a um volume nacional crescente?

## **2.10 Características da Solução**

2.10.1 Nesse tipo de solução, é comum a adoção de interface de gerenciamento de modelos?

2.10.2 Nesse tipo de solução, é comum a adoção de controle e rastreamento de bilhetagem e uso de tokens?

2.10.3 Nesse tipo de solução, em média, quantos modelos podem ser gerados e armazenados simultaneamente?

---

2.10.4 Nesse tipo de solução, é comum a interface mostrar o consumo de recursos por modelo?

## **2.11 Aspectos Gerais**

2.11.1 Nesse tipo de solução, quais serviços complementares costumam ser ofertados? (PoC, diagnóstico, suporte, consultoria etc.)

2.11.2 Nesse tipo de solução, é comum a adoção de Provas de Conceito (PoC)?

2.11.3 Quais SLAs e métricas são aplicadas (tempo de resposta, acurácia, disponibilidade, latência etc.)?

2.11.4 Nesse tipo de solução, há normas ou certificações recomendadas ou exigidas (ISO, ABNT, NIST, OWASP etc.)?

2.11.5 Nesse tipo de solução, quais são os principais riscos associados e quais medidas de mitigação costumam ser implementadas