



**RESPOSTAS DOS QUESTIONAMENTOS**  
**Consulta Pública para Contratação de**  
**Solução de Inteligência Artificial**

Data: 19/01/2023

## 1. Introdução

Este documento destina-se a prestar os esclarecimentos dos questionamentos realizados pelos Licitantes durante Consulta Pública, referente a Contratação de Solução de Inteligência Artificial por um período de 24 (vinte e quatro) meses, incluindo os serviços Suporte Técnico, bem como Capacitação Técnica e 2.856 (duas mil oitocentas e cinquenta e seis) horas de Orientação técnica, a serem utilizados sob demanda.

## 2. Respostas aos Questionamentos da Empresa – PPN TECNOLOGIA

Questionamentos:

- No Capítulo 3. PROVA DE CONCEITO

item b) A LICITANTE deverá prover todo o ambiente necessário (componentes de hardware e software e a conexão de Internet), no endereço acima, para comprovação do atendimento às funcionalidades do ANEXO I – ESPECIFICAÇÃO TÉCNICA....

Pergunta: O ambiente mencionado no item acima poderá ser disponibilizado no laboratório do próprio fabricante/Licitante?

**Resposta: Não, a Prova de Conceitos será realizada nas instalações da Dataprev como descrito no Termo de Referência.**

- No Capítulo 3. PROVA DE CONCEITO

item c) Disponibilizar 01 (um) técnico que se responsabilizará pela comprovação das funcionalidades e requisitos em conformidade com o ANEXO I – ESPECIFICAÇÃO TÉCNICA, por meio de testes práticos ou por comandos de configuração. A aprovação das funcionalidades existentes na Contratação da Solução de Inteligência Artificial apresentada pela LICITANTE serão efetuadas pela Equipe Técnica da DATAPREV.

Pergunta: Levando em conta que a POC contém especificações de diversos alcances podemos considerar a necessidade de mais de 1 perfil de profissional especialista nas áreas de infra, configuração, negócios para comprovação dos requisitos e por isso se faz necessário mais de um 1 técnico disponível para realização da POC. Podemos considerar a frase "Disponibilizar pelo menos 01 técnico que...."

**Resposta: Sim, contudo tendo um técnico como o responsável, como indicado no Termo de Referência, capítulo 3.**

No Capítulo 6. INSTALAÇÃO

Item 6.2. O prazo para conclusão da instalação/disponibilização dos produtos será de 20 (vinte) dias úteis, contados a partir do dia seguinte à ocorrência dos fatos abaixo.

Pergunta: O prazo informado para conclusão da instalação de 20 dias é relativamente pequeno para instalação de uma solução de IA/VC on premise ainda mais sendo instalação "assistida",

levando em conta a complexidade e características de integração com fontes de dados, ldap, etc e a devida homologação. Para esse prazo sugerimos um tempo de 40-45 dias.

**Resposta: A sugestão será avaliada e indicado posicionamento final no Termo de Referência.**

#### No Capítulo 7. ORIENTAÇÃO TÉCNICA

Item 7.1.2 Orientação à equipe da Dataprev quanto ao uso das tecnologias presentes na Solução de Inteligência Artificial, tais como: Estruturas Semânticas, Algoritmos de Aprendizagem de Máquina, Framework Data Science, Linguagem R, Python e Scala.

Pergunta: Entendemos que a orientação técnica mencionada nesse item mais precisamente no que se refere a linguagem R, Python e Scala, são linguagens não da solução contratada e portanto, apesar da solução aceitar o uso, não são focos de apoio do fabricante ou contratada. Para que seja atendido o contratante deverá disponibilizar também profissionais com esses skills o que pode onerar o contrato.

**Resposta: O item 1.3.2.7 faz referência às linguagens open source, Python, R e Scala, portanto o item 7.1.2 está adequado.**

#### No Capítulo 16 SIGILO E INVIOLABILIDADE

Item 16.7 Documentação que comprove a existência de plano de continuidade de negócio com gestão de riscos e resposta à incidentes, bem como comunicar seus recursos, papéis e responsabilidades de segurança da informação para o uso de seus serviços. A gestão de riscos deverá estar alinhada a legislação brasileira.

Pergunta: Entendemos que o plano de continuidade de negócio com gestão de riscos e resposta a incidentes e o plano de continuidade, contendo as ações de recuperação de desastres e contingência de negócio assim como o plano de resposta à incidentes, é pertinente apenas as soluções que tenham o modelo de funcionamento em nuvem pública. Para o caso de soluções on-premise esse plano deverá ser gerado pela CONTRATANTE. Correto nosso entendimento?

**Resposta: Sim, está correto o entendimento. O Termo de Referência deverá ser ajustado.**

Item 16.10 A CONTRATADA deverá apresentar avaliações periódicas, realizadas por empresas independentes no mínimo a cada ano, para apresentação de relatórios tipo I e II de auditoria SOC2 (Service and Organization Controls 2).

Pergunta: Entendemos que os relatórios referentes ao item acima são pertinente apenas as soluções que tenham o modelo de funcionamento em nuvem pública. Correto nosso entendimento?

**Resposta: Sim, está correto o entendimento. O Termo de Referência deverá ser ajustado.**

#### No Capítulo 19. OBRIGAÇÕES DA CONTRATADA

19.9. As CONTRATADAS deverão cumprir os seguintes itens...

Pergunta: Entendemos que os subitens referentes a esse item 19.9 inteiro são pertinentes apenas as soluções que tenham o modelo de funcionamento em nuvem pública. Correto nosso entendimento?

**Resposta: Sim, está correto o entendimento. O Termo de Referência deverá ser ajustado.**

No Capítulo 24. VIGÊNCIA CONTRATUAL

Item 24.1. A vigência do contrato será de 28 (vinte e oito) meses.

Pergunta: O período do contrato informado no item 1.1 OBJETO é de 24 meses sendo diferente da vigência do contrato informado nesse item que é de 28 meses.

**Resposta: O serviço será prestado por 24 meses, pagos em 24 parcelas a partir do termo de aceite da solução. A vigência contratual está com 28 meses para que contemple os prazos antes dos termos de aceite da solução.**

Comentários adicionais:

Sugerimos a revisão da citação no texto que informa o seguinte: "a prova acontecerá no local de entrega do equipamento"

**Resposta: Sim, o texto será corrigido.**

O tópico PLN é parte integrante do Lote 1 mas seu treinamento está contido numa turma de introdução do Lote 2. Sugerimos que o tópico referente a PLN componha um treinamento específico.

**Resposta: O item será corrigido.**

## **Respostas aos Questionamentos da Empresa – LOGICALIS**

Esclarecimento 1

Do item 2.1.13. A SOLUÇÃO DE IA deve permitir que todas as funcionalidades estejam disponíveis nos ambientes de desenvolvimento, teste, homologação e produção, evitando que alguma funcionalidade seja disponibilizada exclusivamente numa versão desktop da solução.

Entendemos que os licenciamentos para ambientes de desenvolvimento, teste e homologação estão inclusos nos quantitativos solicitados, está correto o entendimento?

**Resposta: Sim, o entendimento está correto.**

Esclarecimento 2

Do item 1.3.2.16. A SOLUÇÃO DE IA deve possibilitar a integração com plataformas de hardware para aceleração de treinamento de modelos de Deep Learning com GPU.

Nossa solução possui licenciamento adicional para processamento em GPU.

Solicitamos informar a quantidade ou estimativa de GPUs a serem utilizadas na solução de IA.

Resposta: Não é possível informar a quantidade ou estimativa de GPUs a serem utilizadas na solução de IA. É desejável que a solução a ser ofertada não tenha restrições quanto à quantidade de GPUs presentes na infraestrutura de hardware que deverá suportar a solução.

#### Esclarecimento 3

Do item 1.4.13.4.1. A solução deverá se limitar ao consumo total de recursos de hardware descrito nos itens 2.4.14.5 e 2.4.14.5.2. Os containers deverão ser instanciados na plataforma em uso pela Dataprev em nós worker de clusters Kubernetes com 4 vCPUs e 16GB de memória RAM.

A fim de verificar a compatibilidade com a plataforma que iremos ofertar, questionamos qual a plataforma de kubernetes que a Dataprev utiliza atualmente? Também favor informar a versão atual.

Resposta: Versão Origin Community Distribution of Kubernetes disponibilizada pela Red Hat Enterprise Linux 7.5

Ainda em relação ao item 1.4.13.4.1, caso a plataforma que pretendemos ofertar através de nossos parceiros tenha integração direta com o ambiente existente na Dataprev, entendemos que podemos considerar a integração da Solução de IA com a solução existente na Dataprev, está correto o nosso entendimento?

Resposta: Não está correto o entendimento. Com exceção dos recursos indicados como disponíveis na Especificação Técnica, a solução deverá ser ofertada e instalada integralmente pela contratada.

Também para o correto licenciamento da nossa solução necessitamos saber a estimativa de armazenamento para a solução de IA, por favor podem informar?

Resposta: Não é possível informar a estimativa de armazenamento para a solução de IA. A proponente deverá estimar quaisquer necessidades para este fim com base nas informações de volumetria contempladas no ANEXO I - ESPECIFICAÇÃO TÉCNICA.

#### Esclarecimento 4

Do ANEXO VI – CAPACITAÇÃO TÉCNICA.

Entendemos que a capacitação exigida poderá ser realizada por instrutores da contratada, desde que atendidas as certificações exigidas no item 1.4. Instrutoria.

Está correto nosso entendimento?

Resposta: A CONTRATADA deverá atender as especificações e exigências do item 1.4 do Anexo VI no que diz respeito a Instrutoria.

#### Esclarecimento 5

Do item 2.4.5.A comunicação dos ambientes remotos com a CONTRATANTE deve implementar TLS 1.3 como métodos criptográficos para garantia da confidencialidade e da autenticidade da comunicação.

Solicitamos que também seja aceita a possibilidade de utilizarmos o TLS 1.2.

Nossa solicitação será atendida?

Resposta: A solicitação não será atendida.

## Respostas aos Questionamentos da Empresa – IBM

### Questionamento 1:

Em relação ao item 1.5.2, não identificamos no TR e na Especificação Técnica informações suficientes para avaliar, de forma antecipada, qual será o tempo de resposta estimado da plataforma de IA. Nosso entendimento é que a Dataprev deve especificar de forma detalhada qual será o cenário de teste, igual para todos os fornecedores, que será avaliado na Prova de Conceito. Para o caso dos modelos de Machine Learning e Deep Learning, exemplos de informações a serem adicionadas são:

- Qual o dataset público que será usado para treinamento?
- Qual o algoritmo de Machine Learning que vai ser usado?
- Quais serão as camadas do modelo de Deep Learning que vai ser usado?
- Qual linguagem, biblioteca e framework serão usados para esses modelos?
- De qual forma o tempo de resposta será medido? Qual ferramenta será utilizada para isso?
- Qual o digrama de redes no qual a solução de IA estará inserida? Onde fica o ponto que envia a requisição nesse diagrama?

A Dataprev considera viável essa possibilidade?

**Resposta:** As sugestões indicadas que não tenham descrição correspondente no ANEXO I - ESPECIFICAÇÃO TÉCNICA ou no item 3 do Termo de Referência serão avaliadas para eventual melhoria no requisito.

Esclarecemos que todas as informações sobre recursos e procedimentos suficientes para PROVA DE CONCEITO em sua etapa de comprovação das funcionalidades e requisitos do ANEXO I – ESPECIFICAÇÃO TÉCNICA estão descritas no 3 do Termo de Referência, como por exemplo:

"Para a realização da Prova de Conceito, a LICITANTE fornecerá os recursos de hardware, software e rede necessários para o acesso às funcionalidades disponibilizadas nos limites e especificação estabelecidos no Anexo I -Especificação técnica. A LICITANTE deve considerar que a DATAPREV não irá disponibilizar qualquer recurso de hardware e software para a realização da Prova de Conceito."

"Para comprovação dos itens 1.5.2, 2.5.2 e 2.5.3 do ANEXO I - ESPECIFICAÇÃO TÉCNICA, deverá ser utilizado um dataset com no mínimo 72.000 registros. O dataset deve ter até 50% de campos não tratados, deve ser composto por dados públicos e não deve fazer parte de bibliotecas de motores de Inteligência Artificial (Ex: keras fashion\_mnist)."

### Questionamento 2:

Em relação ao item 1.5.3, não identificamos no TR e na Especificação Técnica as informações necessárias para definir, de forma antecipada, qual volume estimado de usuários simultâneos a plataforma de IA será capaz de suportar. Nosso entendimento é que a Dataprev deve especificar de qual será o cenário de teste, igual para todos os fornecedores, que será avaliado na Prova de Conceito. Para o caso do Ambiente Integrado de Desenvolvimento de Machine Learning e Análise Preditiva, exemplos de informações a serem adicionadas são:

- Qual atividade cada um desses usuários está realizando? Executando notebooks, fazendo tratamento de dados, usando recursos de AutoML?
- De qual forma o tempo de resposta será medido? Qual ferramenta será utilizada para isso?

A Dataprev considera viável essa possibilidade?

Resposta: As sugestões indicadas que não tenham descrição correspondente no ANEXO I – ESPECIFICAÇÃO TÉCNICA ou no item 3 do Termo de Referência serão avaliadas para eventual melhoria no requisito.

O item 1.5.3 do Anexo I do Termo de Referência indica a quantidade mínima de acessos simultâneos nos perfis indicados. Os procedimentos, recursos, volumetria e demais elementos necessários à realização da Prova de Conceitos de Homologação estão detalhados no item 3, PROVA DE CONCEITO do Termo de Referência. Os itens ali descritos remetem aos elementos suficientes de um cenário de teste para o propósito da Prova de Conceitos.

Considerando a pluralidade de soluções e suas características específicas quaisquer parâmetros para comprovação das funcionalidades e requisitos do ANEXO I – ESPECIFICAÇÃO TÉCNICA que não estejam definidos no item 3 do TR são de responsabilidade da LICITANTE.

O item 1.5.3 do Anexo I do Termo de Referência indica a quantidade mínima de acessos simultâneos nos perfis indicados. Os procedimentos, recursos, volumetria e demais elementos necessários à realização da Prova de Conceitos de Homologação estão detalhados no capítulo 3. PROVA DE CONCEITO do Termo de Referência. Os itens ali descritos remetem aos elementos suficientes de um cenário de teste para o propósito da Prova de Conceitos, contudo, as sugestões serão analisadas internamente.

Questionamento 3:

Em relação aos itens 1.4.13.5.1 e 1.4.13.5.2, considerando que a restrição dos requisitos de hardware pode limitar a entrega da melhor configuração da solução, sugerimos a retirada dos limites de cores e vCPUs.

Resposta: A sugestão não será acatada. Os limites de configuração indicados têm o propósito de assegurar solução com consumo razoável de recursos computacionais. Caso haja restrição de atendimento ao requisito solicitamos que a interessada indique os parâmetros de requisitos de hardware necessários, devidamente justificadas, para suportar a solução a ser ofertada.

Questionamento 4:

Em relação ao item 1.1.5. o nosso entendimento é que PII não é um protocolo e sim um tipo de classificação de informação. Dessa forma, se a solução permitir a classificação de conteúdos sensíveis como PII, podemos considerar que atendemos o requisito? Em caso de uma resposta negativa, onde podemos buscar mais informações sobre a aderência ao protocolo PII?

Resposta: Sim, pode ser considerado que atende o requisito.

Questionamento 5:

Em relação ao item 1.2.2.1, entendemos que a bilhetagem tem como objetivo a transparência dos recursos sendo consumidos pela CONTRATANTE que gerarão custo. Nesse sentido, para os recursos que são cobrados por core virtual, devem ser apresentados os respectivos consumos de infraestrutura. Nosso entendimento está correto?

Resposta: Sim, o entendimento está correto, desde que o recurso referenciado seja um insumo que irá compor o custo de utilização da solução de IA, conforme descrito no referente requisito.

Questionamento 6:

Em relação ao 1.3.5.6, a solução de IA permite a criação de modelos de classificação de texto a partir de ferramentas open source ou proprietárias. Entendemos que a CONTRATANTE pode escolher pelos modelos open source quando necessitar de exportação, atendendo ao requisito. Nosso entendimento está correto?

Resposta: Não, o entendimento não está correto. O item indica que os modelos de classificação de texto deverão ser criados a partir da solução de IA atendendo às indicações descritas na especificação.

Questionamento 7:

Em relação ao item 1.3.5.2, entendemos que a extração de entidades, relacionamento e a classificação de parágrafos, são considerados insumos para criar um resumo de texto. Nosso entendimento está correto?

Resposta: Não, o entendimento não está correto. O resumo deve ser criado de forma extrativa ou abstrativa com trechos ou principais conceitos do texto original. O item será ajustado para melhor especificar a descrição.

Questionamento 8:

Em relação ao item 1.3.2.10, entendemos que a solução de AutoML (ou AutoAI) deve também, obrigatoriamente, fazer pré processamento dos dados de entrada (como detectar tipos, criar partições), mapear características mais relevantes e a melhor forma de usá-las no modelo (Feature Engineering), escolher qual o modelo é considerado o mais adequado para o problema e também em quais parâmetros ele deve ser executado (Hyper Parameter Optimization) e, por fim, exibir sua pipeline de escolha e possibilitar o export para notebooks, de modo que o usuário possa adaptar o modelo manualmente. Está correto nosso entendimento?

Resposta: Não, o entendimento não está correto, quanto a obrigatoriedade. O item expressa a necessidade essencial exigida, contudo sem restrição para funcionalidades adicionais.

Questionamento 9:

Em relação ao item 1.3.1.2, entendemos que os projetos também devem mostrar os ativos (arquivos com dados, tabelas, modelos, etc) disponíveis aos usuários através de um catálogo, de forma a associar cada um deles a termos negócio, metadados e uma linhagem, com todos os processamentos e melhorias que sofreram. Está correto nosso entendimento?

Resposta: Esclarecemos que a oferta de solução com os parâmetros indicados atenderá ao requisito. No entanto o questionamento apresenta funcionalidades que excedem ao definido no requisito, os quais não são obrigatórios.



#### Questionamento 10:

Em relação ao item 1.3.2, para manter uma flexibilidade aos usuários compatível com as melhores práticas do mercado, entendemos que a solução de IA também deve disponibilizar ambientes de desenvolvimento na linguagem R, como o R Studio. Nosso entendimento está correto?

**Resposta: Não, o entendimento não está correto. Os recursos para codificação não se restringem à linguagem R, como descreve o item 1.3.2.7.**

#### Questionamento 11:

Em relação ao requisito 1.3.4.5, entendemos que a plataforma de IA deve disponibilizar também os recursos de interface para rotulação de texto, permitindo que usuários de áreas de negócio possam ensinar os modelos de extração de entidades. A solução deve ainda permitir a gestão de anotadores, definindo tarefas de anotação, tempos para conclusão e indicando possíveis conflitos. Nosso entendimento está correto?

**Resposta: Considerando que o item referenciado seja o 1.3.5.5, esclarecemos que a oferta de solução com os parâmetros indicados atenderá ao requisito. No entanto o questionamento apresenta funcionalidades que excedem ao definido no requisito, os quais não são obrigatórios.**

#### Questionamento 12:

Em relação ao item 1.3.3.4, entendemos que cada natureza de modelo possui seu próprio formato de exportação. Nesse sentido, a solução deve gerenciar os modelos provendo um repositório para armazenamento de objetos para os arquivos, de onde eles podem ser baixados para uso futuro (exportados). Nosso entendimento está correto?

**Resposta: Se a solução permitir que o Gerenciador de Modelos efetuar exportação e importação de modelos nos formatos exigidos no referido repositório, esclarecemos que a oferta de solução com os parâmetros indicados atenderá ao requisito.**

#### Questionamento 13:

De acordo com os itens 2.3.1.3, 2.3.1.4, 2.3.1.5, 2.3.1.8 e 2.3.1.9, referente a SOLUÇÃO de IA na funcionalidade OCR, é solicitado um "grau de confiança nos elementos citados e deve atingir no mínimo 80% de acurácia na conversão dos documentos ou imagens processadas."

Considerando que as soluções de captura de imagem (OCR) suportam uma grande variedade de documentos (pdf, tif, jpg, documentos manuscritos...), e que temos o fator "qualidade do documento" (DPis, compressão, conservação, marcas d'água, cores...), sugerimos a retirada do nível de acuracidade em virtude dos fatores citados acima. Para medir um grau de confiança na solução, recomendamos que sejam avaliadas em um teste de homologação (prova de conceito) as seguintes técnicas:

- Des skew: Endireitar uma imagem ligeiramente inclinada;
- Rotação: Fixar a orientação de uma imagem para que fique na vertical;
- Despeckle: remove manchas e ruído de fundo de uma imagem, deixando apenas o texto;
- Remoção e recorte de borda: removendo as bordas de uma imagem para o tamanho real da imagem. Isso é especialmente importante com fotos de cheques;
- Remoção de linha: remoção de linhas dentro de um documento, deixando apenas o texto;

- Correção de texto inverso: O texto inverso (texto branco sobre fundo preto) pode não ser reconhecido, portanto, o texto inverso pode ser corrigido.

Resposta: Não identificamos no questionamento indicação de restrição ao requisito. Solicitamos que a interessada forneça mais insumos sobre o assunto na sessão pública para que a equipe possa avaliar.

#### Questionamento 14:

De acordo com o item 2.5.3, "Tempo máximo de resposta das requisições de OCR: 10 a 40 milissegundos para no mínimo 95% das requisições, definidas no item 3.3.1.1."

Considerando que as soluções de captura de imagem (OCR) suportam uma grande variedade/tamanho de documentos, e que temos o fator "melhoria da qualidade", além de depender de meios externos para realização do processo (scanner, rede e engine de OCR), sugerimos a retirada do percentual mínimo de tempo de leitura. Para medir um grau de confiança na solução, recomendamos que sejam avaliadas em um teste de homologação (prova de conceito), utilizando um conjunto de documentos padronizados entre os participantes.

Resposta: Não identificamos no questionamento indicação de restrição ao requisito. Solicitamos que a interessada forneça mais insumos sobre o assunto na sessão pública para que a equipe possa avaliar.