

ESTUDO TÉCNICO PRELIMINAR (ETP)

				Número do Processo - SISLOG 110374
				Número do Processo - SEI 202400005043096

Em conformidade com a Lei federal nº 14.133, de 01 de abril de 2021 e com o Decreto estadual nº 10.207, de 27 de janeiro de 2023, o Estudo Técnico Preliminar - ETP é o documento constitutivo da primeira etapa do planejamento de uma contratação a fim de atender a uma necessidade administrativa, e tem por objetivo subsidiar a elaboração do Anteprojeto, Termo de Referência ou Projeto Básico, bem como do edital de licitação e da minuta contratual, quando aplicável.

Tópico 1 - DESCRIÇÃO DA NECESSIDADE DA CONTRATAÇÃO

- 1.1. O presente Estudo Técnico Preliminar apresenta os estudos técnicos realizados visando identificar e analisar as soluções disponíveis no mercado, em termos de requisitos, alternativas e justificativas para escolha da melhor solução para alcançar os resultados pretendidos.
- 1.2. Assim, a delimitação da solução nos termos e condições estipulados não é decisão de livre arbítrio desta equipe. Aqui estão pautados elementos que, fundamentadamente, têm a capacidade e potencial para, em tese, considerando o caso concreto, melhor atender ao interesse público.

Previsão no Plano de Contratações Anual:

- 1.3. A demanda a ser contratada está prevista no PCA 2025.

Alinhamento Estratégico:

- 1.4. Esta pretendida contratação apresenta conformidade com os Programas e Ações do PPA 2023-2025, relacionados às atribuições desta Pasta, em conformidade com as suas competências, nos termos da [Lei nº 22.317, 18 de outubro de 2023](#).
- 1.4.1. Atualmente tem-se vigente o planejamento estratégico da ECONOMIA, válido para o período de 2023-2025, onde estão estabelecidas diretrizes que devem ser seguidas e que validam a presente necessidade de modernização, conforme podemos observar abaixo:



- 1.4.2. Para os objetivos estratégicos OE2 (Melhorar os serviços prestados aos contribuintes) e OE7 (Modernizar a Infraestrutura e sistemas de TIC com foco na transformação digital) deve-se buscar iniciativas efetivas, que melhorem a qualidade dos serviços prestados ao cidadão e a TIC é o melhor instrumento/caminho para se alcançar tamanho desafio.
- 1.4.3. Especificamente, sobre o OE7, disponibilizar serviços públicos ao cidadão é um caminho a ser trilhado, conforme demonstrado abaixo:

OBJETIVO ESTRATÉGICO 07: Modernizar a infraestrutura e sistemas de TIC com foco na transformação digital

Desafio 7.2 – Aumentar a disponibilidade de todos os serviços para público interno e externo

Indicador 7.2.1 – Disponibilidade de serviços

Descrição: Mensura o percentual de serviços, para público interno ou externo, com índice de disponibilidade acima de 99%, aferidos no final do ano.

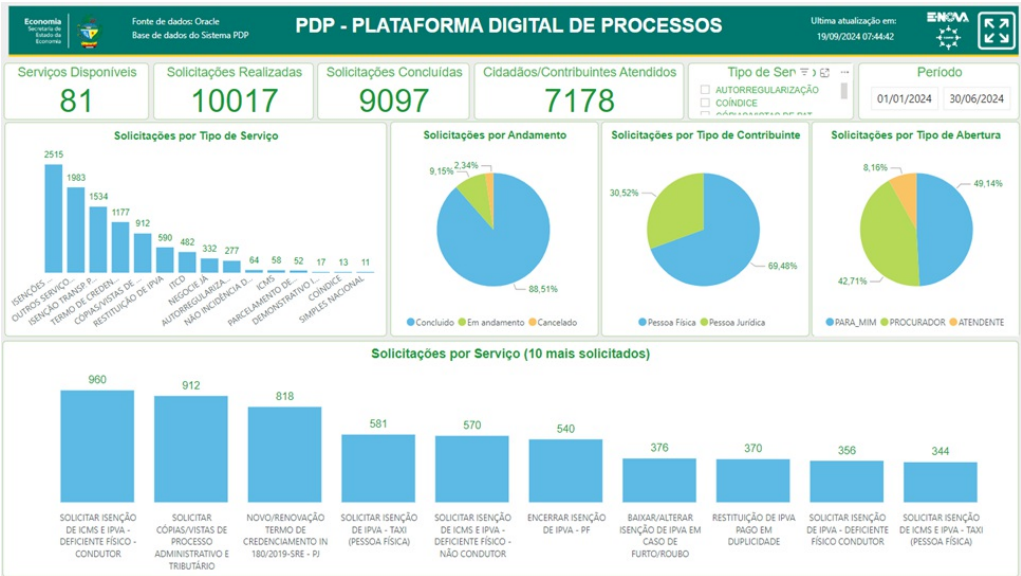
Fórmula de cálculo:
Serviços com disponibilidade maior que 99%/total de serviços*100

Linha de base: ?

Polaridade: Menor melhor

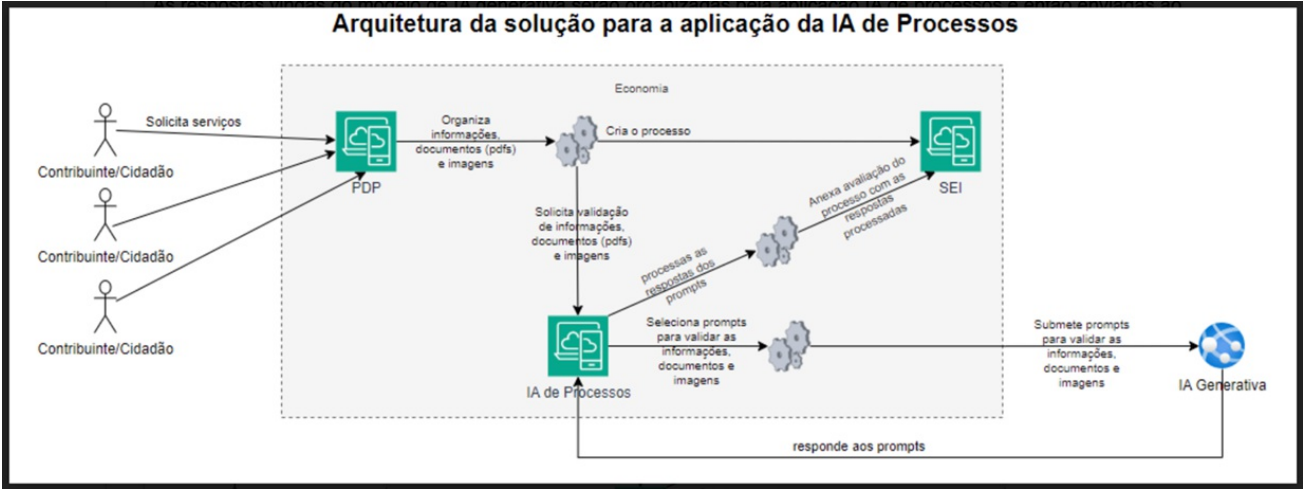


- 1.4.4. Considerando o planejamento estabelecido, temos como principal objetivo permitir a disponibilização de serviços de qualidade para o cidadão, permitindo sua autonomia nas demandas junto à ECONOMIA, acelerando etapas, reduzindo tempos de atendimento, desafogando atendimento presenciais, gerando valor agregado à prestação serviços à sociedade/cidadão/contribuinte.
- 1.4.5. O cidadão/contribuinte demanda dezenas de diferentes tipos de serviços, seja de forma presencial ou virtual.
- 1.4.6. O atendimento de todos esses serviços geram imensas bases de dados no DATACENTER, cujo crescimento é diário e aumenta de forma exponencial.
- 1.4.7. Ao mesmo tempo temos tecnologias inovadoras, como inteligência artificial, que permitem ganhos significativos e estão totalmente aderentes ao planejamento estratégico vigente.
- 1.4.8. Devido a essa quantidade massiva de dados à disposição da ECONOMIA, torna-se imprescindível a utilização de ambiente de processamento que agregue valor ao negócio, fornecendo disponibilidade, segurança, desempenho e inovação.
- 1.5. Cenário Atual
- 1.5.1. Atualmente a Secretaria de Economia de Goiás não faz uso de tecnologias de inteligência artificial, de forma sistematizada e em ambiente de produção nas diversas soluções tecnológicas implantadas para as diversas áreas de negócio.
- 1.5.2. Muito dos serviços disponíveis para o cidadão/contribuinte estão implantados através do PDP (Plataforma Digital de Processos), e necessitam de automação e aplicação de algoritmos de inteligência artificial para melhorar sua acurácia, efetividade e desempenho.
- 1.5.3. Dentro da plataforma PDP podemos citar os seguintes casos de uso nos quais poderemos aplicar estes algoritmos de inteligência artificial em questão, porem não limitados a estes:
- 1.5.3.1. Serviços de IPVA, totalizando mais de 40 diferentes tipos de serviço disponíveis para solicitação do cidadão e contribuinte;
- 1.5.3.2. Reconhecimento de documentos e processamento de imagem para os processos de ITCD, totalizando mais de 5 diferentes tipos de serviço disponíveis para solicitação do cidadão e contribuinte;
- 1.5.3.3. Diversos outros serviços tais como, ICMS, PAT, Autorregularização, termo de credenciamento, coíndice, totalizando mais de 30 diferentes tipos de serviço disponíveis para solicitação do cidadão e contribuinte;
- 1.5.3.4. Pesquisa inteligente em diferentes bases de documentos tais como, pareceres e legislações.
- 1.5.3.5. Elaboração, correção e simplificação de documentos.
- 1.5.4. Apresenta-se alguns dados extraídos da PDP que demonstram o volume crescente dos serviços consumidos pelo cidadão/contribuinte:



- 1.5.5. De forma a contextualizar o funcionamento da PDP podemos descrever seu funcionamento como:
- 1.5.5.1. O contribuinte/cidadão pode solicitar diversos serviços para a Secretária de Economia de Goiás através do sistema PDP (Plataforma Digital de Processos). Para realizar a solicitação de cada serviço o contribuinte/cidadão precisa fornecer uma série de informações, documentos e imagens.
- 1.5.5.2. O PDP transforma essas informações, documentos e imagens em um processo no sistema SEI. Através do SEI cada solicitação de serviço tramita então pelas diversas áreas de negócio necessárias para o atendimento da solicitação.
- 1.5.5.3. Cada informação, documento e imagem precisam ser validadas de acordo com diversas regras de negócio correspondentes a cada tipo de serviço.
- 1.6. Aplicação da IA:
- 1.6.1. Com base no descritivo da situação atual e na busca de evoluir de forma significativa as aplicações disponibilizadas dentro da PDP, necessário se faz, introduzir soluções que venham fazer uso da inteligência artificial para a melhoria da tramitação dos serviços disponibilizados.
- 1.6.2. O uso de IA em aplicações corporativas existentes dentro da Secretaria de Economia tem-se tornado fundamental e estratégico para evoluir com os serviços disponibilizados digitalmente. Com o uso de novas tecnologias envolvendo IA pretende-se:

- 1.6.2.1. Realizar Análise Preditiva: permitir analisar dados históricos para prever tendências futuras, melhorando estratégias;
- 1.6.2.2. Auxiliar a Tomada de Decisão baseada em evidências: ajuda a coletar, analisar e interpretar dados, proporcionando insights relevantes para a tomada de decisões. Isso permite que os gestores façam escolhas baseadas em evidências, evitando “achismos”;
- 1.6.2.3. Melhorar a Eficiência Operacional: tecnologias inovadoras permitem automatizar processos, reduzindo tempos, evitando retrabalho e minimizando perda de produtividade;
- 1.6.2.4. Maximizar as Oportunidades e Mitigar os Riscos: avaliações inteligentes de dados geram oportunidades de melhoria, mitigam riscos potenciais de erros intencionais ou mesmo, não intencionais;
- 1.6.2.5. Gerar Otimização: tais tecnologias inovadoras ajudam na busca de ações eficazes, nos inúmeros contextos que envolvem os processos automatizados;
- 1.6.3. Em resumo, a contratação em questão irá melhorar os diversos serviços digitais disponibilizados ao cidadão/contribuinte.
- 1.6.4. Aplicada ao PDP a IA poderá contribuir com:
- 1.6.4.1. O PDP transforma as informações, documentos e imagens enviadas pelo cidadão/contribuinte em um processo no sistema SEI. Através do SEI cada solicitação de serviço tramita então pelas diversas áreas de negócio necessárias para o atendimento da solicitação.
- 1.6.4.2. Cada informação, documento e imagem precisam ser validadas de acordo com diversas regras de negócio correspondentes a cada tipo de serviço. Essas validações poderão ser realizadas pela aplicação IA de Processos que irá selecionar os prompts (pré-cadastrados) necessários para a validação de cada tipo de serviço e então submete-los (juntamente com o devido contexto formado pelas informações, documentos e imagens informados na solicitação de serviço) para avaliação e resposta por um modelo de IA generativa.
- 1.6.4.3. As respostas vindas do modelo de IA generativa serão organizadas pela aplicação IA de Processos e então enviadas ao processo SEI da solicitação de serviço em questão.
- 1.6.4.4. Essa mesma IA de Processos também poderá ser utilizada para avaliar outros processos bem como documentos de diversos repositórios, tais como pareceres, legislação, ...
- 1.6.4.5. Para esse cenário de aplicação temos a arquitetura de referência abaixo:



- 1.7. Impactos caso a Necessidade Não Seja Atendida:
- 1.7.1. Caso a contratação não seja realizada teremos os seguintes impactos:
- 1.7.2. Dificuldade em melhorar os serviços prestados aos contribuintes/cidadãos;
- 1.7.3. Dificuldade em avançar na transformação digital;
- 1.7.4. Dificuldade em analisar dados históricos;
- 1.7.5. Perda de produtividade devido a reduzida eficiência operacional gerada pela complexidade da redução do tempo e do retrabalho;
- 1.7.6. Dificuldade da geração de otimizações na prestação de serviço e no atendimento ao contribuinte/cidadão;

Tópico 2 - DESCRIÇÃO DA SOLUÇÃO

Definição da solução escolhida

2.1. Abaixo segue a descrição resumida do objeto a ser contratado, definido após a realização de estudo técnico preliminar: **Fornecimento de Bens e Materiais e Serviços - Contratação de solução para processamento de inteligência artificial, incluindo serviços de orientação técnica especializada sob demanda, por 12 (doze) meses.**

Característica do objeto:

- 2.2. O objeto a ser contratado é comum, assim considerado por possuir padrão de desempenho e qualidade que possam ser objetivamente definidos no Termo de Referência, por meio de especificações usuais no mercado, na forma do inciso XIII do art. 6º da Lei Federal nº 14.133 de abril de 2021.
- 2.3. A solução adotada trata-se de objeto comum, pois:
- 2.3.1. é encontrado e praticado no mercado sem maiores dificuldades;
- 2.3.2. é ordinário, sem peculiaridades ou características especiais;
- 2.3.3. é apresentado com identidade e características padronizadas, com perfil qualitativo passível de ser descrito objetivamente; e
- 2.3.4. sua caracterização é garantida tendo por base as exigências detalhadas do Termo de Referência, compatível com o rito procedimental de seleção do fornecedor a ser adotado.
- 2.4. A solução será dividida em itens que compõe um único lote, a serem descritos no Termo de Referência, sendo:
- a. Créditos de serviços de computação de IA na plataforma GCP, para utilização exclusiva em soluções que apliquem inteligência artificial;
- b. Orientação técnica especializada, para utilização em desenvolvimento de soluções envolvendo inteligência artificial, de acordo com os requisitos de negócio e requisitos tecnológicos considerados indispensáveis e/ou importantes para o atendimento das necessidades da STI/ECONOMIA.

2.5. Requisitos da Solução

- 2.5.1. A solução se destina a atender requisitos técnicos de serviços essenciais que necessitam de avanços tecnológicos através da introdução de inteligência artificial, de forma a possibilitar a utilização de Plataforma como Serviço (PaaS) com o intuito de maximizar a disponibilidade dos serviços, sem degradar os outros conceitos de segurança da informação, confidencialidade e integridade.
- 2.5.2. A plataforma tecnológica deverá atender as necessidades de pipeline, tratamento, processamento e governança de Dados;
- 2.5.3. São requisitos tecnológicos da solução:
- 2.5.3.1. Integração com ecossistema tecnológico existente dentro da Economia.
- 2.5.3.2. Ambiente de IA com características:
- 2.5.3.3. Ética e responsabilidade;
- 2.5.3.4. Compreensão contextual;
- 2.5.3.5. Diversidade de aplicações;
- 2.5.3.6. Janela de contexto;
- 2.5.3.7. Tuning de modelos (Indica se o modelo pode ser ajustado para tarefas específicas);
- 2.5.3.8. *Modality* (Refere-se aos tipos de dados que o modelo pode processar);

2.5.3.9. Controlled Generation (indica se a geração de texto pode ser controlada através de esquemas ou outras técnicas);

2.5.3.10. Managed Search Grounding (capacidade de o modelo buscar e utilizar informações externas);

2.5.3.11. Provisioned Throughput (capacidade de provisionar recursos para o modelo);

2.5.3.12. Context Caching (indica se o modelo armazena em cache informações de contextos anteriores);

2.5.3.13. Batch API (permite processar múltiplas solicitações em um único pedido);

2.5.3.14. Machine Learning Integrado;

2.5.3.15. Geoprocessamento nativo;

2.5.3.16. Streaming de dados em tempo real;

2.5.3.17. Integração com outros serviços, por exemplo, o Google Maps;

2.5.3.18. Compactação automática de dados;

2.5.3.19. AutoML com foco em visão computacional.

2.6. Público-alvo a ser atendido:

2.6.1. O público-alvo a ser atendido com a presente contratação compreende:

2.6.1.1. Superintendência de Tecnologia da Informação - STI da Secretaria de Economia do Estado de Goiás – ECONOMIA, suas áreas e seus respectivos times, especialmente aqueles envolvidos na concepção, desenvolvimento, sustentação e gerenciamento do ciclo de vida de soluções de software, mas não se restringindo a estes;

2.6.1.2. Áreas de negócio clientes da ECONOMIA que encaminham demandas às diversas que necessitam de soluções tecnológicas;

2.6.1.3. Usuários gestores e donos dos produtos/soluções de software desenvolvidos e sustentados pela STI/ECONOMIA, os quais atuarão nos processos de priorização de demandas, definições de requisitos, homologação e testes das soluções;

2.6.1.4. Cidadãos e contribuintes que serão beneficiados pela serviços aperfeiçoados através da aplicação das técnicas e tecnologias de inteligência artificial.

Definição da natureza de execução do objeto:

2.7. A execução do objeto contratado pode ser considerado de **natureza continuada**, nos termos do inciso XV do art. 6º da Lei federal nº 14.133, de 01 de abril de 2021, já que são serviços de fornecimentos contínuos aqueles contratados pela Administração Pública para a manutenção da atividade administrativa, decorrentes de necessidades permanentes ou prolongadas.

Regime de fornecimento:

2.8. Para os itens 1 e 2, tendo em vista a necessidade de fornecimento dos bens ou serviços contratados, a entrega será prestada de forma parcelada e sob demanda.

2.8.1. Para o item 1 a entrega será em até 15 (quinze) dias, contados a partir da emissão da Ordem de Fornecimento pela unidade requisitante.

2.8.2. Para o item 1 os créditos de serviços de computação de IA poderão ser utilizados para consumo de todos os componentes da plataforma Google Cloud Platform (GCP) relacionados e necessários para a computação de Inteligência Artificial, assim como para aquisição daqueles comercializados por meio de subscrição, cujos valores devem ser disponibilizados pela fabricante Google a partir de tabela de preços na moeda real.

2.8.3. Para os serviços por subscrição, o valor unitário do crédito de serviços de computação de IA contratado deve ser aplicado ao preço cobrado pela Google em real para efeitos de cobrança.

2.8.4. O pagamento será realizado mensalmente, correspondendo aos créditos utilizados no respectivo mês.

2.8.5. Os serviços deverão ser fornecidos com um índice mensal de disponibilidade conforme a tabela oficial da GCP, disponível em <https://cloud.google.com/terms/sla>. O não cumprimento desses acordos acarretará nas glosas previstas na página web, que são revertidas em créditos para o mês seguinte, sem prejuízo às eventuais penalidades permitidas em Lei.

2.8.6. Para o item 1 os créditos de serviços de computação de IA podem ser utilizados para contratação de serviços na modalidade de Subscription Agreement.

2.8.7. É necessário acesso à solução de "Billing" que contenha as informações necessárias para gerar o relatório de acompanhamento do faturamento e do consumo mensal na moeda real. São informações necessárias os campos list cost, negotiated savings, committed use discounts, spending-based discounts, cost, total savings, subtotal, total, entre outros que venham a compor o cálculo do valor total consumido sob demanda, apurado mensalmente.

2.8.8. Fica definido de forma obrigatória que o fornecimento dos créditos de serviços de computação de IA (item 1) na plataforma GCP, deverão estar vinculados à solução que contemple componentes da inteligência artificial, identificados nos relatórios de prestação de serviços consumidos, sendo de exclusiva responsabilidade da CONTRATADA, que tais créditos de serviço somente sejam consumidos para a finalidade que envolva inteligência artificial. Qualquer uso indevido de serviços dentro da plataforma GCP, que não envolva componentes e/ou soluções aplicadas ao objeto, serão objeto de glosa pelos gestores e fiscais do contrato oriundo deste objeto.

2.8.9. Para o item 2 a entrega será prestada a partir da emissão da Ordem de Serviço. Os serviços de orientação técnica especializada poderão ser realizados remotamente, exceto nos casos em que a contratante determinar que sejam realizados de forma presencial. Os serviços, caso sejam prestados de forma presencial, serão executados no endereço da Superintendência de Tecnologia da Informação (STI) da Secretaria de Estado da Economia de Goiás, preferencialmente em dias úteis e no horário comercial.

2.8.10. Devido às características dos itens a serem contratados, não haverá transporte para nenhum dos itens.

Justificativa da escolha da solução:

2.9. A análise das opções oferecidas pelo mercado, relatada na seção 7 deste ERP, demonstra que a solução escolhida é a que melhor atende à finalidade pública, especialmente pelos seguintes fatos e fundamentos descritos neste documento.

2.10. Foram analisados os critérios de CLOUD (ambiente corporativo e de governança em nuvem), precificação, prompt, grandes modelos de linguagem (LLM), IA/Machine Learning, georeferenciamento e indicadores e posicionamento de mercado.

2.11. Diante dos resultados do trabalho de análise comparativa das soluções demonstrados nas tabelas da seção 7, frente aos principais requisitos e características considerados determinantes para a escolha da ferramenta, conclui-se que a solução GCP (Google Cloud Platform) é a que melhor atende à Superintendência de Tecnologia da Informação - STI da Secretaria de Economia do estado de Goiás, pois atenderá plenamente às necessidades do referido órgão.

2.12. Abaixo, tem-se resumo dos motivos que sustentam a escolha da plataforma GCP para construção de soluções envolvendo IA, dentro da Secretaria de Economia, totalmente alinhada com as diretrizes existentes.

- a) Vantagens em Nuvem, IA e Machine Learning: O GCP oferece uma solução completa e integrada, com alta pontuação em aspectos como multimodalidade, segurança (incluindo adversarial training), flexibilidade no treinamento de modelos, gerenciamento de prompts, regionalização, agentes, ofertas de busca, preço e SLA. Além disso, o Google Vertex AI se destaca por sua integração profunda com o ecossistema Google Cloud, AutoML com foco em visão computacional, Feature Store integrada, experimentos colaborativos no AI Workbench e foco em ética e responsabilidade;
- b) Modelos de Linguagem Avançados: O Gemini 1.5 Pro do Google se destaca com o maior Context Window (2M tokens), capacidade de lidar com cinco modalidades de dados (texto, imagem, documento, vídeo e áudio), controle preciso da geração de texto, busca e uso de informações externas, provisionamento de recursos com SLA e caching de contexto;
- c) Georeferenciamento Completo: O Google Maps oferece a base de dados mais detalhada, recursos de IA avançados para previsões de tráfego e sugestões personalizadas, integração com outros serviços Google como Assistente, além de uma grande comunidade ativa. Diferenciais como o Plus Code, um sistema de endereçamento digital, podem contribuir para a identificação de contribuintes no estado de Goiás;
- d) Liderança no Mercado: O Google é líder em diversos segmentos relevantes, incluindo plataformas de IA, infraestrutura de IA e modelos de linguagem para IA, como demonstrado pelo Gartner® Magic Quadrant™ e The Forrester Wave;
- e) Integração com o Ecossistema Existente: A Secretaria busca integrar o Gemini com suas soluções existentes para criação de documentos, OCR e busca textual, o que é plenamente suportado pelo GCP;
- f) Escalabilidade e Abrangência: O GCP atende desde startups até grandes empresas, o que garante a escalabilidade necessária para a Secretaria;
- g) Em resumo, a escolha pelo GCP se justifica pela sua abrangência, profundidade tecnológica, liderança no mercado e capacidade de integração com o ecossistema da Secretaria da Economia. As análises demonstram que o GCP oferece as ferramentas mais adequadas para atender às necessidades da Secretaria da Economia, impulsionando a transformação digital e a otimização dos serviços públicos.

2.13. A comercialização dos serviços de processamento no ambiente de nuvem da fabricante Google (Google Cloud Platform – GCP) se dá por meio de créditos, que uma vez adquiridos, são abatidos a medida que o recurso computacional é demandado/usado pela Contratante. Esta abordagem proporciona ao gestor do contrato um maior controle sobre a infraestrutura tecnológica disponibilizada, permitindo uma gestão personalizada dos recursos e soluções, além de possibilitar uma adaptação mais flexível e econômica às necessidades da STI.

2.14. Os créditos de processamento de IA não são comercializados pelo fabricante Google, sendo comercializados somente por meio de empresas revendedoras autorizadas pela Google, o que permite ampla concorrência. Por meio do link disponível na página da própria Google (<https://cloud.google.com/find-a-partner/>) é possível ter acesso à extensa lista de revendas autorizadas.

Vigência do contrato:

2.15. O prazo de vigência contratual é de 12 meses, contados imediatamente a partir da assinatura ou retirada de Termo de Contrato, nos termos do Título III, Capítulo V, da Lei federal nº 14.133, de 01 de abril de 2021.

2.15.1. Considerando que o objeto contratado é de natureza continuada, a vigência do contrato é prorrogável nos termos da Lei federal nº 14.133, de 01 de abril de 2021.

Tópico 3 - ESTIMATIVA DA QUANTIDADE A SER CONTRATADA

Identificação dos itens, quantidades e unidades:

3.1. A estimativa da quantidade a ser contratada é justificada nos termos deste ETP, conforme disposto na Lei federal nº 14.133, de 01 de abril de 2021. A descrição com o respectivo quantitativo a ser contratado está apresentado abaixo:

#	Lote	Cod	Descrição	Qtde
001	1	670	licença para uso de software, conforme termo de referência.	636000
002	1	1066	serviços técnicos especializados, consultoria técnica.	1920

Lote	Item	Descrição	Unidade de Serviço	Quantidade
1	1	Créditos de serviços de computação de IA	UCIA (*)	636.000
1	2	Orientação Técnica Especializada	HORA	1.920

(*) UCIA = unidade de crédito de serviço de computação de IA para remuneração dos componentes dentro da plataforma, que devem obrigatoriamente envolver motor para implementação de inteligência artificial, que compreendem os componentes necessários para processamento de IA e os demais componentes básicos para o processamento de tais serviços, incluindo:

- a. Modelos de LLM;
- b. Modelos de especialistas;
- c. Integrações e entrega continuada;
- d. API's complementares necessárias a execução da solução de IA;
- e. Componentes de gerenciamento, conectividade, segurança e escalabilidade;
- f. Armazenamento necessário a execução da solução de IA;
- g. Processamento necessário a execução da solução de IA;

A utilização dessa unidade denominada UCIA deverá consumir os créditos de serviços descritos acima, para execução do motor de IA junto às aplicações da área de negócio da ECONOMIA.

Justificativa de quantitativo:

3.2. Os quantitativos do objeto foram definidos com base no levantamento de demanda existente.

3.3. Memória de cálculo e/ou outros documentos que serviram de fundamentação para definição dos quantitativos:

3.3.1. Considerando as informações recebidas após consulta ao fabricante, tem-se a seguinte memória de cálculo:

3.3.2. Item 01 – Crédito de Serviço de computação de IA: tem-se uma expectativa de consumo mensal na ordem de 53.000 (cinquenta e três mil) créditos de serviço de computação de IA, quando da execução total do objeto contratado.

3.4. Para o cenário de aplicação considerado, estima-se para um período de 12 (doze) meses:

- 50.000 solicitações de serviço (levamos em consideração que no primeiro semestre de 2024 o PDP gerou 10.000 solicitações de serviço. Em 12 meses teríamos então 20.000 solicitações de serviço. Consideramos que com o uso da IA e o ganho de produtividade e a expansão da PDP para outros serviços poderíamos chegar a 40.000 solicitações de serviço. Para a avaliação de outros processos, consultas a bases de documentos e elaboração, correção e simplificação de documentos estimamos que em 12 meses totalizariam 10.000 solicitações de serviço. Dessa forma estimamos ao final 50.000 (40.000+10.000) solicitações de serviço para um período de 12 meses);
- 10 documentos em média por solicitação de serviço do PDP (dados obtidos de solicitações de serviço do PDP);
- 17 páginas em média por solicitação de serviço do PDP (dados obtidos de solicitações de serviço do PDP);
- 2,98MB em média de tamanho por solicitação de serviço do PDP (dados obtidos de solicitações de serviço do PDP);

3.5. Considerando o cenário descrito acima, tem-se os seguintes parâmetros:

ITEM	DESCRIÇÃO	QUANTIDADE	JUSTIFICATIVA
1	Estimativa Total de solicitações de serviços por dia	137	Memória de cálculo: 50.000 solicitações de serviço para um período de 12 meses 50.000/ 365 dias => 137 por dia
2	Estimativa total de requisições por dia	1.370	Memória de cálculo: 1 processo gera 10 requisições em média (1 requisição por página) 50.000 processos em 12 meses ou 365 dias 500.000/365 dias => 1.370 por dia
3	Quantidade de imagens (páginas) por processo	17	Memória de cálculo: Os dados foram retirados do painel de Billing da GCP durante a execução do MVP no mês de julho/2024 Estimativa baseada nos 2.518 processos avaliados com volume de imagens destes processos de 42.813 42.813/2.518 => 17,00 páginas por processo
4	Estimativa de input de caracteres no modelo LLM por processo	2.342	Memória de cálculo: Os dados foram retirados do painel de Billing da GCP durante a execução do MVP no mês de julho/2024 2.518 processos avaliados 5.898.351 inputs destes processos avaliados 5.898.351/2.518 => 2.342 caracteres de input
5	Estimativa de output de caracteres do modelo LLM por processo	855	Memória de cálculo: Os dados foram retirados do painel de Billing da GCP durante a execução do MVP no mês de julho/2024 2.518 processos avaliados

			2.153.230 inputs destes processos avaliados 2.153.230/2.518 => 855 caracteres de output
6	Estimativa de armazenamento de um ano de processos (GB):	145	Memória de cálculo: 1 processo = 2,98 MB 50.000 solicitações (50.000 x 2,98)/1024 => 145 GB
7	Estimativa de requisição de escrita no Object Storage	50.000	Total estimado de solicitações por ano
8	Estimativa de requisição de leitura no Object Storage	4167	500.000 de requisições em 12 meses 500.000/12 =>4167 por mês
9	Número de instâncias – processamento plataforma GCP – modalidade PaaS – Interface de Acesso	3	Volume mínimo disponibilizado pela plataforma considerando a arquitetura e demanda prevista
10	Número de instâncias – processamento plataforma GCP – modalidade PaaS – Processamento da solução	3	Volume mínimo disponibilizado pela plataforma considerando a arquitetura e demanda prevista

3.5.1. Aplicando os parâmetros acima na Calculadora da Google Cloud apura-se um volume estimado mensal de 53.000 (cinquenta e três mil) créditos unitários de nuvem. A *Google Cloud Calculator* tem a função de estabelecer a quantidade de créditos necessários para a realização dos serviços de acordo com o volume estimado. Ela apresenta apenas uma valor de referência que deve ser convertido em créditos, considerando 1 (um) crédito de serviço de computação de IA unitário da plataforma GCP como R\$ 1,00 (um real). O valor apresentado na calculadora não representa o valor final a ser comercializado uma vez que a Google não comercializa diretamente esses produtos, mas sim através de seus representantes que ainda incluirão suas taxas/descontos e demais instrumentos para o calculo do valor final. Os créditos poderão ser utilizados para todos os produtos da *Google Cloud* conforme catalogo da lista <https://cloud.google.com/products> sendo que essa lista não é exaustiva e devido a característica de inovação novos itens podem ser incluídos nesta lista.

3.5.2. Segue o link e as imagens da calculadora Google:

<https://cloud.google.com/products/calculator?e=48754805&dl=CiQzZTQ3ZDc4Yy1lZDY4LTRhYjQ1YTY1MyQwZGRmMDY1YTl4Y2MQEhokRFE0QjAxQjltQzY5NC00N0Y2LTk2QzAtOzBFOFY3QzFzRUM1&hl=en>

Generative AI

R\$1,799.44 / month

Link billing account to view negotiated pricing.

Gemini 1.5 Pro

Context window

<128k tokens

Batch requests

Requests per day

1232

Average input characters

2342

Average output characters

855

Average input images

17

Average input video (seconds)

0

Average input audio (seconds)

0

Cost details

BRL

+ Add to estimate

AI & ML

R\$8,377.79

Multimodal Models

Generative AI

R\$1,799.44

Vertex AI

Conversation

Vertex AI Agent Builder

R\$3,644.89

Vertex AI Search

Vertex AI Agent Builder

R\$2,863.03

ESTIMATED COST

R\$54,959.94 / mo

Share

Vertex AI Agent Builder

R\$3,644.89 / month

Link billing account to view negotiated pricing.

Service type

Vertex AI Conversation

Vertex AI Conversation configuration

Advanced settings

Dialogflow CX Standard Chat

Number of requests per month

0

Dialogflow CX Enterprise Chat

Number of requests per month

30087

Cost details

BRL

+ Add to estimate

AI & ML

R\$8,377.79

Multimodal Models

Generative AI

R\$1,799.44

Vertex AI

Conversation

Vertex AI Agent Builder

R\$3,644.89

Vertex AI Search

Vertex AI Agent Builder

R\$2,863.03

ESTIMATED COST

R\$54,959.94 / mo

Share

Vertex AI Agent Builder

R\$2,863.03 / month

Service type

Vertex AI Search

Vertex AI Search configuration

Advanced settings

Search Standard Edition

Number of requests per month

0

Search Enterprise Edition

Number of requests per month

30087

Search LLM Add-On

Number of requests per month

30087

Link billing account to view negotiated pricing.

Cost details

BRL

Add to estimate

AI & ML

R\$8,377.79

Multimodal Models

Generative AI

R\$1,799.44

Vertex AI Conversation

Vertex AI Agent Builder

R\$3,644.89

Vertex AI Search

Vertex AI Agent Builder

R\$2,863.03

ESTIMATED COST

R\$54,959.94 / mo

Share

Generative AI

R\$70.43 / month

Service type

Language Models

Language Models configuration

Advanced settings

Gemini 1.0 Pro

Requests per day*

989

Average input characters*

1377

Average output characters*

502

Link billing account to view negotiated pricing.

Cost details

BRL

Add to estimate

Vertex AI Search

Vertex AI Agent Builder

R\$2,863.03

Language Models

Generative AI

R\$70.43

COMPUTE

R\$9,204.73

Cloud Run

R\$9,204.73

DATA ANALYTICS

R\$35,435.95

ESTIMATED COST

R\$54,959.94 / mo

Share

Cloud Run

R\$9,204.73 / month

Cloud Run configuration

Advanced settings

Region

southamerica-east1 (São Paulo) - Tier 2

Resource Type

Service

Job

CPU amount per instance

4

Memory amount per instance

16 GiB

16

CPU allocation

CPU allocated only during request processing

CPU always allocated

Number of instances*

6

Committed use discounts

None

1 year

3 years

Link billing account to view negotiated pricing.

Cost details

BRL

Add to estimate

Vertex AI Agent Builder

R\$2,863.03

Language Models

Generative AI

R\$70.43

COMPUTE

R\$9,204.73

Cloud Run

R\$9,204.73

DATA ANALYTICS

R\$35,435.95

Estimate

ESTIMATED COST

R\$54,959.94 / mo

Share

BigQuery

R\$35,435.95 / month

Service type

Editions

Editions configuration

Advanced settings

Slots pricing

Location type

Region

Location

Sao Paulo (southamerica-east1)

Edition

Standard

Enterprise

Enterprise Plus

Maximum slots

Small (100 slots)

Baseline slots

100

Commitment

None

1 Year

3 Years

Slot commitments

100

Cost details

BRL

Add to estimate

COMPUTE

R\$9,204.73

Cloud Run

R\$9,204.73

DATA ANALYTICS

R\$35,435.95

Editions

BigQuery

R\$35,435.95

NETWORKING

R\$1,056.39

Cloud Load Balancing

R\$195.81

ESTIMATED COST

R\$54,959.94 / mo

Share

Networking

R\$195.81 / month

Service type

Cloud Load Balancing

Cloud Load Balancing configuration

Load balancer scope

Regional

Region

Sao Paulo (southamerica-east1)

Number of forwarding rules

5

Amount of inbound data

100

GiB

Amount of outbound data

100

GiB

Cost details

BRL

Add to estimate

DATA ANALYTICS

R\$35,435.95

Editions

BigQuery

R\$35,435.95

NETWORKING

R\$1,056.39

Cloud Load Balancing

Networking

R\$195.81

Cloud VPN

R\$711.38

ESTIMATED COST

R\$54,959.94 / mo

Share

Cloud VPN

R\$711.38 / month

Cloud VPN configuration

Region

Sao Paulo (southamerica-east1)

Number of tunnels

2

Cost details

BRL

Add to estimate

NETWORKING

R\$1,056.39

Cloud Load Balancing

Networking

R\$195.81

Cloud VPN

R\$711.38

Cloud CDN

R\$76.34

NAT Gateway

Networking

R\$72.87

ESTIMATED COST

R\$54,959.94 / mo

Share

Cloud CDN ⓘ

R\$76.34 / month

Cloud CDN configuration

Cache data transfer out ⓘ

Asia Pacific (including Hong Kong)

—

+

GiB ▾

China

—

+

GiB ▾

Europe

—

+

GiB ▾

North America (including Hawaii)

—

+

GiB ▾

Oceania

—

+

GiB ▾

South America

—

100

+

GiB ▾

All other destinations (including Mexico, Central America, and Middle East)

—

+

GiB ▾

Cache fill ⓘ

Within North America or Europe (including Hawaii)

—

+

GiB ▾

Within each of Asia Pacific, South America, Middle East, Africa, and Oceania (including Hong Kong)

—

100

+

GiB ▾

Inter-region cache fill

—

+

GiB ▾

Between any two regions. E.g. Europe and Oceania

Link billing account to view negotiated pricing.

Cost details

BRL ▾

+ Add to estimate

Cloud VPN

R\$711.38

⋮

Cloud CDN

R\$76.34

⋮

NAT Gateway

Networking

R\$72.87

⋮

OPERATIONS

R\$175.41

⌵

Cloud Logging

Cloud Operations

R\$175.41

⋮

ESTIMATED COST

R\$54,959.94 / mo

📄

🔗

📄

Share

Networking ⓘ

R\$72.87 / month

Service type

✔ NAT Gateway ▾

NAT Gateway configuration

Cloud NAT type* ⓘ

Public NAT

Private NAT

Number of assigned VM instances ⓘ

—

3

+

Amount of data processed ⓘ

—

100

+

GiB ▾

Number of IP addresses used ⓘ

—

1

+

Link billing account to view negotiated pricing.

Cost details

BRL ▾

+ Add to estimate

Cloud VPN

R\$711.38

⋮

Cloud CDN

R\$76.34

⋮

NAT Gateway

Networking

R\$72.87

⋮

OPERATIONS

R\$175.41

⌵

Cloud Logging

Cloud Operations

R\$175.41

⋮

ESTIMATED COST

R\$54,959.94 / mo

📄

🔗

📄

Share

Cloud Operations ⓘ

R\$175.41 / month

Service type

✔ Cloud Logging ▾

Cloud Logging configuration

Logging storage amount* ⓘ

—

100

+

GiB ▾

Logging retention duration ⓘ

—

2

+

Months ▾

Link billing account to view negotiated pricing.

Cost details

BRL ▾

+ Add to estimate

NAT Gateway

Networking

R\$72.87

⋮

OPERATIONS

R\$175.41

⌵

Cloud Logging

Cloud Operations

R\$175.41

⋮

Cloud Monitoring

Cloud Operations

R\$0.00

⋮

SECURITY

R\$591.19

⌵

ESTIMATED COST

R\$54,959.94 / mo

📄

🔗

📄

Share

Cloud Armor ⓘ

R\$591.19 / month

Cloud Armor configuration

Service tier*

Standard

ⓘ

Number of requests (globally-scoped) ⓘ

—

+

Million

▼

Number of requests (regionally-scoped) ⓘ

—

10

+

Million

▼

Number of security policies ⓘ

—

15

+

Number of rules ⓘ

—

10

+

Link billing account to view negotiated pricing.

Cost details

BRL ▼

+ Add to estimate

Unsubscribe / Update estimate

SECURITY

R\$591.19 ⓘ

Cloud Armor

R\$591.19 ⓘ

SERVERLESS

R\$50.71 ⓘ

Cloud Functions

R\$50.71 ⓘ

STORAGE

R\$67.76 ⓘ

ESTIMATED COST

R\$54,959.94 / mo

Share

Seção 3 - ESTIMATIVA DA QUANTIDADE A SER CONTRATADA

Cloud Run functions ⓘ

R\$50.71 / month

Cloud Run functions configuration

Advanced settings

Cloud Run functions version*

Cloud Run functions

ⓘ

Region*

southamerica-east1 (São Paulo) - Tier 2

ⓘ

Requests per month (millions)* ⓘ

—

0.04016

+

Average execution time per request (ms)* ⓘ

—

30000

+

Link billing account to view negotiated pricing.

Cost details

BRL ▼

+ Add to estimate

Cloud Armor

R\$591.19 ⓘ

SERVERLESS

R\$50.71 ⓘ

Cloud Functions

R\$50.71 ⓘ

STORAGE

R\$67.76 ⓘ

Cloud Storage

R\$67.76 ⓘ

ESTIMATED COST

R\$54,959.94 / mo

Share

Cloud Storage ⓘ

R\$67.76 / month

Cloud Storage configuration

Advanced settings

Location type*

Region

ⓘ

Location*

Sao Paulo (southamerica-east1)

ⓘ

Storage class*

Standard Storage

ⓘ

Total amount of storage* ⓘ

—

145

+

GIB

▼

Class A operations per month (million) ⓘ

—

1

+

Class B operations per month (million) ⓘ

—

1

+

Data Transfer within Google Cloud ⓘ

—

+

GIB

▼

Source region

North America

▼

Destination region

North America

▼

Link billing account to view negotiated pricing.

Cost details

BRL ▼

+ Add to estimate

Cloud Armor

R\$591.19 ⓘ

SERVERLESS

R\$50.71 ⓘ

Cloud Functions

R\$50.71 ⓘ

STORAGE

R\$67.76 ⓘ

Cloud Storage

R\$67.76 ⓘ

ESTIMATED COST

R\$54,959.94 / mo

Share

3.5.3. Item 02 – Orientação Técnica Especializada: tem-se uma expectativa de consumo de 1.920 (um mil e novecentos e vinte) horas estimadas para a realização de atividades de instalação, provisionamento, configuração, desenvolvimento, integração, customização, manutenção de componentes e treinamento, usando profissionais especializados nos produtos / tecnologias referenciadas no link público do Google Cloud. A estimativa realizada levou em consideração os seguintes tipos de atividade e quantidade de horas:

- Identificação de Oportunidades com IA: 100 horas.
- Desenvolvimento de MVPs do uso de modelos de IA: 100 horas.
- Coleta e Preparação de dados para treinamento de modelos específicos: 220 horas.
- Seleção da tecnologia e ferramentas de IA: 140 horas.
- Desenvolvimento de modelos de IA personalizados: 460 horas.
- Implementação das soluções de IA: 240 horas.

- Integração da IA com aplicações legadas: 220 horas.
- Otimização e Fine-Tuning dos modelos de IA: 200 horas .
- Treinamento e capacitação da equipe em IA: 60 horas.
- Monitoramento, gestão de riscos e conformidade 180 horas.

Tópico 4 - ESTIMATIVAS DO VALOR DA CONTRATAÇÃO

4.1. Os valores referenciais estimados da contratação, unitários e totais, aferidos conforme ampla pesquisa de mercado, são os seguintes:

1	
Descrição do item 001	
Código 670 - Licença para Uso de Software, conforme Termo de Referência.	
Período (Meses)	1
Quantidade	636000
Unidade	unidade
Participação	Ampla Participação
Local de Entrega	superintendência de tecnologia da informação
Diferença Mínima	
Valor Unitário	R\$ 1,21
Valor Total	R\$ 769.560,00

1	
Descrição do item 002	
Código 1066 - Serviços Técnicos Especializados, consultoria técnica.	
Período (Meses)	1
Quantidade	1920
Unidade	servico (s)
Participação	Ampla Participação
Local de Entrega	superintendência de tecnologia da informação
Diferença Mínima	
Valor Unitário	R\$ 357,33
Valor Total	R\$ 686.073,60

4.2. O preço total estimado da contratação é **R\$ 1.455.633,60 (R\$ Um Milhão e Quatrocentos e Cinquenta e Cinco Mil e Seiscentos e Trinta e Três Reais e Sessenta Centavos)**, conforme pesquisa de preços realizada em conformidade com o Decreto estadual nº 9.900, de 07 de julho de 2021.

4.3. O orçamento estimado da presente contratação foi elaborado com base nos parâmetros e calculado em conformidade com o Decreto estadual nº 9.900, de 07 de julho de 2021, cujo documento de Orçamento Estimado, que contém memória de cálculo, será anexado aos autos da contratação, indicando os parâmetros, a metodologia e os preços referenciais utilizados no cálculo estimativo.

Tópico 5 - JUSTIFICATIVA PARA O PARCELAMENTO OU NÃO DA SOLUÇÃO

5.1. Para a contratação pretendida foram consideradas as características técnicas e peculiares de comercialização no mercado, avaliando-se o objeto em conformidade com o Princípio do Parcelamento, nos termos do Art. 40, §§ 2º e 3º da Lei federal nº 14.133, de 01 de abril de 2021.

5.2. A presente contratação será realizada com a adjudicação do objeto **por Lote**.

5.3. A seguir são apresentadas evidências e informações que subsidiaram a decisão de reunião de itens em lote, nos termos do item 5.2:

5.4. Para a contratação pretendida foram consideradas as características técnicas e peculiares de comercialização no mercado, avaliando-se o objeto em conformidade com o Princípio do Parcelamento, nos termos do Art. 40, §§ 2º e 3º da Lei Federal nº 14.133 de abril de 2021.

5.5. A presente contratação será realizada com a adjudicação do objeto por Lote Único.

5.6. A justificativa para agrupar os itens em um único lote, sem perda de competitividade, baseia-se nos seguintes aspectos:

5.6.1. Facilidade de Gestão e Implementação: Agrupar os itens em um único lote simplifica o processo de gestão e implementação, pois reduz a complexidade administrativa de lidar com múltiplos contratos e fornecedores. Isso resulta em uma execução mais eficiente do projeto, com menos burocracia e potencial para uma implementação mais rápida.

5.6.2. Coerência e Coesão na Solução Oferecida: Ao agrupar os itens em um único lote, é possível garantir que todos os componentes necessários para a implementação da solução estejam alinhados e coesos. Isso evita possíveis lacunas ou inconsistências na entrega, proporcionando uma solução mais completa e integrada.

5.6.3. Economia de Escala: A consolidação dos itens em um único lote pode gerar economias de escala significativas, tanto em termos de custos diretos quanto de eficiência operacional. Os fornecedores podem oferecer preços mais competitivos quando contratados para fornecer um conjunto completo de produtos e serviços, em vez de itens individuais.

5.6.4. Redução de Custos Administrativos: Ao optar por um único lote, há uma redução nos custos administrativos associados à condução do processo de licitação, como preparação de documentos, avaliação de propostas e gerenciamento de contratos. Isso otimiza o uso dos recursos disponíveis e permite uma alocação mais eficiente dos recursos da organização.

5.6.5. Estímulo à Concorrência: A abertura da licitação para múltiplos fornecedores que possam fornecer todos os itens do lote único ainda promove a competição saudável entre os participantes. Os fornecedores são incentivados a oferecer propostas competitivas e inovadoras, buscando destacar-se em termos de qualidade, preço e valor agregado.

5.7. Em suma, agrupar os itens em um único lote permite uma abordagem mais coesa, eficiente e econômica para a contratação dos produtos e serviços necessários. Isso garante uma solução integrada, facilita a gestão do projeto e promove a competição entre os fornecedores, sem comprometer a qualidade ou a competitividade do processo de licitação.

5.8. Por último, salientamos que o caráter competitivo da licitação não será afetado, já que as empresas que comercializam produtos das licenças especificadas, por via de regra possuem profissionais aptos a realizarem a prestação dos serviços

5.9. Registre-se que existem diversos fornecedores capazes de executar o objeto com a entrega dos itens do lote deste certame.

Tópico 6 - REQUISITOS DA CONTRATAÇÃO

6.1. Os requisitos necessários à contratação, com vistas ao atendimento da demanda, são os seguintes:

6.2. **Requisitos mínimos de qualidade:**

6.2.1. A presente contratação deverá atender, incluindo os requisitos mínimos do Termo de Referência, a proposta mais vantajosa mediante competição, zelando-se sempre pela contratação da melhor qualidade possível com o menor preço. A descrição dos requisitos no Termo de Referência deve se limitar àqueles requisitos indispensáveis ao atendimento da necessidade, garantindo-se a competitividade da contratação e a maior eficiência possível.

6.3. **Requisitos tecnológicos:**

6.3.1. O objeto a ser contratado deverá ser compatível com os seguintes requisitos tecnológicos:

6.3.1.1. Possuir recursos de processamento e armazenamento adequados à demanda e perfil da solução/software de Inteligência Artificial desenvolvida internamente, inclusive ter capacidade de absorver projetos futuros a médio prazo;

6.3.1.2. Permitir a utilização de recursos e serviços tecnológicos em nuvem pública;

6.3.1.3. Ter suporte técnico continuado com assistência técnica especializada para resolução de problemas, dúvidas e suporte;

6.3.1.4. Oferecer desenvolvimento e manutenção de soluções para criação, implementação e atualização de soluções personalizadas na nuvem para atender às necessidades específicas da STI;

- 6.3.1.5. Contemplar transferência de conhecimento sob demanda para capacitação e treinamento da equipe da STI para utilização eficiente da plataforma em nuvem e das soluções desenvolvidas;
- 6.3.1.6. Para as especificações da solução tecnológica serão observadas as Orientações para Elaboração/Ajuste de Especificações Técnicas de Ativos de TI – Versão 3, do Ministério do Planejamento, Orçamento e Gestão, disponível através do seguinte link: https://www.gov.br/economia/pt-br/assuntos/noticias/planejamento/lancada-nova-versao-do-guia-de-boas-praticas-em-contratacao-de-solucoes-de-ti/guia_de_boas_praticas_v08.pdf/view. Esse documento dispõe sobre boas práticas, orientações e vedações na elaboração de especificações técnicas de bens e serviços de Tecnologia da Informação (TI).
- 6.4. Requisitos Legais:
- 6.4.1. Em relação ao procedimento de compra deverão ser observadas as seguintes leis/decretos:
- 6.4.2. Lei nº 14.133/2021 – Lei de Licitações e Contratos Administrativos;
- 6.4.3. Lei Estadual (Goiás) nº 17.928/2012 – Dispõe sobre normas suplementares de licitações e contratos pertinentes a obras, compras e serviços, bem como convênios, outros ajustes e demais atos administrativos negociais no âmbito do Estado de Goiás.
- 6.4.4. Decreto Estadual (Goiás) nº 9.900/2021 – Dispõe sobre o procedimento administrativo para a realização de pesquisa de preços para aquisição de bens e contratação de serviços em geral, no âmbito da administração pública estadual direta, autárquica e fundacional;
- 6.4.5. Em relação ao objeto, não existem leis que regulamentam o seu uso e manutenção. A solução deverá atender às normas relativas à comercialização de soluções de Tecnologia da Informação e Comunicação – TIC, aos direitos de Propriedade e Autorais, fornecimento de serviços e outras pertinentes, além do Código de Defesa do Consumidor e Código Civil
- 6.5. Requisitos Temporais:
- 6.5.1. O contrato ou instrumento equivalente a ser assinado deverá ter vigência de 12 (doze) meses;
- 6.5.2. O recebimento provisório será emitido no ato da entrega dos produtos e serviços;
- 6.5.3. O recebimento definitivo será emitido pela CONTRATANTE em até 10 dias úteis após o recebimento provisório;
- 6.6. Requisitos de Manutenção e Garantia:
- 6.6.1. A CONTRATADA deve disponibilizar mecanismos para que os técnicos do órgão possam solicitar atendimento e suporte diretamente ao fabricante da solução tecnológica;
- 6.6.2. Deve ser possibilitado o acesso à base de conhecimento do sítio do fabricante e fóruns de discussão;
- 6.6.3. Demais requisitos de garantia serão especificados de acordo com as Orientações para Elaboração/Ajuste de Especificações Técnicas de Ativos de TI – Versão 3, do Ministério do Planejamento, Orçamento e Gestão, disponível através do link: https://www.gov.br/economia/pt-br/assuntos/noticias/planejamento/lancada-nova-versao-do-guia-de-boas-praticas-em-contratacao-de-solucoes-de-ti/guia_de_boas_praticas_v08.pdf/view. Esse documento dispõe sobre boas práticas, orientações e vedações na elaboração de especificações técnicas de bens e serviços de Tecnologia da Informação (TI).
- 6.7. Requisitos de Segurança:
- 6.7.1. A CONTRATADA deverá assegurar na época da implantação da solução, bem como, durante todo o seu ciclo de vida de funcionamento, mecanismos robustos de segurança, aderindo rigorosamente às normativas de segurança estabelecidas pela CONTRATANTE, a fim de proteger a integridade e confidencialidade das informações.
- 6.7.2. Para isso, na instalação e uso dos equipamentos deverá ser observada a seguinte norma de segurança;
- 6.7.3. Durante o contrato, a empresa deverá manter sigilo, sob pena de responsabilidade civil, penal e administrativa, sobre todo e qualquer assunto de interesse da contratante ou de terceiros de que tomar conhecimento em razão da execução do objeto, devendo orientar nesse sentido seus empregados e/ou terceiros sob sua responsabilidade, permanecendo estas obrigações de confidencialidade após o término do contrato. A necessidade de sigilo se aplica no caso das manutenções realizadas no parque computacional onde as licenças serão instaladas, se necessário;
- 6.7.4. Os dados e documentos submetidos ao ambiente de nuvem deverão ficar sob domínio próprio e seguro, não devendo ser acessados por terceiros e nem pela contratada;
- 6.8. Requisitos de Níveis de Serviços:
- 6.8.1. A solução deve incluir uma garantia abrangente que assegure assistência técnica contínua, disponível 24 horas por dia, 7 dias por semana;
- 6.9. Requisitos de Qualificação e Experiência da empresa contratada:
- 6.9.1. Deverá ser comprovada capacitação técnico-operacional através de um ou mais atestados, expedidos por pessoa jurídica de direito público ou privado, mencionando que a empresa forneceu, de forma satisfatória, objeto com características semelhantes a deste projeto;
- 6.10. Requisitos de forma de comunicação:
- 6.10.1. A forma de comunicação com a empresa contratada se dará por:
- Correio eletrônico (e-mail);
 - Telefone;
 - Sistema de envio e entrega de correspondências (empresa dos Correios);
 - Aplicativos de mensagens;
- Tópico 7 - LEVANTAMENTO DE MERCADO
- Identificação de soluções:
- 7.1. Por meio dos estudos realizados, foram analisadas diferentes soluções, em que foi avaliada sua capacidade de solucionar o problema descrito no Tópico 1 deste ETP, e ainda a relação custo-benefício entre as soluções.
- 7.2. Assim, foram identificadas as seguintes possíveis soluções:
- a. Solução 1: Google Cloud Platform;

b. Solução 2: AWS;

c. Solução 3: Azure;

d. Solução 4: Open AI: solução aplicada especificamente para LLM;
- 7.3. Análise comparativa das soluções:
- 7.3.1. Para escolher o melhor tipo de solução a contratar, realizou-se uma análise comparativa entre as soluções disponíveis no mercado, levando em consideração os aspectos técnicos e econômicos, mensurados a partir dos critérios elencados no art. 15 do Decreto Estadual nº 10.207 de janeiro de 2023.
- 7.3.2. A presente análise comparativa foi baseada em documentações e demais informações constantes dos sites oficiais das ferramentas, bem como testes e provas de conceito realizadas internamente pelas equipes técnicas da STI/ECONOMIA.
- 7.3.3. A seguir é apresentado quadro comparativo, com a análise dos principais requisitos, necessidades e características considerados chave para a decisão dentre as soluções identificadas:
- a) Análise sobre aspecto de CLOUD (este aspecto visa avaliar ambiente corporativo e de governança em nuvem) :

Recursos Tecnológicos	Google Cloud	AWS	Microsoft Azure	OpenAI
Modelos	Gemini, modelos de embedding, forte em multimodalidade	Modelos variados, incluindo o Titan Image Model	Oferta abrangente, com foco em personalização	Open AI trabalha especificamente solução de LLM, por isso não considerando para o aspecto de CLOUD
Afinamento e Avaliação	Full-fine tuning, pre-training, OpenAI também oferece full-fine tuning	Opções de personalização, integração com Weights & Biases	Ferramentas de personalização e avaliação	
Segurança	Fortes medidas de segurança, filtros de conteúdo, adversarial training	Medidas de segurança sólidas	Foco em segurança e privacidade	
Gerenciamento de Prompts	Ferramentas avançadas de gerenciamento e otimização	Ferramentas de gerenciamento de prompts	Ferramentas de gerenciamento de prompts	
Regionalização	Disponibilidade global, DRZ (Dual Region Zones)	Disponibilidade global, mas sem garantia de DRZ em todas as regiões	Disponibilidade global, com foco em regiões específicas	
Agentes e Uso de	Ferramentas para criação e	Bedrock Agents para	Ferramentas para criação e	

Ferramentas	gerenciamento de agentes	customização e orquestração	gerenciamento de agentes
Ofertas de Busca	Integração com o Google Search, acesso a diversos conectores de dados	Integração com o Amazon Search, acesso a diversos conectores de dados	Integração com o Bing Search, acesso a diversos conectores de dados
Preço	Modelos de precificação flexíveis, opções de pagamento por uso e pré-pagamento	Modelos de precificação variados, com opções de pagamento por uso e reserva de instâncias	Modelos de precificação competitivos, com opções de pagamento por uso e reserva de instâncias
SLA	SLA abrangente, cobrindo diversos aspectos do serviço	SLA focado em endpoints de modelos	SLA com foco em disponibilidade e latência

b) Sobre o aspecto de precificação, depara-se com 3 (três) processos que contrataram serviços de nuvem, de fabricantes distintos, que podem basilar este ponto de análise, que são:

7.3.4. Solução GOOGLE: Estudo preços GOOGLE CLOUD PLATAFORM – baseado no processo nº 2024/NA/0079 – SEBRAE

Descrição	Valor Unitário	Pregão	Fornecedor
Item 1 – Créditos para serviços da nuvem pública GCP	R\$ 0,78	02/2024	Xértica Brasil Ltda

7.3.5. Solução MULTINUVEM – possibilidade de usar quaisquer um dos seguintes provedores: GOOGLE, MICROSOFT ou ORACLE.), sob o modelo de cloud broker (integrador) de multi-nuvem: Estudo preço MULTINUVEM (GOOGLE CLOUD PLATAFORM, MICROSOFT AZURE ou ORACLE CLOUD) – baseado no Processo Administrativo nº 23172.003606/2023-25 – Instituto Federal de Educação, Ciência e Tecnologia Piauí – Item 3

Descrição	Valor Unitário	Pregão	Fornecedor
Item 3 – Computação em nuvem – Software como Serviço (Saas)	R\$ 1,40	90008/2024	Ready Tecnologia da Informação Ltda

7.3.6. Solução MICROSOFT: Estudo preços MICROSOFT AZURE – baseado no Pregão Eletrônico nº 90001/2024 – Empresa Brasileira de Administração de Petróleo e Gás – Lote 1

Descrição	Valor Unitário	Pregão	Fornecedor
Lote 1 – Unidade de serviço em nuvem	R\$ 1,00	90001/2024	Noventiq International Brasil Comercio e Licenciamento de Software Ltda

7.3.7. Em resumo, para o contexto de CLOUD, depara-se com 3 fabricantes aptos, porém resta evidenciado os diferenciais GOOGLE no tocante multimodalidade, segurança (adversarial training) e oferece mais flexibilidade no treinamento de modelos, gerenciamento de prompts, regionalização, agentes e uso de ferramentas, ofertas de busca, preço e SLA.

c) Análise sobre aspecto de PROMPT:

Característica	Google	AWS	Microsoft (Azure)	OpenAI (GPT)
Integração com Ecossistema	Integração profunda com serviços como Search, Docs, Translate (plataformas abertas)	Integração com serviços como SageMaker, Lex, Polly	Integração com Azure, Office 365	Integração com outros produtos da OpenAI
Investimento em Pesquisa e Desenvolvimento	Alto investimento, foco em ética e responsabilidade	Alto investimento em IA/ML, desenvolvimento de chips de IA	Alto investimento, integração com outras ferramentas da Microsoft	Alto investimento, foco em modelos de linguagem de grande porte
Ética e Responsabilidade	Forte compromisso, mitigação de vieses	Diretrizes éticas, ferramentas para IA responsável	Compromisso com a ética, mas com áreas a serem aprimoradas	Foco em segurança e ética, mas com desafios
Compreensão Contextual	Excelente, capaz de entender nuances e manter o contexto	Boa compreensão contextual, com melhorias contínuas	Boa, com melhorias contínuas	Muito boa, especialmente em textos longos
Diversidade de Aplicações	Ampla gama de ferramentas e APIs	Variedade de serviços de IA/ML para diferentes aplicações	Integração profunda com o Azure, diversas aplicações	Foco em geração de texto criativo e conversacional

7.3.8. Em resumo, o Google se destaca em integração com o ecossistema, compreensão contextual, ética e responsabilidade e diversidade de aplicações. OpenAI se destacou somente em compreensão contextual, enquanto AWS e Microsoft apresentam desempenho moderado em diversas categorias.

d) Análise sobre aspecto de plataforma de IA/MACHINE LEARNING:

Funcionalidades	Google Vertex AI	AWS SageMaker	Azure Machine Learning	OpenAI (GPT)
Integração Profunda	Sim	Não tão profunda	Não tão profunda	Open AI trabalha especificamente solução de LLM, por isso não considerando para o aspecto de IA/MACHINE LEARNING
AutoML com Foco em Visão Computacional	Sim (AutoML Vision)	Sim, mas com menos opções	Sim, mas com menos opções	
Feature Store	Sim	Não possui um recurso equivalente	Não possui um recurso equivalente	
Experimentos Colaborativos AI Workbench	Sim	Sim, mas com menos integração com outros serviços	Sim, mas com menos integração com outros serviços	
Pontos Potenciais	Integração profunda com o Google Cloud, AutoML, modelos de linguagem avançados (PaLM), foco em ética e responsabilidade	Complexidade de configuração, custos variáveis, dependência da AWS	Dependência do Azure, curva de aprendizado, integração com outras ferramentas pode ser complexa	

7.3.9. Em resumo, o Google Vertex AI se destaca como a melhor opção por sua integração profunda com o Google Cloud, AutoML com foco em visão computacional, Feature Store integrada, experimentos colaborativos no AI Workbench e foco em ética e responsabilidade. Comparado com AWS SageMaker e Azure Machine Learning, o Vertex AI oferece uma solução mais completa e integrada, simplificando o desenvolvimento e a implantação de modelos de IA. Enquanto o SageMaker apresenta complexidade de configuração e custos variáveis, e o Azure Machine Learning tem dependência do ecossistema Azure e curva de aprendizado mais acentuada, o Vertex AI oferece uma experiência mais fluida e integrada com o Google Cloud.

e) Análise sobre aspecto de MODELO DE FUNDAÇÃO DE LINGUAGEM PARA IA (LLM)

Categoria	Gemini 1.5 Pro	AWS TITAN	Azure PHI	OpenAI (GPT)
Contexto Window	2M (1-2M private preview)	128K	128K	128K
Tuning: Indica se o modelo				

Finetuning: Indica se o modelo pode ser ajustado para tarefas específicas	Sim	Sim	Sim	Sim
Modality: Refere-se aos tipos de dados que o modelo pode processar	Texto, Imagem, Documento, Vídeo, Audio	Texto, Imagem	Texto, Imagem	Texto, Imagem
Controlled Generation: Indica se a geração de texto pode ser controlada através de esquemas ou outras técnicas	Sim (+ Schema Constraining)	Sim (+ Schema Constraining)	Sim (+ Schema Constraining)	Sim (+ Schema Constraining)
Managed Search Grounding: Relacionado à capacidade de o modelo buscar e utilizar informações externas	Sim	Sim	Não	Não
Provisioned Throughput: Refere-se à capacidade de provisionar recursos para o modelo	Sim com SLA	Sim com SLA	Sim	Sim
Context Caching: Indica se o modelo armazena em cache informações de contextos anteriores	Sim	Não	Não	Não
Batch API: Permite processar múltiplas solicitações em um único pedido	Sim	Não	Sim	Sim

7.3.10. Em resumo, Google se destaca na tabela por apresentar o Gemini 1.5 Pro com o maior Context Window (2M tokens), capacidade de lidar com cinco modalidades de dados (texto, imagem, documento, vídeo e áudio), controle preciso da geração de texto, busca e uso de informações externas, provisionamento de recursos com SLA e caching de contexto. Comparado aos outros modelos (AWS Titan, Azure PHI e OpenAI GPT), o Gemini oferece maior flexibilidade, capacidade multimodal e controle, além de garantir melhor performance e disponibilidade com recursos provisionados e SLA. Embora o AWS Titan se aproxime em algumas características, o Gemini ainda se sobressai na variedade de modalidades e recursos de gerenciamento.

f) Análise sobre aspecto georeferenciamento:

Característica	Google Maps	AWS	Microsoft (Azure)	OpenAI (GPT)
Base de Dados	Mais completa e detalhada	Boa cobertura e precisão, com dados de mapas, pontos de interesse e dados demográficos.	Boa cobertura e precisão, com dados de mapas, pontos de interesse e dados demográficos.	Open AI trabalha especificamente solução de LLM, por isso não considerando para o aspecto de Georeferenciamento
Inteligência Artificial	Avançada, com previsões de tráfego precisas e sugestões personalizadas	Recursos de IA e aprendizado de máquina para georeferenciamento, como Amazon Rekognition e Amazon SageMaker.	Recursos de IA e aprendizado de máquina para georeferenciamento, como Azure Cognitive Services e Azure Machine Learning.	
Integração com Outros Serviços	Integração com outros serviços Google (Assistente, Gmail, Fotos)	Integração com outros serviços da AWS, como Amazon S3 e Amazon EC2.	Integração com outros serviços do Azure, como Azure Blob Storage e Azure Functions.	
Comunidade	Grande comunidade ativa, contribuindo com informações e atualizações	Grande comunidade de usuários e desenvolvedores, com fóruns e recursos online.	Comunidade ativa de usuários e desenvolvedores, com fóruns e recursos online.	
Inovação	Constantes atualizações e novas funcionalidades para desenvolvedores	Investimento contínuo em inovação, com novos recursos e atualizações sendo lançados regularmente.	Investimento em inovação, com novos recursos e atualizações sendo lançados regularmente.	
Visibilidade para Empresas	Google Meu Negócio integrado, alta visibilidade	Amazon Location Service permite integrar mapas e recursos de localização em aplicações.	Azure Maps oferece às empresas a capacidade de integrar mapas e recursos de localização em seus aplicativos e sites.	
Plataforma Aberta	Google Maps Platform, amplamente utilizada por desenvolvedores	APIs e SDKs para integrar os serviços de georeferenciamento da AWS em aplicações.	APIs e SDKs para integração com aplicações de terceiros.	

7.3.11. Em resumo, Google Maps se destaca como a solução mais completa e avançada para georeferenciamento, com a base de dados mais detalhada, recursos de IA avançados para previsões de tráfego e sugestões personalizadas, integração com outros serviços Google como Assistente e Gmail, e uma grande comunidade ativa que contribui com informações e atualizações. Além disso, o Google Maps oferece constante inovação com novas funcionalidades para desenvolvedores, alta visibilidade para empresas através do Google Meu Negócio e uma plataforma aberta amplamente utilizada, o Google Maps Platform. Comparado com AWS e Azure, que também possuem boas bases de dados e recursos de IA, o Google Maps se diferencia pela integração com o ecossistema Google, comunidade ativa e constante inovação. Complementarmente, como diferenciais, tem-se um sistema de endereçamento digital, denominado, PLUS CODE, para identificar qualquer local desejado, mesmo aquele que não possuiu endereço convencional. Tem por funcionalidade converter coordenadas geográficas (longitude e latitude) em um código curto e único. Trata-se de uma solução que poderá contribuir fortemente na identificação de contribuintes espalhados dentro do Estado de Goiás. Essa plataforma possuiu integração nativa e profunda com o ecossistema Google, incluindo Maps, Search, Assistant, Alta precisão, especialmente em áreas urbanas e com dados detalhados no Google Maps, cobertura global completa, graças à extensa base de dados do Google Maps, permitindo diversas funcionalidades, como navegação, cálculo de rotas, compartilhamento de localização.

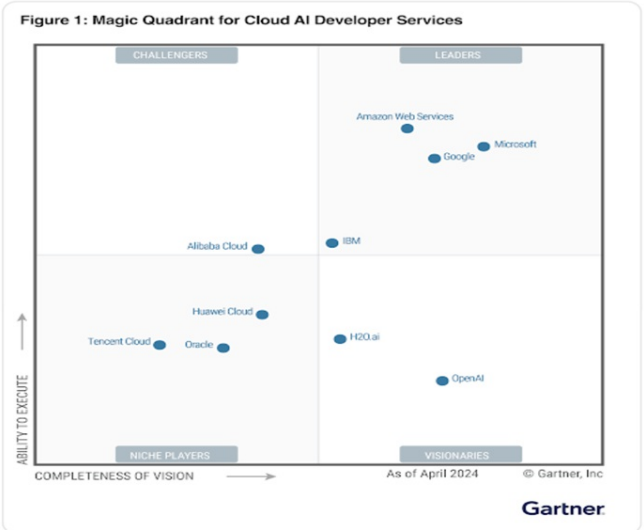
g) Análise sobre os indicadores FORESTER:

Fabricantes	Plataformas de dados de streaming	Data Lakehouses	Infraestrutura de IA	Plataforma de IA	LLM	Somatório Apurada
Amazon Web Services	3,65	2,80	4,08	3,6	2,4	16,53
Cloudera	3,37	3,27	-	2,84	-	9,48
Google	4,05	4,40	4,88	4,2	3,83	21,36
IBM	3,17	2,13	3,52	3,9	3,33	16,05
Microsoft	4,16	3,80	4,46	3,51	2,72	18,65
Oracle	3,99	3,1	2,54	-	-	9,63

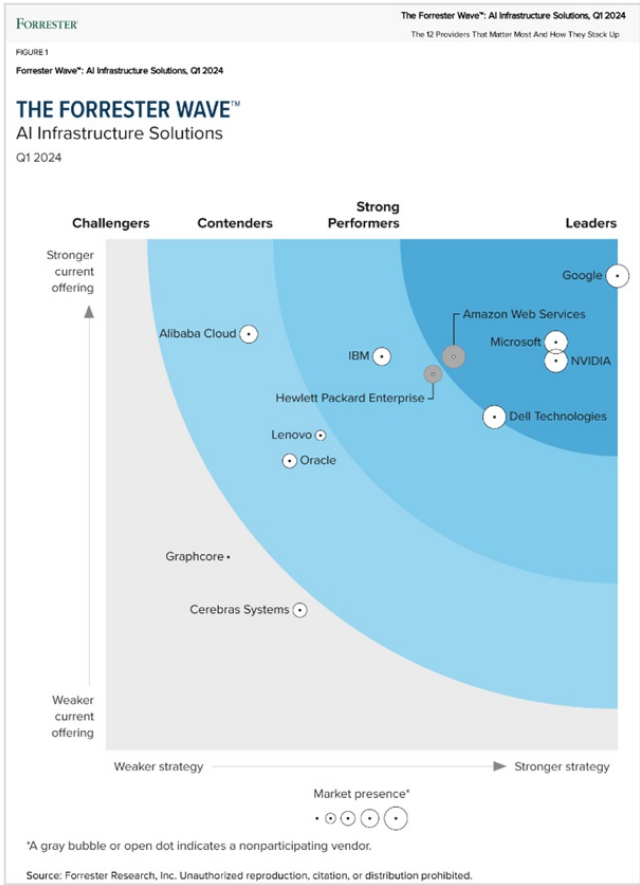
7.3.12. O quadro acima demonstra resultado da pesquisa feita pela FORESTER sobre os itens acima. Os fabricantes Cloudera, IBM e Oracle, não comercializam soluções compatíveis com o objeto, e foram mantidas acima, pois fizeram parte da pesquisa do instituto. A plataforma OpenAI não foi contemplada nessa pesquisa.

h) Análise de posicionamento de mercado

- O Google é líder pelo 5º ano consecutivo, no Gartner® Magic Quadrant™ de 2024 para Cloud AI Developer Services (CAIDS);

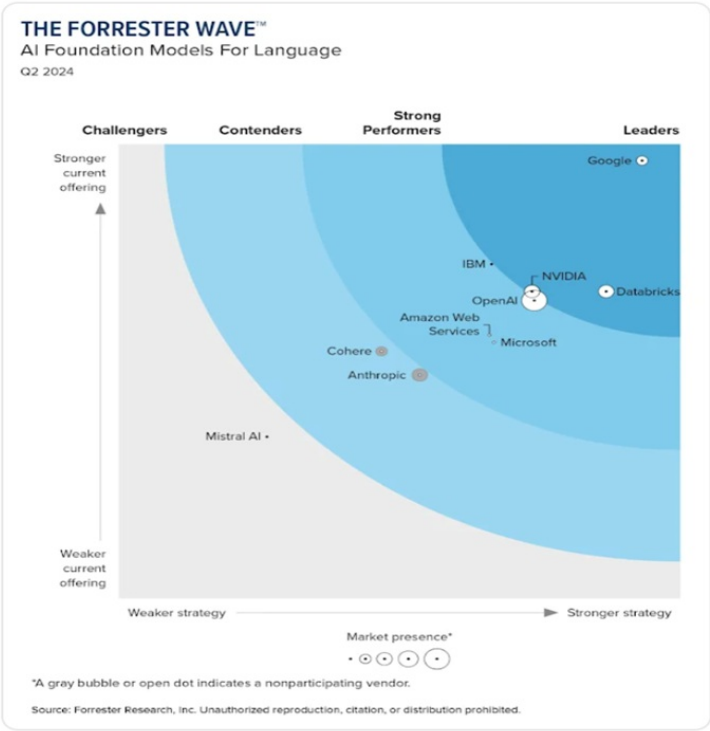


- Google nomeado líder no The Forrester Wave quanto às soluções de infraestrutura de IA, no primeiro trimestre de 2024;



<https://reprints2.forrester.com/#/assets/2/157/RES180430/report>

- Google é líder no The Forrester Wave™: Modelos de base de IA para linguagem, Q2 2024:



Tópico 8 - RESULTADOS PRETENDIDOS

8.1. Considerando que as contratações públicas devem buscar resultados positivos para a Administração, são apontados os resultados pretendidos, em termos de eficiência, eficácia, efetividade e economicidade, em busca do melhor aproveitamento dos recursos humanos, materiais e financeiros disponíveis, bem como de desenvolvimento nacional sustentável.

8.2. Assim, a presente contratação pretende alcançar o(s) seguinte(s) resultado(s):

- Aumentos dos serviços públicos digitais, disponibilizados ao contribuinte/cidadão;
- Melhoria dos processos envolvendo serviços públicos digitais;
- Agilidade da análise e definição dos pedidos feitos através da plataforma digital de processos (PCD);
- Implementação de rotinas que visam incremento de arrecadação e evasão de receitas;

Tópico 9 - POSSÍVEIS IMPACTOS AMBIENTAIS E MEDIDAS MITIGADORAS

9.1. Tendo em vista a natureza do objeto que se pretende contratar, é necessário que o Fornecedor, no âmbito de suas atividades, atenda aos critérios e políticas de sustentabilidade ambiental, sem prejuízo da observância das boas práticas e das normas pertinentes.

9.2. Considerando as particularidades da contratação, não há previsão de possíveis impactos ambientais dela decorrentes.

9.3. Lembrando que a Secretaria de Estado da Economia prioriza, em suas contratações, a promoção de práticas de aquisição responsável e o uso consciente de recursos tecnológicos, minimizando, assim, o impacto ambiental.

Tópico 10 - PROVIDÊNCIAS PRÉVIAS A SEREM ADOTADAS PELA ADMINISTRAÇÃO

10.1. A Administração Pública deverá tomar todas as providências previamente à formalização da contratação, visando à disponibilização da solução contratada em sua plenitude e ao alcance das finalidades da contratação.

10.2. No que tange a necessidade de serem tomadas providências para adequação do ambiente da instituição, frisa-se que não há necessidade de adequação da organização para que a contratação surta seus efeitos.

10.3. Ademais, pela característica do objeto aqui tratado, não há necessidade de capacitação de servidores para fiscalização e gestão contratual.

Tópico 11 - CONTRATAÇÕES CORRELATAS OU INTERDEPENDENTES

11.1. Não existem contratações correlatas e/ou interdependentes para a viabilidade da demanda

AVALIAÇÃO DA VIABILIDADE DA CONTRATAÇÃO

Em virtude de todo o exposto, o presente Estudo Técnico Preliminar evidencia que a contratação da solução: **Fornecimento de Bens e Materiais e Serviços - Contratação de solução para processamento de inteligência artificial, incluindo serviços de orientação técnica especializada sob demanda, por 12 (doze) meses.** Informada neste Estudo Técnico Preliminar, mostra-se necessária e viável tecnicamente, tendo em vista a imprescindibilidade da contratação e o adequado atendimento às demandas apresentadas. Além do mais, os custos previstos são compatíveis e atendem à economicidade; os riscos envolvidos são administráveis; e a área requisitante priorizará o fornecimento de todos os elementos aqui relacionados necessários à consecução dos benefícios pretendidos.

Assim sendo, a Equipe de Planejamento declara a viabilidade desta contratação para o atendimento da necessidade a que se destina, consoante disposto na Lei federal nº 14.133, de 01 de abril de 2021 e no Decreto estadual nº 10.207, de 27 de janeiro de 2023.

DECLARAÇÃO DE VIABILIDADE DA CONTRATAÇÃO

Diante deste Estudo Técnico Preliminar, entende-se como VIÁVEL, técnica e economicamente, a contratação dos bens demandados. Durante a identificação das soluções disponíveis no mercado para a contratação, suas viabilidades técnica e econômica foram analisadas e validadas através de pesquisas junto ao mercado.

Em complemento, os requisitos listados atendem adequadamente às demandas formuladas, os custos previstos são compatíveis e os riscos identificados são administráveis, pelo que RECOMENDAMOS o prosseguimento da contratação.

REFERÊNCIAS

As pesquisas que fundamentam as informações foram obtidas através de diversos sites, dos quais podemos citar:

- <https://reprints2.forrester.com/#/assets/2/157/RES180430/report>
- <https://reprints2.forrester.com/#/assets/2/157/RES180932/report;>
- <https://reprints2.forrester.com/#/assets/2/157/RES178464/report;>
- <https://reprint.forrester.com/reports/the-forrester-wave-ai-ml-platforms-q3-2024-e8e56c78/index.html>
- <https://reprints2.forrester.com/#/assets/2/157/RES180732/report;>
- <https://chatgpt.com/>
- <https://gemini.google.com>

- h) <https://cloud.google.com/?hl=pt-BR>
- i) <https://azure.microsoft.com/pt-br/solutions/ai/#Products>
- j) <https://aws.amazon.com/pt/ai/>
- k) https://cloud.google.com/products/ai/?e=48754805&hl=pt_br
- l) https://cloud.google.com/partners/specializations-expertise?hl=pt_br
- m) <https://cloud.google.com/docs/get-started/aws-azure-gcp-service-comparison?hl=pt-br>
- n) <https://blog.saninternet.com/aws-vs-azure-vs-google>
- o) <https://www.coursera.org/articles/aws-vs-azure-vs-google-cloud>
- p) <https://www.bmc.com/blogs/aws-vs-azure-vs-google-cloud-platforms/>
- q) <https://www.digitalocean.com/resources/articles/comparing-aws-azure-gcp>

EQUIPE DE PLANEJAMENTO RESPONSÁVEL PELA ELABORAÇÃO DESTE ETP:

Responsável	Função	Telefone	Email
ALESSANDRO CRUVINEL MACHADO DE ARAUJO	Integrante Requisitante	62 32692122	alessandro.cruvinel@goias.gov.br
ALESSANDRO CRUVINEL MACHADO DE ARAUJO	Integrante Técnico	62 32692122	alessandro.cruvinel@goias.gov.br
REINALDO PIRES GOUTHIER	Integrante Administrativo	62 32692000	reinaldo.gouthier@goias.gov.br
ALLAN DA COSTA ANDRADE	Integrante Técnico	62 32696260	allan.andrade@goias.gov.br