

Ao Ilmo. Sr. ROGÉRIO SOUZA MASCARENHAS, Titular da Secretaria de Governo Digital do Ministério da Gestão e da Inovação em Serviços Públicos – SGD/MGI

Assunto:

Representação sindical com pedido de diligência, auditoria e medidas preventivas sobre o Contrato IBGE nº 19/2025 (Processo nº 03603.000033/2025-15), especialmente quanto a sigilo estatístico, inteligência artificial, transferência internacional de dados, Data Lake, integração de bases, *multicloud*, propriedade de modelos e governança da execução.

SINDICATO NACIONAL DOS TRABALHADORES EM FUNDAÇÕES PÚBLICAS FEDERAIS DE GEOGRAFIA E ESTATÍSTICA – ASSIBGE/SN, entidade sindical regularmente constituída, dotada de personalidade jurídica de direito privado, inscrita no CNPJ sob o nº 59.954.388/0001-02, com sede na Av. Presidente Wilson, nº 210, 8º andar – CEP 20030-021 – Centro – Rio de Janeiro, e-mail: assibge-sn@uol.com.br, telefone: (21) 3575-5757, com fulcro no art. 8º, III, da Constituição¹, e no art. 9º, IV, da Lei 9.784²,

no exercício de sua função de defesa do IBGE, de seus trabalhadores, da autonomia técnica da estatística pública e da confiança da sociedade brasileira nas pesquisas oficiais, vem apresentar a presente

REPRESENTAÇÃO

para que a Secretaria de Governo Digital examine os fatos abaixo, promova as diligências cabíveis no âmbito do Sistema de Administração dos Recursos de Tecnologia da Informação – SISP e, nas matérias de governança e compartilhamento de dados, **consulte ou provoque o Comitê Central de Governança de Dados – CCGD, sem prejuízo do encaminhamento às demais instâncias competentes.**

¹ “Art. 8º É livre a associação profissional ou sindical, observado o seguinte: (...) III - ao sindicato cabe a defesa dos direitos e interesses coletivos ou individuais da categoria, inclusive em questões judiciais ou administrativas;”

² “Art. 9º São legitimados como interessados no processo administrativo: (...) IV - as pessoas ou as associações legalmente constituídas quanto a direitos ou interesses difusos.

Sumário

I.	INTRODUÇÃO	4
II.	SÍNTESE	4
III.	DA LEGITIMIDADE E DO INTERESSE SINDICAL	6
IV.	DO CONTRATO E DO CONTEXTO PROCESSUAL	6
	1. Planejamento da contratação e escopo originalmente concebido	6
	2. Instrução incompleta, urgência e controle jurídico limitado.....	7
	3. Objeto, valor, vigência e hierarquia documental.....	7
	4. Questionamento sindical e resposta institucional.....	8
	5. Execução iniciada antes da consolidação da governança contratual	8
	6. Reunião técnica IBGE–Serpro de 27 de maio de 2026.....	8
V.	DAS CONSTATAÇÕES DOCUMENTAIS E DOS PROBLEMAS INEQUÍVOCOS	10
	1. Quadro analítico de constatações	10
	2. Contradição entre a vedação de uso secundário e o aprendizado descrito no SerproBots.....	11
	3. A resposta “ <i>hospedado no Serpro</i> ” é materialmente incompleta	11
	4. “Nuvem soberana” não abrange automaticamente todos os componentes contratados	12
	5. Cláusula ampla de propriedade de modelos e ausência de ressalva para derivados.....	13
	6. Data Lake, CNEFE, Receita Federal e risco de transformação indevida do escopo	13
	7. Retenção, descarte e portabilidade sujeitos a regras potencialmente incompatíveis.....	15
	8. Mapa de riscos insuficiente para o objeto efetivamente contratado	15
	9. Ausência de formalização do Termo de Compromisso de Manutenção de Sigilo, sendo que a primeira versão trazia dados desatualizados da IBM, e não do Serpro	16
	10. Agregação de várias soluções sob a denominação de uma única solução de TIC.....	18
	11. Governança tardia e risco de atestes sem fiscalização plenamente constituída	19
	12. A reunião técnica de 27/05/2026 torna concretas questões de arquitetura, classificação e governança.....	21
	1. Fatos documentados.....	25
	2. Inferências tecnicamente justificadas.....	25
	3. Fatos que exigem apuração e medidas preventivas.....	26
VI.	DO REGIME JURÍDICO E NORMATIVO.....	27
	1. Sigilo estatístico e finalidade exclusiva	27
	2. LGPD e papéis funcionais dos agentes.....	27
	3. STF e compartilhamento de dados pelo Poder Público	27
	4. Decreto nº 10.046/2019 e competência do CCGD	28
	5. IN SGD/ME nº 94/2022 e Portaria SGD/MGI nº 5.950/2023	28

VII. DAS DILIGÊNCIAS E DOS DOCUMENTOS REQUERIDOS.....	28
VIII. DAS PERGUNTAS QUE DEVEM SER RESPONDIDAS SEM REMISSÃO GENÉRICA	30
IX. CONCLUSÃO	31
X. DAS MEDIDAS PREVENTIVAS	32
XI. DOS PEDIDOS À SGD E AO CCGD.....	33

I. INTRODUÇÃO

Na presente quadra histórica, a ASSIBGE — Sindicato Nacional não se opõe, assim como não se opôs no passado, à modernização tecnológica do IBGE. Ao longo de décadas de existência, a entidade sindical assistiu a profundas transformações e as saudou, como fez diante do emprego de computadores portáteis na coleta de dados. Justamente por reconhecer a importância da inovação e da necessidade de robustez das novas ferramentas, reivindica ativamente que a área de tecnologia da informação não seja tratada como atividade meramente acessória, passível de terceirização ou de outras formas de precarização.

A ASSIBGE, portanto, não se opõe ao emprego da computação em nuvem nem ao uso responsável da inteligência artificial.

Contudo, o IBGE está submetido ao compromisso com o sigilo estatístico, estabelecido pela legislação brasileira e por princípios internacionais. Esse compromisso exige cuidado minucioso com as informações, especialmente com aquelas desagregadas ou passíveis de reidentificação. Desse modo, não se pode admitir hipótese em que microdados sigilosos sejam utilizados para treinar modelos gerais de inteligência artificial alheios ao próprio Instituto, sejam eles fornecidos ou operados pelo Serpro ou por terceiros.

Diante dos riscos identificados, apresenta-se esta representação, fundamentada em fatos documentalmente verificáveis: **o contrato admite tecnologias e fluxos de alto risco; seus próprios anexos contêm disposições conflitantes ou contraditórias; o planejamento não individualizou riscos essenciais; a execução começou sem que a governança contratual estivesse tempestivamente constituída e suas responsabilidades claramente atribuídas; e não tardaram a surgir aparentes inobservâncias dos termos pactuados.**

A gravidade da situação impõe à ASSIBGE o recurso à presente via, a fim de exigir que a direção de turno do IBGE, assim como as vindouras, demonstre, mediante documentos e controles auditáveis, que cada operação respeita a finalidade estatística, a proteção legal das informações, a legislação relativa à proteção de dados pessoais, as normas aplicáveis à contratação de serviços de nuvem e as obrigações expressamente assumidas pelo IBGE e pelo Serpro, na hipótese de manutenção do contrato ora examinado.

II. SÍNTESE

O exame do Processo nº **03603.000033/2025-15** permite destacar, desde logo, as seguintes constatações:

1. o **Contrato IBGE nº 19/2025**, assinado em 28 de agosto de 2025, tem valor estimado de **R\$ 76.768.423,08**, vigência de 36 meses, possibilidade de prorrogação por até dez anos e objeto amplo, que reúne *multicloud*, segurança, relacionamento com o cidadão, SerproBots, tokens GPT, consultoria, serviços profissionais, Propostas Técnicas de Atendimento – PTA e **“demais serviços”**;
2. o contrato vincula o Termo de Referência, a proposta do Serpro e seus anexos, estabelecendo que, em caso de divergência, **prevalecem os parâmetros do Termo de Referência**;
3. o **Termo de Referência nº 80/2025, item 4.11**, veda o uso de informações do IBGE para propaganda, “otimização de mecanismos de inteligência artificial” ou qualquer uso secundário não autorizado;

4. em sentido que exige compatibilização formal e técnica, o **Anexo 3 da Proposta Comercial – SerproBots informa que mensagens podem ser enviadas às nuvens da IBM e da Microsoft Azure/OpenAI, armazenadas nos Estados Unidos por 30 dias e utilizadas para “aprendizado” do serviço (Item 3.8.1 do ANEXO 2 – GOVSHIELD);**
5. a proposta contempla o **Data IBGE**, descrito como Data Lake capaz **de consolidar dados em forma bruta ou processada e acessível a públicos variados, além de API que combina referência CNEFE com dados cadastrais de pessoas físicas ou jurídicas da Receita Federal**; esses serviços constam como “em concepção” ou “concebido”, não havendo, no material examinado, prova de sua entrada em produção;
6. o contrato **atribui ao Serpro a propriedade exclusiva da “tecnologia e modelos desenvolvidos direta ou indiretamente”** para a prestação dos serviços, **sem ressalva textual específica** para modelos, *embeddings*, índices, *datasets*, ajustes ou outros artefatos derivados de **informações protegidas do IBGE**;
7. o Mapa de Gerenciamento de Riscos contém riscos genéricos de contratação e execução, mas não individualiza sigilo estatístico, treinamento de IA, transferência internacional, reidentificação, integração de bases, suboperadores, propriedade de artefatos derivados, retenção em nuvens parceiras, portabilidade e dependência tecnológica;
8. a Procuradoria Federal reconheceu a regularidade jurídica **com ressalvas**, registrando expressamente que sua análise do Estudo Técnico Preliminar (ETP) e do Termo de Referência era formal e que os aspectos técnicos, econômicos e financeiros permaneciam sob responsabilidade da Administração;
9. **a resposta do IBGE ao pedido sindical informou que o “ChatBot (IA)” poderia ser utilizado “estando hospedado no Serpro”, mas não esclareceu que o anexo técnico prevê processamento e retenção temporária por IBM e OpenAI nos Estados Unidos;**
10. em fevereiro de 2026, **meses depois do início da execução** e quando já havia cobranças relativas a outubro, novembro e dezembro de 2025, unidades do próprio IBGE registraram que **não localizavam portaria de designação de gestores e fiscais e discutiam a quem caberia gerir o contrato. Também foram apontadas pendências de registro e publicidade contratual.**
11. **a gravação da Reunião Técnica IBGE-Serpro de 27 de maio de 2026 registra participação direta da Amazon Web Services – AWS** no levantamento de requisitos e no desenho da arquitetura, inclusive com representante da AWS apresentando componentes; discussão de arquitetura híbrida entre a nuvem soberana do Serpro e *cloud pública/multicloud*; referência ao uso, em IA, apenas de dados então qualificados como “não sensíveis”; integração de bases por CPF, Cadastro Único, endereço e *joins*; hipótese de Data Lake institucional; discussão de anonimização; alternativa de software como serviço (SaaS) no contrato Multicloud; e dúvidas relevantes sobre escopo, aprovação e cadeia decisória.

Esses elementos tornam necessária uma apuração técnica que distinga três planos:

- **FATOS COMPROVADOS PELO PRÓPRIO PROCESSO ADMINISTRATIVO**, como a contratação de tokens GPT, a possibilidade de transferência internacional, o armazenamento de conversas, a previsão de aprendizado, o Data Lake em concepção e a demora na estruturação da gestão contratual;

- **INCOMPATIBILIDADES E LACUNAS QUE JÁ EXIGEM CORREÇÃO**, como a **contradição entre a vedação do Termo de Referência e o mecanismo de aprendizado descrito no SerproBots**, a insuficiência do mapa de riscos e a falta de delimitação da propriedade sobre artefatos derivados;
- **HIPÓTESES QUE DEPENDEM DE PROVA**, como a utilização efetiva de microdados sigilosos em IA, a ativação do Data Lake, a integração operacional CNEFE–Receita Federal e o uso de qualquer informação protegida para aperfeiçoar produto geral do Serpro ou de terceiro.

III. DA LEGITIMIDADE E DO INTERESSE SINDICAL

O trabalho estatístico depende de uma relação singular de confiança. Pessoas, famílias, empresas e instituições prestam informações ao Estado sob a garantia de que os dados individuais serão protegidos e utilizados nas finalidades legalmente autorizadas. **Uma eventual falha nessa governança não atingiria apenas sistemas de tecnologia.** Poderia afetar a confiança dos informantes, a qualidade das respostas, as condições de trabalho das equipes de coleta e análise, a credibilidade das estatísticas oficiais, a autonomia técnica do Instituto e a própria capacidade do País de conhecer sua realidade.

Os trabalhadores do IBGE conhecem, em sua atividade cotidiana, a diferença entre dado público, informação agregada, registro administrativo, dado pessoal, dado anonimizado e microdado protegido por sigilo estatístico. A participação sindical, portanto, seja nesta pauta ou em outras de defesa institucional do IBGE, não é acessória: integra a defesa do patrimônio informacional público e das condições institucionais indispensáveis à produção estatística.

IV. DO CONTRATO E DO CONTEXTO PROCESSUAL

1. Planejamento da contratação e escopo originalmente concebido

O Documento de Formalização da Demanda e o **Estudo Técnico Preliminar nº 16/2025 (ETP)** estruturaram uma contratação abrangente de serviços de computação em nuvem, relacionamento com o cidadão, segurança e consultoria. O ETP menciona, entre as cargas ou usos possíveis, migração de bancos de dados, infraestrutura para o Censo Agropecuário, assistentes com inteligência artificial, chatbots para trabalhadores e cidadãos, e “inteligência artificial com GitHub”.

O ETP estimou o atendimento de aproximadamente 6 mil servidores e 32 mil trabalhadores temporários do Censo Agropecuário por chatbot, além de 12 mil unidades anuais de mil tokens GPT. Registrou também que a migração de bancos de dados poderia abranger informações confidenciais e sigilosas, justificando a utilização da chamada Nuvem de Governo. *(ETP nº 16/2025, itens 4, 6 e 7; processo, arq. 1, PDF pp. 130–155.)*

Apesar dessa complexidade, o **ETP afirmou não serem necessários treinamento dos gestores e fiscais**, e tampouco adaptações no ambiente institucional, sob o fundamento de que o IBGE já possuía contratos similares.

Tal conclusão, como vem apontando a ASSIBGE, merece reavaliação diante das especificidades de IA generativa, suboperadores estrangeiros, integração de bases e sigilo estatístico.

2. Instrução incompleta, urgência e controle jurídico limitado

Em maio de 2025, o processo tramitou com urgência e com registro de que nem todos os artefatos exigidos pela IN SGD/ME nº 94/2022 estavam assinados, porque integrantes da equipe se encontravam em férias. E-mails juntados aos autos mencionam a necessidade de máxima urgência antes de paralisação da área de compras. *(Despacho DTI/GPG/IBGE nº 78/2025 e comunicações subsequentes; arq. 1, PDF pp. 271–285.)*

A Procuradoria Federal devolveu inicialmente os autos para que fossem demonstradas a adaptação à Portaria SGD/MGI nº 5.950/2023, a habilitação do Serpro, uma pesquisa efetivamente referenciada ao mercado e a análise crítica dos preços. Em nova manifestação, registrou que a pesquisa se apoiava apenas em contratos similares do próprio Serpro e reiterou a necessidade de comparação com outras empresas. Também questionou a definição aberta do item “Serviços Especiais”. *(Despachos nº 00227/2025 e nº 00250/2025/COACON/PFE-IBGE/PGF/AGU; arq. 1, PDF p. 287, e arq. 3, PDF p. 562.)*

No Parecer nº 00136/2025, a Procuradoria concluiu pela **regularidade jurídica com ressalvas**, condicionada ao atendimento de recomendações. O próprio parecer enfatizou que não avaliava conclusivamente os aspectos técnicos, mercadológicos, econômicos ou de conveniência; considerou o ETP e o Termo de Referência formalmente adequados e remeteu à Administração a responsabilidade pela suficiência material das especificações, riscos e preços. *(Arq. 3, PDF pp. 622–635.)*

Essa limitação é importante: a existência do parecer jurídico não substitui a necessária avaliação técnica de sigilo estatístico, IA, arquitetura, transferência internacional, suboperadores e artefatos derivados.

3. Objeto, valor, vigência e hierarquia documental

O Contrato nº 19/2025 reuniu dezoito itens, organizados em cinco grupos:

- Serpro Multicloud, gerenciamento de nuvem, conectividade e última milha;
- Dialoga, SerproBots, tokens GPT, WhatsApp Business e conversações;
- GovShield;
- serviços profissionais genéricos de nuvem com gerenciamento de migração e consultoria técnica;
- Serviços Especiais por PTA.

O valor total estimado é de **R\$ 76.768.423,08, com pagamento por consumo**. A vigência é de 36 meses, prorrogável até o limite de dez anos. A subcontratação parcial é admitida mediante aprovação prévia na respectiva Ordem de Serviço, desde que se limite a serviços acessórios e instrumentais e não prejudique a segurança dos dados nem as razões da contratação direta. *(Contrato nº 19/2025, cláusulas 1ª, 2ª, 4ª e 5ª; arq. 4, PDF pp. 1–6.)*

O contrato incorpora, independentemente de transcrição, o Termo de Referência, a proposta do Serpro e seus anexos. A cláusula 1.5 é expressa: se houver divergência entre a proposta, o Termo de Referência e o contrato, **prevalecem os parâmetros do Termo de Referência**.

Essa regra torna juridicamente relevante a incompatibilidade descrita a seguir.

4. Questionamento sindical e resposta institucional

Em 30 de outubro de 2025, a ASSIBGE-SN apresentou o Ofício nº 70/25, solicitando esclarecimentos sobre os dados a serem transferidos, os riscos à soberania estatística, a confiança dos informantes, a participação das áreas técnicas, o alinhamento ao planejamento, a autonomia tecnológica, a continuidade financeira e a hipótese de privatização do Serpro.

Em resposta juntada ao processo, o IBGE remeteu genericamente ao ETP, informou que bancos de dados poderiam ser migrados para a Nuvem de Governo e afirmou que o “ChatBot (IA)” poderia ser utilizado “estando hospedado no Serpro”. **Sobre os riscos de vazamento e ordens de governo estrangeiro, declarou que diretores estavam cientes e haviam “avaliado os impactos e aceitado o contrato”, sem apresentar avaliação específica ou ata correspondente.** (*Processo Administrativo nº 03603.000033/2025-15, pp. 974–979.*)

Ora, a ciência de dirigentes não equivale à identificação, mensuração, tratamento e aceitação formal de riscos. **Tampouco a expressão “hospedado no Serpro” descreve integralmente o fluxo previsto no Anexo SerproBots, que alcança IBM, Microsoft Azure/OpenAI e armazenamento temporário nos Estados Unidos.**

5. Execução iniciada antes da consolidação da governança contratual

Em fevereiro de 2026, a **Gerência de Planejamento e Gestão da DTI registrou que já recebia documentos de cobrança, inclusive relativos ao GovShield nos meses de outubro, novembro e dezembro de 2025, mas não localizava a portaria de designação dos gestores e da equipe de fiscalização. Apontou também pendências de registro no Contratos.gov e de publicidade no PNCP e no Diário Oficial da União.**

Seguiu-se troca de manifestações entre DTI e unidades de compras sobre quem deveria gerir a contratação. Apenas em 12 e 13 de fevereiro de 2026 foram expedidas orientações para que o CDDI indicasse gestores e, em conjunto com a DTI, fiscais técnicos, além de providenciar registro, acompanhamento e recepção das notas fiscais. (*Processo Administrativo nº 03603.000033/2025-15, pp. 179–197.*)

Não se trata de falha meramente formal. Em contrato que admite IA, transferência internacional, múltiplas nuvens, subcontratação e tratamento de informação protegida, **a gestão e a fiscalização precisam existir ANTES da ativação dos serviços, com atribuições, competências e evidências claramente definidas.**

6. Reunião técnica IBGE–Serpro de 27 de maio de 2026

A ASSIBGE-SN dispõe da gravação integral da reunião técnica realizada em 27 de maio de 2026. Conforme consta da gravação, integrante da arquitetura corporativa do Serpro afirmou apoiar o projeto, **“junto com a AWS” [Amazon],** no levantamento de requisitos e na arquitetura a ser proposta. Em seguida, confirmou que os componentes então exibidos eram apresentados por “Carlos da AWS”. A relação foi descrita como parceria ligada ao contrato Multicloud do Serpro e os participantes da AWS como parceiros no desenvolvimento da solução. Em passagem posterior, o Serpro e a AWS foram também qualificados como consultores para o desenvolvimento do projeto. (*Gravação de 27/05/2026, aproximadamente 1h46min–1h49min e 2h04min.*)

Questionado se os dados seriam levados para a nuvem da Amazon, o interlocutor respondeu que não necessariamente todos e afirmou que a arquitetura previa que apenas

dados considerados “não sensíveis” fossem usados na inteligência artificial disponível naquele ambiente. Na sequência, descreveu-se **arquitetura híbrida**, com utilização da denominada nuvem soberana do Serpro e de recursos da cloud pública/multicloud, pois certas ferramentas, sobretudo de IA, não estariam disponíveis no ambiente interno contratado. Ressalvou-se, contudo, que ainda seria necessário definir o que seria o Data Lake, onde ficaria e quais dados reuniria. *(Gravação, aproximadamente 1h48min–1h51min.)*

A expressão “não sensíveis” exige esclarecimento específico. Ela pode não coincidir com “dados pessoais sensíveis” da LGPD e não afasta, por si só, o sigilo estatístico, o regime jurídico da base de origem, a indispensável checagem da possibilidade de reidentificação nem outras restrições de acesso. Devem ser identificados o conceito adotado na reunião, os critérios, a autoridade responsável pela classificação e os controles que impeçam que dados protegidos, pseudonimizados ou reidentificáveis alcancem ambientes externos de IA.

A discussão não se limitou a consultas unitárias por API. Participantes do IBGE explicitaram necessidade de integração entre CPF, Cadastro Único e informações de endereço, para localizar correspondências entre pessoas e registros de bases distintas, e **foi afirmado que os dados deveriam estar disponíveis para realização de joins.** Também se cogitou estrutura de Data Lake com vocação institucional, reutilizável em projetos além do Censo Agropecuário, embora outros participantes contestassem que essa finalidade estivesse definida. Houve, ainda, discussão sobre anonimização e seus limites no contexto de cruzamentos e integração de bases. *(Gravação, aproximadamente 1h51min–2h02min.)*

Em outro momento, diante da observação de que o contrato não continha item de desenvolvimento por ponto de função, foi mencionada como alternativa a entrega da solução na forma de software como serviço – SaaS no âmbito do contrato Multicloud. **A própria reunião evidencia que escopo, requisitos, instrumento contratual e critérios de aceite precisavam ser formalmente esclarecidos antes de qualquer implementação.** *(Gravação, aproximadamente 3h11min–3h15min.)*

Por fim, a gravação registra questões de governança e cadeia decisória: **afirmações de que a demanda teria partido da Presidência do IBGE**, apresentações e tratativas em nível de diretores e coordenadores e **caracterização do projeto como prioritário ou estratégico coexistiram com manifestações de equipes técnicas que desconheciam o documento de visão, os requisitos, a aprovação e até a finalidade precisa da solução.** Também se observou a ausência, na reunião, de responsável apontado como gestor geral do contrato. São elementos de relevo que tornam indispensável distinguir exploração técnica, proposta, aprovação institucional, contratação, ordem de serviço, homologação e entrada em produção. *(Gravação, aproximadamente 1h57min–2h00min e 3h18min–3h29min.)*

A ASSIBGE-SN entende, assim, ser inadiável uma apuração documental e técnica do alcance efetivo da participação da AWS e de quaisquer outros provedores, com a distinção entre soluções tão somente cogitadas daquelas eventualmente autorizadas ou implementadas.

V. DAS CONSTATAÇÕES DOCUMENTAIS E DOS PROBLEMAS INEQUÍVOCOS

1. Quadro analítico de constatações

Constatação documental	Referência no Processo Administrativo nº 03603.000033/2025-15 ou outra base documental	Implicação que exige providência
O Termo de Referência proíbe uso de informações do IBGE para otimização de IA ou uso secundário não autorizado.	TR nº 80/2025, item 4.11; arq. 1, PDF p. 189.	A vedação deve prevalecer e ser traduzida em controles técnicos auditáveis.
O SerproBots prevê envio de mensagens à IBM e Azure/OpenAI, retenção de 30 dias nos EUA e “aprendizado” com interações.	Proposta nº 20250140v002 (<i>que integra o Contrato nº 19/2025 por força de suas cláusulas 1.3 e 1.5</i>), Anexo 3, itens 3.6 a 3.10; arq. 4, PDF pp. 133–135.	Há colisão direta com o item antecedente ; o serviço não deve receber dado protegido.
O contrato prevê tokens GPT e o ETP cogita chatbot para o Censo Agropecuário.	Contrato, itens 9 e 10; ETP, itens 4 e 7; arq. 1, PDF pp. 131, 139 e 149; arq. 4, PDF pp. 1–4.	O risco não é abstrato; IA generativa está efetivamente dentro do objeto contratado.
A proposta contempla Data Lake com dados brutos ou processados e API CNEFE–Receita .	Proposta, Anexo 8; arq. 4, PDF pp. 163–167.	É necessário impedir que descrição genérica seja tratada como autorização de acesso , integração ou compartilhamento.
A propriedade de tecnologia e modelos desenvolvidos para o serviço é atribuída exclusivamente ao Serpro.	Contrato, cláusula 20.1; arq. 4, PDF p. 12.	Falta ressalva expressa para dados, modelos e artefatos derivados de informações protegidas do IBGE.
A proposta afasta responsabilidade do Serpro por políticas de continuidade, segurança e privacidade em nuvens parceiras , embora o cliente possa delegar parte da gestão.	Anexo Multicloud, item 3.2.1; arq. 4, PDF p. 94.	As responsabilidades precisam ser individualizadas por carga, provedor e Ordem de Serviço.
O encerramento se sujeita também às políticas de retenção dos provedores parceiros, e o Serpro não garante recuperação.	Anexo Multicloud, itens 10.2 e 10.3; arq. 4, PDF p. 115.	A eliminação e a portabilidade não podem depender de políticas externas mutáveis e não anexadas.
A gestão contratual foi objeto de disputa interna meses após o início das cobranças.	Arq. 4, PDF pp. 179–197.	A fiscalização pretérita e atual deve ser auditada e documentalmente recomposta.
A reunião técnica de 27/05/2026 registra participação da AWS no levantamento de requisitos e arquitetura, hipótese híbrida/multicloud, uso de IA com dados qualificados como ‘não sensíveis’ , integrações por joins, Data Lake institucional, anonimização, alternativa SaaS e dúvidas de governança.	Gravação integral da reunião técnica IBGE–Serpro de 27/05/2026, especialmente aprox. 1h46min–2h04min e 3h11min–3h29min.	É necessário esclarecer papéis, classificação, bases, ambientes, decisões e execução, de modo a impedir eventual envio de dados protegidos à AWS ou uso em treinamento.

2. Contradição entre a vedação de uso secundário e o aprendizado descrito no SerproBots

O Termo de Referência determina:

É vedado o uso de informações do CONTRATANTE pelo provedor de serviço de nuvem para propaganda, otimização de mecanismos de inteligência artificial ou qualquer uso secundário não autorizado.

A proposta do Serpro, porém, descreve que:

- o IBM Watson interpreta mensagens em servidores externos da IBM;
- o componente GPT conecta o chatbot à nuvem Azure e ao serviço OpenAI;
- históricos de conversas e mensagens são armazenados e podem ser usados para auditoria e curadoria;
- *logs* precisam ser submetidos a ambiente externo de fornecedor terceiro;
- **os dados permanecem fora do território nacional por um mês;**
- **esse armazenamento tem por objetivo prover “aprendizado de máquina”,** utilizando interações dos usuários para aprendizado pelo serviço.

O contrato resolve a hierarquia jurídica em favor do Termo de Referência, mas não resolve a arquitetura. Permanecem perguntas objetivas: o “aprendizado” descrito é apenas inferência transitória, curadoria pelo cliente, ajuste de instância dedicada ou melhoria do serviço de terceiro? **Que configuração impede que conteúdo do IBGE seja utilizado para treinamento ou aperfeiçoamento geral? Como se concilia retenção de 30 dias com uma obrigação de não otimização?**

Enquanto essas questões não forem documentalmente sanadas e tecnicamente verificadas, **o SerproBots não deve receber microdados, registros administrativos protegidos, informações submetidas a sigilo estatístico nem conteúdo que permita acesso indireto a essas bases.**

3. A resposta “*hospedado no Serpro*” é materialmente incompleta

Ao responder à ASSIBGE — Sindicato Nacional, o IBGE afirmou que o ChatBot/IA poderia ser usado “*estando hospedado no Serpro*”. O anexo contratual, entretanto, prevê processamento por IBM e OpenAI, por meio de IBM Cloud e Microsoft Azure, com possibilidade de armazenamento nos Estados Unidos, conforme se afere em trecho transcrito abaixo:

3.10.1 A Plataforma Serprobots utiliza a **solução “Watson Assistant” através da nuvem “Cloud” da empresa IBM** - International Business Machines e a solução “GPT” da empresa OpenAI através da nuvem “Azure” da Microsoft. Nesse contexto, o usuário fica ciente de que os **dados digitados no chat poderão ser transferidos internacionalmente e ficam armazenados na infraestrutura da empresa por um período de 30 dias**. Após este período os dados são excluídos em definitivo. Tal armazenamento tem o objetivo de prover o aprendizado de máquina da ferramenta, onde as interações dos usuários no chat são utilizadas para “aprendizado” pelo serviço, que envia as respostas automáticas quando o usuário está sendo atendido por meio de chatbots que o utilizem.

[...]

3.10.3.1 **País: Estados Unidos da América**, Organizações: “IBM - International Business Machines” e “OpenAI”. (g.n.)

(Referência: ARQ. 4, PDF, p. 134 – Anexo 3 da Proposta Comercial SERPRO – nº 20250140v002, que passou a integrar o Contrato nº 19/2025 por força de sua cláusula 1.3, onde se lê: “Vinculam esta contratação, independentemente de transcrição: [...] 1.3.3. A Proposta do contratado;” o que é reiterado na cláusula 1.5, com expressa referência ao número da Proposta Comercial.)

Portanto, hospedar a camada de orquestração no Serpro não significa manter todo o tratamento no Serpro, ou, em termos mais mezinhos, não basta que a camada mais aparente se dê através do Serpro para garantir as condições adequadas de tratamento dos dados. É necessário, isso sim, informar separadamente:

- onde ficam a interface, a orquestração, o banco de histórico, os logs e as chaves;
- para onde seguem mensagens, prompts, anexos e metadados;
- qual provedor recebe cada conteúdo;
- **que região e jurisdição se aplicam;**
- **qual política de retenção e treinamento está ativada;**
- **quem possui acesso administrativo e de suporte;**
- se há mecanismos de *egress filtering* (controle de tráfego/fluxo de saída), anonimização, pseudonimização e prevenção de exfiltração (saída não autorizada de dados, direta ou indiretamente, como para nuvens, APIs, serviços de IA, telemetria ou suporte de terceiros).

É indispensável, portanto, que a resposta institucional seja corrigida e completada.

4. “Nuvem soberana” não abrange automaticamente todos os componentes contratados

O ETP justificou parte da contratação pela necessidade de uma nuvem soberana para informações confidenciais e sigilosas. Essa justificativa pode ser pertinente para cargas alocadas em infraestrutura governamental compatível. Contudo, ao se confrontarem os termos do contrato, seus anexos e os fatos supervenientes, tal justificativa não subsiste senão como propaganda ou devaneio.

O objeto é multicloud. A proposta admite nuvens parceiras e informa que o Serpro não responde pelo mau uso do cliente nem pela definição ou implementação das políticas de continuidade, segurança e privacidade das aplicações hospedadas. **O próprio processo registra serviços na Azure e, em junho de 2026, despacho requerendo autorização para a dispensa de licitação relativa ao SEI em provedor Huawei.** (Arq. 4, PDF p. 200), *in verbis*:

SOLICITO AUTORIZAÇÃO para a aquisição por meio de DISPENSA de LICITAÇÃO, [...] junto à empresa pública Serviço Federal de Processamento de Dados (SERPRO) atendendo as necessidades específicas da Administração [...] no PI INFO, do item 16 (Serpro Multicloud - Cloud Generic Professional Services com Cloud Migration Management CATSER: 27332) do grupo 4 - Consultoria, correspondente à Solicitação de Serviço nº 2026SS/2202611942X (Atualização Versão - SEI, **considerando o Serpro MultiCloud, em provedor Huawei**).

É indispensável classificar cada workload e demonstrar, por Ordem de Serviço, onde ela está, por que aquele ambiente é compatível com seu regime jurídico e quais fluxos externos subsistem. A soberania do data center principal não elimina dependências de plano de controle, telemetria, suporte, atualização, APIs de modelo, chaves, backups ou fornecedores estrangeiros.

5. Cláusula ampla de propriedade de modelos e ausência de ressalva para derivados

A cláusula 20.1 atribui ao Serpro a propriedade intelectual da tecnologia e dos modelos desenvolvidos direta ou indiretamente para a prestação dos serviços. Em paralelo, o anexo Multicloud garante ao órgão exclusividade sobre as informações tratadas, inclusive cópias e backups.

Essas previsões não são equivalentes. Pois, de um lado, os dados podem permanecer formalmente pertencentes ao IBGE, ao passo que, de outro lado, embeddings (*representações numéricas ou vetoriais geradas a partir dos dados, utilizadas por sistemas de inteligência artificial para busca, comparação e recuperação semântica*), índices vetoriais, datasets de avaliação, regras de curadoria, adapters, pesos ajustados, classificadores, modelos ou outros artefatos derivados são reivindicados como tecnologia do prestador.

É indispensável pontuar que **um *embedding* não deve ser tratado automaticamente como dado anônimo ou inofensivo**. Sendo um derivado do dado original, pode permanecer associado a registros ou indivíduos, além de permitir recuperação, inferência ou correlação, dependendo da arquitetura, pois textos com sentidos próximos tendem a gerar *embeddings* “próximos” entre si, o que permite buscas por significado, recuperação de documentos, classificação e sistemas RAG (Retrieval-Augmented Generation, ou geração aumentada por recuperação de informação).

O problema, portanto, não está em o Serpro conservar sua tecnologia preexistente. Está na falta de uma fronteira expressa que:

- exclua dados do IBGE e quaisquer derivados protegidos da propriedade exclusiva do Serpro;
- proíba reutilização entre clientes e produtos;
- discipline titularidade, licença, portabilidade e eliminação dos artefatos criados sob encomenda;
- impeça que informação sigilosa se incorpore, de modo irreversível, a modelo geral ou ativo reutilizável.

Ainda que se trate de cláusula aplicada a outros contratos celebrados pelo Serpro, não se pode ignorar que os dados coletados pelo IBGE, enquanto desagregados, exigem excepcional cuidado, sob o risco de prejudicar a confiança das pessoas físicas e jurídicas quanto à coleta dos dados.

A cláusula precisa ser aditada ou, de outro modo, ser interpretada restritivamente.

6. Data Lake, CNEFE, Receita Federal e risco de transformação indevida do escopo

O Anexo 8 (*arq. 4, PDF, fl. 163*) descreve o “Data IBGE” como plataforma de armazenamento massivo para consolidar dados coletados pelo Instituto, em forma bruta ou processada, e torná-los acessíveis a governos, empresas, pesquisadores e público em geral. **Também descreve API**

de geocodificação que combina referência CNEFE com dados cadastrais de pessoas físicas ou jurídicas da Receita Federal.

Vejamos abaixo os trechos do anexo:

2. Data IBGE

Plataforma de armazenamento massivo de dados, projetada para consolidar e organizar o grande volume de dados coletados pelo IBGE, permitindo o acesso e a análise de informações em sua forma bruta ou processada. Este Data Lake irá suportar múltiplas áreas de inteligência, como demografia, economia, agropecuária, geociências e educação, sendo acessível para usuários variados, como governos, empresas, pesquisadores e o público em geral.

4.1. API - Geocodificação com Referência CNEFE

Solução de interoperabilidade (API) que permite aos interessados acesso em tempo real a dados georreferenciados da base de dados do IBGE combinados com dados cadastrais de pessoas físicas e/ou jurídicas da base de dados da Receita Federal do Brasil.

Neste ponto, destaca-se que o próprio apêndice afirma que produtos do Serpro seriam baseados em **dados públicos**, dependeriam de autorização do IBGE, Portarias específicas e instrumentos próprios. Essa ressalva é indispensável e não deve ser flexibilizada. Contudo, a linguagem excessivamente genérica sobre dados “brutos”, a associação com CNEFE e Receita Federal e a possibilidade de atualizar serviços do contrato por apostilamento (*item 7 do Anexo 8*), ou seja, a alteração do contrato sem a necessidade de um termo aditivo, criam risco de expansão do escopo sem o exame exigido para cada base e finalidade.

É necessário registrar de forma inequívoca que:

1. a descrição do Anexo 8 não autoriza, por si só, acesso a microdados ou registros protegidos;
2. cada integração exige base jurídica, finalidade específica, minimização, classificação, gestor de dados, instrumento de compartilhamento quando aplicável e **avaliação de reidentificação**;
3. **apostilamento não pode ser usado para introduzir obrigação material nova, novo tratamento de risco ou nova finalidade que em verdade demande termo aditivo, nova contratação ou procedimento próprio**;
4. **produtos destinados a terceiros devem utilizar exclusivamente dados públicos ou resultados estatísticos cuja divulgação esteja formalmente autorizada pelo IBGE**;
5. **nenhum produto pode permitir reconstrução, inferência ou reidentificação de informação individual protegida.**

7. Retenção, descarte e portabilidade sujeitos a regras potencialmente incompatíveis

O Termo de Referência afirma que não há tolerância para manutenção de dados após o encerramento da vigência. O anexo Multicloud, contudo, condiciona a manutenção em ambientes desativados às políticas vigentes dos provedores parceiros e declara que o Serpro não garante recuperação de dados ou ambientes. O Anexo de Proteção de Dados prevê eliminação em até 30 dias, ressalvadas hipóteses legais, e admite que o Serpro retenha dados diante de base legal ou contratual própria, inclusive para resguardo de seus interesses legítimos.

Além disso, o SerproBots prevê retenção internacional por 30 dias. Essas regras precisam ser harmonizadas por tipo de dado e serviço. Para sigilo estatístico, não basta remissão genérica à LGPD ou a políticas de terceiros. Devem existir:

- prazo de retenção contratualmente fixado;
- retorno e portabilidade em formato utilizável;
- eliminação verificável de produção, logs, caches, backups e derivados;
- tratamento específico de incidentes e litígios;
- vedação de retenção para interesse próprio quando incompatível com sigilo legal ou finalidade estatística;
- certificado de destruição e direito de auditoria.

8. Mapa de riscos insuficiente para o objeto efetivamente contratado

O Mapa de Gerenciamento de Riscos examinado lista riscos como atraso, especificação incompleta, baixa qualidade, incapacidade do fornecedor, documentação da contratação direta, níveis de serviço e fiscalização inadequada. **Portanto, não individualiza os riscos mais relevantes revelados pelos demais artefatos:**

- violação do sigilo estatístico e perda de confiança dos informantes;
- uso de dados ou interações para treinamento, ajuste ou melhoria de IA;
- transferência internacional e **incidência de jurisdição estrangeira;**
- **reidentificação pela combinação de bases;**
- suboperadores e cadeias de fornecimento multicloud;
- **incompatibilidade entre propriedade dos dados e propriedade de modelos derivados;**
- **retenção, eliminação e recuperação em provedores parceiros;**
- **dependência tecnológica, portabilidade e saída;**
- **uso de apostilamentos ou PTAs** para ampliar tratamentos e finalidades;
- gestão de **acessos privilegiados e segregação entre clientes;**
- **indisponibilidade de gestores e fiscais qualificados.**

O mapa deve ser refeito e acompanhado durante toda a execução, conforme exige a IN SGD/ME nº 94/2022.

9. Ausência de formalização do Termo de Compromisso de Manutenção de Sigilo, sendo que a primeira versão trazia dados desatualizados da IBM, e não do Serpro

Uma sequência de documentos do processo enseja esclarecimentos específicos quanto à elaboração e à efetiva formalização dos instrumentos destinados à proteção das informações sigilosas do IBGE. Trata-se de ponto essencial que parece ter sido tratado com açodado desmazelo.

Na primeira versão apresentada do Termo de Referência, o **Anexo II-A – Termo de Compromisso de Manutenção de Sigilo**, embora inserido na contratação que se destinava ao Serpro, identificava como **CONTRATADA a IBM Brasil Indústria, Máquinas e Serviços Ltda.**, com razão social, CNPJ nº 33.372.251/0001-56 e endereço na Avenida Pasteur, 138/146, Botafogo, Rio de Janeiro. O texto chegava a declarar que, em razão do futuro “CONTRATO PRINCIPAL”, a IBM “poderá ter acesso a informações sigilosas do CONTRATANTE”. Ao mesmo tempo, permaneciam em aberto campos como o número do contrato, foro, nomes, qualificações, data, assinaturas e testemunhas.

A presença da IBM não se limitava, portanto, a uma referência genérica ou a um exemplo abstrato: o modelo fora preenchido com sua **razão social, CNPJ, endereço e qualificação expressa como CONTRATADA**. Além disso, o endereço indicado correspondia às antigas instalações da empresa no Bairro Botafogo, **deixadas pela IBM ainda em 2021**, indicando, portanto, que tal minuta provavelmente foi elaborada antes de tal marco, quando o modelo de IA por LLM era ainda incipiente.

Nota-se, assim, a hipótese de **reaproveitamento de instrumento oriundo de relação contratual anterior sem revisão integral de seu conteúdo**, embora esse elemento, isoladamente, não permita excluir a hipótese de que a IBM estivesse sendo então considerada como participante da cadeia tecnológica da solução (ainda que inexistisse minuta quanto aos demais integrantes de tal cadeia).

A tramitação posterior, contudo, fornece elemento adicional relevante. Na **versão 4.0 do Termo de Referência nº 80/2025, concluída em 15 de agosto de 2025**, a identificação da IBM foi substituída pela da **Empresa Nacional de Inteligência em Governo Digital e Tecnologia da Informação – Serpro**, agora corretamente qualificada como CONTRATADA, com seu CNPJ e endereço em Brasília, reforçando a hipótese de minuta anterior sem revisão adequada.

A correção, porém, **não corresponde à formalização do Termo de Compromisso**. Mesmo na versão final do Termo de Referência, o documento continua expressamente estruturado como modelo: o número do contrato permanece como “<nº do contrato>”; o foro como “<CIDADE DA CONTRATANTE>”; e os campos destinados aos representantes da contratante e da contratada, às testemunhas, ao local e à data permanecem sem preenchimento, como se demonstra dos recortes abaixo:

Pelo presente instrumento a FUNDAÇÃO INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, sediado Avenida Franklin Roosevelt, 166, Centro, na cidade do Rio de Janeiro, RJ, CNPJ nº 33.787.094 /0001-40, doravante denominado **CONTRATANTE**, e, de outro lado, a Empresa Nacional de Inteligência em Governo Digital e Tecnologia da Informação (SERPRO), sediada à Setor de Grande Área Norte, Quadra 601, Lote V – Asa Norte, em Brasília / DF – CEP 70836-900, inscrita no CNPJ - M.F. sob o nº 33.683.111/0001-07 doravante denominada **CONTRATADA**; CONSIDERANDO que, em razão do CONTRATO N.º <nº do contrato> doravante denominado **CONTRATO PRINCIPAL**, a **CONTRATADA** poderá ter acesso a informações sigilosas do **CONTRATANTE**; CONSIDERANDO a necessidade de ajustar as condições de revelação destas informações sigilosas, bem como definir as regras para o seu uso e proteção;
CONSIDERANDO o disposto na Política de Segurança da Informação e Privacidade da **CONTRATANTE**;

de celebrarem qualquer outro acordo entre si.

9 – FORO

A CONTRATANTE elege o foro da <CIDADE DA CONTRATANTE>, onde está localizada a sede da CONTRATANTE, para dirimir quaisquer dúvidas originadas do presente TERMO, com renúncia expressa a qualquer outro, por mais privilegiado que seja.

10 – ASSINATURAS

CONTRATADA	CONTRATANTE
<Nome>	<Nome>
<Qualificação>	Matrícula: xxxxxxxx

TESTEMUNHAS	
<Nome>	<Nome>
<Qualificação>	<Qualificação>

<Local>, <dia> de <mês> de <ano>.

(referência das reproduções: arq. 4, pp. 58–63 – Termo de Referência; “ANEXO II-A — TERMO DE COMPROMISSO DE MANUTENÇÃO DO SIGILO”)

A mesma situação se verifica no **Termo de Ciência** subsequente, no qual continuam em aberto o número do contrato (“xxxx/aaaa”), objeto, identificação da contratada, preposto, gestor e empregados que deveriam declarar ciência das obrigações de sigilo. O próprio documento

determina que os empregados diretamente envolvidos na contratação assumam formalmente esse conhecimento e que novos empregados também apresentem termos assinados.

Essa distinção é juridicamente relevante. O Contrato nº 19/2025 estabelece que o Termo de Referência integra a contratação e disciplina sua relação com a proposta comercial. Assim, **o modelo de Termo de Compromisso constante do Termo de Referência compõe a documentação contratual e estabelece obrigações a serem observadas**. Isso, todavia, não autoriza concluir que o instrumento autônomo nele previsto tenha sido efetivamente celebrado.

O próprio modelo associa seus efeitos à formalização: estabelece que o compromisso permanece em vigor **“desde a data de sua assinatura”**, refere-se às consequências assumidas pela contratada **“ao assinar o presente instrumento”** e contém seção específica destinada às assinaturas da contratante, da contratada e de testemunhas.

Compulsado o Processo nº 03603.000033/2025-15, não é possível localizar **versão posterior do Termo de Compromisso de Manutenção de Sigilo preenchida com o número do contrato, datada e assinada pelo IBGE e pelo Serpro**, nem foram localizados Termos de Ciência devidamente preenchidos e assinados pelos empregados abrangidos.

Não se pode, portanto, afirmar que o Termo de Compromisso propriamente dito tenha sido celebrado e passado a integrar o Contrato nº 19/2025. O que se pode afirmar é que seu modelo integra o Termo de Referência, o que a ASSIBGE não considera suficiente.

O episódio é particularmente significativo porque envolve justamente o instrumento concebido para formalizar a proteção de informações sigilosas. *Primeiro*, o processo contém modelo que identifica como contratada empresa diversa daquela que seria contratada, **inclusive com indícios de ter sido formulado em data em que tecnologias sensíveis ao contrato em comento sequer eram cogitadas pela Administração**.

Posteriormente, essa identificação é corrigida para o Serpro, mas o documento permanece com cláusulas cuja redação é idêntica à do documento anterior, e, apesar de alguns ajustes, permaneceu como mera minuta, sem número do contrato, data ou assinaturas.

A sucessão desses fatos revela deficiência objetiva na revisão, rastreabilidade e comprovação dos próprios mecanismos documentais de proteção do sigilo, questão que se torna especialmente relevante em contratação que contempla multicloud, inteligência artificial, subcontratação e atuação de provedores tecnológicos externos.

Deve-se, assim, exigir a apresentação de **todos os Termos de Compromisso de Manutenção de Sigilo e Termos de Ciência efetivamente assinados no âmbito do Contrato nº 19/2025**, com indicação das respectivas datas, signatários, empregados abrangidos e empresas às quais se aplicam, bem como esclarecimento sobre a origem da minuta anteriormente preenchida com os dados da IBM e sobre os controles adotados para assegurar que terceiros, subcontratados, suboperadores e provedores que eventualmente tenham acesso a informações protegidas estejam submetidos a obrigações de confidencialidade compatíveis com o regime especial de sigilo aplicável ao IBGE.

10. Agregação de várias soluções sob a denominação de uma única solução de TIC

O Termo de Referência afirma que todo o objeto seria uma única solução de Tecnologia da Informação e Comunicação (TIC), apesar de ela reunir vários componentes diferentes, como

nuvem (multicloud), segurança, chatbots, IA, consultoria, conectividade, serviços especiais e novos produtos de dados.

A IN SGD/ME nº 94/2022 veda, como regra, contratar mais de uma solução de TIC em um único contrato. A Procuradoria registrou a questão, mas aceitou formalmente a declaração da área técnica. Cabe à SGD verificar se existe integração funcional suficiente para sustentar essa unidade **ou se a agregação reduziu a precisão de requisitos, riscos, responsabilidades, preços, fiscalização e critérios de aceite.**

11. Governança tardia e risco de atestes sem fiscalização plenamente constituída

O processo documenta cobranças e pedidos de ateste relativos a serviços de 2025 enquanto, em fevereiro de 2026, ainda se discutiam gestor, fiscais, registro e publicidade do contrato.

Tal situação fica evidenciada em e-mail colacionado nos autos administrativos 03603.000033/2025-15 (Arq. 4, fl. 179-181), onde servidora da área técnica de TI se dirige ao Diretor de Tecnologia da Informação do IBGE (*e ex-presidente do Serpro*), Sr. Marcos Vinicius Ferreira Mazoni, adicionando em cópia outros agentes, expressando o seguinte teor:

Enviado: quarta-feira, 4 de fevereiro de 2026 09:21

Prezado Senhor Diretor,

A Diretoria de Tecnologia da Informação (DTI) vem recebendo documentos de cobrança relativos a serviços prestados no âmbito do Contrato nº 19/2025, celebrado em 28/08/2025 entre o IBGE e a Empresa Nacional de Inteligência em Governo Digital e Tecnologia da Informação (SERPRO), cuja formalização foi realizada pelo Senhor Presidente desta Instituição.

Diante disso, **esta Gerência DTI/GPG procedeu à análise do Processo nº 03603.000033/2025-15, ocasião em que foram identificadas as seguintes pendências: ausência de divulgação do contrato no sistema Contratos do Comprasnet, inexistência de publicação no Portal Nacional de Contratações Públicas (PNCP) e no Diário Oficial da União (D.O.U.), bem como a não localização de portaria de designação dos gestores e da equipe de fiscalização do contrato.**

Ressalta-se que o Senhor Presidente do IBGE autorizou a contratação em sua totalidade, conforme registrado às fls. 2006 a 2007 do processo, e atuou como parte na formalização da contratação. À Diretoria de Tecnologia da Informação coube apenas a autorização da contratação de parte um dos itens contratados.

Diante do exposto, considerando que, inicialmente, a DTI teria responsabilidade técnica apenas pelos Grupo 1 (sem SEI e sem e-Commerce), 3 e 4, solicitamos seu pronunciamento quanto à sua responsabilidade na função de Ordenador de Despesa e da DTI na Gestão do referido processo. (*g.n.*)

O e-mail foi respondido pelo diretor, que acaba por confirmar as pendências apontadas, senão vejamos:

Enviado em: sexta-feira, 6 de fevereiro de 2026 10:02

Gostaria de informar que não exerço a função de Ordenador de Despesas e, com isso, não tenho competência para designar os gestores e equipe de fiscalização, porém, como DTI possui a responsabilidade técnica pelos Grupos 1 (sem o sistema SEI e sem o e-Commerce), 3 e 4, indicamos os fiscais técnicos para estes grupos e os relatórios de cobrança recebidos em decorrência desses serviços serão devidamente atestados pelos fiscais técnicos, por nós indicados, e encaminhados a DE para os atestes finais das Notas Fiscais correspondentes. **Como trata-se de serviços demandados e acompanhados diretamente por cada diretoria responsável, incluindo-se a presidência em instância superior a DTI, não cabe a DTI analisar e validar** a pertinência e o correto atendimento destas demandas das referidas áreas.

Quanto às pendências de publicação e divulgação no sistema comprasnet, acredito que devemos solicitar que a GECOS/GCOD conclua essa tarefa, pois o Ato que autoriza a Contratação Direta no PNCP e cadastro do contrato no SDA, foi realizado por essas Gerências.

Obrigado

Marcos Vinicius Ferreira Mazoni

A resposta do titular da Diretoria de Tecnologia da Informação confirma ainda que tal diretoria, que detém a expertise nos pontos nevrálgicos do contrato, não autorizou e não fará o acompanhamento do Grupo 2 do contrato, que é justamente o grupo em que se encontram o SerproBots, os tokens GPT e o WhatsApp.

É preciso esclarecer se a responsabilidade técnica e o acompanhamento da execução do contrato no que concerne ao Grupo 2 ficarão diretamente atrelados à Presidência do IBGE, que autorizou a contratação na totalidade, ou ao Centro de Documentação e Disseminação de Informações (CDDI), que consta no Documento de Formalização da Demanda que inaugura os autos administrativos.

Num ou noutro caso, parecem cabíveis preocupações mais acentuadas, não quanto à aferição da entrega dos produtos, mas sim quanto aos cuidados com a segurança dos dados tratados anteriormente.

Diante de tais inconsistências, devem ser auditados:

- a data efetiva de ativação de cada item;
- quem autorizou cada Ordem de Serviço;
- quem verificou segurança, privacidade e aderência ao TR;
- quais relatórios de medição foram produzidos;
- quem atestou ou deixou de atestar cobranças;
- se houve execução antes da publicação eficaz e da designação formal;
- quais correções foram adotadas e se alcançaram o período pretérito.

12. A reunião técnica de 27/05/2026 torna concretas questões de arquitetura, classificação e governança

O objetivo da reunião, apresentado pelo Serpro, foi avançar no levantamento de requisitos e no desenho de uma solução baseada em integração de registros administrativos, inicialmente tomando o Censo Agropecuário como projeto-piloto. Segundo informado, o Serpro e a AWS estariam trabalhando no **“levantamento dos requisitos e da arquitetura que vai ser proposta para a solução”**.

A reunião é caracterizada como a primeira reunião de alinhamento com a equipe operacional depois de discussões realizadas em nível de diretores e coordenadores.

O objetivo concreto que o Serpro dizia ter recebido era, em síntese, identificar bases administrativas externas que pudessem contribuir para o Censo Agro, mapear informações do questionário às bases existentes e construir uma solução que permitisse integração e uso desses registros. O representante do Serpro afirmou que “o Censo Agro é um piloto de uma necessidade maior, que é o IBGE” e que o projeto seria de criação de uma estrutura de **Data Lake para integrar e gerir dados para vários projetos**, não apenas para aquela edição do Censo Agropecuário. Falou-se inclusive em um Data Lake institucional para o IBGE.

Contudo, a reunião demonstrou que a execução colide com o desenho contratual, diante da participação direta da AWS (Amazon) antes mesmo que houvesse definição clara e prévia sobre dados, arquitetura e governança.

A gravação da reunião técnica de 27 de maio de 2026 não constitui mero elemento contextual. Seu conteúdo revela **divergências relevantes entre o modo pelo qual a solução vinha sendo formalmente concebida e projetada para execução e as salvaguardas, os procedimentos e os limites aplicáveis.**

A primeira das divergências diz respeito à própria escolha e ao papel do provedor de nuvem. A Proposta Comercial estabelece, para o serviço de Cloud Architecture Design, que o planejamento deve resultar em cenários técnicos comparativos entre provedores e dispõe expressamente:

PROPOSTA COMERCIAL Nº 20250140v002:

(incorporada ao contrato por força das cláusulas 1.3 e 1.5)

3. CARACTERÍSTICAS ESPECÍFICAS DO SERVIÇO

3.1 O atendimento às demandas do CLIENTE será feito por projeto.

[...]

3.2.2.3.1 Antes de planejar a arquitetura dos serviços em nuvem, identifica-se junto ao CLIENTE quais as suas necessidades, quer seja migração de ambientes existentes, criação de ambientes para uma nova necessidade de negócio ou mudanças em ambientes na nuvem.

3.2.2.3.3 Será elaborado um relatório exclusivamente baseado em critérios técnicos (desempenho, qualidade, custos, entre outros) considerando os requisitos identificados no item 3.2.2.3.1.

3.2.2.3.2 Uma vez identificadas as necessidades da infraestrutura atual, são elaboradas propostas de solução e orçamento baseadas nos requisitos elencados.

Na reunião, porém, não se apresentava arquitetura neutra construída pelo Serpro para posterior escolha do IBGE. Representante do Serpro afirmou que estava trabalhando **“junto com a AWS nesse**

processo de levantamento dos requisitos e da arquitetura que vai ser proposta para a solução [01:47:44], confirmou que os desenhos então apresentados eram de profissional da AWS e qualificou a empresa como **“parceira no desenvolvimento dessa solução”**.

A própria AWS foi além da condição de simples fornecedora eventual de infraestrutura.

Seu representante afirmou expressamente: **“Aqui a AWS está como consultora. E nós estamos aqui para fazer um projeto para atender vocês”** [02:00:15]. Assim, uma empresa estrangeira, que não é parte signatária do Contrato nº 19/2025, participava diretamente do levantamento de requisitos, do desenho da arquitetura e da definição da solução, antes que estivesse claramente demonstrado, nos documentos examinados, **qual era sua posição jurídica específica naquele projeto — provedora selecionada, consultora, parceira do Serpro, subcontratada ou suboperadora —, quais obrigações de sigilo lhe eram aplicáveis e qual instrumento autorizava e delimitava essa atuação.**

A situação é especialmente relevante porque o Termo de Referência proíbe o compartilhamento de dados pessoais com terceiros fora das hipóteses legais, determina que a Administração seja informada sobre todos os contratos de suboperação e impõe ao Serpro o dever de exigir de suboperadores e subcontratados o cumprimento das obrigações de proteção de dados.

Não basta, portanto, denominar a AWS informalmente de “parceira” ou “consultora”. Se sua participação alcança desenho de fluxos de dados, arquitetura, ferramentas de IA ou eventual processamento, **sua função deve encontrar correspondência em instrumentos contratuais, registros de suboperação, deveres de confidencialidade, responsabilidades e controles auditáveis.**

A contradição torna-se ainda mais sensível quando a discussão passa aos dados. Questionado sobre a possibilidade de utilização da nuvem da Amazon, representante do Serpro respondeu que a arquitetura previa que **“só dados que não são sensíveis possam ser usados na IA, que eles têm disponível lá”**. Logo depois, explicou que se cogitava uma arquitetura híbrida utilizando tanto a nuvem soberana do Serpro quanto aspectos da **cloud pública da AWS**, porque determinadas funcionalidades de inteligência artificial e outras ferramentas não estariam disponíveis no ambiente contratado dentro do Serpro. Ao mesmo tempo, admitiu que ainda era necessário definir o que seria o Data Lake, onde ele ficaria e quais dados reuniria.

Essa forma de discussão **não é compatível com a precisão exigida pelo próprio contrato para uma operação dessa natureza**. A Proposta Multicloud determina que o provedor não faça uso de informações do órgão para **“otimização de mecanismos de inteligência artificial ou qualquer uso secundário não-autorizado”**, garante exclusividade do órgão sobre as informações tratadas e estabelece salvaguardas contra uso, transferência ou liberação indevida de dados e informações. Nesse contexto, dizer apenas que serão utilizados dados “não sensíveis” é insuficiente. A expressão não define se a referência trata de dado pessoal sensível na acepção estrita da LGPD, de dado protegido por sigilo estatístico, de informação submetida a outro sigilo legal, de dado anonimizado ou de dado simplesmente considerado de baixo risco pelos arquitetos da solução.

No âmbito do IBGE, essa ambiguidade é particularmente grave. **Informação protegida pelo sigilo estatístico não se torna livremente utilizável apenas por não constituir “dado pessoal sensível” nos termos da LGPD.** Antes de qualquer utilização em ferramenta externa de IA seria indispensável estabelecer formalmente quais dados podem ser processados, para qual finalidade, em que ambiente, por quais agentes, sob qual regime de retenção, com quais registros de acesso e se haverá ou não utilização para treinamento, ajuste, avaliação, otimização ou melhoria de modelos.

A reunião mostra, ao contrário, que questões dessa magnitude estavam sendo tratadas enquanto a própria arquitetura permanecia indefinida. Discutia-se que algumas funções poderiam rodar na

nuvem pública, outras na nuvem soberana, enquanto ainda se perguntava **“esse data lake vai ser o quê? Vai ser aonde? Vai juntar os dados de quem?”**. Ou seja, a seleção tecnológica avançava tal como a carroça que passa à frente dos bois, pois inexistia ainda definição institucional do conjunto de dados, de sua localização, finalidade e regime jurídico.

O descompasso é bastante ilustrado por um segmento da reunião (por volta de 3h17min) em que, após vários técnicos do IBGE questionarem a ausência de *“ponto de função”* no contrato - jargão usado para quantificar o tamanho funcional de um software a partir da visão do usuário -, com apontamentos sobre a generalidade das soluções, o representante do Serpro, Sr. Marcos Martins Melo, adota um tom bastante ríspido e exaltado, inclusive interrompendo repetidamente os representantes do IBGE na reunião, ocorrendo, no ponto 3h22min05s, o seguinte diálogo:

MARCOS MARTINS MELO – SERPRO:

- “Os meninos da AWS fizeram um puta trabalho de arquitetura em relação à demanda existente. Então, assim, algumas falas desqualificam o trabalho das pessoas. (...) Então, assim, você como membro do IBGE (...), tem que se comportar como membro do IBGE, então, o IBGE fez uma demanda pro SERPRO, e esta demanda teve uma devolutiva neste momento”

TÉCNICA DO IBGE:

- “A gente está tentando apurar quem fez essa demanda, porque, por exemplo, ele é coordenador do Censo Agro, o Vladimir é o Diretor Adjunto e aqui eu estava tentando *[inaudível]* essa demanda.”

MARCOS MARTINS MELO – SERPRO:

- “Foi feita pelo presidente *[do IBGE]* junto com o Mazoni *[DTI do IBGE]*”

Não se busca uma integração meramente abstrata ou de dados agregados. Os servidores do IBGE explicaram que a necessidade discutida envolvia relacionar bases por identificadores individuais, indagando, por exemplo, quais CPFs e quais pessoas do CadÚnico correspondiam a determinados endereços. Em outro momento, foram discutidas combinações entre bases por chaves como CPF e CNPJ e sua territorialização por endereço e setor censitário. Trata-se, portanto, de arquitetura potencialmente capaz de operar sobre conjuntos altamente individualizados de registros administrativos, o que torna ainda menos aceitável que expressões genéricas como “não sensíveis” funcionem como critério suficiente para decidir quais informações podem alcançar ferramentas externas de IA ou ambientes de nuvem pública.

No transcorrer da reunião ficou claro que o escopo debatido extrapolava claramente uma solução pontual para o Censo Agropecuário. Em determinado momento, falou-se em **“lago de dados para o IBGE”**, de caráter institucional e passível de expansão para outros projetos e pesquisas. A própria discussão interna imediatamente questionou qual problema concreto estava sendo resolvido e registrou que inserir mudanças naquele estágio da preparação do Censo Agro representava risco operacional. O representante da AWS descreveu o Data Lake como mecanismo para trazer bases mantidas em outros órgãos, integrá-las no IBGE, convertê-las e permitir a manipulação conjunta de grande volume de informações.

Há ainda indício particularmente preocupante de **adequação da solução ao instrumento contratual, em vez de adequação do instrumento contratual à solução efetivamente necessária**. Quando se

pontuou que o contrato não possuía item de desenvolvimento por *ponto de função*, a alternativa mencionada foi entregar a solução como **SaaS — Software as a Service — “dentro do contrato do Multicloud”**. Na sequência, outro participante ressaltou que se tratava de algo específico para o IBGE, provavelmente dependente de desenvolvimento, e que sequer estava claramente definido o que se pretendia construir. Mais adiante, participante foi ainda mais explícito ao afirmar que **“o contrato não cobre desenvolvimento”** e que determinada aprovação não deveria ter ocorrido justamente por essa razão.

Esse ponto não pode ser reduzido a discussão terminológica entre “desenvolvimento” e “SaaS”. Se uma solução específica precisa ser projetada e desenvolvida para atender necessidade própria do IBGE, **não é juridicamente indiferente requalificá-la como serviço pronto de nuvem apenas para acomodá-la ao instrumento existente.**

A reunião justifica diligência específica para apurar se houve tentativa de fazer caber no Contrato nº 19/2025 atividades que demandariam contratação, ordem de serviço, especificação, análise de riscos ou instrumento jurídico próprio.

A gravação revela, ademais, problemas relevantes de governança interna. Servidores afirmaram que aquela era sua **primeira interação** com o projeto e que não possuíam informações anteriores sobre o que estava sendo apresentado. Em outro trecho, representantes do próprio IBGE solicitaram acesso ao documento que originara a demanda para poder avaliar internamente se existia efetivamente uma necessidade a ser atendida.

A reunião foi marcada, portanto, pelo desalinhamento entre as necessidades técnicas – naquele momento sequer conhecidas pela contratada - e as soluções apresentadas, evidência que, quando apontada, era rebatida pelo contratado com a alegação de que a demanda foi apresentada pela “presidência do IBGE”, e que o técnico que o questionasse não estaria, portanto, falando em nome do IBGE. Ao final da pantomima, foi observado que **o gestor do contrato sequer estava presente.**

Esses fatos tornam inadequado tratar a reunião como simples etapa informal de *brainstorming*. O que se observa é que **Serpro e AWS já haviam desenvolvido trabalho arquitetural, selecionado componentes tecnológicos e apresentado hipóteses de processamento em cloud pública e inteligência artificial enquanto setores técnicos do próprio IBGE ainda discutiam qual era o problema, quem havia formulado a demanda, qual era o escopo, quais dados seriam utilizados e se o contrato sequer comportava o desenvolvimento pretendido.**

O contrato, por sua vez, pressupõe lógica inversa: identificação das necessidades do cliente, elaboração de alternativas técnicas, comparação entre provedores, decisão da arquitetura pelo cliente e validação dos planos de trabalho. A documentação do Multicloud expressamente coloca no cliente a tomada de decisão sobre a arquitetura e prevê avaliação técnica das alternativas, e não a apresentação de arquitetura previamente orientada a determinado provedor.

A reunião de 27 de maio de 2026 deve, portanto, ser considerada **elemento probatório relevante de possível desconformidade na execução contratual**, num contrato, deve-se dizer, já marcado por problemas. A reunião demonstra, ao menos, a necessidade de apurar:

- a participação formal da AWS e o instrumento que a autorizou;
- a razão de uma arquitetura baseada em componentes AWS quando o contrato previa comparação entre provedores e escolha pelo IBGE;
- os critérios empregados para classificar dados como “não sensíveis”;
- quais bases foram apresentadas à AWS durante o levantamento de requisitos;

- se houve acesso a amostras, metadados, esquemas, dicionários, registros ou informações protegidas;
- quais serviços de IA foram considerados e para quais operações;
- se houve POC, teste, upload ou processamento efetivo;
- e se a solução discutida como SaaS foi utilizada para contornar a inexistência de item contratual de desenvolvimento.

O ponto crítico é que a própria reunião documenta uma execução tecnicamente avançada em aspectos que ainda não haviam sido institucional e juridicamente delimitados pelo IBGE. A presença de terceiro não signatário na definição da arquitetura, a orientação concreta para componentes de um provedor específico, a cogitação de cloud pública e IA com base na categoria indeterminada de dados “não sensíveis”, a ampliação do debate para um Data Lake institucional e a tentativa de enquadrar desenvolvimento específico como SaaS revelam riscos objetivos de desvio do desenho contratual, tratamento inadequado de informações protegidas e diluição das responsabilidades de decisão, fiscalização e auditoria.

V – DA QUALIFICAÇÃO PRECISA DOS FATOS

A ASSIBGE-SN entende como adequado traçar uma delimitação dos fatos.

1. Fatos documentados

São comprovados pelo processo:

- a contratação de SerproBots e tokens GPT;
- a previsão de IBM Watson e OpenAI por meio de nuvens externas;
- a possibilidade de transferência internacional e retenção por 30 dias;
- a descrição de aprendizado a partir das interações;
- a vedação do Termo de Referência ao uso para otimização de IA e uso secundário;
- a previsão de Data Lake, integração de bases e API CNEFE-Receita em estágio de concepção;
- a propriedade exclusiva do Serpro sobre tecnologia e modelos;
- a insuficiência temática do mapa de riscos;
- a demora na definição da gestão e fiscalização do contrato.
- a gravação registra participação do Serpro e da AWS no levantamento de requisitos e arquitetura e identifica representante da AWS apresentando componentes;
- foram discutidas arquitetura híbrida, nuvem soberana, cloud pública/multicloud e eventual destinação à IA apenas de dados qualificados na ocasião como “não sensíveis”;
- foram discutidos integração de CPF, Cadastro Único e endereço, disponibilização de dados para joins, hipótese de Data Lake institucional, anonimização e alternativa SaaS no contrato Multicloud;
- a reunião revelou divergências e graves lacunas quanto a escopo, requisitos, aprovação, gestão contratual e cadeia decisória.

2. Inferências tecnicamente justificadas

Esses fatos permitem inferir que:

- existe risco real de incompatibilidade entre o comportamento padrão do SerproBots e as obrigações do Termo de Referência;
- a expressão “*hospedado no Serpro*” não descreve todo o tratamento;
- cláusulas genéricas de LGPD não abrangem, por si sós, o regime especial do sigilo estatístico;
- a propriedade de modelos pode alcançar artefatos derivados se não houver ressalva;
- a arquitetura multicloud exige decisão por *workload*³, e não garantia genérica de soberania;
- o contrato permite expansão significativa por ordem de serviço, proposta técnica de atendimento (PTA), portaria e apostilamento, revelando excesso de permissividade que demanda governança reforçada.
- a participação de provedor externo no desenho torna necessária a individualização dos papéis de consultor, parceiro, fornecedor, subcontratado e eventual suboperador;
- a classificação genérica de dados como “não sensíveis” é insuficiente sem definição jurídica e técnica, critérios, autoridade competente e teste de reidentificação;
- a integração por *joins* e a reutilização institucional ampliam o risco de vinculação, inferência e reidentificação e exigem governança por base e finalidade;

3. Fatos que exigem apuração e medidas preventivas

Embora não restem demonstrados pelas evidências conhecidas, é preciso apurar:

- o possível envio de microdados sigilosos para a IBM, Microsoft, OpenAI, Meta, Huawei ou outro terceiro;
- o uso de dados sigilosos para treinar, ajustar ou avaliar modelo geral;
- se o Data IBGE/Data Lake está em produção;
- se a API CNEFE-Receita foi operacionalizada;
- se os serviços previstos no Anexo 8 foram ofertados a terceiros;
- se a cláusula de propriedade intelectual foi invocada sobre artefato derivado de dado protegido.
- o possível envio de dados protegidos do IBGE para a AWS ou a outra cloud pública em decorrência da arquitetura discutida;
- que qualquer dado do IBGE tenha sido usado pela AWS, pelo Serpro ou por terceiro para treinamento, ajuste ou aperfeiçoamento de modelo;
- que a arquitetura híbrida, o Data Lake institucional, as integrações por *joins* ou a alternativa SaaS tenham sido formalmente aprovados, contratados ou implementados.

Tais pontos precisam ser respondidos com evidências, e não com afirmações vagas como a já adotada pela direção do IBGE, segundo a qual determinado serviço seria seguro por estar “hospedado no Serpro”.

³ Termo que designa uma carga de trabalho computacional, que pode ser um sistema, aplicação, banco de dados, um serviço de IA, dentre outros. Neste caso, sendo serviços que precisam rodar em nuvem, cada serviço precisaria de uma avaliação sobre o tipo de ambiente adequado, se exige uma nuvem soberana ou se pode admitir a flexibilização de uma nuvem pública.

VI. DO REGIME JURÍDICO E NORMATIVO

1. Sigilo estatístico e finalidade exclusiva

O parágrafo único do art. 1º da **Lei nº 5.534/1968** determina que as informações prestadas ao IBGE têm caráter sigiloso, são usadas exclusivamente para fins estatísticos e não podem servir como certidão ou prova, ressalvada a hipótese legal.

O núcleo da proteção é a vinculação de finalidade. Utilizar informação individual em processamento estritamente necessário à produção estatística é juridicamente diferente de incorporá-la a produto, modelo ou ativo reutilizável do prestador. **Nem consentimento individual, nem localização do data center, nem a natureza pública do fornecedor afastam essa limitação legal.**

2. LGPD e papéis funcionais dos agentes

Quando houver dado de pessoa natural identificada ou identificável, incidem também os princípios de finalidade, adequação, necessidade, segurança, prevenção e responsabilização da **Lei nº 13.709/2018 (LGPD)**.

O contrato qualifica o Serpro como operador, mas os papéis são funcionais. Em cada operação é necessário identificar quem determina finalidade e meios essenciais. Se o Serpro ou um terceiro retém interações para melhorar um produto geral, atua para finalidade própria naquele ponto e não pode apoiar-se apenas no rótulo contratual de operador.

O consentimento mencionado no Anexo SerproBots não resolve o tratamento de informação submetida a sigilo estatístico, nem necessariamente constitui a base adequada para serviço público. A transparência ao usuário é necessária, mas não substitui a competência legal do IBGE, a finalidade pública e as restrições especiais de sigilo.

3. STF e compartilhamento de dados pelo Poder Público

No julgamento conjunto da ADPF 695 e da ADI 6.649, concluído em 15 de setembro de 2022, o Supremo Tribunal Federal conferiu interpretação conforme ao regime federal de compartilhamento de dados e assentou que o trânsito de dados pessoais entre órgãos e entidades públicas pressupõe propósitos legítimos, específicos e explícitos, compatibilidade entre o tratamento e a finalidade informada e limitação dos dados ao mínimo necessário, além da observância integral das garantias pertinentes da LGPD.

O Tribunal exigiu, ainda, publicidade sobre a previsão jurídica, finalidade, procedimentos e práticas de compartilhamento e, especificamente quanto à formação de base integradora estatal, estabeleceu a necessidade de justificativa formal, prévia e minuciosa para a inclusão de novos dados e para a escolha das bases que a compõem, acompanhada de mecanismos rigorosos de controle e registro de acessos.

Esses parâmetros devem ser aplicados a cada API, cópia, integração, Data Lake, acesso de suboperador ou uso de IA. O que é especialmente preocupante é que, na reunião técnica de 27 de maio de 2026, tenham sido discutidos um Data Lake institucional, integração de múltiplas bases administrativas, processamento em arquitetura híbrida e eventual utilização de recursos de inteligência artificial enquanto ainda permaneciam em aberto questões elementares acerca de

quais bases integrariam a solução, onde seriam processadas, quais dados seriam considerados aptos a cada ambiente e qual problema institucional concreto se pretendia resolver.

A definição dessas matérias não constitui mera implementação técnica posterior: envolve finalidade, necessidade, minimização, seleção das bases, controle de acesso e governança do tratamento.

Por fim, convém destacar que a ADPF 695 foi ajuizada em junho/2020 para questionar uma conduta do próprio Serpro, consistente no compartilhamento dos dados de CNH de mais de 76 milhões de brasileiros para a ABIN, com base no Decreto 10.046/2019.

4. Decreto nº 10.046/2019 e competência do CCGD

O **Decreto nº 10.046/2019** disciplina a governança do compartilhamento de dados na Administração Pública Federal e atribui ao CCGD competências sobre categorização, diretrizes e integração de dados.

A ASSIBGE-SN não presume que toda hospedagem técnica pelo Serpro seja compartilhamento autônomo. Requer que sejam qualificados, caso a caso, os fluxos que envolvam Receita Federal, CNEFE, registros administrativos, APIs, Data Lake, interessados externos ou outros órgãos. A SGD é legitimada a consultar o CCGD sobre políticas e diretrizes de governança, sem prejuízo das competências de IBGE, ANPD, GSI, CGU e demais órgãos.

5. IN SGD/ME nº 94/2022 e Portaria SGD/MGI nº 5.950/2023

A IN nº 94/2022 exige planejamento, gerenciamento de riscos durante todas as fases e gestão do contrato, além de vedar mais de uma solução de TIC no mesmo contrato, salvo as hipóteses normativas aplicáveis.

A Portaria nº 5.950/2023 estabelece modelo obrigatório de contratação de software e nuvem e exige decisões compatíveis com classificação, segurança e gestão das cargas. O fato de a contratação por dispensa do art. 75, IX, estar dispensada da aprovação prévia de alçada prevista na IN SGD/MGI nº 6/2023 não afasta a competência normativa e supervisora da SGD, nem o dever de cumprimento do modelo de nuvem.

VII. DAS DILIGÊNCIAS E DOS DOCUMENTOS REQUERIDOS

Requer-se que o IBGE e o Serpro apresentem, observadas as restrições de segurança e com juntada sigilosa quando necessário:

1. relação de todas as Ordens de Serviço, Propostas Técnicas de Atendimento (PTAs), apostilamentos, portarias e instrumentos celebrados sob o Contrato nº 19/2025;
2. inventário das workloads e bases previstas, em homologação e em produção, com classificação jurídica e informacional;
3. diagramas lógicos e físicos “as built”, incluindo plano de dados, plano de controle, APIs, egressos, backups, suporte e telemetria;
4. lista completa de provedores, suboperadores, parceiros e regiões de processamento;

5. documentação das configurações do IBM Watson, Azure OpenAI, SerproBots, WhatsApp/Meta e demais componentes;
6. Termos de Uso, Políticas de Privacidade, contratos de suboperação e cláusulas de não treinamento aplicáveis na data de cada ativação;
7. registros de tratamento e evidências de envio, retenção, exclusão e transferência internacional;
8. inventário de prompts, respostas, logs, embeddings, bases vetoriais, datasets de avaliação, adapters, checkpoints, pesos, caches e backups;
9. model registry, histórico de versões e registros de jobs de treinamento, fine-tuning, curadoria automatizada ou avaliação;
10. evidências técnicas de bloqueio de treinamento e uso secundário, inclusive configurações de zero data retention quando disponíveis;
11. avaliação de impacto à proteção de dados e avaliação específica de impacto ao sigilo estatístico, à confiança dos informantes e à reidentificação;
12. novo mapa de riscos que abranja todos os pontos desta representação;
13. instrumentos de classificação das cargas e decisões de alocação em Nuvem de Governo, nuvem pública ou ambiente parceiro;
14. relação de acessos privilegiados, matriz de perfis e logs de administração;
15. plano de saída, portabilidade, continuidade, recuperação, retenção e destruição certificada;
16. interpretação formal da cláusula 20.1 e proposta de ressalva para dados e artefatos derivados do IBGE;
17. instrumentos assinados de sigilo, inclusive os aplicáveis a IBM, Microsoft, OpenAI, Meta, Huawei e demais terceiros;
18. atas e decisões das instâncias do IBGE que avaliaram sigilo, IA, transferências, integração de bases e riscos à confiança;
19. portarias de gestores e fiscais, relatórios de fiscalização, atestes e pagamentos desde a ativação do contrato;
20. comprovação das datas de publicação e registro no PNCP, Contratos.gov, sítio oficial e demais sistemas obrigatórios.
21. lista de participantes, convites, pauta, atas, chats, apresentações, diagramas e demais materiais exibidos ou mencionados na reunião técnica de 27 de maio de 2026;
22. documento de visão, planilhas de bases e requisitos, versões da arquitetura e registros de sua elaboração, revisão, aprovação ou rejeição, com identificação de autores, participantes e instâncias decisórias;
23. instrumentos que disciplinam a atuação da AWS no levantamento de requisitos e arquitetura, inclusive contrato Multicloud, parceria, consultoria, subcontratação, confidencialidade, escopo, custos e responsabilidade por dados, metadados e materiais recebidos;
24. taxonomia e critérios usados para qualificar dados como “sensíveis” ou “não sensíveis”, incluindo relação com sigilo estatístico, LGPD, classificação da informação, regime da base de origem, anonimização, pseudonimização e risco de reidentificação;
25. documentação de integração das bases mencionadas, incluindo CPF, Cadastro Único, endereços e joins, com finalidade, base jurídica, variáveis, chaves, minimização, qualidade, segregação, perfis de acesso e testes de reidentificação;

26. estudos, decisões e instrumentos relativos a eventual Data Lake institucional e à alternativa SaaS no contrato Multicloud, inclusive justificativa contratual, precificação, propriedade intelectual, critérios de aceite, portabilidade e plano de saída;

27. cadeia decisória completa do projeto, com identificação do demandante, patrocinador, responsáveis por requisitos, segurança, privacidade, sigilo estatístico, arquitetura, contratação, gestão, fiscalização, homologação e autorização de produção.

VIII. DAS PERGUNTAS QUE DEVEM SER RESPONDIDAS SEM REMISSÃO GENÉRICA

1. Quais itens do contrato foram efetivamente ativados e em que data?
2. Alguma mensagem, prompt, anexo, log ou telemetria do IBGE foi enviada à IBM, à Microsoft, à OpenAI, à Meta ou a outro terceiro?
3. Algum desses elementos permaneceu fora do Brasil? Por quanto tempo e sob qual garantia de eliminação?
4. O “aprendizado” descrito no item 3.10.1 do Anexo SerproBots ocorreu com interações do IBGE? Qual é a operação técnica exata?
5. Houve pré-treinamento, fine-tuning, ajuste, avaliação, benchmark, geração sintética, curadoria automatizada ou melhoria de produto a partir de informação do IBGE?
6. Que controle torna impossível ou detectável o uso vedado pelo item 4.11 do Termo de Referência?
7. Alguma solução de IA possui acesso direto ou indireto a microdados, CNEFE, registros administrativos ou bases internas?
8. O Data IBGE/Data Lake saiu da fase de concepção? Quais bases o integram e quem autorizou cada uma?
9. A API que combina CNEFE e dados cadastrais da Receita Federal foi implementada, testada ou disponibilizada?
10. Algum produto previsto no Anexo 8 foi oferecido a terceiro? Qual instrumento, base e finalidade?
11. Como a cláusula de propriedade intelectual é aplicada a embeddings, índices, datasets, ajustes e modelos derivados?
12. Quais cargas estão em Nuvem de Governo, Azure, Huawei, IBM Cloud ou outro ambiente?
13. Quem foram os gestores e fiscais desde a primeira ativação, e como foram validados os serviços anteriores às designações localizadas em fevereiro de 2026?
14. Quais recomendações da Procuradoria foram atendidas antes da assinatura e quais foram postergadas para a formalização ou execução?
15. Qual foi o escopo exato da participação da AWS antes, durante e depois da reunião de 27 de maio de 2026, e sob qual instrumento jurídico e contratual ela atuou?
16. A AWS ou outro provedor recebeu amostras, esquemas, metadados, dicionários, diagramas, logs ou qualquer conteúdo derivado de bases do IBGE para elaborar, testar ou validar a arquitetura?

17. O que significava “não sensíveis” na arquitetura discutida, quem estabeleceu essa classificação e como ela se relaciona com sigilo estatístico, LGPD, regime da base de origem e risco de reidentificação?
18. Quais funcionalidades de IA dependeriam da cloud pública/multicloud, quais dados ou artefatos poderiam alcançá-las e quais controles impediriam retenção, treinamento, telemetria ou acesso administrativo indevido?
19. Quais bases seriam integradas por CPF, Cadastro Único, endereço ou outras chaves, quais joins seriam realizados e para quais finalidades estatísticas específicas?
20. Que técnica de anonimização ou desidentificação foi considerada, em que estágio do fluxo seria aplicada e que testes demonstrariam a ausência de reidentificação após os cruzamentos?
21. O Data Lake discutido teria vocação institucional? Quais projetos poderiam reutilizá-lo e que ato autorizaria cada nova finalidade e cada nova base?
22. A alternativa SaaS foi formalmente analisada no âmbito do Contrato nº 19/2025? Que item, Ordem de Serviço, PTA ou outro instrumento a suportaria, e como seriam definidos desenvolvimento, preço, propriedade, aceite e saída?
23. Quem demandou, aprovou e patrocinou o projeto; quem aprovou o documento de visão e a arquitetura; e quem poderia autorizar testes, homologação e produção?
24. Quais decisões da reunião permaneceram exploratórias e quais produziram encaminhamentos, documentos, contratações, configurações, testes ou execução posterior?

IX. CONCLUSÃO

Embora não nos pareça que microdados sigilosos já tenham alimentado modelo geral, vemos como absolutamente sério o presente contrato, que nos parece abrir tal possibilidade de risco sem demonstrar, de forma suficiente, como ela seria tecnicamente fechada.

O Termo de Referência proíbe otimização de IA e uso secundário. O Anexo SerproBots descreve aprendizado, retenção e transferência internacional. O contrato protege as informações, mas atribui ao Serpro propriedade exclusiva sobre modelos. O planejamento invoca nuvem soberana, mas o objeto é multicloud e envolve parceiros. O Anexo 8 limita produtos a dados públicos, mas descreve dados brutos, Data Lake, CNEFE, Receita Federal e ampliação por atos posteriores. O mapa de riscos não enfrenta esses pontos. E a execução começou antes de a governança contratual estar claramente consolidada.

A reunião de 27 de maio de 2026 reforça a necessidade dessa apuração: registra participação direta da AWS no levantamento de requisitos e na arquitetura, discussão de cloud pública para IA, classificação ainda indefinida de dados como “não sensíveis”, integração de bases por joins, possível Data Lake institucional, anonimização, alternativa SaaS e cadeia decisória controversa. Embora não demonstre que dados protegidos tenham sido enviados, evidencia que essas alternativas foram concretamente discutidas e devem ser esclarecidas por documentos e controles auditáveis.

Esse conjunto não deve ser tratado como detalhe burocrático. O sigilo estatístico é condição de existência da estatística pública. A confiança dos informantes não pode depender de fórmulas genéricas, consciência difusa da alta administração ou presunção de segurança fundada apenas na natureza pública do fornecedor.

A ASSIBGE-SN requer, por isso, apuração efetiva, correção das incompatibilidades, demonstração dos controles e definição institucional inequívoca de que informação protegida do IBGE e seus derivados não podem alimentar modelo geral, produto reutilizável ou finalidade estranha à produção estatística legalmente autorizada.

X. DAS MEDIDAS PREVENTIVAS

A ASSIBGE-SN **reivindica** que, até que se conclua a instrução, sejam tomadas as seguintes providências cautelares:

1. proibição formal de enviar a motores externos de IA informação protegida por sigilo estatístico, dado de acesso restrito ou registro administrativo sigiloso;
2. suspensão de qualquer treinamento, ajuste, avaliação persistente ou aprendizado com conteúdo do IBGE, salvo operação dedicada, previamente autorizada e demonstradamente necessária à finalidade estatística;
3. preservação segura dos logs, configurações, versões, ordens de serviço, registros de acesso e demais evidências;
4. revisão imediata de acessos privilegiados, suboperadores, egressos, telemetria e retenção;
5. bloqueio de reutilização entre clientes de dados ou artefatos derivados do IBGE;
6. exigência de autorização específica e parecer de segurança, privacidade e sigilo para cada workload de alto risco;
7. impedimento de ativar Data Lake, API CNEFE-Receita ou novo produto de dados apenas com fundamento genérico no Anexo 8;
8. saneamento da gestão contratual, com reconstrução da fiscalização do período anterior às designações;
9. aditamento ou instrumento equivalente que harmonize TR, contrato e proposta e estabeleça regra inequívoca de *no training / no secondary use*;
10. comunicação à ASSIBGE-SN, em nível compatível com a segurança, das medidas adotadas.
11. não disponibilização à AWS ou a outro provedor externo de dados, amostras, metadados ou artefatos protegidos enquanto não estiverem formalmente definidos papéis, finalidade, classificação, ambiente, retenção, acesso, suboperação e controles de não treinamento;
12. vedação de tratar como “*não sensível*” qualquer dado apenas por não se enquadrar no art. 5º, II, da LGPD, exigindo avaliação cumulativa de sigilo estatístico, regime de origem, identificabilidade, inferência e reidentificação;
13. não ativação de arquitetura híbrida, Data Lake institucional, integração por *joins* ou alternativa SaaS sem ato específico da autoridade competente, participação das áreas técnicas responsáveis, avaliação de impacto e instrumento contratual adequado.

XI. DOS PEDIDOS À SGD E AO CCGD

Diante do exposto, a ASSIBGE Sindicato Nacional **REQUER**:

1. **autuação da presente representação**, a fim de que sejam tomadas as diligências cabíveis no âmbito do Sistema de Administração dos Recursos de Tecnologia da Informação;
2. **remessa da presente representação ao Comitê Central de Governança de Dados – CCGD**, a fim de que analise as matérias de governança, compartilhamento de dados, e outras que lhes caibam;
3. **requisição dos documentos e respostas** indicados nas Seções VII e VIII;
4. **auditoria técnica** da aderência do Contrato nº 19/2025 à IN SGD/ME nº 94/2022 e à Portaria SGD/MGI nº 5.950/2023;
5. **verificação material, e não apenas formal**, do Estudo Técnico Preliminar - ETP, do mapa de riscos, do Termo de Referência e da gestão do contrato;
6. **manifestação sobre a compatibilidade** entre a vedação do item 4.11 do Termo de Referência e os itens 3.9 e 3.10 do Anexo SerproBots;
7. determinação de que **nenhuma informação protegida seja processada** em fluxo cuja retenção, aprendizado, suboperação ou transferência não esteja integralmente esclarecida e autorizada;
8. análise da cláusula 20.1, com exclusão expressa de dados e artefatos derivados protegidos da propriedade exclusiva do prestador;
9. qualificação, com consulta ao CCGD quando cabível, dos fluxos relativos a Data Lake, CNEFE, Receita Federal, APIs e oferta de produtos a terceiros;
10. **avaliação da necessidade de encaminhamento à ANPD, ao GSI, à CGU, ao TCU e às demais autoridades competentes**;
11. análise da **tempestividade da designação dos gestores e fiscais**, dos atestes, da publicidade e da eficácia contratual;
12. adoção das **medidas preventivas** da Seção IX da presente manifestação;
13. **participação da ASSIBGE-SN e de trabalhadores das áreas técnicas** na instrução, preservados os sigilos legais;
14. **resposta conclusiva, motivada e acompanhada das evidências** passíveis de divulgação.
15. **auditoria específica dos fatos e encaminhamentos da reunião de 27 de maio de 2026**, com confronto entre gravação, degravação, documentos apresentados, instrumentos contratuais e execução posterior;
16. **manifestação sobre o conceito e os critérios de “sensível” e “não sensível”** aplicáveis ao IBGE, distinguindo dados pessoais sensíveis, sigilo estatístico, sigilo da base de origem, acesso restrito, anonimização e risco de reidentificação;
17. determinação de **que a participação de AWS ou de qualquer provedor externo seja formalmente qualificada e submetida aos controles aplicáveis** antes de eventual acesso a dados, metadados, artefatos ou ambientes do IBGE.

Dos elementos probatórios:

A ASSIBGE apresenta como anexo os autos do Processo Administrativo IBGE nº 03603.000033/2025-15, divididos em quatro arquivos PDF.

Quanto à gravação **da reunião**, considerando que se trata de arquivo MP4 com 1,36 GB, e, portanto, inviável de ser remetido como anexo comum pelo sistema de protocolo, a ASSIBGE-SN pede que seja indicada, no trâmite desta representação, ou através do e-mail assibge-sn@uol.com.br, a forma mais adequada de envio do arquivo, por meio digital ou mediante entrega física.

Termos em que,
pede deferimento.

Rio de Janeiro, 26 de agosto de 2026.

Executiva Nacional da ASSIBGE – Sindicato Nacional