



## GOVERNO DO ESTADO DE MINAS GERAIS

### Ata

Comitê Gestor de Governança de Dados (CGGD)	Reunião realizada em 25 de novembro de 2025
Evento	3ª Reunião Ordinária/Extraordinária do Comitê Gestor de Governança de Dados (CGGD)
Convocada por	Secretaria Executiva do CGGD
Registrador	Ausier Vinícius de Oliveira Santos
Local e hora	Reunião realizada presencialmente na sala 06 do 11º andar do Prédio Gerais da Cidade Administrativa das 10:00 às 11:35

Participantes	Órgão	T/S
Rodrigo Diniz Lara	Seplag	T
Fabício de Barros Salum	Seplag	S
Roberto Tostes Reis	Prodemge	T
Clara Pinheiro Oliveira Costa	Secretaria-Geral	T
Cynthia Martins Vieira	CGE	T
Raphaela Hytomi Pianchão Aihara	SEGOV	S
Mariana Villela Nunes	Seplag	Convidado
Felipe da Cruz Rocha	Prodemge	Convidado
Alexandre Dieguez Mesquita	Prodemge	Convidado

### 1. PAUTA DA REUNIÃO

Apresentação demandas Secretaria Geral  
Apresentação Data Lake MG  
Minuta de dados de referência

### 2. RESUMO DA REUNIÃO

No dia 25 de novembro de 2025, foi realizada a 3ª Reunião Ordinária do Comitê Central de Governança de Dados - CGGD, presencialmente na sala 06 do 11º andar do Prédio Gerais da Cidade Administrativa. Houve participação de convidados, os quais foram registrados. Rodrigo Diniz (Seplag) iniciou a reunião às 10:05 informando a pauta da reunião (apresentações da Secretaria-Geral, Prodemge e minuta de deliberação de dados de referência).

Clara Costa (Secretaria-Geral), Superintendente de Assessoramento Temático, iniciou sua apresentação destacando o papel de articulação da Secretaria-Geral entre órgãos e entidades do governo, atuando como um filtro de informações para governador e vice-governador para obtenção de informações estratégicas de diversos assuntos. Foram explicadas as principais frentes de trabalho, que incluem a triagem de e-mails e comunicações, o apoio técnico para agendas e o monitoramento de frentes estratégicas (distintas dos

projetos estratégicos gerenciados pela Seplag), bem como os principais tipos de documentos produzidos (*briefings*, *papers* e estudos técnicos) e suas características, voltados para subsidiar interações com atores externos. Quanto aos *briefings* (elaborado por equipes técnicas) e *papers* (elaborados pela comunicação), citou que há demandas mais comuns relacionadas ao desenvolvimento econômico, que há uma padronização e as informações são resumidas para o público-alvo, entretanto reforçou a variedade de temas tratados. Já os estudos técnicos apresentam formato mais diversos, geralmente voltados para avaliação de normativos, benchmarking com outros estados. Foram exemplificados casos recentes e diversos de agendas/documentos realizados. A produção destes documentos é realizada de forma manual, sendo dependente da coleta de informações junto aos órgãos, e enfrenta dificuldades relacionadas ao tempo de resposta necessário para atendimento à demanda (dependência de respostas em tempo hábil pelos órgãos), contextualização para dimensão real do problema (clareza sobre ordem de grandeza), ausência de histórico de informações e coerência dos dados recebidos em lapsos temporais distintos (envio dos dados atrelado mais às pessoas que respondem às demandas do que aos processos de obtenção dos dados). Entre os impactos, têm-se que as decisões no contexto político não irão aguardar uma eventual necessidade de qualificação do dado recebido.

Para minimizar os impactos supracitados, fizeram um exercício de identificar os principais dados demandados e os organizaram em quatro eixos: 1. Cobertura e qualidade dos serviços/atendimentos (equipamentos, municípios, cidadãos atendidos e indicadores de qualidade); 2. Pessoal (servidores, concursos realizados e benefícios pagos); 3. Orçamento (receitas e orçamento previstos versus executado); 4. Dados econômicos (investimentos e empregos gerados). Clara Costa (Secretaria-Geral) destacou que as áreas mais demandadas envolvem saúde, educação, infraestrutura e desenvolvimento econômico, relatou dificuldade de acesso a algumas informações para acesso à completude dos dados e que, na maturidade atual, não se justifica a contratação de uma pessoa com formação específica em análise de dados na equipe. Quanto aos dados de pessoal, Cynthia Vieira (CGE) exemplificou que BO não mostra o histórico de pagamento de férias-prêmio, enquanto Rodrigo Diniz (Seplag) citou a informação existe e não estaria organizada e que a agência Invest Minas poderia ser abordada para contribuir no painel de dados econômicos.

Clara Costa (Secretaria-Geral) informou que, para atender ao encaminhamento da 2ª Reunião Ordinária do CGGD, procederam com análises dos decretos de competências dos órgãos, PPAG e mapas estratégicos (quando disponíveis) para montar a lista de dados a serem levantados apresentada ao Comitê. Ressaltou que seria uma lista inicial que poderia ser incrementada e que geralmente as informações recebidas são focadas em recursos, cujo detalhamento maior depende dos órgãos. Rodrigo Diniz (Seplag) sugeriu que a lista de dados seja analisada para realizar a identificação de sistemas corporativos e transacionais que possam fornecer os dados, citando como exemplo que dado de emprego tem que ser obtido do Caged, e identificar quais já estão disponíveis no Data Lake MG. Roberto Reis (Prodemge) sugeriu proceder com esta análise e realizar o acionamento dos órgãos para questioná-los, enquanto Cynthia Vieira (CGE) questionou qual seria o modo de recebimento destes dados, dados que há painéis para alguns assuntos demonstrados. Clara Costa (Secretaria-Geral) ressaltou que costuma solicitar acesso aos painéis, entretanto, há casos em que a interpretação de alguns casos demanda conhecimento técnico da área. Cynthia Vieira (CGE) já sugeriu vincular os dados deste levantamento com os dados de referência da minuta.

Roberto Reis (Prodemge) iniciou a apresentação da Prodemge destacando a nova estratégia da companhia, que focará na coordenação técnica e atuação em parcerias com o mercado, em oposição à visão antiga de tentar atuar para resolver todos os problemas, o que seria impossível. Reforçou que se trata de uma mudança a ser realizada gradualmente e a atuação da Prodemge sobre os dados é frequentemente questionada pelo Vice-Governador, logo, há um desejo que a companhia assuma um papel relevante no CGGD para atender à demanda por políticas públicas baseadas em dados. Julga importante que o Comitê defina e alinhe premissas, modelo de arquitetura, possibilitando que órgãos mantenham sua autonomia em seus projetos de *data lake* e inteligência artificial. Sugeriu apresentar propostas para regulamentar o compartilhamento de dados entre órgãos, que apesar do Data Lake MG estar hospedado *on premise*, não há impedimento para integração de soluções em nuvem, mas há uma demanda de catalogação de dados e criação de camadas de transação para não replicar dados.

Alexandre Dieguez (Prodemge) e Felipe Rocha (Prodemge) conduziram a apresentação sobre o Data Lake MG. Destacaram a essencialidade da gestão de dados para uso de IA e os desafios enfrentados (dados fragmentados, formação de silos de dados, segurança e LGDP).

Foi apresentada a arquitetura do Data Lake MG, baseada na metodologia *medallion*, que utiliza três camadas para tratar os dados:

- Bronze/Staging: Dados brutos, como vêm dos sistemas de origem.
- Silver/Master: Camada intermediária para validação, enriquecimento e padronização.
- Gold/Reporting: Camada final com dados modelados e prontos para consumo (BI, IA, relatórios).

A arquitetura centraliza os dados na primeira camada para preservar os sistemas de origem, garantir segurança e reduzir riscos (extrair e carregar). Na camada intermediária ocorre validação, enriquecimento e padronização dos dados (transformação). Assim, os dados estão “otimizados para a escrita”. Na camada final, que entrega valor ao órgão por já apresentar as regras de negócio, voltada para governança e disponibilização, os dados estão prontos para consumo, de forma genérica ou específica, de IA, *machine learning* e painéis em formas diferentes de tratamento (“otimizado para leitura”). Também foi apresentada a arquitetura de compartilhamento, sugerindo que a camada de bronze seja armazenada no Data Lake MG, com uma cama de compartilhamento por virtualização para evitar a replicação de dados, manter o lastro dos dados e criar um ponto único de acesso. O Data Lake MG já contém dados de áreas como segurança pública, planejamento, trânsito, educação e previdência. Foi apresentado um case da Segurança Pública, onde uma metodologia de acompanhamento de egressos do sistema prisional, desenvolvida para Polícia Militar com dados do Data Lake, resultou em uma queda na reincidência criminal.

Cynthia Vieira (CGE) pontuou que há dois níveis de usuários, os que precisam dos dados como informação na última camada (analista de negócio) e usuários que precisam de dados brutos (usuário analítico), como a CGE, por exemplo. Alexandre Dieguez (Prodemge) citou que solução está apta para atender aos dois públicos.

A governança de dados é central, definindo quem pode e deve acessar quais dados. A partir da definição de um padrão, a gestão de acesso será de responsabilidade do órgão gestor dos dados. Foi apresentada a ferramenta "Ranger", utilizada para criar e gerenciar políticas de acesso, permitindo mascarar ou ocultar dados com base no perfil do usuário, que chega em nível de coluna e linha de determinada base de dados. O sistema possui uma funcionalidade de auditoria que registra todas as consultas realizadas, incluindo acessos negados. Integrada na base do SSC, a ferramenta evita liberações por CPF, sendo explicado o modelo da Segurança Pública, em que houve criação de nível tático, operacional e estratégico e de inteligência na segurança pública, sugerindo assim a replicação deste modelo em todos os órgãos. Destacou-se a importância destas medidas para segurança dos dados. Cynthia Vieira (CGE) pontuou a dificuldade de definição destas regras a partir da experiência na CGE.

Roberto Reis (Prodemge) suscitou uma discussão sobre a governança de custos em ambientes de nuvens híbridas, especialmente, sobre quem arcará com os custos quando um órgão consumir dados de outro que estão em uma plataforma paga (ex: Databricks). O consumo na nuvem é tarifado por consulta e pode gerar custos elevados. Cynthia Vieira (CGE) informou que há órgãos que negam o acesso a dados justificados sobre este tipo de consumo. Alexandre Dieguez (Prodemge) citou que na BISP ocorre a transmissão de 23 terabytes por semana e que, caso ocorresse em plataforma de nuvem representariam um custo relevante. Rodrigo Diniz (Seplag) sugeriu que nesta hipótese, o custo seja arcado sempre pelo órgão de origem dos dados, enquanto Roberto Reis (Prodemge) sugeriu que os órgãos consumam os dados diretamente do Data Lake MG (camada bronze), evitando a tarifação por consulta da nuvem e a movimentação de dados entre diferentes provedores, que gera custos adicionais.

Rodrigo Diniz (Seplag) sugeriu como encaminhamentos a análise da planilha dos dados apresentados pela Secretaria-Geral, identificar fontes de informação existentes e o que já está no Data Lake, consolidar e verificar meios de recebimento destes dados pela Secretaria Geral.

Rodrigo Diniz (Seplag) lembrou o exemplo apresentado pela SEE em reunião anterior em que a consulta de dados pela alta gestão ocorria por interface com LLM no whatsapp, sugerindo que sirva de inspiração para promover uma entrega palpável do CGGD ao governador. Roberto Reis (Prodemge) citou que fizeram prova de conceito desta natureza na Fundação João Pinheiro da FJP. Foi proposto focar em um piloto com dados de educação para demonstrar valor rapidamente, consolidando os dados sinalizados na apresentação para criar algo palpável para a Secretaria Geral.

Sobre a minuta de dados de referência, a apresentação foi retirada da pauta para manter o horário de encerramento da reunião, que ocorreu às 11:30.

### 3. ENCAMINHAMENTOS

Encaminhamentos	Responsável	Prazo
Análise da planilha dos dados apresentados pela Secretaria-Geral	Secretaria Executiva CGGD	Próxima reunião
Encaminhamento Minuta de dados de referência aos membros do CGGD por meio eletrônico	Secretaria Executiva CGGD	Próxima reunião
Organizar um piloto com a Secretaria Geral usando os dados disponíveis de Educação no Data Lake MG para consulta via LLM	Prodemge	Não definido



Documento assinado eletronicamente por **Raphaela Hytomi Pianchao Aihara**, **Assessor(a)**, em 02/12/2025, às 14:52, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 47.222, de 26 de julho de 2017](#).



Documento assinado eletronicamente por **Cynthia Martins Vieira**, **Coordenador(a)**, em 02/12/2025, às 22:48, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 47.222, de 26 de julho de 2017](#).



Documento assinado eletronicamente por **Clara Pinheiro Oliveira Costa**, **Superintendente**, em 11/12/2025, às 13:47, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 47.222, de 26 de julho de 2017](#).



Documento assinado eletronicamente por **Rodrigo Diniz Lara**, **Subsecretário**, em 12/12/2025, às 09:14, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 47.222, de 26 de julho de 2017](#).



A autenticidade deste documento pode ser conferida no site [http://sei.mg.gov.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](http://sei.mg.gov.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **128035118** e o código CRC **120ED75D**.

Referência: Processo nº 1500.01.0130573/2025-28

SEI nº 128035118