

# Plataforma de Integração de Sistemas Estruturantes da Execução Financeira Federal

Isaque Alves<sup>1</sup>, João Egewarth<sup>2</sup>, Laila Bellix<sup>2</sup>, Alex Reis<sup>2</sup>, Carla Rocha<sup>2</sup>

<sup>1</sup>Instituto de Matemática e Estatística – Universidade de São Paulo (IME/USP)  
São Paulo – SP – Brasil

<sup>2</sup>Faculdade de Ciências e Tecnologias em Engenharia – Universidade de Brasília (FCTE/UnB)  
Brasília – DF – Brasil

<sup>3</sup>Instituto Cidade Democrática  
São Paulo – SP – Brasil

isaque.alves@ime.usp.br, {egewarth, llbellix}@gmail.com

{alexreis, c.aguiar}@unb.br

**Abstract.** *This paper presents an experience report on implementing an open-source platform for integrating data from Structuring Systems involved in federal budget execution (SIOP, SIAFI, SIAPE, TransfereGov, and ContratoGov). Given the challenges of promoting interoperability among systems managed by different agencies, a data-driven approach based on the Medallion architecture was adopted. The solution enables automated report generation, decentralized data governance, and the identification of requirements for future interoperability. As a result, it offers a scalable and reusable alternative with centralized data, documented APIs, and interactive dashboards, along with lessons learned on process improvement and enhanced public management.*

**Resumo.** *Este artigo apresenta o relato de experiência na implementação de uma plataforma open source para integrar dados de diferentes Sistemas Estruturantes da execução orçamentária federal (SIOP, SIAFI, SIAPE, TransfereGov e ContratoGov). Diante da dificuldade de promover interoperabilidade entre sistemas geridos por diferentes órgãos, adotou-se uma abordagem orientada por dados, baseada na arquitetura Medallion. A solução permite a geração automatizada de relatórios, governança descentralizada e mapeamento de requisitos para futura interoperabilidade. Como resultado, oferece uma alternativa escalável e reutilizável, com dados centralizados, APIs documentadas e dashboards interativos, além de lições aprendidas sobre a qualificação de processos e a melhoria da gestão pública.*

## 1. Introdução

A evolução da administração pública para um modelo mais eficiente, transparente e acessível tem sido impulsionada pela transformação digital, um processo que vai além da simples informatização de serviços. Nesse contexto, o Governo Digital representa um estágio mais avançado dessa transformação, em que as tecnologias digitais são elementos centrais para a criação de valor público [Gil-Garcia *et al.* 2018].

A adoção do Governo Digital ocorre por meio de estágios evolutivos, nos quais a administração pública incorpora progressivamente tecnologias para aprimorar a prestação de serviços, a eficiência operacional e a participação cidadã [Gil-Garcia *et al.* 2018, Dunleavy *et al.* 2006]. Diferente do conceito de governo eletrônico, que se concentra na digitalização de processos e na automação de serviços, o Governo Digital implica uma nova lógica de atuação, fundamentada na transparência, na interoperabilidade e no uso estratégico de dados. Instituições como a OCDE (Organização para a Cooperação e Desenvolvimento Econômico) <sup>1</sup> e a ONU (Organização das Nações Unidas) <sup>2</sup> destacam o Governo Digital como um avanço em relação ao governo eletrônico, promovendo uma administração mais aberta, interativa e orientada ao cidadão.

O Governo Digital se sustenta em quatro pilares principais [Viana 2021, Gil-Garcia *et al.* 2018, da Silva Cristóvam *et al.* 2020]: (1) cocriação e participação cidadã, na qual os cidadãos deixam de ser apenas usuários de serviços para se tornarem agentes ativos no processo decisório; (2) adoção de novas tecnologias, como *Big Data* e Inteligência Artificial, permitindo modelos inovadores de governança; (3) integração e interoperabilidade, fundamentais para conectar sistemas de diferentes órgãos e proporcionar maior eficiência administrativa; e (4) abordagem centrada no cidadão, que prioriza a experiência do usuário na concepção dos serviços digitais, garantindo acessibilidade e usabilidade.

Este artigo foca nos pilares (2) e (3) e nos desafios enfrentados na transição para o Governo Digital, especialmente no que diz respeito à interoperabilidade entre Sistemas Estruturantes (SE) <sup>3</sup> [Viana 2021, Gil-Garcia *et al.* 2018, da Silva Cristóvam *et al.* 2020]. A fragmentação desses sistemas ainda constitui um obstáculo à modernização do setor público, dificultando a automação de processos, e a gestão baseada em dados qualificados. Assim, o desenvolvimento de soluções inovadoras para a interoperabilidade é essencial para consolidar um modelo de Governo Digital que seja eficiente, transparente e orientado ao cidadão [Viana 2021].

Uma alternativa viável e de impacto imediato tem sido promover a integração entre esses sistemas a partir dos dados. Ao estruturar, qualificar e cruzar informações extraídas de diferentes fontes, é possível criar mecanismos de gestão orientada a dados, possível diagnosticar falhas em processos e estabelecer requisitos de interoperabilidade, mesmo em um cenário com sistemas legados e APIs restritas [Nadal *et al.* 2019]. Essa abordagem permite ganhos concretos em governança, rastreabilidade, reprodutibilidade e uso estratégico da informação. Um exemplo é o Programa Capacidades Analíticas (PCA), uma iniciativa que aplica ciência de dados sobre a base pública do SICOM para estruturar e qualificar informações contábeis e financeiras dos municípios mineiros, possibilitando a criação de ferramentas para detecção de irregularidades e apoio à fiscalização [Costa *et al.* 2024].

Diante desse cenário, este artigo apresenta um relato de experiência sobre o desenvolvimento do **GovHub**, uma solução *open source* voltada à integração de dados dos

---

<sup>1</sup>Disponível em: <https://www.oecd.org>, Acesso em: 7 maio 2025.

<sup>2</sup>Disponível em: <https://www.un.org> Acesso em: 7 maio 2025.

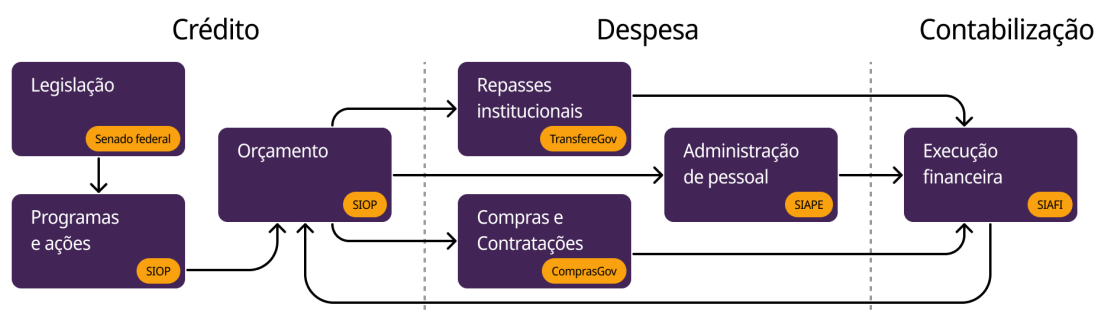
<sup>3</sup>Sistemas com suporte de tecnologia da informação responsáveis pelo planejamento, coordenação, execução, descentralização, delegação de competência, controle ou auditoria das ações do Estado, além de outras atividades auxiliares. (Norma Complementar nº 19/IN01/DSIC/GSIPR)

SE responsáveis pela execução orçamentária do Governo Federal. Os sistemas integrados incluem: SIOP, SIAFI, TransfereGov, ComprasGov e SIAPE. O objetivo é demonstrar como uma abordagem orientada por dados pode aprimorar a gestão pública, tornando-a mais estratégica e fundamentada em evidências. Como principal resultado, propomos uma solução baseada na *Medallion Architecture* [Russom *et al.* 2017], que oferece dados qualificados e integrados desses sistemas, disponibilizados para relatórios, *dashboards* e iniciativas de governança de dados descentralizada. Por fim, discutimos as lições aprendidas com a integração dessas bases, evidenciando seu valor imediato para uma gestão orientada a dados e seu potencial como etapa intermediária rumo à interoperabilidade, ao mapear tanto os desafios processuais quanto os requisitos técnicos de integração entre os sistemas.

Este relato está organizado da seguinte forma: a Seção 2 apresenta o contexto teórico e os principais desafios relacionados à interoperabilidade de sistemas. A Seção 3 detalha o processo de diagnóstico e desenvolvimento. Por fim, a Seção 4 discute as lições aprendidas durante o projeto.

## 2. Problemas de Interoperabilidade nos Sistemas Estruturantes

Interoperabilidade pode ser definida como a capacidade de diferentes sistemas trocarem informações de forma eficiente, segura e compreensível, independentemente de suas tecnologias ou arquiteturas<sup>4</sup>. No contexto da administração pública, essa troca de informações é essencial para assegurar o fluxo contínuo de dados entre plataformas governamentais, viabilizando a automação de processos, a rastreabilidade e a governança de dados. Entre essas plataformas, destacam-se os SEs: Sistema Integrado de Planejamento e Orçamento (SIOP), o Sistema Integrado de Administração Financeira (SIAFI), o ComprasGov e o TransfereGov.



**Figura 1. Fluxo de informações entre os sistemas orçamentários e financeiros, desde a definição do orçamento até sua execução.**

Como ilustrado na Figura 1, no SIOP é feito o planejamento orçamentário do Governo Federal, sendo a plataforma onde são definidos o Plano Plurianual (PPA) e a Lei Orçamentária Anual (LOA). Após a aprovação da LOA pelo Senado Federal, o orçamento é liberado para execução pelos órgãos governamentais. As despesas, por sua vez, são geradas em sistemas distintos, como o ComprasGov, utilizado para processos de compras e contratações, e o TransfereGov, voltado à gestão de transferências institucionais [Ministério do Planejamento 2010]. Esses sistemas registram e processam

<sup>4</sup> Adaptado da definição da ISO/IEC 2382:2015

informações financeiras, que são posteriormente repassadas ao SIAFI, responsável pela contabilização e controle da execução orçamentária e financeira, bem como pelo registro da movimentação dos recursos públicos.

Contexto	Sistema	Problema
TEDs	TransfereGov	Não possui registro de NEs (notas de empenho) para acompanhamento da execução orçamentária e financeira.
TEDs	TransfereGov	Não possui registro do número de transferência (identificador do TED no SIAFI).
TEDs	TransfereGov	TEDs firmados antes de 2022 não estão registrados no sistema.
TEDs	TransfereGov	Valores de NCs (notas de crédito) e PFs (programação financeira) não são registrados no sistema, apenas no SIAFI, sendo necessário o cruzamento entre as bases para acompanhar os valores.
TEDs	SIAFI	NEs não possuem campo estruturado para registro do número de transferência, NCs ou PFs correspondentes, impedindo o acompanhamento da execução orçamentária e financeira dos TEDs.
TEDs	SIAFI	Registro do número de transferência, NCs e/ou PFs é realizado no campo de descrição e/ou observação sem nenhuma padronização.
Contratos	ComprasGov	Não possui campo para registro do valor planejado para o contrato no ano, sendo necessária a utilização de planilhas internas do órgão para realização do planejamento.
Contratos	ComprasGov	Campos necessários para o acompanhamento interno como a unidade requisitante, data de proposta comercial e de publicação não são obrigatórios.
Contratos	ComprasGov	Campo para registro de parcelas permite a inserção tanto da quantidade de parcelas anuais, quanto a quantidade de parcelas totais do contrato, gerando despachonização dos dados.
Contratos	ComprasGov	NEs registradas diretamente no SIAFI não são vinculadas automaticamente aos contratos, sendo necessária uma inserção manual.
Contratos	ComprasGov	Ausência de regras para impedir o registro de NEs em mais de um contrato.
Contratos	ComprasGov	A integração existente com o SIAFI não consegue identificar a finalização de um contrato mesmo após o pagamento completo do valor contratado.
Contratos	SIAFI	NEs não possuem campo estruturado para registro do identificador do contrato e/ou da fatura correspondente.
Contratos	SIAFI	Registro do número do contrato e/ou número e mês da fatura é realizado no campo de descrição e/ou observação sem nenhuma padronização.

**Tabela 1. Problemas de interoperabilidade nos TEDs e contratos administrativos.**

A interoperabilidade garante rastreabilidade e transparência na execução orçamentária, mas ainda é uma das principais barreiras para a consolidação de um Governo Digital. Os SEs, em sua maioria legados, foram desenvolvidos em contextos distintos e com tecnologias que não previam integração entre plataformas. Como resultado, suas bases são distribuídas e não compartilham metadados, dificultando consultas integradas e o processamento das informações. A Tabela 1 apresenta os principais problemas relacionados à interoperabilidade e à estruturação dos dados. Essa fragmentação impede que os órgãos públicos disponham de informações qualificadas para subsidiar decisões estratégicas, afetando diretamente a qualidade da gestão, por exemplo, o planejamento financeiro e a análise consolidada dos gastos públicos.

### 3. Centralizando Dados Estruturantes do Governo Federal

Alinhado com o pilar (1) Cocriação e participação, do Governo Digital, o Laboratório de Competências em Software Livre (Lab Livre) da Universidade de Brasília (UnB), em parceria com o Instituto de Pesquisa Econômica Aplicada (IPEA), deu início a um projeto com o objetivo de implementar uma solução de código aberto para integração de dados entre SEs. A partir da colaboração direta entre os técnicos do IPEA — que vivenciam diariamente as limitações de interoperabilidade dos sistemas — e a equipe de pesquisadores, cientistas de dados e professores do laboratório, o projeto se configurou como uma iniciativa de construção conjunta, centrada na experiência prática de quem sente na rotina a ausência de governança de dados e informação qualificada.

O ciclo do projeto foi dividido em duas fases principais. A primeira foi dedicada ao diagnóstico e ao planejamento, incluindo o mapeamento dos sistemas, dos dados e dos

gargalos operacionais. A segunda fase concentrou-se na implementação da arquitetura, na modelagem dos dados e na construção dos fluxos de integração, com base em boas práticas de engenharia de dados e governança. A seguir detalhamos cada uma dessas fases.

### **3.1. Diagnóstico e Planejamento**

A primeira etapa teve início com o mapeamento do problema, a partir da análise dos relatórios do IPEA, identificamos os indicadores e informações utilizadas em diferentes áreas, como gestão de contratos, TEDs e orçamentos. Esse trabalho permitiu compreender quais dados eram necessários para gerar os relatórios.

Após a análise do problema, conduzimos análises exploratórias dos dados [Data *et al.* 2016] e examinamos a documentação das APIs, com foco na identificação dos *endpoints* disponíveis e na avaliação da estrutura e qualidade das informações, com base nos critérios de completude, consistência, precisão e atualização. Em paralelo, mapeamos as relações entre as entidades de cada base de dados e entre os diferentes SEs, processo que envolveu a identificação de chaves de integração e a modelagem de fluxogramas para representar as conexões e a organização dos dados. Como resultado dessas atividades, elaboramos um dicionário de dados com a descrição das variáveis, seus tipos e padrões, além de definir os fluxos de coleta, estruturando o processo de Extração, Transformação e Carga (*ETL*).

Por fim, foi realizada a etapa de modelagem e design da arquitetura, que consistiu na definição dos requisitos não funcionais do sistema, com base na análise das ferramentas já utilizadas pelo IPEA e na avaliação de tecnologias de software livre com comunidades ativas. O principal objetivo foi desenvolver uma plataforma compatível com o ecossistema tecnológico do órgão, garantindo aderência institucional, viabilidade de adoção e sustentabilidade técnica a longo prazo. A proposta também visou reduzir custos de manutenção, facilitar a capacitação das equipes envolvidas e assegurar a evolução contínua da solução.

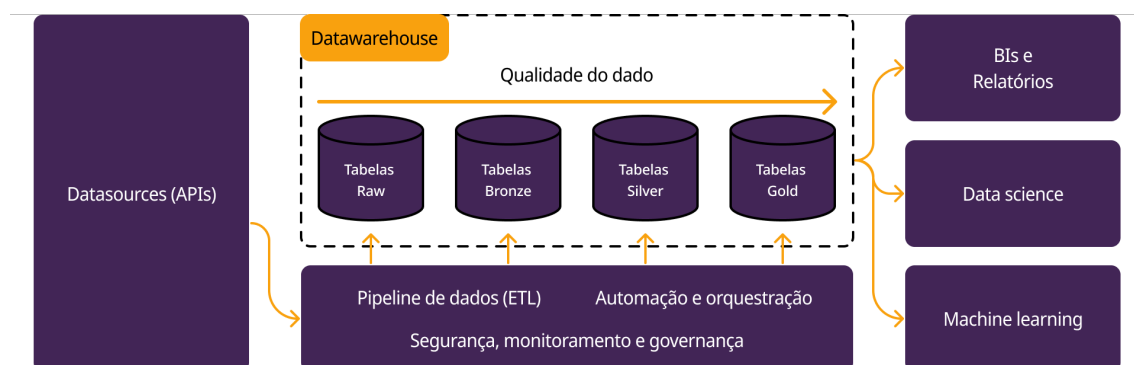
Essa etapa organizou os dados de forma eficiente e escalável, apoiando uma gestão pública orientada a dados e a colaboração entre diferentes interessados. Para isso, foi adotada a Arquitetura Medallion, que estrutura os dados em camadas progressivas de qualidade — *bronze*, *silver* e *gold* —, permitindo seu tratamento incremental e o reúso inteligente das informações [Russom *et al.* 2017].

A modelagem adotada visa garantir que os dados estejam acessíveis e qualificados, atendendo às demandas de diversas áreas e promovendo o uso colaborativo das informações. Embora centralizados em um único repositório, os dados seguem um modelo que favorece a governança descentralizada: qualquer equipe pode consumir bases existentes, adaptá-las às suas necessidades ou derivar novas a partir delas [Machado *et al.* 2022]. Essa abordagem estimula a construção coletiva do conhecimento, amplia a reutilização das informações e proporciona maior autonomia a gestores, analistas e pesquisadores.

### **3.2. Implementação da plataforma**

Na segunda fase, desenvolvemos a arquitetura apresentada na Figura 2, estruturada para permitir o tratamento e cruzamento progressivo dos dados, desde sua extração bruta até

a consolidação final em dados qualificados (*Gold*). Na etapa inicial, o sistema realiza a ingestão automática de dados provenientes das APIs dos SEs. Esses dados são armazenados em sua forma original na camada *Raw*, garantindo fidelidade e auditabilidade da origem.



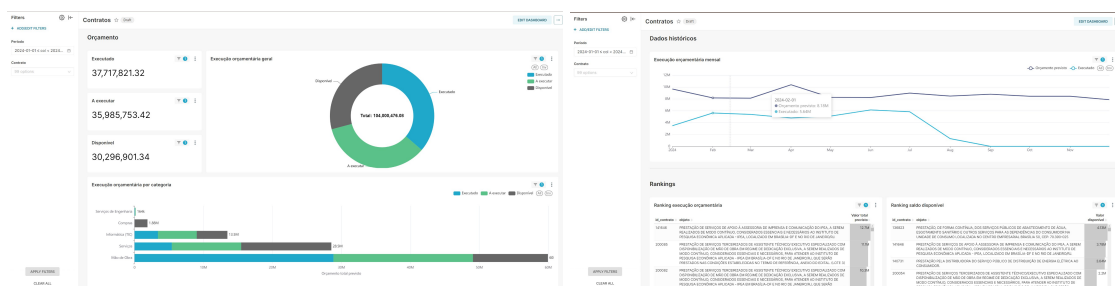
**Figura 2. Arquitetura Medallion com camadas de dados e integração com BI e Machine Learning.**

A partir dessa base, inicia-se o *pipeline* de dados (*ETL*), responsável por transformar, padronizar e qualificar as informações. Na camada *Bronze*, os dados passam por um processo de limpeza, padronização e estruturação inicial. Em seguida, na camada *Silver*, são enriquecidos a partir do cruzamento entre diferentes bases. Por fim, a camada *Gold* organiza e reúne os dados mais qualificados, prontos para consumo analítico.

Durante o *ETL*, são aplicadas práticas de segurança, automação, monitoramento e governança, visando rastreabilidade, integridade e reprodutibilidade. Além disso, o sistema conta com mecanismos de orquestração automática, que garantem a execução contínua e escalável dos fluxos de dados. Por fim, os dados qualificados são consumidos por ferramentas de *Business Intelligence* (BI), geração de relatórios analíticos e aplicações de *Machine Learning*, viabilizando o uso efetivo da informação por diferentes perfis de usuários, como gestores, analistas e pesquisadores.

Como resultado, desenvolvemos sete *dashboards* interativos voltados ao acompanhamento orçamentário e financeiro. No contexto dos TEDs, foram elaborados dois painéis: um dedicado ao monitoramento dos destaques orçamentários (NCs), da execução orçamentária (NEs) e dos repasses financeiros relacionados (PFs); e outro consolidando essas informações para todos os TEDs. Para os contratos, criamos cinco painéis, abrangendo desde estatísticas gerais — como quantidade, valores totais, categorias e status — até aspectos relacionados à execução financeira e à transparência das faturas. A Figura 3 apresenta uma visão consolidada dos contratos, com valores totais, execução e pagamentos.

O GovHubBR é uma solução de código aberto em desenvolvimento (sob licença MIT), documentada e replicável. Sua arquitetura foi concebida para ser adaptável a outros contextos da administração pública. Espera-se que a publicação do repositório e das instruções incentive o reúso da solução por outras equipes públicas que enfrentam desafios semelhantes. O repositório e mais informações estão disponíveis em: [www.gov-hub.io](http://www.gov-hub.io).



**Figura 3. Dashboards gerados a partir das tabelas *golds* apresentando valores contratuais e execução orçamentária.**

## 4. Lições Aprendidas

O primeiro desafio e principal motivador deste projeto é que os SEs não foram projetados para interoperabilidade. Isso se reflete na ausência de uma chave única de integração entre as bases de dados, o que torna a conexão entre sistemas um desafio estrutural. Para contornar essa limitação, **criamos regras empíricas — ou heurísticas de integração — baseadas na observação de como os dados são registrados e utilizados no dia a dia.**

Essas heurísticas partem da observação do padrão de uso dos sistemas, sendo continuamente refinadas conforme novas exceções são descobertas. Esse processo retroalimenta o aprimoramento tanto dos sistemas quanto dos processos internos. Quando uma heurística se torna robusta o suficiente, ela pode embasar uma mudança no sistema — como a criação de um campo estruturado — ou orientar ajustes no processo que corrijam o uso do sistema, impactando positivamente na padronização dos dados.

**Adotar a arquitetura Medallion garantiu uma estrutura robusta para a ingestão e qualificação dos dados.** Diferentemente de abordagens em que os dados são transformados diretamente em *dashboards* ou por meio de consultas *ad hoc*, o modelo Medallion organiza os dados em camadas progressivas de qualidade, promovendo maior confiabilidade, rastreabilidade e padronização [Russom *et al.* 2017]. Ainda que os dados estejam tecnicamente centralizados, a arquitetura favorece uma governança descentralizada e viabiliza a evolução contínua do ecossistema de dados, garantindo controle, autonomia e escalabilidade. Equipes e usuários podem criar indicadores e conjuntos derivados sem comprometer a integridade do repositório central. Com boas práticas de governança e documentação, o modelo incentiva uma cultura de dados abertos e qualificados, ampliando sua aplicação em diferentes níveis de decisão.

Um dos impactos mais relevantes da solução foi **viabilizar a abertura e a qualificação dos dados da execução orçamentária e financeira federal.** Ao centralizar, padronizar e integrar dados de SEs como SIAFI, ComprasGov e TransfereGov, a plataforma oferece uma visão consolidada e transparente de todo o fluxo orçamentário, do planejamento à execução. Ao tornar os dados acessíveis, auditáveis e reutilizáveis, fortalece o controle institucional, a fiscalização e a atuação de pesquisadores e jornalistas. A iniciativa reforça os princípios do Governo Digital ao promover o uso efetivo da informação pública para análise, decisão e participação social.

O projeto também gerou **documentações atualizadas das APIs governamentais.** Durante o processo de qualificação e integração dos dados, realizamos uma análise aprofundada das APIs disponíveis nos SEs, incluindo a verificação de *endpoints* e parâmetros,

bem como a testagem dos dados retornados e sua compatibilidade com outras fontes. Em muitos casos, as documentações oficiais estavam desatualizadas, incompletas ou ausentes, especialmente em sistemas legados.

Diante dessas lacunas, foi necessário reconstruir parte da lógica das APIs por meio de engenharia reversa, mapeando o tráfego de dados, os campos utilizados e as transformações aplicadas no *backend*. Esse processo resultou em documentações revisadas com propostas de melhoria e indicou caminhos mais acessíveis para futuras iniciativas de interoperabilidade. Além disso, a equipe enfrentou barreiras operacionais e institucionais, como exigência de certificados digitais, credenciamentos e ausência de orientações técnicas e contratuais. Para mitigar esses entraves e facilitar futuras integrações, **elaboramos um guia prático de acesso a dados públicos via APIs** com instruções passo a passo, pontos de atenção e cobrindo aspectos técnicos, administrativos e jurídicos.

## Referências

- Costa, L., Dutra, M. T., Oliveira, G., Silva, M., Soares, D. C., de Cássia Silva Faria, L., Jr., W. M., and Pappa, G. (2024). Ciência de dados e transparência: Experiências com dados públicos do sicom. In *Anais Estendidos do XXXIX Simpósio Brasileiro de Bancos de Dados*, pages 246–252, Porto Alegre, RS, Brasil. SBC.
- da Silva Cristóvam, J. S., Saikali, L. B., and de Sousa, T. P. (2020). Governo digital na implementação de serviços públicos para a concretização de direitos sociais no brasil. *Seqüência (Florianópolis)*, 43(89):209–242.
- Data, M. C., Komorowski, M., Marshall, D. C., Saliccioli, J. D., and Crutain, Y. (2016). Exploratory data analysis. *Secondary analysis of electronic health records*, pages 185–203.
- Dunleavy, P., Margetts, H., Bastow, S., and Tinkler, J. (2006). New public management is dead—long live digital-era governance. *Journal of Public Administration Research and Theory*, 16(3):467–494.
- Gil-Garcia, J. R., Dawes, S. S., and Pardo, T. A. (2018). Digital government and public management research: Finding the crossroads. *Public Management Review*, 20(5):633–646.
- Machado, I. A., Costa, C., and Santos, M. Y. (2022). Data mesh: concepts and principles of a paradigm shift in data architectures. *Procedia Computer Science*, 196:263–271.
- Ministério do Planejamento (2010). Introdução ao siop – sistema integrado de planejamento e orçamento. Acesso em: 27 mar. 2025.
- Nadal, S., Romero, O., Abelló, A., Vassiliadis, P., and Vansummeren, S. (2019). An integration-oriented ontology to govern evolution in big data ecosystems. *Information systems*, 79:3–19.
- Russom, P. et al. (2017). Data lakes: Purposes, practices, patterns, and platforms. *TDWI white paper*.
- Viana, A. C. A. (2021). Transformação digital na administração pública: do governo eletrônico ao governo digital. *Revista Eurolatinoamericana de Derecho Administrativo*, 8(1):115–136.