

FUNDAÇÃO GETULIO VARGAS
ESCOLA DE POLÍTICAS PÚBLICAS E GOVERNO

MARIA TEREZA TRINDADE VELOSO

**A GOVERNANÇA DE DADOS NOS ESTADOS
BRASILEIROS: UM ESTUDO SOBRE
RESTRITIVIDADE, ENTOPIA E DIVERSIDADE
NORMATIVA NOS DECRETOS**

BRASÍLIA

2024

MARIA TEREZA TRINDADE VELOSO

**A GOVERNANÇA DE DADOS NOS ESTADOS
BRASILEIROS: UM ESTUDO SOBRE
RESTRITIVIDADE, ENTOPIA E DIVERSIDADE
NORMATIVA NOS DECRETOS**

Monografia apresentada à Escola de
Políticas Públicas e Governo da Fundação
Getulio Vargas, como requisito para
obtenção do título de Bacharel em
Administração Pública

Área de concentração: Administração
Pública

Orientador: Prof. Dr. Lizandro Lui

BRASÍLIA

2024

MARIA TEREZA TRINDADE VELOSO

**A GOVERNANÇA DE DADOS NOS ESTADOS BRASILEIROS: UM ESTUDO
SOBRE RESTRITIVIDADE, ENTOPIA E DIVERSIDADE NORMATIVA NOS
DECRETOS**

Monografia apresentada à Escola de Políticas Públicas e Governo da Fundação Getúlio Vargas, como requisito para obtenção do título de Bacharel em Administração Pública.

Área de Concentração: Administração Pública

Data de aprovação: ____/____/____

Banca examinadora:

Prof. Dr. Lizandro Lui
FGV-EPPG

Prof. Dr. Jeferson Teodorovicz
FGV-EPPG

DEDICATÓRIA

*Dedico este trabalho ao meu pai,
Lourenço Cristovam Veloso, por me
ajudar a entender os meus sonhos, e
à minha mãe, Gislane Magalhães
Trindade Veloso, por me permitir
correr até eles.*

AGRADECIMENTOS

Gostaria, primeiramente, de agradecer a Deus e aos Orixás, que iluminaram meus caminhos e me trouxeram até Brasília. À minha família, pelo apoio incondicional e amor inabalável durante toda a graduação. Em especial, agradeço ao meu pai, Lourenço Cristovam Veloso, e à minha mãe, Gislene Magalhães Trindade Veloso, cuja força e carinho foram minha base, e aos meus três irmãos, que foram o combustível necessário para seguir em frente.

Aos amigos que tornaram minha jornada mais leve e cheia de momentos inesquecíveis: Sofia, Gabriela, Geovanna, Akila e Miguel, vocês foram fundamentais para as risadas e a alegria nos dias em Brasília. Minha gratidão também ao Douglas Ataides, que esteve ao meu lado em todos os momentos, me amparando e oferecendo apoio nas minhas maiores dificuldades, inclusive até a conclusão deste trabalho.

Aos meus mestres, sou profundamente grata. Dr. Lizandro Lui, pela paciência e dedicação, sendo uma verdadeira inspiração como profissional. À Professora Dra. Tatiana Soster, por me apresentar uma área pela qual hoje me sinto realizada profissionalmente. Por fim, ao Professor Fernando Filgueiras, que, junto ao Professor Lui, significativamente para o desenvolvimento de pesquisas de dados no Brasil.

RESUMO

Este trabalho analisa a governança de dados nos estados brasileiros à luz da Lei Geral de Proteção de Dados (LGPD), utilizando a Gramática Institucional (IGT) como ferramenta analítica. A pesquisa investigou a restritividade, entropia e diversidade normativa dos decretos estaduais, evidenciando padrões e lacunas nas estratégias de governança. Identificou-se que a maioria dos estados adota abordagens centradas em controle e conformidade, priorizando obrigações e proibições, o que limita a inovação e a flexibilidade necessárias para implementar políticas públicas mais adaptáveis. A análise reforça a importância de uma maior integração entre os níveis de governo, a promoção de estratégias de compartilhamento seguro de dados e a ampliação da participação de diferentes stakeholders.

PALAVRAS-CHAVE: Governança de Dados, LGPD, Gramática Institucional, Restritividade, Entropia, Diversidade Normativa.

ABSTRACT

This paper analyzes data governance in Brazilian states in light of the General Data Protection Law (LGPD), using Institutional Grammar (IGT) as an analytical tool. The research investigated the restrictiveness, entropy, and normative diversity of state decrees, highlighting patterns and gaps in governance strategies. It was identified that most states adopt approaches centered on control and compliance, prioritizing obligations and prohibitions, which limits the innovation and flexibility needed to implement more adaptable public policies. The analysis reinforces the importance of greater integration between levels of government, the promotion of secure data sharing strategies, and the expansion of the participation of different stakeholders.

KEYWORDS: Data Governance, LGPD, Institutional Grammar, Restrictiveness, Entropy, Normative Diversity.

LISTA DE TABELAS

Tabela 1- Elementos sintáticos da gramática institucional	23
Tabela 2 - Classificação dos Enunciados Institucionais das Regulamentações Estaduais	24
Tabela 3 - Número de Enunciados Institucionais nos Documentos Estaduais	28
Tabela 4 - Frequência da categoria operador deôntico em todos os estados.....	29
Tabela 5 - Frequência da categoria atributo em todos os Estados.	31
Tabela 6 - Frequência da categoria alvo em todos os estados.....	33
Tabela 7 - Frequência da categoria objeto em todos os estados.	36

LISTA DE GRÁFICOS

Gráfico 1 - Restritividade por Estado	29
Gráfico 2 - Entropia do atributo nos municípios brasileiros	30
Gráfico 3- Entropia da categoria objeto nos estados brasileiros	33
Gráfico 4 - Entropia da categoria objeto nos estados brasileiros	44

SUMÁRIO

1	INTRODUÇÃO	11
2	DESENVOLVIMENTO DA LGPD	13
3	GOVERNANÇA DE DADOS.....	16
4	METODOLOGIA	19
4.1	Texto como Dado.....	19
4.2	Gramática Institucional	21
4.3	Procedimentos de pesquisa	25
5	ANÁLISE DE DADOS.....	27
5.1	Restritividade	28
5.2	Entropia do atributo.....	30
5.3	Entropia do Alvo	32
5.4	Entropia do Objeto	35
6	DISCUSSÃO.....	37
7	CONCLUSÃO E SUGESTÕES PARA FUTURAS PESQUISAS	39
	REFERÊNCIA	41

1 INTRODUÇÃO

Nos últimos anos, a proteção da privacidade e dos dados pessoais emergiu como uma preocupação global crescente, refletida na adoção de legislações específicas como o Regulamento Geral sobre a Proteção de Dados (GDPR) na União Europeia e, mais recentemente, na Lei Geral de Proteção de Dados (LGPD) no Brasil. A LGPD, sancionada em agosto de 2018 e efetivada em setembro de 2020, estabelece um marco importante na regulamentação brasileira do tratamento de dados pessoais. Esta legislação, inspirada em padrões internacionais, proporciona diretrizes claras para a coleta, uso e proteção das informações pessoais, para assegurar a privacidade dos cidadãos e fortalecer a segurança jurídica das organizações que lidam com dados. A LGPD não só afeta o ambiente corporativo, mas também modifica o cotidiano dos cidadãos ao garantir maior controle sobre a utilização de suas informações pessoais.

No setor público, a governança de dados vai além da simples aplicação de normas e políticas estabelecidas pela LGPD. Ela exige um esforço coordenado para integrar dados de diferentes fontes e garantir que sejam utilizados de maneira ética e eficiente, tanto no nível federal quanto nos estados e municípios (Filgueiras et.al, 2025). Os governos precisam modificar o comportamento de administradores e analistas de dados por meio de regras formais e informais, desenvolvendo novas dinâmicas que enfrentem os desafios relacionados à coleta, armazenamento, processamento, compartilhamento e uso de dados ao longo de seu ciclo de vida. A interoperabilidade entre sistemas de dados é outro aspecto fundamental, facilitando o compartilhamento de informações entre órgãos governamentais, promovendo uma administração mais integrada e eficiente (Micheli et al. 2020).

A LGPD também se aplica aos estados brasileiros, exigindo que estes estejam conforme suas disposições, promovendo a proteção da privacidade e dos direitos individuais em todas as suas atividades. Essa exigência eleva a importância da governança de dados no setor público, onde a necessidade de garantir a qualidade dos dados ao longo de todo o ciclo de vida se torna essencial para evitar decisões baseadas em informações inconsistentes ou incorretas. O desenho de uma governança eficaz de dados emerge de estruturas institucionais que moldam as interações dos atores envolvidos na gestão de dados em larga escala para apoiar políticas públicas.

Os estados brasileiros, no exercício de suas funções, são responsáveis por uma série de áreas críticas, como segurança pública, gestão de regiões metropolitanas, redes de ensino fundamental e médio, regionalização da saúde, além de funções administrativas como o registro

de veículos e arrecadação do ICMS. Esses domínios não demandam apenas uma administração tradicional, mas também uma governança sofisticada de dados, adequada à complexidade institucional e às demandas crescentes por eficiência e transparência. A literatura acadêmica destaca que uma governança de dados estruturada é essencial para a formulação de políticas públicas baseadas em evidências. Segundo Janssen et al. (2017), a governança de dados envolve não apenas o controle e a proteção das informações, mas também a capacidade de os governos integrarem, processarem e utilizarem dados de forma estratégica para enfrentar desafios sociais e econômicos. Assim, a eficácia das políticas públicas depende diretamente de como os estados conseguem gerir e utilizar seus dados de maneira eficiente.

Neste Trabalho de Conclusão de Curso, investigaremos as estratégias adotadas pelos estados brasileiros em relação à governança de dados e à diversidade de atores envolvidos na implementação da LGPD. Nosso objetivo é analisar o nível de diversidade linguística e o grau de restritividade presentes nas regulamentações estaduais, utilizando a Gramática Institucional como ferramenta analítica central para oferecer uma visão detalhada das implicações da LGPD. Este instrumento, desenvolvido por Crawford e Ostrom (1995), tem sido ampliado ao longo dos anos por diversos pesquisadores, com destaque para Siddiki, que o aplicou na análise de políticas regulatórias.

Inspirando-se na abordagem de Filgueiras e Lui, que investigaram a restritividade, entropia e diversidade institucional em decretos municipais de cidades com mais de 500.000 habitantes, este trabalho também incorpora técnicas de análise textual adaptadas a partir da proposição de Valle et al. (2022), conforme utilizadas por esses autores. Por meio desta pesquisa, buscamos não apenas compreender as estratégias estaduais de conformidade à LGPD, mas também oferecer insights valiosos para organizações e legisladores, promovendo uma discussão informada sobre proteção de dados e governança em um contexto federativo diversificados.

2 DESENVOLVIMENTO DA LGPD

O Brasil vem consolidando sua tradição no desenvolvimento de processos de coleta, armazenamento, processamento e compartilhamento de dados eletrônicos. De acordo com Filgueiras e Almeida (2021) e Palotti et al. (2020), esse processo iniciou com a criação do Serviço Federal de Processamento de Dados (SERPRO) em 1964, que estabeleceu a infraestrutura tecnológica necessária para o uso de informações na formulação de políticas públicas. Posteriormente, em 1974, foi criada a Empresa de Tecnologias e Informações em Proteção Social (DATAPREV), ampliando a capacidade computacional e reforçando o papel da tecnologia na gestão de dados do governo federal.

Esse desenvolvimento facilitou a implementação de um governo eletrônico entre as décadas de 1970 e 2000, período em que diversos serviços de informação passaram a ser oferecidos, permitindo o armazenamento e processamento de dados em larga escala. No entanto, conforme Filgueiras et al. (2019), a incorporação dessa infraestrutura tecnológica se deu de forma descentralizada, com cada agência ou departamento federal operando suas próprias estruturas de dados. Essa ausência de uma estratégia centralizada para todo o setor público resultou em um aprendizado restrito às necessidades específicas de cada órgão, sem uma coordenação abrangente para o uso de tecnologias da informação e comunicação (TIC) em todas as esferas governamentais.

As bases institucionais do governo eletrônico no Brasil, como apontam Filgueiras et al. (2019), foram formadas por meio de uma série de políticas que buscavam facilitar a digitalização e apoiar a infraestrutura de TIC. Um marco importante foi a promulgação da Lei 5.172, de 25 de outubro de 1966, durante o regime militar, que estabeleceu regras sobre a proteção e confidencialidade dos dados fiscais, conforme definido nos artigos 198 e 199 do Código Tributário Nacional (CTN). Essas normas foram interpretadas pelos gestores de dados como altamente restritivas em relação ao compartilhamento de informações, criando um ambiente onde os dados coletados por órgãos públicos eram considerados recursos internos, cujo compartilhamento entre agências era, por norma, proibido, exceto por ordem judicial.

Em 2016, houve uma reorientação significativa na política de dados do governo brasileiro com o lançamento da Estratégia de Governança Digital (EGD). Essa nova abordagem visava promover a inovação nos serviços públicos, construindo uma plataforma única que integrasse todos os serviços prestados aos cidadãos e empresas, culminando na criação da Plataforma Gov.br. A EGD representou uma transição de um modelo de governo eletrônico (e-

governo) para uma concepção de governança digital, incorporando ferramentas de big data e redesenhando a estrutura de serviços públicos. Essa mudança foi reforçada pelo Decreto 10.046, que instituiu o Comitê Central de Governança de Dados (CCGD), responsável por desenhar a política de dados, facilitar o compartilhamento de informações e resolver conflitos entre as diferentes organizações do setor público.

Nesse contexto, a promulgação da Lei Geral de Proteção de Dados (LGPD), em 2018, marcou um ponto de virada crucial na política de dados no Brasil. A LGPD, aprovada em meio a debates sobre modernização estatal e privacidade, trouxe novas diretrizes regulatórias para a coleta, armazenamento, processamento e compartilhamento de dados pessoais, afetando tanto o setor público quanto o privado. A partir de 2020, a LGPD introduziu um conjunto de regras que visam proteger a privacidade e garantir maior transparência no uso de dados, exigindo que os gestores e organizações públicas adotem práticas que assegurem a proteção das informações pessoais. Como apontado por Filgueiras e Lui (2023) As bases institucionais do governo eletrônico no Brasil foram estabelecidas a partir de uma série de políticas destinadas a promover a digitalização e a sustentação da infraestrutura de tecnologia da informação e comunicação (TIC), ou seja, até então não tinham um padrão nacional estabelecido.

No que diz respeito às responsabilidades dos estados no contexto da LGPD, eles desempenham um papel crucial na implementação da legislação, uma vez que a lei se aplica a todas as esferas de governo — federal, estadual e municipal. Os estados brasileiros são responsáveis por um conjunto significativo de políticas públicas, incluindo segurança pública, gestão de regiões metropolitanas, educação básica e saúde. Em todas essas áreas, os estados lidam diretamente com grandes volumes de dados pessoais, tornando essencial a conformidade com a LGPD.

A lei exige que os estados estabeleçam mecanismos adequados para garantir a proteção dos dados pessoais sob sua gestão. Isso inclui a criação de comitês ou órgãos de governança de dados que possam implementar políticas de privacidade, monitorar o cumprimento da LGPD e garantir que os dados sejam tratados de maneira adequada e segura. Além disso, os estados devem nomear um encarregado de proteção de dados (DPO), responsável por supervisionar o tratamento dos dados e atuar como ponto de contato entre as organizações públicas e a Autoridade Nacional de Proteção de Dados (ANPD) (Filgueiras et. al, 2023).

As obrigações dos estados vão além do cumprimento técnico da lei. Eles precisam desenvolver uma cultura de proteção de dados, garantindo que os servidores públicos compreendam e sigam os princípios da LGPD no tratamento das informações pessoais dos cidadãos. Isso é particularmente importante em áreas sensíveis, como a segurança pública e a

saúde, onde o manejo inadequado de dados pode gerar graves violações de privacidade.

A LGPD também impõe aos estados a responsabilidade de garantir que os sistemas de coleta, armazenamento e processamento de dados sejam interoperáveis entre diferentes órgãos governamentais. O que é fundamental para facilitar o compartilhamento de dados de maneira segura e eficiente, promovendo uma administração pública mais integrada. Ao mesmo tempo, os estados devem assegurar que o compartilhamento de dados esteja conforme as normas de segurança e privacidade estabelecidas pela LGPD, evitando o uso indevido de informações pessoais.

Essas responsabilidades tornam os estados atores centrais na implementação da LGPD, exigindo uma governança de dados robusta, capaz de balancear a eficiência administrativa com a proteção dos direitos dos cidadãos. A introdução da LGPD e a criação de novos padrões de governança de dados no Brasil representam um passo fundamental na modernização do estado, mas também impõem desafios significativos aos gestores públicos, que precisam se adaptar a um novo cenário de maior controle e transparência sobre o uso de dados pessoais.

3 GOVERNANÇA DE DADOS

A governança de dados vai muito além de um conjunto de regras formais; trata-se de um processo complexo e contínuo que visa garantir o gerenciamento eficaz, ético e seguro dos dados em uma organização. Esse sistema não apenas define como os dados são coletados, usados, protegidos e compartilhados, mas também fomenta uma cultura organizacional que valoriza a transparência, a responsabilidade e a integridade na gestão de informações. Ao lidar com o crescente volume e a importância estratégica dos dados, tanto no setor público quanto no privado, a governança de dados se estabelece como um dos pilares fundamentais da era digital, exigindo um equilíbrio entre inovação, segurança e conformidade regulatória (Micheli et al., 2020).

Um dos principais pilares da governança de dados é a clara definição de autoridade e responsabilidade nas organizações. Esta governança define papéis e estruturas que regulam quem pode acessar os dados, como esses dados devem ser protegidos contra ameaças internas e externas, e como as políticas de gestão de dados são implementadas, monitoradas e atualizadas ao longo do tempo. A governança de dados estabelece uma estrutura multifuncional para gerenciar dados como um ativo estratégico da organização. Define claramente os direitos de decisão e as responsabilidades associados ao gerenciamento e uso dos dados, promovendo a tomada de decisão eficiente e alinhada aos objetivos organizacionais (Abraham et al., 2019).

No contexto brasileiro, a Lei Geral de Proteção de Dados (LGPD), promulgada em 2018, desempenha um papel central nesse processo, ao impor diretrizes sobre a coleta, armazenamento, processamento e compartilhamento de dados pessoais. A criação do Comitê Central de Governança de Dados (CCGD) é um exemplo do esforço governamental para garantir que as normas da LGPD sejam aplicadas consistentemente em diferentes esferas da administração pública, promovendo a interoperabilidade e a segurança dos dados compartilhados entre órgãos governamentais.

A governança de dados também implica na criação de diretrizes rigorosas sobre o uso adequado de dados, estabelecendo restrições claras sobre como as informações podem ser compartilhadas, para que finalidades específicas podem ser utilizadas, e quem tem a autorização necessária para acessar tais dados. Por exemplo, no contexto da LGPD, dados pessoais devem ser utilizados apenas para os fins explicitamente informados aos titulares, garantindo que as práticas organizacionais respeitem a privacidade e os direitos individuais, além de assegurarem a conformidade com os padrões internacionais.

Outro aspecto fundamental da governança de dados é a implementação de auditorias regulares e independentes para garantir que as políticas de dados sejam seguidas e que não ocorram violações inadvertidas. Esses mecanismos permitem monitorar o cumprimento das políticas e práticas de proteção de dados, prevenindo e identificando falhas que poderiam resultar em incidentes de segurança. As auditorias regulares também reforçam a responsabilidade (accountability) das organizações e indivíduos que lidam com os dados, assegurando que todos os envolvidos sejam devidamente monitorados e responsabilizados por suas ações (Abraham et al., 2019).

Além das auditorias, ferramentas de monitoramento contínuo e relatórios periódicos são essenciais para rastrear o uso de dados em tempo real e garantir conformidade constante. Isso permite que as organizações mantenham visibilidade sobre o ciclo de vida dos dados, garantindo que os acessos e operações realizados sejam autorizados e conforme as normas. O fortalecimento da accountability é particularmente importante em ambientes regulados, onde as infrações às normas podem resultar em sanções financeiras ou jurídicas severas. Como destacam Dawes, Cresswell e Pardo (2009), a capacidade de rastrear, auditar e corrigir problemas de conformidade em tempo real é um elemento chave para promover a transparência e construir confiança tanto entre as partes interessadas internas quanto externas.

A segurança e a integridade dos dados são componentes essenciais da governança de dados. Isso inclui o desenvolvimento e a implementação de mecanismos robustos, como criptografia avançada, uso de firewalls e auditorias periódicas para prevenir acessos não autorizados e assegurar a integridade dos dados. No Brasil, a governança de dados enfrenta desafios significativos, como a falta de interoperabilidade entre sistemas de diferentes instituições e a relutância de alguns órgãos em compartilhar informações. Isso ocorre, em parte, devido a percepções divergentes sobre a segurança dos dados e a desconfiança em relação às capacidades tecnológicas de outras instituições.

Além do aspecto técnico, a governança de dados também se depara com complexidades políticas e institucionais. Como aponta o estudo de Filgueiras, Lui e Veloso (2025), o contexto federativo brasileiro introduz um grau significativo de diversidade institucional, onde as práticas de governança variam entre os entes subnacionais. Em muitos municípios, a regulamentação da LGPD tem se limitado a um controle burocrático estrito, com foco em conformidade, em detrimento de uma abordagem mais estratégica e inovadora de governança de dados.

Por outro lado, os dados são vistos cada vez mais como recursos estratégicos essenciais para a transformação digital e a formulação de políticas públicas eficientes. Em iniciativas de

idades inteligentes, por exemplo, grandes volumes de dados são coletados e analisados para informar decisões relacionadas a mobilidade urbana, segurança pública e serviços de saúde, entre outros. No entanto, como apontam Drapalova e Weigrich (2020), erros na governança de dados podem comprometer o sucesso dessas iniciativas, reforçando a importância de estruturas de governança robustas que promovam não apenas a proteção dos dados, mas também a inovação e a eficiência na prestação de serviços públicos.

A partir desses questionamentos, investigaremos a gramática institucional, com base na perspectiva da entropia e na abordagem do texto como dado. Este trabalho segue a metodologia proposta por Filgueiras et al., Para isso, examinaremos 35 decretos de 20 estados que, até agosto de 2024, haviam regulamentado a LGPD, definindo estratégias operacionais para a governança de dados. Essa análise permitirá identificar padrões e variações na formulação dessas políticas, explorando como os textos institucionais refletem as escolhas estratégicas e os desafios da governança de dados em diferentes contextos regionais.

4 METODOLOGIA

4.1 Texto como Dado

As ciências sociais aplicadas têm adotado cada vez mais o uso de textos brutos como fonte de dados. Essa prática reflete a tendência de combinar métodos qualitativos e quantitativos para obter análises mais abrangente e detalhadas. A utilização de grandes volumes de texto amplia o escopo das pesquisas, oferecendo uma perspectiva mais rica e multifacetada que contribui para a produção de um conhecimento mais profundo e completo.

O conceito de "Texto como Dado" (Text as Data) envolve a aplicação de técnicas automatizadas, oriundas das ciências computacionais e da estatística, para a análise sistemática de grandes volumes de conteúdo textual. Esse enfoque permite que os textos sejam explorados quantitativamente, facilitando a identificação de tendências, temas recorrentes e relações entre variáveis que, muitas vezes, passam despercebidas em análises exclusivamente qualitativas. A combinação dessas abordagens possibilita uma análise mista, onde a interpretação contextual e a objetividade dos números se complementam, proporcionando uma leitura mais completa dos fenômenos sociais.

No artigo "Do 'texto como texto' ao 'texto como dado'" (2022), Davi Moreira, em colaboração com Medeiros e Pires, explora o potencial dessa abordagem nas pesquisas em Relações Internacionais, destacando a discrepância na proporção de publicações que adotam o "texto como dado" em comparação ao "texto como texto". Eles identificam que o "texto como texto" é predominantemente qualitativo, focado na análise interpretativa do conteúdo e seu contexto. Em contraste, o "texto como dado" trata o conteúdo textual como um conjunto de dados quantitativos, estruturando-o para ser analisado sistematicamente por meio de ferramentas estatísticas, ampliando assim as possibilidades de análise e descobertas dentro do campo (Moreira, 2022).

O uso de algoritmos de aprendizado de máquina (Machine Learning) e de ferramentas de Processamento de Linguagem Natural (NLP) demonstra como o "Texto como dado" também é utilizado em outros campos. O aprendizado de máquina, desenvolvido pelas ciências computacionais, permite que sistemas aprendam e melhorem automaticamente a partir de dados, sem a necessidade de programação explícita. Essa técnica, amplamente utilizada em inteligência artificial (IA), gera conhecimento a partir de exemplos, em vez de soluções algorítmicas convencionais. Para muitos desses problemas, não há uma fórmula programática clara que resolva a questão, no entanto, a disponibilidade abundante de dados possibilita treinar

algoritmos e a aplicação eficaz do aprendizado de máquina. Dessa forma, os sistemas se tornam capazes de identificar padrões, tomar decisões e fazer previsões com base em informações previamente processadas (Ludermir, 2021)

O Processamento de Linguagem Natural (PLN), por sua vez, é uma subárea do aprendizado de máquina que possibilita que a IA entenda e manipule a linguagem humana. Como os sistemas de IA operam naturalmente com códigos e números, sem compreensão inata da comunicação humana, o PLN faz a ponte entre esses dois mundos. Ele permite que computadores "compreendam" textos e se comuniquem de maneira mais natural com os seres humanos. Ou seja, as máquinas podem, interpretar e analisar dados textuais baseados na linguagem humana.

Essas inovações tecnológicas revolucionaram como os dados textuais são analisados, expandindo o uso do conceito de "texto como dado" em áreas como a governança pública. Em 2023, Filgueiras e Lui (2023) trouxeram para o cenário nacional a discussão sobre a Institutional Grammar Tool (IGT), uma ferramenta analítica proposta por Crawford e Ostrom. A IGT decompõe as instituições em enunciados sintáticos, permitindo uma análise detalhada do desenho institucional e, assim, aplicando o conceito de "texto como dado" de maneira sistemática. A introdução dessa metodologia abre novas perspectivas para a análise das políticas públicas brasileiras.

Filgueiras, Lui e Veloso (2024) utilizaram essa abordagem para investigar a estruturação institucional da proteção de dados no Brasil, com foco na Lei Geral de Proteção de Dados (LGPD). Usando a IGT, os pesquisadores desconstruem os decretos estaduais que regulamentam a LGPD, classificando diferentes elementos sintáticos que compõem as estratégias regulatórias. A pesquisa analisa como os conceitos de proteção de dados, governança de dados e privacidade foram incorporados nas regulamentações estaduais, revelando a estrutura institucional subjacente

Os resultados indicam que as estratégias adotadas pelos estados brasileiros priorizam o controle burocrático e organizacional sobre o processamento de dados, em detrimento da criação de mecanismos efetivos para garantir os direitos de privacidade dos cidadãos. As administrações públicas estaduais desempenham um papel central na regulamentação, com pouca participação de empresas privadas e cidadãos no processo de tomada de decisões. Esse foco nas estruturas burocráticas reflete um desequilíbrio entre controle estatal e a implementação prática dos direitos de privacidade, destacando a necessidade de um maior envolvimento de múltiplos atores na governança de dados

4.2 Gramática Institucional

O presente artigo visa compreender as estratégias adotadas pelos estados brasileiros para gerir a governança de dados no contexto da Lei Geral de Proteção de Dados (LGPD). Essas estratégias são delineadas por instituições que emergiram a partir da LGPD e incluem enunciados institucionais que moldam aspectos importantes da governança de dados. Utilizando a metodologia de Análise Institucional, o estudo examina os desenhos institucionais presentes nos decretos estaduais, explorando elementos como a diversidade de atores envolvidos e o grau de similaridade e restritividade nas estratégias adotadas pelos estados. A análise não se restringe às normativas e regulamentações estabelecidas, mas também investiga como essas políticas são implementadas, refletindo as demandas e características específicas de cada estado. Compreender essas estratégias é fundamental para avaliar a eficácia das medidas adotadas e identificar boas práticas que possam ser replicadas em outras jurisdições, contribuindo para o aprimoramento contínuo da governança de dados no país.

O uso de "texto como dado" é uma ferramenta essencial para a análise institucional, especialmente no contexto de Elinor Ostrom, que enfatiza a importância de decompor as instituições. Ao tratar decretos estaduais e documentos legais como fontes de dados quantitativos, essa abordagem permite a desagregação dos elementos institucionais em suas partes constituintes, facilitando a análise sistemática e comparativa. Crawford e Ostrom (1995) descrevem as instituições como padrões de ação estáveis, organizados por meio de estratégias, normas ou regras que direcionam o comportamento humano de forma previsível. Essas instituições oferecem diretrizes claras sobre o que é permitido, exigido ou proibido em determinados contextos. As normas, conforme as autoras, são prescrições coletivas reforçadas por incentivos, enquanto as estratégias são planos de ação que indivíduos elaboram em uma estrutura de incentivos gerada por regras, normas e expectativas.

No contexto dos decretos estaduais, esses podem ser compreendidos como estratégias operacionais que os governos locais utilizam para definir como os dados devem ser compartilhados e geridos. Elinor Ostrom (2005) divide as instituições em formais (leis, regulamentos) e informais (normas sociais, costumes), destacando que ambas são essenciais para a coordenação de ações coletivas e a gestão de recursos comuns. Ostrom enfatiza que as instituições se organizam em múltiplos níveis, adaptando-se às condições sociais e ecológicas de cada contexto, e que a diversidade institucional é crucial para promover uma governança eficaz e resiliente. No caso da governança de dados no Brasil, conforme discutido por Filgueiras (2024), essa diversidade institucional também é evidente. Tal como Ostrom propõe, as

regulamentações da LGPD nos municípios refletem uma combinação de regras operacionais e estratégias de implementação adaptadas a diferentes contextos locais.

Considerando que as instituições são embutidas em estruturas linguísticas interpretadas e cristalizadas pelos atores a partir dos enunciados prescritos, conhecidos como “institutional statements” (Filgueiras; Lui; Veloso, 2025) e Crawford e Ostrom (1995, p. 583) definem esses enunciados institucionais como “a restrição ou oportunidade linguística compartilhada que prescreve, permite ou aconselha ações ou resultados para os atores (individuais ou coletivos)”. Segundo as autoras, esses enunciados podem ser falados, escritos ou tacitamente compreendidos de maneira inteligível para os atores em um cenário empírico. Assim, as instituições podem ser formais ou informais, dependendo de onde, quando e por quem os enunciados institucionais são prescritos aos atores.

Para isso, as autoras desenvolveram ferramenta analítica voltada ao estudo da gramática institucional, visando descrever e decompor instituições, facilitando o entendimento de como regras, normas e estratégias moldam o comportamento dos atores em diferentes cenários. Essa gramática decompõe as instituições em cinco componentes principais: Atributos, Deontico, Objetivo, Condições e Ou Então. Por exemplo, o inciso “Esta Lei dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público, ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural”. Os atributos (A) referem-se aos participantes específicos a quem a declaração institucional se aplica. Estes podem ser indivíduos ou organizações e são definidos por características como idade, posição, gênero ou outras variáveis relevantes. Por exemplo, a lei em questão dispõe sobre “tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado”. O deontico (D) indica a modalidade da regra, definindo se a ação especificada é permitida, obrigatória ou proibida. Em termos práticos, usa-se os verbos “pode” (permissão), “deve” (obrigação) ou “não deve” (proibição) para descrever o que os atores podem ou não fazer, nesse caso “dispor sobre o tratamento de dados pessoais”. O objetivo (I) descreve a ação ou resultado específico que está sendo regulado. É o foco central da regra, norma ou estratégia, determinando o que deve ser feito ou evitado pelos participantes. Por exemplo, “com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural” é o objetivo do inciso. As condições (C) definem o contexto em que a regra se aplica, incluindo quando, onde e como a ação ou resultado deve ocorrer. Tal como, “inclusão dos meios digitais” e “aplicação tanto para pessoas naturais quanto jurídicas”. Por fim, o então (O) descreve as sanções ou

consequências aplicáveis em caso de descumprimento de uma regra. Apenas as regras contêm este componente, que define as penalidades que serão impostas se a regra for violada, como multas ou outras formas de punição.

Tabela 1- Elementos sintáticos da gramática institucional

Elementos de sintaxe	Conceito
Atributo (Attribute)	O componente Atributo captura o ator responsável por um determinado enunciado institucional, seja de forma explícita ou implícita. Ou seja, a quem ou o que é atribuída uma ação. Além disso, em alguns enunciados, um Atributo e suas propriedades podem ser apresentados juntos.
Operador deôntico (Deontic)	O componente operador deôntico baseia-se nas operações modais usadas na lógica deôntica para distinguir enunciados prescritivos de não prescritivos. O conjunto completo de operadores deônticos, D, consiste em P permitido, O obrigatório, e F proibido.
Alvo (Aim)	O alvo é a descrição específica de uma parte ativa em uma situação de ação à qual se refere um enunciado institucional. Ou seja, qual ação é obrigatória, permitida ou requerida do atributo?
Objeto (Object)	O componente Objeto é definido como o receptor de uma ação (descrita pelo Objetivo) e executada pelo ator (descrita pelo Atributo). Objetos podem ser entidades animadas e inanimadas em enunciados institucionais, e cada enunciado pode conter vários objetos.
Condição (Condition)	Condições indicam o conjunto de variáveis que definem quando e onde um enunciado institucional se aplica. Por exemplo, as condições para um enunciado institucional podem indicar quando ele se aplica, como durante certas condições meteorológicas, em um horário definido ou em uma etapa específica de algum processo.
Consequência (Or else)	O componente final de nossa sintaxe institucional é a consequência que um enunciado institucional atribui ao descumprimento detectado com os outros componentes desse enunciado. Em alguns casos, a consequência especifica uma série de punições possíveis se uma regra não for seguida.

Fonte: Filgueiras et al., 2025

Como apresentado anteriormente as estratégias, conforme a análise institucional de Elinor Ostrom, representam os mecanismos pelos quais os estados definem suas abordagens para a governança de dados. Elas servem como ponte entre as regras formais—leis e regulamentos—e as práticas cotidianas, traduzindo princípios teóricos em ações concretas. Por meio dessas estratégias, os estados estabelecem estruturas de incentivos, normas de comportamento e mecanismos de monitoramento que orientam como os dados são coletados, compartilhados e utilizados. Filgueiras, Lui e Veloso (2025) realizaram uma análise quantitativa dos decretos estaduais para identificar os componentes mais frequentes do modelo da Gramática Institucional. Com base nessa análise, os autores categorizaram esses

componentes, o que permitiu uma melhor compreensão dos decretos e a aplicação do IGT.

Tabela 2 - Classificação dos Enunciados Institucionais das Regulamentações Estaduais da LGPD

Componente	Categorias
Atributo	<ul style="list-style-type: none"> - Cidadãos e usuários - Empresas - Encarregado de dados - Controlador de dados - Operador de dados - Governos municipais - Governos estaduais - Governo federal - Organizações da sociedade civil - Agência reguladora - Conselhos, comissões ou organizações equivalentes - Organizações internacionais
Operador deôntico	<ul style="list-style-type: none"> - Permissão - Obrigação - Proibição - Revogação/anulação
Alvo	<ul style="list-style-type: none"> - Competência - Coleta e processamento de dados - Compartilhamento de dados - Certificação e conformidade - Autorização - Monitoramento - Padronização - Direitos
Objeto	<ul style="list-style-type: none"> - Governança - Gestão e processos - Produto - Serviço - Recursos naturais
Condição	<ul style="list-style-type: none"> - Prazos e condições determinados - Prazos e condições indeterminados
Consequência	<ul style="list-style-type: none"> - Multas - Suspensão de atividade - Cassação

Fonte: Filgueiras et al., 2025

4.3 Procedimentos de pesquisa

Este trabalho baseou-se na abordagem apresentada por Filgueiras et al., que investigaram a restritividade, entropia e diversidade institucional nos decretos que regulamentam a LGPD em municípios brasileiros com mais de 500.000 habitantes. Para isso, foram analisados 35 decretos de 20 estados que, até agosto de 2024, já haviam regulamentado a LGPD e definido estratégias operacionais de governança de dados.

Inicialmente, os enunciados institucionais desses decretos foram fragmentados em sentenças para identificar seus elementos sintáticos. Com base nisso, foi conduzida uma análise de restritividade usando a métrica adaptada de Valle et al. (2022), expressa pela fórmula:

$$\text{Restritividade} = \frac{\text{enunciados institucionais restritivos}}{\text{enunciados institucionais totais}}$$

Nesse contexto, palavras classificadas como restritivas, ou seja, aquelas que expressam obrigações ou proibições, foram identificadas por meio do “operador deôntico” do IGT. Além disso, a permissão foi incluída como uma terceira categoria relevante para a análise, considerando sua presença nos decretos. A métrica foi calculada pelo número de palavras restritivas em relação ao total de declarações institucionais em cada decreto, permitindo determinar o número médio de palavras restritivas por ato normativo.

A segunda métrica utilizada foi a entropia, aplicada aos elementos sintáticos de atributo, alvo e objeto. A metodologia de Valle et al. (2022) foi adaptada para calcular a entropia dos decretos, com base na teoria da informação proposta por Shannon e Weaver (1949). A entropia mede a complexidade linguística dos enunciados, sendo maior quando há maior diversidade e distribuição uniforme dos elementos analisados. A fórmula utilizada para calcular a entropia de Shannon é:

$$D = - \sum_{i=1}^N p_i \cdot \log_2(p_i)$$

Em que **D** é a medida de diversidade, **N** é o número total de categorias analisadas, e **p_i** representa a frequência relativa de cada categoria no conjunto de dados. Essa análise foi

aplicada às palavras correspondentes a atributos, alvos e objetos nos decretos estaduais. Quanto maior a variedade de palavras e sua distribuição uniforme, maior a entropia, refletindo uma complexidade linguística mais elevada nos enunciados institucionais.

A análise da entropia é relevante porque decreta a complexidade dos atos normativos, evidenciando o envolvimento de múltiplos atores e ações no sistema de governança de dados. Assim, pressupõe-se que maior entropia implica estratégias mais robustas de governança. Já a análise da restritividade permite compreender os limites impostos pelos governos locais, facilitando a identificação de padrões regulatórios e a arquitetura institucional subjacente.

Por fim, a desconstrução dos enunciados institucionais em suas partes sintáticas possibilitou identificar as regras, seus sujeitos, os operadores deônticos, alvos, objetos, condições e consequências. Isso permitiu adaptar o índice de Shannon à análise legislativa, destacando a relação entre restritividade e complexidade linguística no contexto da governança de dados nos estados brasileiros, a partir disso será possível discutir a gramática institucional de dados nos entes subnacionais brasileiros.

5 ANÁLISE DE DADOS

A regulamentação da Lei Geral de Proteção de Dados (LGPD) nos estados brasileiros apresenta uma grande diversidade, tanto na quantidade de enunciados institucionais quanto na complexidade dos decretos estaduais. Observa-se na Tabela 3, que o número de enunciados institucionais dos estaduais que regulamentam LGPD variam de estado para estado. Em alguns estados, como Santa Catarina e Tocantins, observa-se uma regulamentação extremamente simplificada e curta. Por outro lado, outras unidades federativas como o caso do Rio de Janeiro elaboraram documentos mais robustos, com definições detalhadas e regras claras sobre os procedimentos para a proteção de dados pessoais.

Até agosto de 2024, 20 estados brasileiros haviam regulamentado a LGPD, evidenciando um panorama de adesão heterogêneo à legislação nacional. A análise regional revela disparidades marcantes.

- Sul e Sudeste, todos os estados já regulamentaram a LGPD, demonstrando maior alinhamento às diretrizes nacionais. Este cenário pode ser reflexo de uma maior capacidade institucional e administrativa nessas regiões.
- Região Norte, estados como Acre, Pará e Roraima não disponibilizaram decretos regulamentares em seus sites oficiais, indicando uma adesão limitada à legislação. Esse quadro pode estar relacionado a desafios específicos de governança.
- Região Centro-Oeste, apenas Mato Grosso não regulamentou a LGPD. Goiás e Mato Grosso do Sul já apresentaram decretos regulamentares, enquanto o Distrito Federal destacou informações sobre conformidade com a lei federal, detalhando competências dos cargos e autoridades envolvidas na aplicação da norma.
- Região Nordeste apresenta um cenário diversificado. Estados como Bahia, Rio Grande do Norte e Maranhão ainda não regulamentaram a LGPD, enquanto Piauí, Pernambuco, Paraíba, Ceará, Alagoas e Sergipe já contam com regulamentações em vigor. Essa análise demonstra que a regulamentação da LGPD nos estados está associada a fatores regionais e estruturais. Regiões com maior infraestrutura administrativa e recursos, como o Sul e o Sudeste, apresentam maior adesão à lei, enquanto regiões como o Norte enfrentam maiores desafios, que podem incluir limitações tecnológicas e menor priorização da pauta de proteção de dados.

Essas diferenças podem gerar desigualdade no tratamento de dados pessoais e na proteção aos cidadãos em todo o território nacional. Além disso, a falta de regulamentação em alguns estados compromete a implementação de políticas públicas eficientes e integradas no âmbito da governança de dados. Por outro lado, estados com regulamentações robustas têm maior potencial para garantir segurança jurídica, proteger os direitos dos cidadãos e alinhar suas políticas às melhores práticas de governança de dados.

Tabela 3 - Número de Enunciados Institucionais nos Documentos Estaduais de Regulamentação

UF	Frequência	Porcentagem
AL	34	4,34
AM	19	2,42
AP	51	6,51
CE	28	3,57
DF	70	8,93
ES	44	5,61
GO	22	2,80
MG	31	3,95
MS	21	2,68
PB	19	2,42
PE	30	3,83
PI	59	7,53
PR	51	6,51
RJ	128	16,34
RO	38	4,85
RS	37	4,72
SC	12	1,53
SE	46	5,87
SP	26	3,32
TO	17	2,17

Fonte: Elaboração Própria

5.1 Restritividade

A análise se fundamenta na adaptação da metodologia de Valle et al. (2022) para calcular a restritividade em decretos estaduais. A fórmula utilizada divide o número de enunciados institucionais classificados como "obrigação" (ou seja, elementos normativos restritivos) pelo número total de enunciados institucionais presentes nos atos normativos de cada estado. Essa métrica produz um índice entre 0 e 1, onde valores mais próximos de 1 indicam maior proporção de restrição.

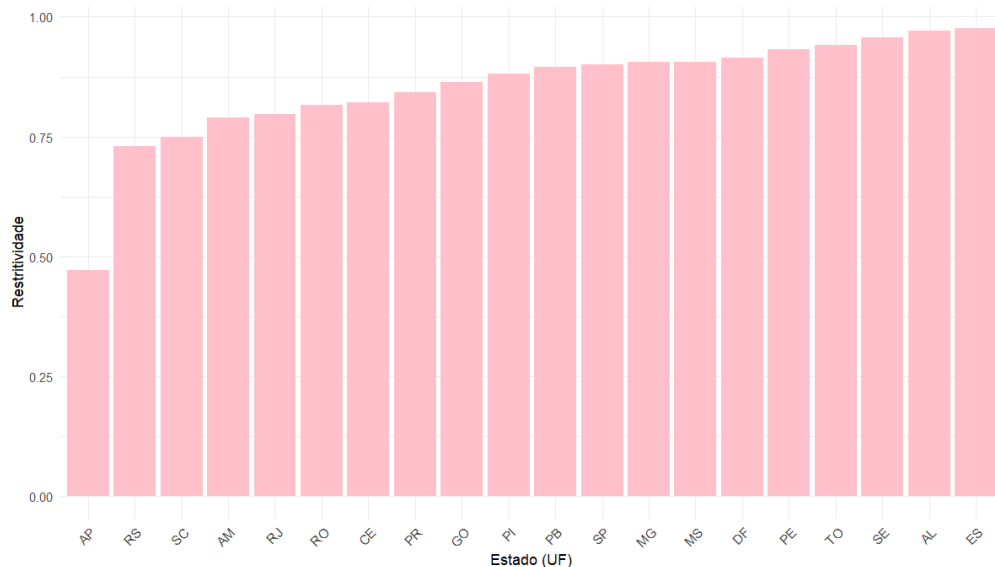
Tabela 4 - Frequência da categoria operador deôntico em todos os estados.

Categoria	Frequência
Obrigação	660
Permissão	98
Proibição	25

Fonte: Elaboração Própria

Os resultados do gráfico revelam que os estados brasileiros, de forma geral, adotam estratégias marcadas por altos níveis de restritividade em suas regulamentações relacionadas à governança de dados. Mesmo com variações entre os estados, a maioria apresenta índices de restritividade elevados, com alguns estados, como AL, ES e PE, atingindo proporções próximas de 1, evidenciando um forte foco na imposição de obrigações e proibições. Mesmo estados com índices mais baixos, como AP (0,471), demonstram níveis significativos de restrição, reforçando uma abordagem normativa centrada no controle.

Gráfico 1 - Restritividade por Estado



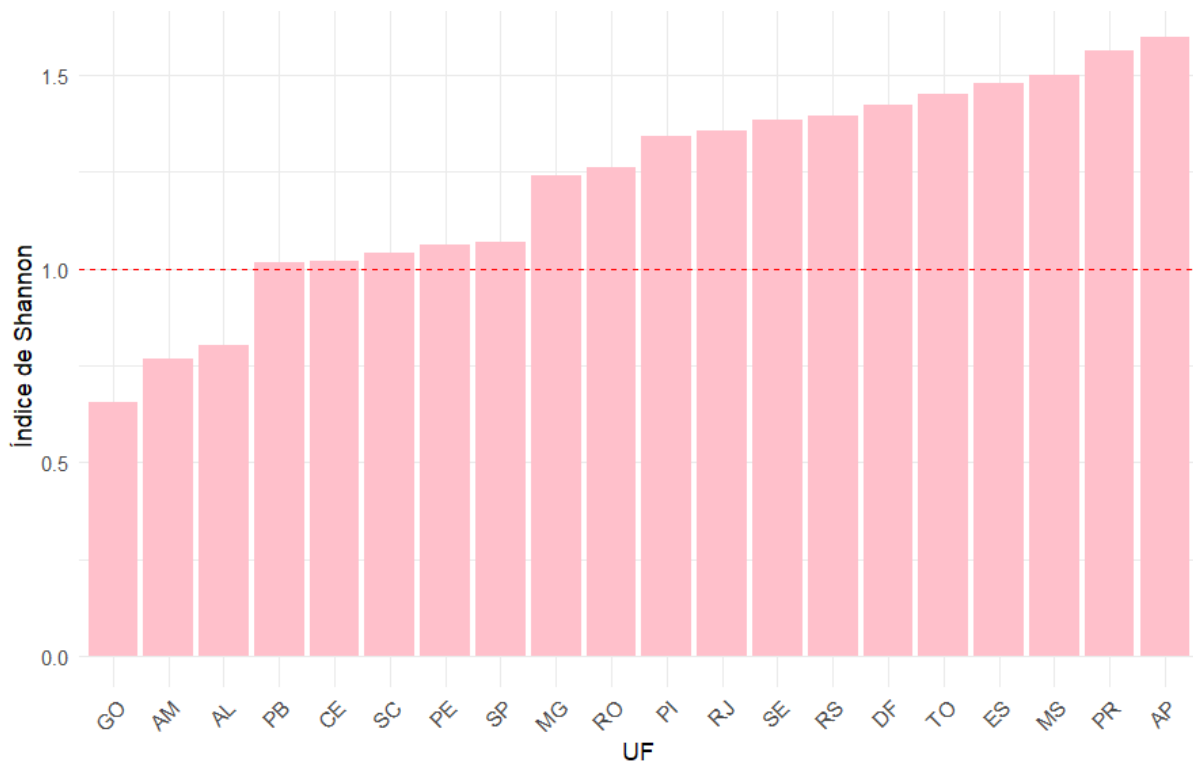
Fonte: Elaboração Própria

Esse padrão reflete uma gramática institucional que prioriza o estabelecimento de limites e condições para as ações dos atores regulados, deixando menor ênfase para permissões. Enquanto essa lógica de controle garante maior segurança jurídica e conformidade com a LGPD, ela pode criar barreiras à inovação e à flexibilidade, dificultando a adaptação às demandas locais e às rápidas mudanças tecnológicas. A imposição excessiva de restrições pode comprometer o dinamismo necessário para implementar políticas públicas mais inovadoras e adaptáveis.

5.2 Entropia do atributo

Após a análise dos atores presentes, procedeu-se ao cálculo do Índice de Shannon, a fim de mensurar a entropia dos atores nas regulamentações de cada município de forma individual. Nessa abordagem, os valores são escalonados de 0 a 1: quanto mais próximo de 1, maior a entropia, e valores superiores a 1 indicam uma alta entropia, enquanto valores próximos de zero refletem uma baixa entropia.

Gráfico 2 - Entropia do atributo nos municípios brasileiros



Fonte: Elaboração Própria

O Amapá (AP) apresenta o maior Índice de Shannon, ultrapassando 1,5, o que reflete uma alta diversidade de atributos regulados. Isso sugere que o estado adota uma abordagem inclusiva, contemplando uma ampla gama de categorias em sua regulação. Em contraste, estados como Goiás (GO) e Amazonas (AM), com índices abaixo de 1, indicam menor diversidade e uma regulação mais concentrada em poucas categorias. Já UFs como Piauí (PI), Rondônia (RO) e Rio de Janeiro (RJ), com índices próximos de 1, apresentam uma diversidade intermediária de atributos regulados, demonstrando alguma inclusão de diferentes categorias, mas sem alcançar o mesmo grau de heterogeneidade de estados como o AP.

A tabela de frequências fornece uma visão adicional sobre as categorias de atributo. A Administração Pública Estadual lidera em frequência, com 365 ocorrências, evidenciando que a maioria das regulações estaduais está concentrada nesse atributo. Essa concentração explica os baixos índices de Shannon em estados como GO e AM, que possuem menos diversidade de atributos regulados. Outras categorias, como Conselhos, Comissões ou Organizações Equivalentes (146 ocorrências) e Encarregados de Dados (108 ocorrências), também têm relevância significativa e contribuem para a diversidade moderada observada em estados como Minas Gerais (MG) e Rio de Janeiro (RJ).

Tabela 5 - Frequência da categoria atributo em todos os Estados.

Categorias	Frequência
Administração Pública Estadual	365
Concelhos, Comissões ou Organizações Equivalentes	146
Encarregado de Dados	108
Operador de Dados	50
Cidadãos e Usuários	36
Controlador de Dados	34
Empresas	32
Administração pública Federal	3
Organizações Internacionais	1
Agências Reguladoras	1

Fonte: Elaboração Própria

As categorias menos frequentes, como Operadores de Dados (50), Cidadãos e Usuários (36), Controladores de Dados (34) e Empresas (32), indicam uma menor participação em

comparação com as categorias predominantes. Além disso, a baixa frequência de atributos como Administração Pública Federal (3 ocorrências), Organizações Internacionais (1 ocorrência) e Agências Reguladoras (1 ocorrência) revela uma influência limitada desses atores na governança estadual, o que pode restringir a diversidade em algumas UFs.

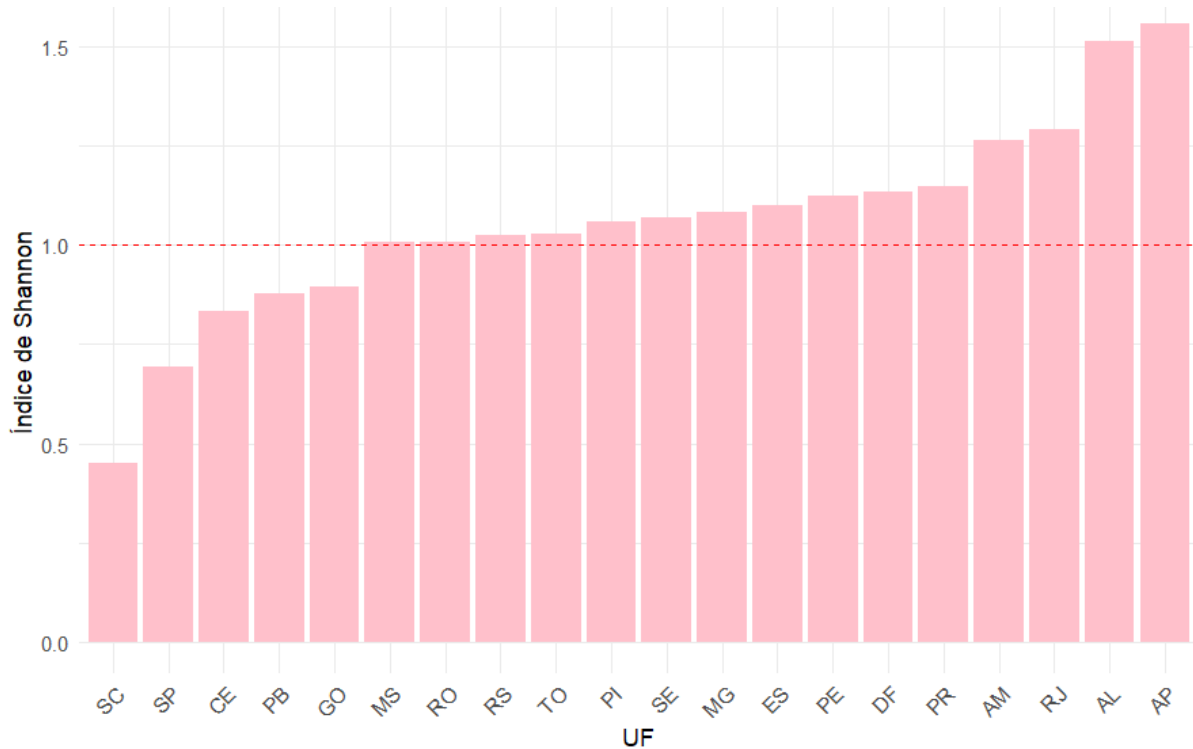
A relação entre o gráfico e a tabela reforça que estados com altos índices de Shannon, como o AP, conseguem equilibrar a predominância de categorias tradicionais com a inclusão de outras menos frequentes, ampliando sua diversidade. Em contrapartida, estados com baixos índices, como GO e AM, concentram-se principalmente na Administração Pública Estadual, o que reduz a entropia. Estados intermediários, como PI e RO, mostram uma regulação mais diversificada, incorporando várias categorias com maior equilíbrio, embora ainda predominem atributos tradicionais.

Esses resultados evidenciam que a diversidade de atributos regulados varia amplamente entre as UFs. Enquanto estados como o AP demonstram uma governança mais abrangente e inclusiva, outros, como GO e AM, permanecem centrados em estruturas tradicionais. A predominância da Administração Pública Estadual reflete uma tendência de concentração, mas há oportunidades para ampliar a inclusão de outras categorias, como Operadores de Dados, Empresas e Organizações Internacionais.

5.3 Entropia do Alvo

A análise do Índice de Shannon por Unidade da Federação (UF), com base na categorização de alvo, evidencia a diversidade dos aspectos regulatórios adotados pelos estados brasileiros. O índice mede a entropia, indicando o grau de diversidade dos alvos regulados em cada estado. O gráfico e a tabela de frequências complementam essa análise, destacando os padrões predominantes e as variações entre as UFs.

Gráfico 3- Entropia da categoria objeto nos estados brasileiros



Fonte: Elaboração Própria

No gráfico, estados como Amapá (AP) e Alagoas (AL) apresentam os maiores índices de Shannon, ultrapassando 1,5. Isso demonstra uma alta diversidade de alvos regulatórios, sugerindo que esses estados adotam uma governança inclusiva e abrangente, regulando uma ampla variedade de categorias. Em contrapartida, estados como Santa Catarina (SC) e São Paulo (SP), com índices bem abaixo de 1, indicam menor diversidade e uma concentração em poucos alvos regulatórios. Já UFs como Rio de Janeiro (RJ), Paraná (PR), Amazonas (AM) e Distrito Federal (DF), com índices próximos de 1, refletem uma diversidade moderada, equilibrando esforços entre diferentes categorias sem atingir a alta diversidade de estados como AP e AL.

Tabela 6 - Frequência da categoria alvo em todos os estados

Categoria	Frequência
Certificação e Conformidade	414
Competência	166
Padronização	85
Compartilhamento de Dados	50

Coleta e Processamento de Dados	35
Autorização	14
Direitos	4
Monitoramento	2

Fonte: Elaboração Própria

A tabela de frequências complementa a análise, evidenciando a predominância de algumas categorias de alvo sobre outras. A Certificação e Conformidade, com 414 ocorrências, é o principal alvo regulatório, representando a prioridade da maioria dos estados. Essa predominância pode explicar os baixos índices de Shannon em estados como SC e SP, que concentram seus esforços regulatórios nessa categoria.

Outras categorias relevantes incluem Competência (166 ocorrências) e Padronização (85 ocorrências), que aparecem como prioridades secundárias em estados com diversidade moderada, como PR e RJ. Entretanto, categorias como Compartilhamento de Dados (50 ocorrências), Coleta e Processamento de Dados (35), Autorização (14), Direitos (4) e Monitoramento (2) apresentam frequências significativamente menores. Essa baixa inclusão dessas categorias contribui para a menor diversidade observada em muitos estados, especialmente aqueles com índices mais baixos no gráfico.

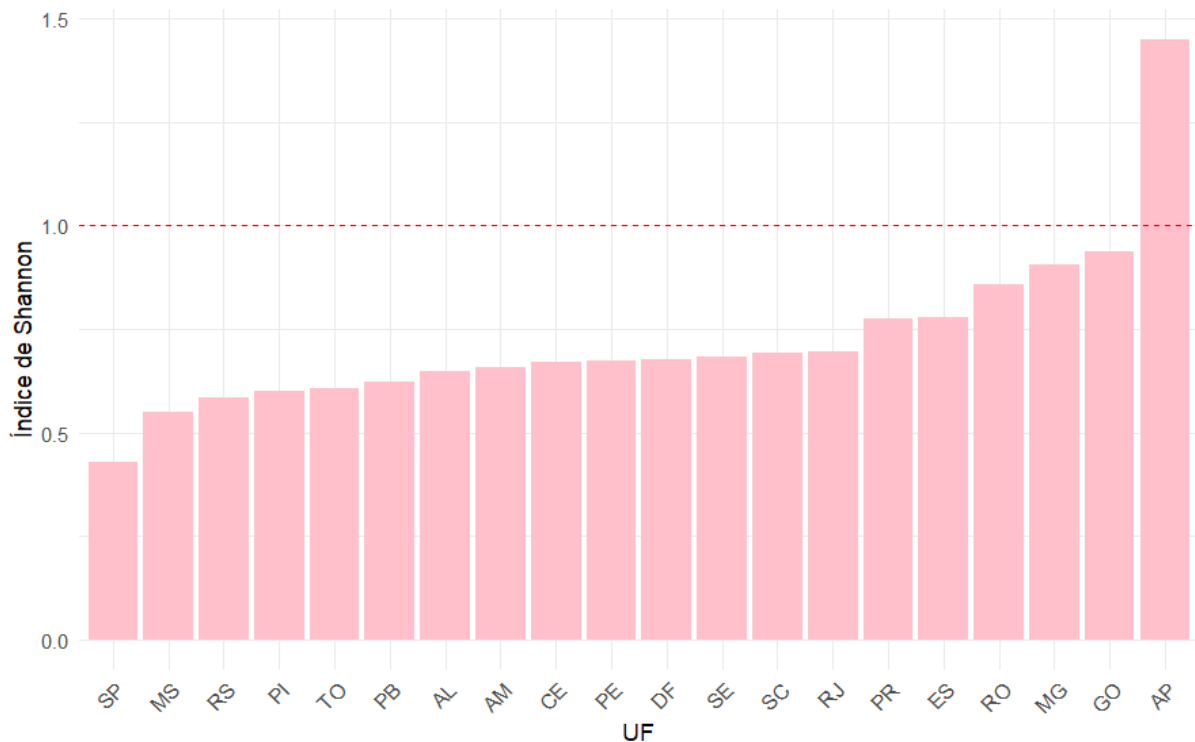
A relação entre o gráfico e a tabela evidencia que estados com altos índices de Shannon, como AP e AL, regulam uma combinação mais diversificada de categorias, abrangendo tanto as predominantes, como Certificação e Conformidade, quanto as menos frequentes, como Compartilhamento de Dados e Monitoramento. Por outro lado, estados com baixos índices, como SC e SP, concentram seus esforços em categorias dominantes, resultando em uma regulação menos diversa. Estados com índices intermediários, como RJ e DF, apresentam uma governança mais equilibrada, distribuindo seus esforços entre categorias predominantes e secundárias.

O Índice de Shannon e as frequências das categorias de alvo revelam variações significativas na abordagem regulatória das UFs. Estados como AP e AL demonstram uma governança mais inclusiva e diversa, enquanto SC e SP permanecem focados em categorias tradicionais, como Certificação e Conformidade. A baixa frequência de alvos como Direitos, Monitoramento e Coleta de Dados aponta para a necessidade de estratégias que ampliem a diversidade regulatória, garantindo maior representatividade e eficácia.

5.4 Entropia do Objeto

A análise do Índice de Shannon por Unidade da Federação (UF), com base na categorização de objeto, revela importantes informações sobre a diversidade na regulação de diferentes categorias pelos estados brasileiros. O índice mede a entropia dos objetos regulados, indicando o grau de diversidade em cada UF. A linha vermelha de referência, posicionada no valor 1, estabelece um marco para diferenciar entre níveis moderados e altos de diversidade.

Gráfico 4 - Entropia da categoria objeto nos estados brasileiros



Fonte: Elaboração Própria

O Amapá (AP) apresenta o maior Índice de Shannon, ultrapassando 1,5, o que demonstra uma alta diversidade de objetos regulados. Esse resultado sugere que o estado incorpora uma ampla variedade de categorias na governança, refletindo uma abordagem mais inclusiva e abrangente. Por outro lado, estados como São Paulo (SP) e Mato Grosso do Sul (MS) exibem índices bem abaixo de 1, indicando uma menor diversidade e uma possível concentração em categorias específicas. UFs como Goiás (GO), Minas Gerais (MG), Rondônia (RO), Espírito Santo (ES) e Paraná (PR) possuem índices moderados, próximos de 1, apontando para uma regulação mais equilibrada, com inclusão de diferentes categorias, embora sem

alcançar altos níveis de diversidade.

Tabela 7 - Frequência da categoria objeto em todos os estados.

Categoria	Frequência
Governança	407
Gestão e Processo	346
Serviço	15
Produto	2

Fonte: Elaboração Propria

A tabela de frequências das categorias de objeto complementa essa análise, dividindo os objetos em quatro grupos principais: Governança, Gestão e Processos, Serviço e Produto. A categoria Governança domina a regulação, com 407 ocorrências, evidenciando que a maioria dos estados concentra esforços em aspectos institucionais e organizacionais. Essa predominância pode explicar os baixos índices de entropia em estados como SP e MS, onde há pouca diversidade na escolha dos objetos regulados.

A categoria Gestão e Processos, com 346 ocorrências, aparece como a segunda mais frequente. Estados com índices intermediários, como GO e MG, provavelmente mantêm um equilíbrio entre governança e gestão, promovendo uma regulação mais diversa em relação a essas categorias. Em contraste, Serviço e Produto têm uma presença significativamente menor, com apenas 15 e 2 ocorrências, respectivamente. Essa baixa frequência reflete uma limitada atenção regulatória a esses aspectos, sugerindo que as prioridades estaduais estão voltadas principalmente para processos e estruturas organizacionais, enquanto a regulação de serviços e produtos é marginalizada.

A relação entre o gráfico e a tabela indica que estados com alta entropia, como o AP, provavelmente regulam uma combinação mais ampla e equilibrada de governança, gestão, serviços e produtos, o que explica seus altos índices de Shannon. Já estados com baixa entropia, como SP e MS, demonstram uma clara concentração em governança e gestão, com pouca ou nenhuma inclusão de outras categorias.

6 DISCUSSÃO

A regulamentação da Lei Geral de Proteção de Dados (LGPD) nos estados brasileiros demonstra uma significativa diversidade nas estratégias de governança de dados, revelando tanto avanços quanto lacunas no alinhamento às exigências legais. Esta análise, baseada nos decretos regulamentares estaduais, evidencia padrões distintos de restritividade, diversidade normativa e entropia.

Os estados brasileiros apresentam uma ampla variação no nível de restritividade dos enunciados institucionais em seus decretos regulamentares. A metodologia empregada, adaptada de Valle et al. (2022) e seguindo os pressupostos lançados pela ciência política brasileira com Moreira (2022), mediu a proporção de palavras normativas restritivas, como obrigações e proibições, nos textos analisados. Observou-se que estados como Alagoas, Espírito Santo e Pernambuco possuem índices de restritividade elevados, próximos de 1, o que reflete um foco acentuado em estabelecer controles rígidos e burocráticos sobre o uso de dados. Embora essa abordagem garanta maior conformidade com a LGPD e ofereça segurança jurídica, ela pode restringir a inovação e a adaptabilidade necessária para implementar políticas públicas mais dinâmicas e sensíveis às particularidades locais.

Por outro lado, a análise da entropia dos atributos, revelou disparidades na diversidade dos atores contemplados nas regulamentações estaduais. Estados como o Amapá destacaram-se por adotar uma abordagem mais inclusiva, com alta entropia, indicando uma ampla variedade de categorias de atores regulados. Essa diversidade sugere um esforço para integrar múltiplas perspectivas e interesses na governança de dados. Em contraste, estados como Goiás e Amazonas apresentaram baixa entropia, demonstrando uma regulação concentrada em poucos atores, predominantemente vinculados à administração pública estadual. Essa concentração reflete uma centralização excessiva e aponta para a necessidade de incluir categorias sub-representadas, como operadores de dados, cidadãos e organizações internacionais.

A análise dos alvos regulatórios reforçou a tendência de predominância de certas categorias, como Certificação e Conformidade, que registrou 414 ocorrências, indicando ser a principal prioridade nos decretos estaduais. Estados como Alagoas e Amapá, com os maiores índices de entropia nessa dimensão, demonstraram uma maior diversidade regulatória, abordando uma combinação mais ampla de alvos. Ao contrário de estados como Santa Catarina e São Paulo que evidenciaram baixa diversidade, com foco quase exclusivo em categorias tradicionais.

A análise dos objetos regulados nos decretos estaduais revelou uma predominância das

categorias Governança e Gestão e Processos, com 407 e 346 ocorrências, respectivamente. Estados com altos índices de entropia, como o Amapá, apresentaram uma regulação mais equilibrada e abrangente, incorporando categorias como serviços e produtos, ainda que de forma limitada. Já estados como São Paulo e Mato Grosso do Sul mostraram baixa diversidade, concentrando-se quase exclusivamente em aspectos de governança, o que pode reduzir a eficácia das regulamentações ao negligenciar aspectos cruciais para a promoção da inovação e a eficiência no uso de dados.

Essa análise evidencia um cenário regulatório marcado por contrastes entre controle burocrático e abertura à diversidade. Embora a priorização de conformidade seja essencial para garantir a segurança jurídica e o respeito aos direitos de privacidade, o foco excessivo em restrições pode limitar a capacidade dos estados de explorar plenamente o potencial dos dados para inovação e eficiência. Estados com regulamentações mais inclusivas e diversificadas têm maior potencial para criar um ambiente de governança que não apenas protege os dados pessoais, mas também promove seu uso estratégico em políticas públicas.

Portanto, os resultados obtidos destacam a necessidade de uma maior inclusão de diferentes categorias de atores e objetivos nas regulamentações estaduais. Essa abordagem pode garantir que as estratégias de governança de dados sejam ao mesmo tempo eficazes na proteção dos direitos dos cidadãos e adaptáveis às demandas de uma sociedade em constante transformação tecnológica. Assim, ao equilibrar segurança, inovação e inclusão, os estados brasileiros poderão construir uma governança de dados mais robusta e alinhada às melhores práticas internacionais.

7 CONCLUSÃO E SUGESTÕES PARA FUTURAS PESQUISAS

Este trabalho buscou compreender as estratégias adotadas pelos estados, explorando a restritividade, entropia e diversidade normativa presentes nos decretos regulamentares por meio da Gramática Institucional (IGT). A análise demonstrou não apenas os avanços alcançados, mas também as lacunas e os desafios que persistem na implementação de uma governança de dados eficaz, inovadora e inclusiva.

A LGPD marca um divisor de águas na regulação do uso de dados no Brasil, impondo diretrizes rigorosas para proteção da privacidade e segurança dos cidadãos. No entanto, a resposta dos estados a essa legislação variou amplamente, refletindo disparidades regionais e estruturais. Regiões com maior capacidade administrativa e infraestrutura, como Sul e Sudeste, apresentaram regulamentações mais completas e alinhadas às diretrizes nacionais. Por outro lado, estados de regiões como o Norte e Nordeste enfrentaram desafios significativos, incluindo limitações tecnológicas e menor priorização da proteção de dados, evidenciando desigualdades estruturais que precisam ser endereçadas para promover uma governança de dados equitativa.

A predominância de abordagens restritivas nos decretos estaduais aponta para um foco em controle e conformidade, muitas vezes em detrimento da inovação e da eficiência na gestão de dados. Embora esse enfoque garanta segurança jurídica e evite violações à LGPD, ele também restringe o potencial de exploração estratégica dos dados para formulação de políticas públicas. Estados com regulamentações mais diversas e inclusivas, como o Amapá, demonstraram maior capacidade de incorporar múltiplas perspectivas e interesses, evidenciando que a diversidade normativa pode ser um diferencial importante para uma governança mais adaptável e robusta.

Além disso, a análise destacou a necessidade de maior integração e interoperabilidade entre os diferentes níveis de governo. A governança de dados exige uma coordenação mais efetiva entre estados, municípios e o governo federal, considerando o contexto federativo brasileiro e os desafios impostos por essa estrutura. A criação de mecanismos institucionais que promovam o compartilhamento seguro de informações, sem comprometer a privacidade e os direitos dos cidadãos, é essencial para fortalecer a governança de dados no país.

Outro ponto crítico identificado foi a limitada inclusão de atores externos à administração pública estadual nos processos regulatórios. Categorias como cidadãos, organizações internacionais e empresas privadas tiveram pouca representação nos decretos analisados, evidenciando a centralização das decisões na esfera governamental. Essa abordagem pode limitar a inovação e o engajamento de diferentes stakeholders na construção

de uma governança de dados mais abrangente e participativa.

No que diz respeito às contribuições para futuras pesquisas, este trabalho abre diversas possibilidades para aprofundar o estudo sobre governança de dados no Brasil. Primeiramente, a análise comparativa entre as práticas de governança de dados no Brasil e em outros países com estruturas federativas semelhantes. Essa abordagem pode identificar boas práticas e estratégias bem-sucedidas que poderiam ser adaptadas ao contexto brasileiro. Ademais, investigações que incorporem a perspectiva dos cidadãos e outros atores externos à administração pública podem trazer contribuições relevantes para entender como diferentes grupos percebem e interagem com as regulamentações existentes.

Além de, pesquisas que utilizem abordagens tecnológicas avançadas, como machine learning e processamento de linguagem natural, poderiam enriquecer a análise de dados normativos, identificando padrões e tendências que podem passar despercebidos em abordagens mais tradicionais. Tais estudos poderiam ainda contribuir para o desenvolvimento de ferramentas analíticas mais eficazes para auxiliar gestores e legisladores na formulação de políticas públicas baseadas em dados.

A governança de dados no Brasil está em um momento decisivo. Os estados têm a oportunidade de construir arranjos institucionais que não apenas assegurem a conformidade com a LGPD, mas que também aproveitem o potencial dos dados para transformar políticas públicas, promover eficiência administrativa e fomentar a inovação. Para isso, é necessário equilibrar controle e flexibilidade, incluir diferentes atores no processo decisório e adotar práticas que fortaleçam a confiança entre o governo e a sociedade. Ao avançar nesse sentido, e com o suporte de estudos futuros, o Brasil poderá estabelecer-se como um exemplo de governança de dados eficaz e inclusiva em contextos federativos complexos.

REFERÊNCIA

ABRAHAM, Rene; SCHNEIDER, Johannes; VOM BROCKE, Jan. Data governance: a conceptual framework, structured review, and research agenda. **International Journal of Information Management**, v. 49, p. 424-438, 2019. DOI: <https://doi.org/10.1016/j.ijinfomgt.2019.07.008>. Disponível em: <https://doi.org/10.1016/j.ijinfomgt.2019.07.008>.

BENJAMIN, Ruha. Assessing risk, automating racism. **Science**, v. 366, n. 6464, p. 421-422, 2019.. Disponível em: <https://www.science.org/doi/abs/10.1126/science.aaz3873>.

BLANKE, Jordan; HILLER, Janine. Predictability for privacy in data-driven government. **Minnesota Journal of Law, Science & Technology**, v. 20, n. 1, p. 32-76, 2019. Disponível em: <https://scholarship.law.umn.edu/mjlst/vol20/iss1/3>.

BOXEBAUM, Eva; JONSSON, Stefan. **Isomorphism, diffusion and decoupling: concept evolution and theoretical challenges**. In: The Sage Handbook of Organizational Institutionalism, v. 2, p. 79-104, 2017.

CRAWFORD, Sue E. S.; OSTROM, Elinor. A grammar of institutions. **The American Political Science Review**, v. 89, n. 3, p. 582-600, 1995. Disponível em: <https://doi.org/10.2307/2082975>.

DAWES, Sharon S; CRESSWELL, Anthony; PARDO, Theresa . From "need to know" to "need to share": tangled problems, information boundaries, and the building of public sector knowledge networks. **Public Administration Review**, v. 69, n. 3, p. 392-402. 2009. Disponível em: https://doi.org/10.1111/j.1540-6210.2009.01987_2.x.

DRAPALOVA, Eliska; WEGRICH, Kai. Who governs 4.0? Varieties of smart cities. **Public Management Review**, v. 22, n. 5, p. 668-686, 2020. DOI: <https://doi.org/10.1080/14719037.2020.1718191>. Disponível em: <https://doi.org/10.1080/14719037.2020.1718191>.

FILGUEIRAS, Fernando; ALMEIDA, Virgílio A. F. **Governance for the digital world: neither more state nor more market**. 1. ed. London: Palgrave Macmillan, 2021. 117 p.
FILGUEIRAS, Fernando; FERNANDES, Cireno; PALLOTTI, Pedro. Digital transformation and public service delivery in Brazil. **Latin American Policy**, v. 10, n. 2, p. 195-219, nov. 2019. DOI: <https://doi.org/10.1111/lamp.12169>. Disponível em: <https://doi.org/10.1111/lamp.12169>.

FILGUEIRAS, Fernando; LUI, Lizandro. A gramática institucional como ferramenta de análise do desenho institucional. **Revista Brasileira de Ciências Sociais**, v. 38, n. 111, p. 1-38, 2023. DOI: <https://doi.org/10.1590/3811032/2023>. Disponível em: <https://doi.org/10.1590/3811032/2023>.

FILGUEIRAS, Fernando; LUI, Lizandro; VELOSO, Maria Tereza Trindade. A gramática institucional da proteção de dados e da privacidade no Brasil. **DADOS - Revista de Ciências Sociais**, Rio de Janeiro, v. 68, n. 1, p. e20220169, 2025. Disponível em: <https://doi.org/10.1590/dados.2025.68.1.346>. Acesso em: 27 nov. 2024.

FILGUEIRAS, Fernando; SILVA, Bruno. Desenhando políticas e governança de dados para cidades inteligentes: ensaio teórico com o uso da IAD Framework para analisar políticas orientadas por dados. **Revista de Administração Pública**, v. 56, n. 4, p. 508–528, 2022. Disponível em: <https://doi.org/10.1590/0034-761220220078>.

JANSSEN, M., KONOPNICKI, D., SNOWDON, J. L., & OJO, A. (2017). Driving public sector innovation using big and open linked data (BOLD). **Information Systems Frontiers**, 19(2), 189–195. Disponível em: <https://doi.org/10.1007/s10796-017-9746-2>

LUDERMIR, Teresa Bernarda. Inteligência artificial e aprendizado de máquina: estado atual e tendências. **Estudos Avançados**, v. 35, n. 101, p. 85-94, jan./abr. 2021. Disponível em: <https://doi.org/10.1590/s0103-4014.2021.35101.007>.

MICHELI, Marina; PONTI, Marisa; CRAGLIA, Max; SUMAN, Anna Berti. Emerging models of data governance in the age of datafication. **Big Data & Society**, v. 7, n. 2, p. 1–15, 2020. Disponível em: <https://journals.sagepub.com/home/bds>.

MICHELI, Marina; PONTI, Marisa; CRAGLIA, Max; SUMAN, Anna Berti. Emerging models of data governance in the age of datafication. **Big Data & Society**, v. 7, n. 2, p. 1-15, 2020. Disponível em: <https://doi.org/10.1177/2053951720948087>.

MOREIRA, D.; PIRES, A.; MEDEIROS, M. A. de A. Do "texto como texto" ao "texto como dado": o potencial das pesquisas em Relações. **Revista de Sociologia e Política**, v. 30, 2022. MOREIRA, Davi. **Text as Data para Ciências Sociais: guia prático com aplicações**. 1. ed. [S.l.: s.n.], 2019. Disponível em: https://bookdown.org/davi_moreira/txt4cs2/.

OSTROM, Elinor. **Understanding institutional diversity**. Princeton: Princeton University Press, 2005. Disponível em: <https://doi.org/10.2307/j.ctt7s7wm>.

PALOTTI, Pedro; FILGUEIRAS, Fernando; NASCIMENTO, Maricilene. Policy design e múltiplas evidências: proposta analítica da dinâmica da política de transformação digital dos serviços públicos da administração pública federal brasileira. **Boletim de Análise Político-Institucional**, v. 24, p. 79-89, 2020.

SIDDIKI, Saba. Assessing policy design and interpretation: an institutions-based analysis in the context of aquaculture in Florida and Virginia, United States. **Review of Policy Research**, 2014. Disponível em: <https://doi.org/10.1111/ropr.12075>.

SIDDIKI, Saba; BASURTO, Xavier; WEIBLE, Christopher M. Using the institutional grammar tool to understand regulatory compliance: the case of Colorado aquaculture. **Regulation & Governance**, v. 6, n. 2, p. 167–188, 2012. Disponível em: <https://doi.org/10.1111/j.1748-5991.2012.01132.x>.

VALLE, L. M.; DANTAS, S. G.; SILVA, D. G.; DIAS, U. S.; MONASTÉRIO, L. M. RegBR: A novel Brazilian government framework to classify and analyze industry-specific regulations. **PLOS ONE**, v. 17, n. 9, e0275282, 2022. Disponível em: <https://doi.org/10.1371/journal.pone.0275282>.