

RELATÓRIO INTELIGÊNCIA REGULATÓRIA

Mapeamento de tendências para a Agenda Regulatória
2027/2028

ANPD

Agência Nacional de Proteção de Dados

Diretor-Presidente

Waldemar Gonçalves Ortunho Junior

Diretores

Iagê Zendron Miola

Lorena Giuberti Coutinho

Miriam Wimmer

Coordenação de Planejamento Regulatório (COPR)

Coordenação-Geral de Planejamento e Análise Regulatória (CGPAR)

Superintendência de Regulação (SRE)

Denis Marcelo de Oliveira

Rodrigo Santana dos Santos

Lucas Borges de Carvalho

Equipe de elaboração

Caroline Rodrigues Lopes

Maria Gabriela Grings



ANPD

Agência
Nacional de
Proteção de Dados

Inteligência Regulatória

A Inteligência Regulatória é uma atividade do planejamento regulatório da ANPD, voltada à identificação e análise de lacunas/problemas regulatórios que podem compor as iniciativas da Agenda Regulatória, bem como apoiar a definição da estratégia regulatória da Agência.

Ela representa uma boa prática para a construção de uma atuação regulatória baseada em evidências, capaz de responder de forma estratégica e proativa aos desafios do ambiente digital no Brasil. Por meio da coleta, análise e sistematização de informações qualificadas — provenientes de fontes internas e externas, como estudos técnicos, monitoramento de tendências, demandas sociais e benchmarks internacionais — a inteligência regulatória permite identificar, com maior precisão, os temas prioritários que devem compor a Agenda Regulatória da Agência.

Histórico de versões

Versão 1.0	Junho 2026
------------	------------

1. Sumário Executivo..... 5

2. Introdução..... 6

3. Metodologia 7

3.1. Aspectos Gerais 7

3.2. Premissas 7

3.3. Experiência comparada 8

3.4. Fontes de Informação 10

3.4.1. Brasil..... 10

3.4.2. Global 11

3.4.3. Europa..... 12

3.4.4. Estados Unidos 12

3.5. Palavras-chave da pesquisa..... 13

4. Análise 14

4.1. Metodologia de análise 14

4.2. Resultados..... 16

5. Conclusão..... 22

1. Sumário Executivo

O presente relatório apresenta os resultados do mapeamento de tendências realizado pela ANPD, no âmbito das atividades de Inteligência Regulatória, com o objetivo de subsidiar o acompanhamento contínuo e antecipatório de temas emergentes relacionados ao ambiente digital e potenciais desafios regulatórios.

A iniciativa teve como propósito aplicar técnicas de coleta e análise automatizada de dados, por meio raspagem de dados, para identificar padrões, tendências e conexões temáticas a partir de informações disponíveis em fontes nacionais e internacionais. O monitoramento foi estruturado em três eixos temáticos prioritários: (i) proteção de dados e privacidade; (ii) inteligência artificial; e (iii) proteção de dados de crianças e adolescentes.

A metodologia adotada envolveu a coleta de milhares de artigos e notícias, seguida da aplicação de técnicas de processamento de linguagem natural, modelagem de tópicos e análise de redes semânticas, permitindo a identificação de relações entre assuntos recorrentes e a compreensão dos principais debates em curso no ecossistema digital.

Os resultados indicam que a inteligência artificial constitui atualmente o principal eixo estruturante das discussões analisadas, funcionando como elemento de conexão entre temas tradicionalmente tratados de forma independente, como proteção de dados pessoais, governança digital, segurança da informação e proteção de públicos vulneráveis. As análises evidenciam uma crescente integração entre as agendas de proteção de dados e inteligência artificial, com destaque para temas como governança de dados, tomada de decisão automatizada, discriminação algorítmica, IA generativa, criança e adolescentes e cibersegurança.

Os achados apontam para a necessidade de abordagens regulatórias cada vez mais transversais, capazes de considerar simultaneamente aspectos tecnológicos, jurídicos e sociais associados ao uso intensivo de dados e ao desenvolvimento de novas tecnologias. Além disso, o monitoramento demonstra a permanência da relevância das discussões relacionadas à proteção de crianças e adolescentes, especialmente diante dos impactos de sistemas algorítmicos, mecanismos de recomendação de conteúdo, processos de verificação etária e riscos associados à crescente digitalização das interações sociais.

Dessa forma, a aplicação da Inteligência Regulatória contribui para ampliar a capacidade institucional da ANPD de antecipar desafios, identificar temas prioritários e subsidiar a formulação de estratégias regulatórias baseadas em evidências, fortalecendo o processo de definição e atualização da Agenda Regulatória da Agência.

2. Introdução

A crescente complexidade do ambiente digital e a velocidade de transformação das tecnologias emergentes têm ampliado os desafios enfrentados pelas autoridades reguladoras, especialmente em setores marcados por forte dinamismo tecnológico, assimetria informacional e constante surgimento de novos riscos e de novos modelos de negócio. Nesse contexto, a Inteligência Regulatória surge como instrumento voltado à identificação antecipada de tendências, lacunas normativas, riscos regulatórios e temas emergentes relevantes para a atuação institucional da ANPD. Trata-se de atividade inserida no âmbito do planejamento regulatório da Agência destinada a subsidiar, de forma contínua, a definição de prioridades, apoiar a construção da Agenda Regulatória e fortalecer a capacidade institucional de resposta estratégica e baseada em evidências.

A abordagem fundamenta-se na coleta, sistematização e análise de informações provenientes de múltiplas fontes, incluindo estudos técnicos, produção acadêmica, *benchmarks* internacionais, monitoramento de tendências tecnológicas, sinais emergentes e demandas sociais. Seu objetivo não se limita à identificação de problemas regulatórios já consolidados, mas também à antecipação de possíveis impactos decorrentes da evolução tecnológica e das transformações do ambiente regulado. Nesse sentido, a inteligência regulatória aproxima-se das abordagens contemporâneas de governança antecipatória e inteligência estratégica desenvolvidas por organismos internacionais, especialmente pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), que vêm defendendo a adoção de modelos regulatórios mais adaptativos, responsivos e preventivos.

A experiência comparada demonstra que diferentes países e organizações internacionais já vêm utilizando mecanismos de inteligência regulatória em seus processos de formulação de políticas públicas. Reino Unido, Austrália, Portugal, Coreia do Sul, Alemanha e União Europeia desenvolveram estruturas voltadas ao monitoramento contínuo de tendências e sinais emergentes, utilizando ferramentas de prospecção regulatória para subsidiar decisões estratégicas e antecipar desafios associados às tecnologias emergentes.

Nesse contexto, o presente relatório busca estruturar metodologia e mecanismos de monitoramento e análise capazes de apoiar a atuação regulatória da ANPD, inclusive mediante utilização de ferramentas automatizadas de coleta e processamento de informações, como

scripts de raspagem de dados, definição de palavras-chave temáticas e acompanhamento sistemático de fontes estratégicas nacionais e internacionais.

3. Metodologia

3.1. Aspectos Gerais

Na regulação tecnológica, a inteligência regulatória e a governança antecipatória têm sido entendidas como instrumentos centrais para que governos consigam responder adequadamente aos desafios impostos pelas tecnologias emergentes e pela transformação digital.

Nesse contexto, o *Regulatory Policy Outlook 2025*, da OCDE, ressalta que governos devem adotar processos regulatórios mais ágeis, adaptáveis e responsivos, substituindo abordagens meramente reativas por mecanismos de antecipação e aprendizado contínuo¹. A instituição também destaca que, em cenários de fragilidade da confiança social nas instituições públicas, a regulação baseada em evidências torna-se essencial para transmitir legitimidade às decisões regulatórias².

3.2. Premissas

A inteligência regulatória pode ser definida como a *produção e utilização de conhecimento destinado a auxiliar formuladores de políticas na compreensão dos impactos presentes e potenciais futuros da ciência, tecnologia e inovação, especialmente em contextos de elevada incerteza tecnológica*³.

Para isso, a abordagem utiliza ferramentas como monitoramento de tendências, análises avançadas de dados, previsões e avaliações tecnológicas, com o objetivo de subsidiar o desenvolvimento de visões, planos e roteiros estratégicos⁴.

A metodologia proposta estrutura-se em duas etapas interconectadas. A primeira corresponde ao monitoramento de horizonte (*horizon scanning*), definido como *exploração regular de novos desenvolvimentos tecnológicos e sinais que indicam quais tecnologias ou questões sociotécnicas estão se tornando áreas de interesse (vistas como uma oportunidade ou uma preocupação*

¹ OCDE. *Regulatory Policy Outlook 2025*, OCDE Publishing, Paris, <https://doi.org/10.1787/56b60e39-en>, p. 06.

² OCDE. *Regulatory Policy Outlook 2025*, OCDE Publishing, Paris, <https://doi.org/10.1787/56b60e39-en>, p. 124.

³ OCDE, *Framework for Anticipatory Governance of Emerging Technologies*, OCDE Science, Technology and Industry Policy Papers, No. 165, OECD Publishing, Paris, <https://doi.org/10.1787/0248ead5-en>, p. 18.

⁴ OCDE, *Framework for Anticipatory Governance of Emerging Technologies*, OCDE Science, Technology and Industry Policy Papers, No. 165, OECD Publishing, Paris, <https://doi.org/10.1787/0248ead5-en>, p. 18.

potencial). Essa etapa envolve uma exploração “360º” de todos os domínios tecnológicos em estágio inicial⁵.

A segunda etapa envolve o uso de ferramentas analíticas destinadas a estruturar e interpretar informações provenientes de diferentes fontes. Entre elas, destaca-se a metodologia PESTEL (Política, Econômica, Social, Tecnológica, Ambiental e Legal), amplamente utilizada para análise de ambientes complexos e identificação de fatores externos capazes de influenciar processos decisórios. Ao permitir a organização sistemática de informações segundo múltiplas dimensões estratégicas, a abordagem contribui para uma compreensão mais abrangente do contexto regulatório, favorecendo a identificação de riscos, oportunidades e interdependências entre fenômenos tecnológicos, sociais, econômicos e normativos⁶. Em atividades de inteligência regulatória, sua utilização possibilita não apenas mapear temas relevantes, mas também avaliar sua transversalidade e potencial impacto sobre a atuação regulatória.

3.3. Experiência comparada

Diversos países já vêm incorporando essas metodologias em seus processos regulatórios. O Reino Unido utiliza *horizon scanning* na elaboração de regulação relacionada às tecnologias quânticas. Portugal estabeleceu o Centro de Competências para Planeamento, Políticas e a Prospetiva da Administração Pública. A Austrália, através do Escritório Regulador Nacional de Segurança Ferroviária, utiliza inteligência regulatória para, entre outros, fundamentar decisões regulatórias e atividades de fiscalização, além de antecipar e responder a riscos e desafios de segurança atuais e futuros, via abordagem regulatória proativa⁷.

Em conjunto com outros órgãos de governo, o Comitê de Reforma Regulatória da Coreia do Sul elaborou um “Roteiro Preventivo de Inovação Regulatória”, voltado à detecção antecipada e ao tratamento preventivo de desafios regulatórios associados às tecnologias emergentes. Nos últimos anos a Alemanha desenvolveu capacidade prospetiva, com enfoque nos efeitos das

⁵ OCDE, *Framework for Anticipatory Governance of Emerging Technologies*, OCDE Science, Technology and Industry Policy Papers, No. 165, OECD Publishing, Paris, <https://doi.org/10.1787/0248ead5-en>, p. 20.

⁶ INSTITUTE FOR INNOVATION AND TECHNOLOGY. Environment Analysis: recognizing the future with the pestel framework, disponível em <https://www.iit-berlin.de/en/environment-analysis-recognizing-the-future-with-the-pestel-framework/>

⁷ OFFICE OF THE NATIONAL RAIL SAFETY REGULATOR. *Regulatory Intelligence Strategy*, 2025, p. 04, disponível em <https://nraspricms01.blob.core.windows.net/assets/documents/Publication/ONRSR-Regulatory-Intelligence-Strategy-September-2025.pdf>

inovações tecnológicas, como ciberataques, uso de meios militares não convencionais e o papel da manipulação das redes sociais nas campanhas eleitorais⁸.

O Sistema Europeu de Análise de Estratégias e Políticas (ESPAS), em conjunto com outras instituições, realiza desde 2022, o projeto *Horizon Scanning* com o intuito de auxiliar os reguladores a anteciparem e lidarem com os desafios e as oportunidades enfrentados pela União Europeia. O objetivo não seria somente identificar riscos, mas também, oportunidades, que surgem dos chamados ‘sinais de novidade’, percebidos a partir de fontes de informação variadas, captadas por uma rede de scanners.

De acordo com o site da iniciativa, o passo seguinte é interpretar os (...) sinais *por meio de workshops, identificando possíveis tendências emergentes e os chamados "fatores de mudança". Finalmente, por meio de nossas sessões de construção de significado, buscamos compreender as possíveis consequências e implicações políticas. Esses aspectos são detalhados em nossos boletins periódicos (...)*⁹. A iniciativa já publicou 10 relatórios semestrais com os temas apontados como os mais destacados a cada período¹⁰.

A Superintendência de Regulação pretende empregar a metodologia durante o planejamento regulatório, na fase de inteligência regulatória. Para tanto, a primeira etapa consiste na escolha de palavras-chave, relacionadas à temática da proteção de dados pessoais que, na sequência, serão buscadas em fontes pré-determinadas, sites oficiais de autoridades brasileiras e estrangeiras, de entidades do terceiro setor, de centros de pesquisa, de repertórios de notícias e de empresas do setor de tecnologia.

Na sequência, com o uso de *scripts* elaborados para esta finalidade pela equipe do projeto, será empregada a técnica de raspagem de dados, durante intervalo de tempo pré-determinado. Os dados serão coletados e os resultados analisados, com descarte daqueles de menor relevância, seja pela sua pertinência temática com a proteção de dados pessoais, seja pela sua baixa reincidência nos resultados.

Os temas aferidos que, por sua centralidade, complexidade ou recorrência durante a etapa de inteligência regulatória integrarão a Agenda Regulatória da ANPD, instrumento de planejamento

⁸ OCDE. *Regulatory Policy Outlook 2025*, OCDE Publishing, Paris, <https://doi.org/10.1787/56b60e39-en>, p. 95.

⁹ JOINT RESEARCH CENTRE. *Horizon Scanning: Spotting change for future-oriented policymaking*, disponível em https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/horizon-scanning-spotting-change-future-oriented-policymaking-2024-12-04_en#:~:text=By%20regularly%20and%20systematically%20'scanning,setting%2C%20and%20broader%20foresi ght%20work.

¹⁰ Disponíveis em <https://espas.eu/horizon.html>

da atividade regulatória, que compõe o plano de gestão anual e se alinha com o plano estratégico quadrienal, ambos previstos nos arts. 17 e 21 da Lei 13.848/2019, que regulamenta, entre outros, o processo regulatório das agências reguladoras.

3.4. Fontes de Informação

3.4.1. Brasil

- Instituto Alana (www.alana.org.br)

Uma das principais entidades de defesa de direitos da infância no Brasil. Lidera o programa Criança e Consumo e foi protagonista na aprovação do ECA Digital.

- SaferNet Brasil (<https://new.safernet.org.br/>)

Central de denúncias de crimes online contra crianças. Desenvolve pesquisa e tecnologia antiabusos, incluindo detecção por IA em português.

- Data Privacy Brasil (www.dataprivacybr.org)

Referência em LGPD e temas relacionados a regulação de tecnologias digitais. Em 2026 lidera agenda sobre ECA Digital, regulação de IA e governança algorítmica — temas que caminham juntos.

- ITS Rio — Instituto de Tecnologia e Sociedade (www.itsrio.org)

Think tank de regulação digital. Produz análises sobre LGPD, plataforma da educação e impacto da IA em direitos de crianças e adolescentes.

- CETIC.br — TIC Kids Online Brasil (www.cetic.br/pesquisa/kids-online)

Pesquisa anual sobre uso de internet por crianças.

- Plan International Brasil (www.plan.org.br)

ONG global com foco especial na proteção de meninas online. Atuou diretamente na articulação do ECA Digital no Congresso Nacional.

- CGI.br — Comitê Gestor da Internet no Brasil (www.cgi.br)

Coordena e integra as iniciativas de serviços de internet no país.

- NIC.br – OBIA (<https://obia.nic.br/>)

Define políticas para a internet brasileira. Realiza o Simpósio de Crianças e Adolescentes na Internet e pauta a governança da IA no contexto nacional.

3.4.2. Global

- WeProtect Global Alliance (www.weprotect.org)

Aliança de mais de 100 países contra exploração sexual infantil online. Publica o *Global Threat Assessment*, a edição 2025 focou no crescimento exponencial de *Child Sexual Abuse Material* (CSAM) gerado por IA.

- 5Rights Foundation (www.5rightsfoundation.com)

ONG britânica que promove os cinco direitos digitais das crianças. Influenciou o *Children's Code* do ICO e defende design centrado na criança para sistemas de IA.

- Internet Watch Foundation – IWF (www.iwf.org.uk)

Líder global na remoção de CSAM. Em 2025 registrou aumento de 26.385% em vídeos de abuso gerados por IA — referência em pesquisa e tecnologia de detecção.

- Digital Policy Alert (www.digitalpolicyalert.org)

Monitoramento em tempo real de regulações digitais globais — DSA, *Online Safety Acts*, proteção de menores e legislações de IA em mais de 40 jurisdições.

- Relatório Internacional de Segurança da IA (www.internationalaisafetyreport.org)

Publicação científica de referência global sobre riscos de IA. Documenta CSAM gerado por IA, *deepfakes* e violações de privacidade como *harms* já estabelecidos.

- CIPL - Centre for Information Policy Leadership
(<https://www.informationpolicycentre.com/>)

Think tank global dedicado à privacidade de dados, uso responsável da informação e governança digital.

- OECD (<https://oecd.ai/en/>)

Fórum internacional que busca soluções para problemas comuns e identifica e promove boas práticas políticas, econômicas e sociais.

- Thorn: Digital Defenders of Children (<https://www.thorn.org/>)

Organização sem fins lucrativos focada em desenvolver tecnologia para proteger crianças contra o abuso e a exploração sexual.

- TrustArc (<https://trustarc.com>)

Plataforma global de software e serviços focada na gestão e conformidade de privacidade de dados.

3.4.3. Europa

- ICO — UK Children's Code (www.ico.org.uk)

O *Children's Code* britânico inspirou o ECA Digital, a Califórnia e a Austrália. Referência em *privacy by design* e em como aplicar essas regras a sistemas de IA.

- Future of Life Institute (www.futureoflife.org)

Publica o *AI Safety Index* semestral avaliando as principais empresas de IA em segurança, privacidade e proteção contra uso indevido — inclui riscos de CSAM.

- European Data Protection Board (https://www.edpb.europa.eu/edpb_en)

Organismo independente da União Europeia encarregado de garantir que as leis de privacidade sejam aplicadas de forma uniforme e consistente em todos os países membros.

- European Data Protection Supervisory (https://www.edps.europa.eu/_en)

Autoridade independente da União Europeia responsável por fiscalizar se as instituições e agências europeias respeitam as leis de privacidade ao tratar dados pessoais.

- CNIL – Autoridade Francesa de Proteção de Dados (<https://www.cnil.fr/en/home>)

Órgão público independente da França que regula e fiscaliza a proteção de dados pessoais.

- RIPD - Rede Ibero Americana de Proteção de Dados (<https://www.redipd.org/en>)

Fórum internacional que reúne autoridades de proteção de dados e especialistas da região, com o objetivo de promover a cooperação, o intercâmbio de experiências e a criação de leis e políticas que garantam o direito à privacidade e à proteção de dados.

3.4.4. Estados Unidos

- Stanford Center for AI Safety (<https://aisafety.stanford.edu/>)

Centro acadêmico líder em segurança de IA. Desenvolve o AIR-Bench, benchmark de risco que inclui CSAM, privacidade e vigilância como domínios críticos de avaliação.

- Center for AI Safety (www.safe.ai)

Reduz riscos de IA em escala societária. Pesquisa técnica e de governança com foco em mitigação de *harms* como *deepfakes*, vigilância e conteúdo abusivo infantil.

- Partnership on AI (www.partnershiponai.org)

Coalização de empresas e OSCs para IA responsável. Publica o *Synthetic Media Framework* e casos de uso sobre CSAM gerado por IA — referência para plataformas e reguladores.

- Center for Democracy & Technology (<https://cdt.org/>)

Equilibra liberdade de expressão e proteção infantil. Referência no debate sobre *age assurance*, vigilância parental e regulação de *chatbots* e IA para menores.

- DW Observatory — Child Safety Online (www.dig.watch/topics/child-safety-online)

Observatório da ONU/*DiploFoundation* que agrega debates globais sobre crianças, IA e o *Global Digital Compact*.

- All Tech Is Human (www.alltechishuman.org)

Coautora dos Princípios *Safety by Design* para IA generativa (com Thorn, Amazon, Google, Meta, OpenAI). Publica relatório anual de impacto de IA responsável.

- Pew Research Center (<https://www.pewresearch.org/topic/internet-technology/>)

Instituto de pesquisa norte-americano neutro e sem filiação partidária, tem como objetivo coletar e divulgar dados para compreender as grandes questões, atitudes e o comportamento da sociedade nos Estados Unidos e ao redor do mundo.

- WSGR (<https://www.wsgr.com/en/>)

É um dos maiores e mais prestigiados escritórios de advocacia do mundo, famoso por sua forte ligação com o Vale do Silício. Atua na consultoria jurídica e proteção de dados pessoais, auxiliando empresas no cumprimento de leis de privacidade e na gestão de incidentes de segurança.

3.5. Palavras-chave da pesquisa

Os termos de busca foram categorizados em três eixos temáticos principais: proteção de dados e privacidade, inteligência artificial, e crianças e adolescentes. As palavras-chave foram pesquisadas tanto em português quanto em inglês, respeitando as especificidades de cada fonte de informação:

- Proteção de Dados e Privacidade:

Proteção de dados, privacidade digital, dados pessoais, dados sensíveis, governança de dados, segurança da informação, cibersegurança, vazamento de dados, compartilhamento de dados, transferência internacional de dados, consentimento digital, rastreamento online, monitoramento de usuários, vigilância digital, vigilância algorítmica, perfilamento, publicidade comportamental, economia de dados, corretores de dados, anonimização de dados, pseudonimização, identificação biométrica, reconhecimento facial, geolocalização, privacidade infantil.

- Inteligência Artificial:

Inteligência artificial, IA, IA generativa, sistemas algorítmicos, algoritmos, automação decisória, tomada de decisão automatizada, aprendizado de máquina, modelos de linguagem, dados de treinamento, *web scraping*, conteúdo sintético, *deepfakes*, reconhecimento emocional, análise preditiva, transparência algorítmica, explicabilidade, viés algorítmico, discriminação algorítmica, governança de IA, auditoria algorítmica, supervisão humana, sistemas autônomos, assistentes virtuais, moderação automatizada.

- Crianças e Adolescentes:

Proteção infantil online, segurança digital infantil, crianças no ambiente digital, adolescentes online, controle parental, verificação etária, estimativa de idade, privacidade infantil, design persuasivo, padrões obscuros (*dark patterns*), padrões manipulativos, manipulação comportamental, recomendação de conteúdo, amplificação algorítmica, moderação de conteúdo, dependência digital, tempo de tela, gamificação, notificações *push*, reprodução automática, redes sociais e menores, educação midiática, alfabetização digital, cyberbullying, exploração infantil online, publicidade infantil digital, bem-estar digital.

4. Análise

4.1. Metodologia de análise

A solução desenvolvida integra técnicas de *web scraping*, processamento de linguagem natural, modelagem de tópicos, análise de redes semânticas e detecção de tendências emergentes.

A primeira etapa, de extração e limpeza dos dados, consistiu na coleta automatizada de conteúdos publicados em fontes previamente selecionadas. Os dados coletados foram armazenados em banco de dados estruturado, formando uma base consolidada de documentos para análise.

Após a coleta, os documentos passaram por um processo de tratamento para remoção de elementos que poderiam introduzir ruído analítico, como termos institucionais recorrentes, URLs,

expressões de navegação (*read more, subscribe, click here, etc.*), pontuação, *stopwords* em português e inglês. Essa limpeza visa garantir que as análises sejam realizadas sobre o conteúdo semântico efetivamente relevante.

Em seguida foi desenvolvido um dicionário temático estruturado contendo conceitos em português e inglês com o objetivo de padronizar terminologias e permitir a identificação automática de conceitos equivalentes em diferentes idiomas. Os conceitos foram organizados em três macrotemas: Proteção de Dados e Privacidade; Inteligência Artificial; Crianças e Adolescentes.

Os textos foram analisados utilizando correspondência semântica baseada no dicionário construído. Cada documento foi transformado em um conjunto estruturado de conceitos identificados automaticamente.

Além disso, foi realizada a análise de N-gramas, técnica que permite identificar expressões compostas relevantes que dificilmente seriam capturadas por análises em palavras isoladas. Nesse estudo, foi feita a extração de bigramas, que são combinações de duas palavras consecutivas, por exemplo, reconhecimento facial.

Na etapa de modelagem de tópicos a alocação latente de *Dirichlet* (*Latent Dirichlet Allocation - LDA*), um algoritmo de modelagem de tópicos que identifica conjuntos de conceitos que tendem a aparecer juntos em diferentes documentos. Em vez de analisar palavras individualmente, o modelo identifica estruturas temáticas latentes presentes na coleção documental. Dessa forma, é possível identificar os principais temas discutidos no conjunto de documentos.

Da mesma forma, foram construídas redes semânticas, com uso do método *Louvain*, iniciada com a elaboração de uma rede de coocorrência entre conceitos. Essa técnica permite observar conceitos centrais, assuntos frequentemente associados e ecossistemas temáticos. Na rede cada nó representa um conceito e cada conexão representa conceitos que aparecem simultaneamente em um mesmo documento.

Após a construção da rede, foi aplicado o algoritmo *Louvain* para identificação automática de comunidades dentro da rede semântica. O algoritmo identifica agrupamentos naturais de conceitos com forte relacionamento entre si, permitindo revelar ecossistemas temáticos presentes no conjunto de documentos analisados.

Na sequência, foram calculadas métricas de centralidade na rede que indicam o grau de conexões que um conceito possui. Dessa forma, é possível identificar quais conceitos atuam como pontes entre diferentes grupos temáticos.

Por fim, para análise temporal, os conceitos identificados foram organizados de acordo com a data de publicação dos documentos. Em seguida, foi calculada a frequência mensal de ocorrência de cada termo, considerando sua presença relativa no conjunto de documentos coletados em cada período. A partir desse procedimento, foi possível mensurar a participação percentual dos conceitos ao longo do tempo e identificar tendências de crescimento, estabilidade ou declínio dos temas monitorados.

4.2. Resultados

Foram analisados 2.030 artigos e notícias coletados em 1º de junho de 2026, provenientes de 31 fontes nacionais e internacionais especializadas em tecnologia, proteção de dados, inteligência artificial e direitos digitais.

Não foi estabelecido um recorte temporal prévio para a data de publicação dos conteúdos. A estratégia de coleta priorizou a captura das publicações mais recentes disponibilizadas por cada fonte, considerando, conforme a estrutura de navegação de cada site, as dez primeiras páginas de resultados ou os cem conteúdos mais recentes disponíveis no momento da extração.

A coleta de dados foi realizada por meio de procedimentos automatizados de raspagem (*web scraping*), observando as políticas de acesso e utilização definidas por cada fonte. O processo foi configurado para respeitar limites de requisições, intervalos de acesso e demais boas práticas de coleta automatizada, de modo a evitar sobrecarga dos servidores e garantir a obtenção dos dados de forma ética e responsável.

Na tabela abaixo estão os 10 conceitos mais frequentes no ambiente monitorado:

Macrotema	Conceitos	Nº de documentos	Frequência
Inteligência Artificial	Inteligência artificial	1358	67%
Proteção de dados e Privacidade	Proteção de dados	832	41%
Proteção de dados e Privacidade	Dados pessoais	713	35%
Inteligência Artificial	Discriminação algorítmica	360	18%
Proteção de dados e Privacidade	Cibersegurança	274	13%
Inteligência Artificial	IA generativa	259	13%
Proteção de dados e Privacidade	Governança de dados	250	12%

Inteligência Artificial	Algoritmos	199	10%
Proteção de dados e Privacidade	Dados sensíveis	188	9%
Inteligência Artificial	Tomada de decisão automatizada	164	8%

Tabela 1 – Conceitos mais frequentes

A análise de frequência evidencia a forte centralidade do tema inteligência artificial no conjunto de documentos monitorados, estando presente em 67% dos conteúdos analisados. Esse resultado sugere que a IA atua como tema transversal aos debates sobre governança digital, proteção de dados e direitos de crianças e adolescentes, configurando-se como o principal eixo organizador do ecossistema informacional observado.

Entre os temas de proteção de dados e privacidade, destacam-se proteção de dados, dados pessoais e cibersegurança, indicando que as preocupações relacionadas ao tratamento, à segurança e ao uso de informações pessoais permanecem centrais nas discussões contemporâneas. A presença expressiva de conceitos como discriminação algorítmica e IA generativa sugere, ainda, um deslocamento gradual do debate para os impactos sociais e éticos dos sistemas de inteligência artificial, especialmente em contextos de tomada de decisão automatizada.

Observa-se também que conceitos tradicionalmente associados à governança de dados aparecem frequentemente articulados a temas de inteligência artificial, indicando uma crescente convergência entre agendas regulatórias antes tratadas de forma relativamente independente.

Quanto à análise por modelagem de tópicos (LDA), o algoritmo requer a definição prévia do número de tópicos (k) a serem identificados. Neste projeto foram realizados testes exploratórios com diferentes valores de k, avaliando a coerência temática e a interpretabilidade dos agrupamentos gerados. O valor k = 4 foi selecionado por apresentar a melhor separação entre os temas identificados, produzindo tópicos mais consistentes e facilmente interpretáveis.

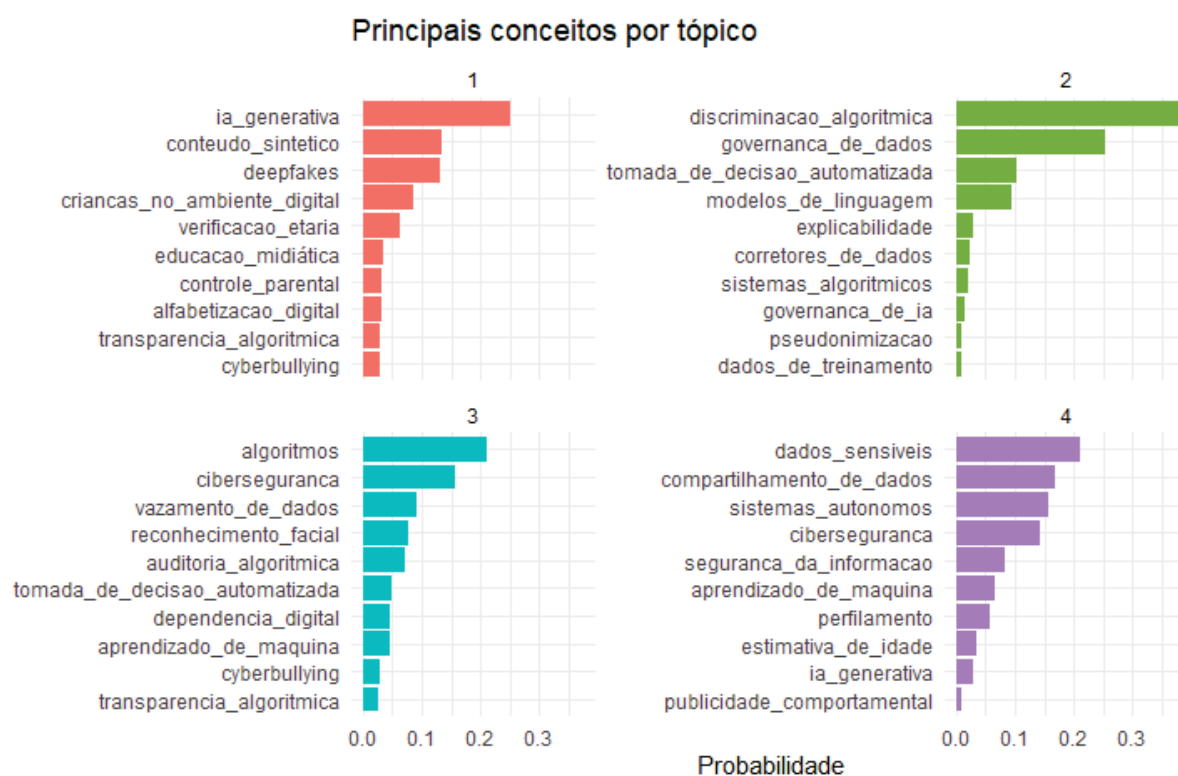


Figura 1 – Principais conceitos por tópico

A análise dos tópicos identificados pelo modelo LDA evidencia que os quatro tópicos apresentam estrutura temática híbrida, com diferentes graus de concentração entre inteligência artificial, proteção de dados e privacidade e, questões relacionadas a crianças e adolescentes.

O tópico 1 apresenta forte centralidade do tema crianças e adolescentes no ambiente digital, articulado a discussões sobre tecnologias de inteligência artificial. Esse eixo é evidenciado por um conjunto consistente de termos como crianças no ambiente digital, verificação etária, educação midiática, controle parental, alfabetização digital e cyberbullying, que estruturam a dimensão principal do tópico e indicam uma preocupação com mediação, proteção e letramento digital de públicos infantojuvenis. Embora a inteligência artificial esteja presente de forma relevante, especialmente por meio de termos como IA generativa, conteúdo sintético e *deepfakes*, sua função no tópico parece ser mais instrumental e contextual, relacionada aos riscos e transformações que essas tecnologias produzem no ambiente digital frequentado por crianças e adolescentes. Nesse sentido, a IA aparece como tecnologia estruturante do problema, enquanto o foco analítico central recai sobre seus impactos sociais e regulatórios sobre o público infantojuvenil.

O tópico 2 apresenta forte predominância de conceitos associados à inteligência artificial, especialmente relacionados a aspectos de impacto e governança, como discriminação algorítmica, tomada de decisão automatizada, explicabilidade e modelos de linguagem. Embora

a inteligência artificial seja central, observa-se também uma contribuição relevante de termos ligados à proteção de dados e privacidade, como governança de dados, corretores de dados e pseudonimização, indicando uma articulação entre sistemas algorítmicos e regulação de dados.

O tópico 3 apresenta uma configuração mais distribuída, com destaque para termos como algoritmos, cibersegurança, vazamento de dados e reconhecimento facial, combinando tanto dimensões técnicas da inteligência artificial quanto preocupações com segurança e proteção de dados. A presença de termos como dependência digital e cyberbullying sugere também uma dimensão social e comportamental associada ao uso dessas tecnologias, ainda que de forma menos central.

O tópico 4 é o que apresenta maior concentração relativa em proteção de dados e privacidade, com termos como dados sensíveis, compartilhamento de dados, segurança da informação e perfilamento. Ainda assim, há presença de elementos de inteligência artificial, como sistemas autônomos, aprendizado de máquina e IA generativa, indicando que mesmo o eixo de privacidade é atravessado por tecnologias de IA.

De forma geral, os resultados indicam que os tópicos não se organizam como compartimentos temáticos isolados, mas como estruturas interdependentes, nas quais a inteligência artificial atua como eixo transversal que atravessa os diferentes debates. A proteção de dados e privacidade aparece como um segundo eixo estruturante, frequentemente articulado às discussões sobre sistemas algorítmicos. Já o tema de crianças e adolescentes surge de forma mais localizada, com maior concentração no tópico 1, mas ainda presente em outros tópicos, indicando que ele opera como um subeixo temático conectado aos impactos sociais das tecnologias digitais.

No modelo LDA, cada tópico é representado por uma distribuição de probabilidades sobre os conceitos presentes no vocabulário da coleção documental. As probabilidades atribuídas indicam o peso relativo de cada termo na composição do tópico, permitindo identificar quais elementos são mais representativos de cada estrutura temática. Termos com maiores probabilidades desempenham papel central na interpretação e caracterização dos temas extraídos pelo modelo, funcionando como indicadores das principais dimensões semânticas associadas a cada tópico.

Posteriormente, foi realizada a análise de redes semânticas, que permite identificar de forma automática a estrutura de agrupamento dos conceitos com base em suas relações de coocorrência. Embora o modelo tenha retornado quatro comunidades, observa-se que a estrutura resultante não corresponde a uma separação temática estritamente equivalente entre os grupos. Em particular, uma das comunidades concentra simultaneamente dois macrotemas, indicando forte sobreposição semântica e articulação entre diferentes dimensões analíticas. As demais

Em contrapartida, temas como crianças no ambiente digital, tomada de decisão automatizada, algoritmos e cyberbullying apresentaram maior peso relativo até aproximadamente 2023, passando posteriormente a ocupar uma posição menos central na distribuição temática dos documentos. Esse movimento sugere um deslocamento do foco analítico, com redução relativa da sua centralidade em favor de temas emergentes ligados à governança e aos impactos das tecnologias digitais.

Por outro lado, observa-se o surgimento e a crescente consolidação de conceitos associados à proteção e ao gerenciamento de dados e às novas aplicações da inteligência artificial. Temas como dados pessoais, IA generativa, cibersegurança, compartilhamento de dados e governança de dados passam a ganhar maior relevância a partir de 2024, mantendo presença consistente nos períodos subsequentes. Esse movimento reflete o amadurecimento das discussões sobre riscos, regulação, transparência e segurança no contexto da transformação digital.

A partir de 2025 e, sobretudo, em 2026, verifica-se uma ampliação da diversidade e intensidade temática, com múltiplos conceitos apresentando alta recorrência relativa nos documentos analisados. Além da continuidade das discussões sobre inteligência artificial e governança de dados, ganham destaque tópicos como discriminação algorítmica, reconhecimento facial, vazamento de dados, *deepfakes* e conteúdo sintético, indicando uma crescente complexidade das preocupações éticas, sociais e regulatórias associadas às tecnologias digitais avançadas.

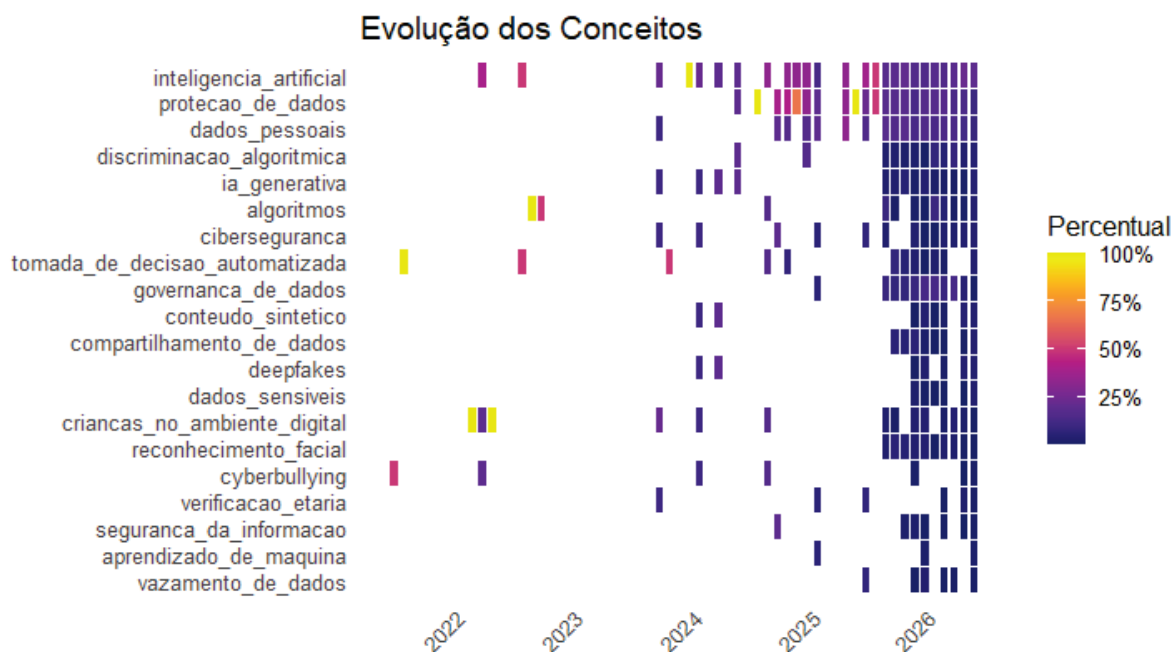


Gráfico 1 – Evolução temporal dos conceitos

Os resultados desse projeto indicam que inteligência artificial constitui atualmente o principal eixo estruturante do ecossistema informacional monitorado, conectando debates sobre proteção de dados, governança digital e direitos de crianças e adolescentes. As análises de tópicos e redes semânticas revelam uma crescente convergência entre esses domínios, sugerindo que futuras agendas regulatórias tendem a abordar tais questões de forma integrada. A metodologia demonstrou capacidade de identificar padrões temáticos relevantes e apresenta potencial para evolução em um sistema contínuo de monitoramento prospectivo.

5. Conclusão

Os resultados do primeiro monitoramento estruturado de tendências realizado pela ANPD, com base em milhares de artigos e notícias coletados, com uso de *web scraping*, processamento de linguagem natural, modelagem de tópicos e análise de redes semânticas, indicam que a inteligência artificial constitui atualmente o principal eixo estruturante dos debates observados. O tema funciona, atualmente, como elemento de conexão entre discussões tradicionalmente tratadas de forma separada, como proteção de dados, governança digital, segurança da informação e proteção de públicos vulneráveis.

As análises convergem para um mesmo diagnóstico: uma crescente integração entre as agendas de proteção de dados e inteligência artificial. Temas como dados pessoais, governança de dados, tomada de decisão automatizada, discriminação algorítmica, IA generativa e cibersegurança

aparecem de forma recorrente e fortemente associados entre si, sugerindo que futuras iniciativas regulatórias tenderão a demandar abordagens transversais, capazes de lidar simultaneamente com aspectos tecnológicos, jurídicos e sociais. Os resultados também revelam que questões relacionadas a crianças e adolescentes permanecem relevantes, especialmente quando associadas aos impactos de sistemas algorítmicos, mecanismos de recomendação de conteúdo, verificação etária e riscos decorrentes da crescente digitalização das interações sociais.

A análise temporal reforça esse cenário ao demonstrar uma mudança gradual de foco nos debates monitorados. Enquanto temas relacionados ao uso da internet por crianças e adolescentes, cyberbullying e tomada de decisão automatizada apresentavam maior destaque nos anos anteriores, observa-se, a partir de 2024, o fortalecimento de discussões relacionadas à governança de dados, IA generativa, compartilhamento de dados, cibersegurança e transparência algorítmica. Mais recentemente, ganham relevância tópicos associados a *deepfakes*, conteúdo sintético, reconhecimento facial, discriminação algorítmica e vazamentos de dados, indicando crescente preocupação com os impactos sociais, éticos e regulatórios das tecnologias digitais avançadas.

Os resultados demonstram o potencial da inteligência regulatória como mecanismo de apoio ao planejamento institucional e à definição de prioridades regulatórias. A metodologia empregada mostrou-se capaz de identificar tendências, mapear convergências temáticas e destacar temas emergentes que podem demandar acompanhamento contínuo pela ANPD. Mais do que um retrato do estado atual dos debates, o exercício fornece subsídios para uma atuação regulatória prospectiva, orientada por evidências e apta a acompanhar a rápida evolução do ambiente digital.

A continuidade e o aperfeiçoamento dessa iniciativa permitirão à ANPD manter acompanhamento sistemático das transformações do ambiente digital e regulatório, produzindo insumos continuamente atualizados para apoiar a identificação de temas emergentes, a definição de prioridades institucionais e a construção de futuras Agendas Regulatórias baseadas em evidências.