



Briefing sobre a Cúpula de Impacto de IA de 2026 ("AI Impact Summit")

Rafael A. F. Zanatta
Codiretor da Data Privacy Brasil
Dezembro de 2025

Sobre a Data

A Data Privacy Brasil é uma organização que nasce da união entre uma escola e uma associação civil em prol da promoção da cultura de proteção de dados e direitos digitais no Brasil e no mundo. Fundada em 2018, a Data Privacy Brasil Ensino surgiu como um espaço para difundir e inovar no conhecimento sobre privacidade e proteção de dados no país. Com conteúdo adaptado para uma linguagem mais prática, com exercícios e estudos de caso, trata-se de uma escola para todos aqueles que se interessam e querem se aprofundar na temática da privacidade, proteção de dados e novas tecnologias.

A Associação Data Privacy Brasil de Pesquisa é uma organização da sociedade civil, sem fins lucrativos e suprapartidária, que promove a proteção de dados pessoais e outros direitos fundamentais a partir de uma perspectiva de justiça social e das assimetrias de poder.

A partir de 2023, as duas instituições se uniram para formar uma única organização, mantendo os mesmos princípios e atividades. Com o apoio de uma equipe multidisciplinar, realizamos formações, eventos, certificações, consultorias, conteúdos multimídia, pesquisas de interesse público e auditorias cívicas para a promoção de direitos em uma sociedade datificada marcada por assimetrias e injustiças. Por meio da educação, da sensibilização e da mobilização da sociedade, buscamos uma sociedade democrática em que as tecnologias estejam a serviço da autonomia e dignidade das pessoas.

Ficha técnica

Direção

Bruno Bioni, Mariana Rielli e Rafael Zanatta

Coordenação

Carla Rodrigues, Jaqueline Pigatto, Pedro Martins, Pedro Saliba e Victor Barcellos

Equipe

Barbara Yamasaki, Bianca Marques Eduardo Mendonça, Gabriela Vergili, Giovana Andrade, Isabelle Santos, João Paulo Vicente, Larissa Pacheco, Louise Karczeski, Luize Ribeiro, Matheus Arcanjo, Natasha Nóvoa, Pedro Henrique, Rafael Guimarães, Rennan Willian, Rodolfo Rodrigues e Vinicius Silva

Licença

Creative Commons

É livre a utilização, circulação, ampliação e produção de documentos derivados desde que citada a fonte original e para finalidades não comerciais.

Imprensa

Para esclarecimentos sobre o documento e entrevistas, entrar em contato pelo e-mail imprensa@dataprivy.org

Briefing sobre a Cúpula de Impacto de IA de 2026 (“AI Impact Summit”)

A Cúpula de Impacto da Inteligência Artificial, internacionalmente conhecida como AI Impact Summit, é um evento que será realizado pelo governo da Índia em fevereiro de 2026, em Nova Déli, em continuidade às cúpulas anteriormente promovidas pelos governos do Reino Unido e da França.

Trata-se de um evento diplomático de grande porte. Fundamentada nos princípios de “Pessoas, Planeta e Progresso”, a Cúpula vislumbra um futuro em que a inteligência artificial impulse a humanidade, fomente o crescimento inclusivo e contribua para a proteção do planeta como bem comum.

As Cúpulas de IA constituem um tipo relativamente novo de evento internacional, caracterizado por um baixo grau de institucionalização formal. Tanto a Cúpula do Reino Unido, realizada em Bletchley Park, quanto a da França, em Paris, não contaram com estruturas institucionais previamente consolidadas. Esses encontros funcionam, em certa medida, como uma “feira diplomática”, na qual os governos dispõem de ampla margem de autonomia para a organização e execução do evento, ao mesmo tempo em que se desenvolvem negociações estratégicas entre Estados sobre o compromisso de sediar e conduzir edições futuras.

ÍNDICE

1. Características da Cúpula da Índia	6
2. Os debates sobre IA em Bengaluru	9
3. Os debates sobre AI Safety em Nova Delhi	13
4. Perspectivas de participação da sociedade civil na Cúpula	17
5. Líderes dos grupos de trabalho	19
Referências úteis	23

1. Características da Cúpula da Índia

A Cúpula da Índia desloca o foco do risco existencial e da segurança sistêmica (central no Reino Unido) e da governança normativa (central na França) para os efeitos concretos da IA sobre desenvolvimento, trabalho, inclusão e serviços públicos. O termo “impacto” é central e orienta o enquadramento político do evento.

A Índia se posiciona como porta-voz dos países em desenvolvimento, buscando inserir no debate internacional temas como assimetrias globais de capacidade tecnológica, dependência de infraestrutura e dados, desigualdade na distribuição de benefícios da IA, soberania digital e desenvolvimento econômico. Esse enquadramento reforça a leitura da IA como questão de justiça global, não apenas de regulação técnica.

O governo indiano propõe sete temas principais como mobilizadores da cúpula (“Summit”):



Os três princípios principais são chamados de *sutras*, e significam guias normativos para sabedoria e ação.

Com base nos três *sutras* fundamentais, as deliberações da Cúpula de Impacto da IA serão organizadas em torno de sete *chakras* interconectados, áreas de cooperação multilateral concebidas para canalizar a energia coletiva em direção a uma transformação social holística. Esses sete *chakras* representam domínios específicos de colaboração internacional que traduzem os princípios mais amplos de Pessoas, Planeta e Progresso em áreas concretas de ação:

1] Capital Humano: “A IA está transformando o cenário global do emprego, criando novas oportunidades e, ao mesmo tempo, impactando empregos e habilidades tradicionais. Respostas fragmentadas, a ausência de padrões globais de alfabetização em IA e o investimento insuficiente na transição da força de trabalho ameaçam aprofundar a ‘divisão da IA’. O Human Capital Chakra aborda esses desafios promovendo a capacitação equitativa, a transição inclusiva da força de trabalho e o amplo compartilhamento dos ganhos de produtividade impulsionados pela IA”.

2] Inclusão para o Empoderamento Social: As tecnologias de IA possuem um imenso potencial para promover avanços socioeconômicos, desde que sejam tomadas as medidas adequadas para garantir que seu design, desenvolvimento e implementação sejam inclusivos. O Chakra da Inclusão para o Empoderamento Social na Cúpula de Impacto da IA 2026 visa incentivar o design, o desenvolvimento e a implementação de sistemas de IA que considerem e acomodem ativamente as diversas necessidades, identidades e experiências de todos os usuários.

3] IA Segura e Confiável: À medida que a IA remodela as economias e as sociedades, as preocupações com a segurança, a transparência e a confiança exigem salvaguardas robustas, baseadas na tecnologia, que possam acompanhar a sofisticação e a escala dos modernos sistemas de IA. O Princípio da IA Segura e Confiável busca preencher essa lacuna de governança, capacitando todas as nações a participarem de forma significativa na supervisão da IA, ao mesmo tempo que apoia a inovação contínua e o avanço tecnológico.

4] Resiliência, Inovação e Eficiência: À medida que a infraestrutura de IA se expande, seu grande impacto ambiental e energético ameaça a resiliência climática e o crescimento inclusivo em todo o mundo. O Chakra da Resiliência, Inovação e Eficiência promove IA frugal, adaptável e sustentável, com sistemas modulares leves que funcionam em contextos de poucos recursos sem comprometer o impacto e o desempenho.

5] Ciência: A inteligência artificial está remodelando a forma como a ciência é feita, da genômica e descoberta de medicamentos à pesquisa climática e de materiais. O Science Chakra visa promover a IA na pesquisa científica de ponta e traduzir descobertas em impacto inclusivo e concreto no mundo real, fortalecendo a colaboração científica global e estabelecendo normas científicas claras e padrões de transparência para a IA na pesquisa.

6] Democratização dos Recursos de IA: As crescentes demandas do desenvolvimento de IA exigem uma vasta infraestrutura — computação e hardware, infraestrutura de dados — que hoje está concentrada em poucas regiões, reforçando as desigualdades globais na capacidade de inovação. O projeto Democratizing AI Resources Chakra busca democratizar esses recursos de infraestrutura fundamentais, melhorando a acessibilidade, o acesso e a demanda para todas as nações, a fim de fomentar um progresso tecnológico inclusivo.

7] IA para o Crescimento Econômico e o Bem Social: A IA possui um potencial significativo para catalisar o crescimento econômico e promover o progresso social inclusivo. O Chakra do Crescimento Econômico e do Bem Social na Cúpula de Impacto da IA 2026 busca promover caminhos impulsionados pela IA que não apenas fortaleçam a prestação de serviços e o acesso em setores críticos, mas também gerem contribuições mensuráveis para a produtividade, a inovação e o crescimento, promovendo assim um desenvolvimento inclusivo e equitativo.

Cada chakra representa uma esfera distinta, porém interconectada, de engajamento multilateral por meio da qual o potencial transformador da IA pode ser aproveitado para gerar benefícios tangíveis alinhados aos três sutras. Juntos, eles oferecem uma estrutura abrangente que permite que países, organizações internacionais e partes interessadas coordeneem estratégias de implementação de IA que possibilitem a maximização dos benefícios coletivos, ao mesmo tempo que abordam desafios comuns.

Para a Data Privacy Brasil, a Cúpula de Impacto é estratégica por alguns motivos:

- É a primeira Cúpula a ser realizada no Sul Global, em um país que tem exercido forte influência no Brasil na temática de transformação digital centrada na vida das pessoas e infraestruturas públicas digitais;
- A Cúpula pode oferecer uma linguagem distinta do mainstream dominante do Norte Global com relação à regulação de IA e modificar as relações geopolíticas de poder e governança de IA na esfera internacional mais sensíveis a prioridades sociais e capacidades estatais assimétricas;
- A Cúpula pode se conectar, em termos de linguagem sobre princípios, regulação e políticas públicas, a documentos que a Data Privacy Brasil auxiliou na construção, de forma mais ou menos direta, como a Declaração de São Luís de 2024 e a posição do Brasil na Declaração de Inteligência Artificial dos BRICS de 2025.

Este briefing aprofunda esses elementos e traz informações sobre como a Data Privacy Brasil está se preparando para a Cúpula de Impacto de IA em Nova Delhi em fevereiro de 2026.

2. Os debates sobre IA em Bengaluru

Em 2025, nos envolvemos com dois eventos importantes que foram contabilizados como eventos oficiais de preparação para a Cúpula. O primeiro foi o evento “Inteligência Artificial Regenerativa”, realizado em Bengaluru nos dias 31 de outubro e 01 de novembro de 2025 pela [organização IT for Change](#).

O evento “Inteligência Artificial Regenerativa” foi realizado no prédio da InfoSys em Bengaluru, um dos principais polos de inovação e tecnologia da Índia. O encontro reuniu organizações civis de várias partes do mundo para um diagnóstico sobre quais componentes sociais e éticos devem fazer parte da conversa sobre IA para que valores humanos sejam preservados, bem como “economias sociais e endógenas”. Partindo de uma forte tradição de ativismo em economia política e democracia, o encontro traçou um diagnóstico sobre combate aos monopólios digitais, ampliação das infraestruturas digitais compartilhadas, investimento em centros de pesquisa e universidades públicas e uma agenda de direitos fundamentais.



Ao longo de dois dias, os participantes mergulharam em trabalhos colaborativos em diversos laboratórios comunitários que abordaram interseções temáticas cruciais, incluindo danos aos dados, sustentabilidade e economias de IA lideradas pelo Sul.

Por meio de apresentações catalisadoras, trabalhos em grupo e “apresentações relâmpago”, os participantes colaboraram em possíveis caminhos e propostas para a construção de uma economia de IA justa. As discussões se concentraram nos elementos de um paradigma de inovação intencional que priorizaria o **conhecimento pluralista, o trabalho significativo e os limites ambientais**.

O encontro “Towards Regenerative AI” teve como método a construção de “laboratórios” que elaboraram diagnósticos sobre o que pode ser avançado na Cúpula de Impacto da Índia.

A diversidade de legislações de proteção de dados e de inteligência artificial, bem como seus distintos níveis de aplicação e fiscalização, produz impactos desiguais sobre os diferentes atores envolvidos (Estados, empresas, sociedade civil e indivíduos). Essa heterogeneidade regulatória evidencia assimetrias de informação e de poder, além de desafios práticos para a eficácia da aplicação da lei, sobretudo em contextos marcados por capacidades institucionais limitadas. Nesse cenário, torna-se central adotar abordagens de regulação contextual, capazes de reconhecer especificidades setoriais, sociais e territoriais, evitando tanto a subregulação quanto a imposição de modelos normativos pouco aderentes às realidades locais¹.

- A partir desse diagnóstico, emerge a necessidade de identificar e mensurar, com base em evidências empíricas, os danos concretos causados pelo uso de sistemas de IA. A ausência de métricas compartilhadas e de metodologias robustas de avaliação dificulta o reconhecimento institucional desses danos e, consequentemente, a formulação de respostas regulatórias proporcionais. A construção de evidências sobre impactos sociais, econômicos e ambientais da IA é, portanto, condição indispensável para sustentar decisões políticas informadas e socialmente legítimas.

Esse esforço de mensuração deve estar articulado a controles mais rigorosos sobre a coleta não regulamentada de dados, ao mesmo tempo em que se promove a coleta significativa, orientada por finalidades claramente delimitadas. O fortalecimento dos princípios de limitação da finalidade e minimização de dados não apenas reduz riscos de abusos e externalidades negativas, como também contribui para a construção de ecossistemas de dados mais confiáveis e alinhados ao interesse público. A proteção de dados pessoais é um pilar de um ecossistema informacional justo.

Nesse contexto, como foi discutido no seminário na Índia, ganha relevância a institucionalização de mecanismos de consulta e supervisão multisectorial sobre o uso e a governança de tecnologias de grande escala baseadas em dados e IA. A participação coordenada de governos, setor privado, academia e sociedade civil permite ampliar a legitimidade das decisões, incorporar múltiplas formas de conhecimento e antecipar impactos que, de outro modo, permaneceriam invisíveis nos processos decisórios exclusivamente técnicos ou estatais.

Entretanto, a efetividade dessa supervisão enfrenta um desafio estrutural: a opacidade algorítmica imposta por empresas que mantêm dados, modelos e processos decisórios sob sigilo comercial. O debate sobre transparência algorítmica, portanto, não pode ser reduzido à divulgação irrestrita de códigos, mas deve avançar na definição de obrigações proporcionais de explicabilidade, prestação de contas e acesso regulatório, compatíveis com a proteção de segredos comerciais e, ao mesmo tempo, com a salvaguarda de direitos e interesses coletivos. Especificamente sobre esse ponto, a Data Privacy Brasil tem defendido que projetos de datacenters não podem mobilizar o argumento de segredo comercial (ou industrial) para não provimento de informações

1 Sobre o tema, ver Data Privacy Brasil ([2023](#)).

ambientais relevantes para comunidades ([Bioni; Rielli; Zanatta, 2025](#)).

Para que esses mecanismos sejam operacionais, é imprescindível investir em capacitação contínua de servidores e funcionários públicos, bem como fomentar formas significativas de colaboração com a sociedade civil organizada. A assimetria técnica entre reguladores, empresas e usuários finais compromete a capacidade do Estado de exercer supervisão efetiva, tornando a formação técnica e a cooperação intersetorial elementos estruturantes da governança da IA.

Complementarmente, a criação de incentivos positivos pode desempenhar papel estratégico na indução de boas práticas. Iniciativas de reconhecimento público (como a certificação de “provedores de dados confiáveis” ou prêmios nacionais, a exemplo dos Prêmios Índia Digital) podem estimular governos, empresas e startups a adotarem padrões elevados de governança de dados e IA, deslocando parte do debate do campo exclusivamente sancionatório para uma lógica de indução e reputação.

Esse instrumento devem ser articulados a uma abordagem organizacional de baixo para cima, orientada à construção democrática da governança tecnológica. **A criação de um repositório público de danos causados por sistemas de IA**, alimentado por múltiplos atores, pode funcionar como ferramenta de aprendizado coletivo, transparência e formulação de políticas baseadas em evidências, além de ampliar a visibilidade social dos impactos negativos da tecnologia².

No campo regulatório, esse conjunto de medidas aponta para a necessidade de tornar obrigatorias auditorias algorítmicas, especialmente em sistemas utilizados em contextos de alto risco ou de interesse público. Auditorias independentes, com critérios claros e metodologias auditáveis, podem funcionar como ponte entre transparência, responsabilização e proteção de direitos, sem inviabilizar a inovação tecnológica.

O debate feito na Índia, em Bengaluru, evidenciou a necessidade de insistir em princípios específicos de governança da inteligência artificial que não se limitem a diretrizes abstratas, mas que se traduzam em práticas institucionais concretas, incluindo **a participação ativa de membros da comunidade por meio de processos de cocriação**. A incorporação de saberes locais, experiências vividas e perspectivas interseccionais é condição fundamental para que a governança tecnológica responda a contextos sociais desiguais e evite soluções impostas de forma verticalizada.

Nesse sentido, a construção de estruturas de incentivo adequadas emerge como elemento estratégico para a formação de um ecossistema de IA verdadeiramente plural, que não seja dominado pelas grandes empresas de tecnologia. Políticas públicas e mecanismos regulatórios precisam criar condições para a participação de pequenos desenvolvedores, iniciativas comunitárias, startups e instituições públicas,

2 Projeto semelhante está sendo articulado pela “Biblioteca de Danos em IA” do projeto “IA com Direitos” da Data Privacy Brasil.

reduzindo a concentração de poder econômico e decisório que hoje molda de forma desproporcional os rumos da inovação tecnológica.

Essa reconfiguração institucional exige, simultaneamente, uma mudança de mentalidade no próprio ritmo da inovação, deslocando o foco do “crescer rápido” para a lógica de “ir devagar, construir melhor”. Considerar o intervalo temporal entre a implementação de uma tecnologia e a possibilidade efetiva de contestação pública é crucial, assim como incorporar avaliações contínuas ao longo de todo o ciclo de inovação, desenvolvimento e implementação, e não apenas em momentos posteriores, quando os impactos negativos já se consolidaram.

A educação desempenha papel central nesse processo. Tornar acessível, em todos os níveis de ensino e formação profissional, a compreensão sobre a natureza, o funcionamento e as implicações da IA é condição para reduzir assimetrias de conhecimento e fortalecer a capacidade coletiva de discernimento. Esse esforço educativo deve também promover a conexão entre discursos hoje fragmentados, questionando, por exemplo, se os engenheiros e desenvolvedores compreendem plenamente os efeitos sociais e políticos dos sistemas que criam.

O debate destacou, ainda, que as salvaguardas não podem ser tratadas como camadas externas ou corretivas adicionadas a posteriori, mas **como elementos constitutivos da própria infraestrutura tecnológica**. A responsabilidade, nesse sentido, precisa ser codificada nos sistemas desde sua concepção, incorporando mecanismos de prestação de contas, rastreabilidade e explicabilidade como **requisitos estruturais**, e não opcionais.

Para sustentar essa abordagem, propõe-se a criação de instâncias permanentes de escuta e supervisão, como ouvidorias ou comitês de IA com representação interseccional. Esses espaços podem funcionar como canais institucionais de mediação entre desenvolvedores, gestores públicos e comunidades afetadas, contribuindo para reduzir a distância entre decisões técnicas e seus impactos sociais concretos.

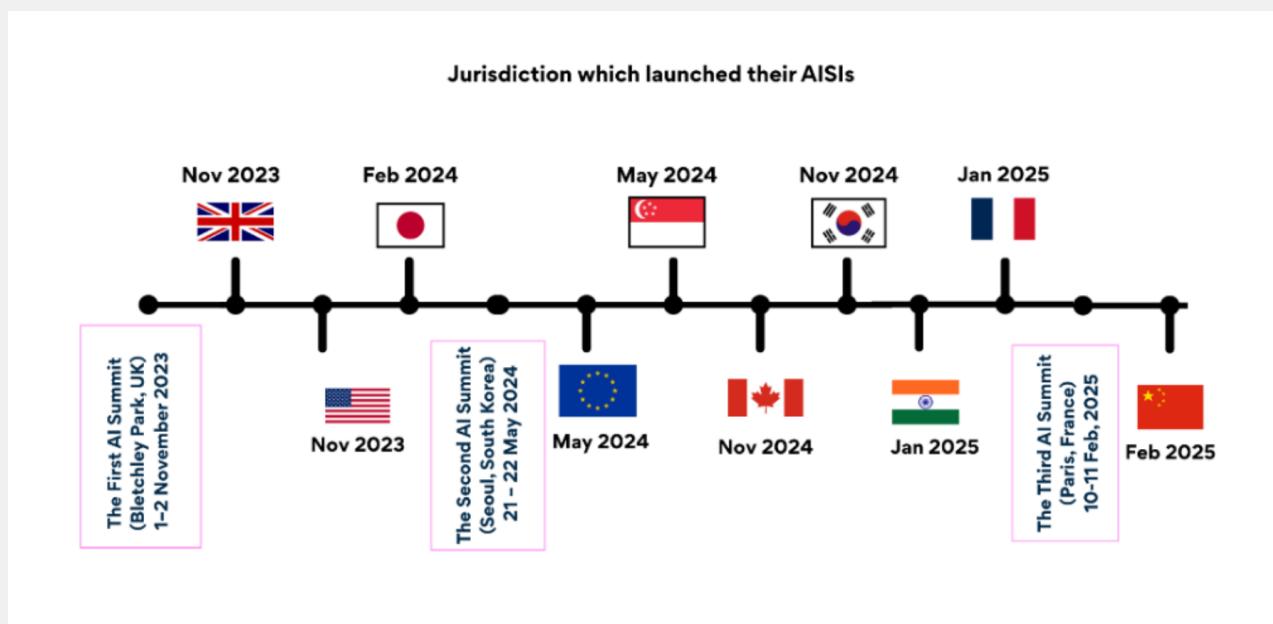
No campo das ferramentas regulatórias, a exploração de modelos de certificação de segurança em IA aparece como uma possibilidade promissora para estabelecer padrões mínimos de confiabilidade e proteção, especialmente em contextos de alto risco. No entanto, o debate alertou que tais mecanismos só serão efetivos se acompanhados de capacidade crítica e institucional para avaliar não apenas a conformidade técnica, mas também a adequação social e ética das soluções certificadas.

3. Os debates sobre AI Safety em Nova Delhi

O segundo evento de preparação que participamos foi o encontro "Operationalising AI Safety" no dia 06 de novembro de 2025, com apoio do Center for Communication Governance (CCG) e da Konrad Adenauer Stiftung.

O painel "AI Safety in the Global South" contou com a participação do Dr. Samuel Segun (Global Center on AI Governance), Edrine Wanyama (CIPESA), Dr. Jun E-Tan (Khazanah Research Institute), Mert Chandana (LIRNEasia) e Kavitzha Kuhhi (OpenAI).

O evento destacou a emergência dos "AI Safety Institutes" como um fenômeno acoplado ao surgimento das Cúpulas.



Algumas jurisdições no Norte Global, como o Reino Unido e a UE, estabeleceram seus Institutos de Segurança da IA (AISIs) dentro de órgãos governamentais existentes. Em contraste, algumas nações do Sul Global adotaram uma abordagem ligeiramente diferente para as estruturas de governança de seus AISIs. Por exemplo, a Associação Chinesa de Segurança e Desenvolvimento de IA (CNAISDA) é um consórcio de várias organizações de pesquisa chinesas existentes, incluindo think tanks, universidades e instituições apoiadas pelo Estado, garantindo contribuições de uma ampla gama de participantes.

Pioneira no Sul Global, a AISI da Índia, moldada em um modelo de centro e periferia, foi concebida para operar de forma descentralizada, facilitando o engajamento com uma rede de partes interessadas, incluindo startups, indústria, organizações da sociedade civil, divisões governamentais e academia, com supervisão do Ministério da Eletrônica

e Tecnologia da Informação. No âmbito do pilar de IA Segura e Confiável da “Missão IndiaAI”, o governo indiano visa concentrar-se na pesquisa nacional e promover a governança da IA com o estabelecimento do Instituto de Segurança. Por meio deste Instituto, o governo pretende desenvolver ferramentas, diretrizes e padrões destinados a promover o desenvolvimento e a implantação de IA segura e confiável ([CCG, 2025](#)).

O encontro debateu a inexistência de “AI Safety Institutes” em países do Sul Global e discutiu os resultados de uma investigação empírica conduzida pelo CCG:

Country	Research & Development	Testing & evaluation of AI systems	Developing policies and guidelines	Developing standards and benchmarks	International collaborations*	Outputs **
CANADA (Canadian AI Safety Institute)	Yes, and is developing research projects ⁸¹ on combating misinformation and ensuring safety in AI systems	Yes, recently contributed ⁸² to a joint testing exercise to evaluate AI agents	Unclear	Unclear	Collaborating with Australia on a research agenda focused on mitigating risks ⁸³ from AI-generated content *	Published results for the joint testing exercises on agentic safety ⁸⁴ and multilingual safety ⁸⁵ evaluations performed with other members from the International Network
INDIA (IndiaAI Safety Institute)	Yes, aims to drive R&D ⁸⁶ on indigenous tools, technologies, benchmarks, measurement systems, and other outputs to mitigate AI risks	Unclear	Yes, intends to prepare India-specific guidelines ⁸⁷ on AI governance	Yes, developing standards ⁸⁸ contextualised to India's unique challenges and opportunities	Yes, aims to collaborate with other jurisdictions ⁸⁹	Upcoming AISI. Hence, there have been no publications at this stage.
BRAZIL	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
KENYA	Not applicable	Not applicable	Not applicable	Not applicable	Conducted joint testing exercises on agentic safety ⁹⁰ and multilingual safety ⁹¹ evaluations with other jurisdictions from the International Network of AISIs	Published results for the joint testing exercises on agentic safety ⁹² and multilingual safety ⁹³ evaluations performed with other members from the International Network
CHINA (China AI Safety and Development Association)	Yes, aims to conduct research ⁹⁴ on AI safety.	Unclear ⁹⁵ , does not appear to be carrying out these functions	Unclear ⁹⁶ , does not appear to be carrying out these functions	Unclear ⁹⁷ , does not appear to be carrying out these functions	Yes ⁹⁸ , as per the vision statement	Unclear

* In addition to other collaborations, they have internationally collaborated with the members of the International Network of the AISIs

** Has also produced a joint output as an outcome of the international collaboration with the members of the International Network of the AISIs

Conforme discutido no evento do CCG, as funções de um AISI normalmente incluem pesquisa e desenvolvimento (“P&D”), o desenvolvimento de políticas ou diretrizes e padrões de segurança, e colaboração e coordenação global e nacional por meio de parcerias e redes. As funções de P&D, especialmente as atividades técnicas, como testes, avaliação e benchmarking, exigem infraestrutura e expertise técnica de ponta, reservas de armazenamento massivas, mecanismos de acesso a dados proprietários e medidas de segurança para proteger dados sensíveis. A eficiência de um AISI depende, portanto, de sua capacidade em termos de recursos humanos, financeiros e de infraestrutura.

Alguns AISIs do Norte Global estão concentrando a pesquisa e o desenvolvimento em prioridades de seus países, em domínios como segurança cibernética e riscos químico-biológicos. Da mesma forma, esses AISIs também estão priorizando preocupações específicas do Norte Global no desenvolvimento de políticas e padrões de segurança. Por exemplo, influenciados pelas áreas de foco nas discussões de Bletchley Park, houve uma forte ênfase nos riscos decorrentes da IA de ponta, como a futura perda de controle sobre os sistemas de IA e a aquisição de armas biológicas.

A produção recente do Centre for Communication Governance (CCG) permite formular uma crítica consistente à concepção de AI Safety consolidada a partir da Cúpula do Reino Unido, em Bletchley Park, sem descartá-la por completo. O ponto central não é negar a relevância dos riscos associados a modelos de fronteira ou a cenários de dano catastrófico, mas demonstrar que essa abordagem, quando tomada como eixo dominante da governança global da inteligência artificial, produz um estreitamento conceitual e político do próprio significado de safety. Ao privilegiar riscos futuros, altamente abstratos e concentrados em poucas tecnologias de ponta, **essa concepção tende a marginalizar danos já materializados**, especialmente nos países do Sul Global, onde sistemas de IA são empregados em **contextos de desigualdade estrutural, fragilidade institucional e assimetrias profundas de poder** ([CCG, 2025](#)).

Os documentos do CCG insistem que superar essa concepção exige uma mudança de foco da segurança existencial para uma noção de segurança sociotécnica. Nessa perspectiva, AI Safety passa a abranger não apenas a contenção de eventos extremos, mas a **prevenção de danos cotidianos e sistêmicos**, como discriminação algorítmica, vigilância desproporcional, exclusão de populações de políticas públicas e **erosão de direitos fundamentais**. A segurança deixa de ser compreendida como um atributo técnico dos sistemas e passa a ser entendida como um resultado institucional e social, dependente de contextos específicos, práticas de governança e capacidades regulatórias. Essa ampliação conceitual é central para tornar o debate sobre AI Safety relevante fora do eixo euro-atlântico ([CCG, 2025](#)).

Outro deslocamento fundamental proposto pelo CCG diz respeito à arquitetura de governança que sustenta as iniciativas de AI Safety. A abordagem promovida pela Cúpula do Reino Unido tende a concentrar autoridade decisória em Estados com alta capacidade tecnológica e em grandes empresas desenvolvedoras de IA, reforçando uma lógica tecnocrática e centralizada. Em contraste, os textos do CCG defendem uma governança multisectorial substantiva, na qual sociedade civil, pesquisadores de diferentes campos do conhecimento e comunidades diretamente afetadas pelos sistemas de IA participem de forma estruturada da definição de riscos, prioridades e mecanismos de mitigação. Essa abertura não é apresentada como um ideal normativo abstrato, mas como uma condição prática para a legitimidade e a eficácia das políticas de segurança em contextos marcados por diversidade social e institucional.

A Data Privacy Brasil participou dos debates e trouxe dois argumentos para o debate sobre AI Safety, com vistas à Cúpula de Impacto da Índia.

O primeiro argumento sustenta que *AI Safety* deve necessariamente abranger a segurança do trabalho de dados, incluindo a dignidade, a integridade física e psíquica e os direitos laborais das pessoas envolvidas nas **etapas invisibilizadas da cadeia produtiva da inteligência artificial**. A Data Privacy Brasil evidencia, em documentos apresentados no âmbito do T20 Brasil e na [Declaração de São Luís sobre IA \(2024\)](#), que sistemas de aprendizado de máquina dependem estruturalmente de atividades humanas como rotulagem de dados, moderação de conteúdo e processos de “higiene de dados”, frequentemente realizados em condições precárias, terceirizadas

e desprovidas de proteção adequada. Esses trabalhadores são expostos de forma recorrente a conteúdos violentos, sexualmente explícitos, discriminatórios ou traumáticos, sem garantias suficientes de apoio psicológico, transparência contratual ou reconhecimento social de sua função.

Ao incorporar essas dimensões, pode-se romper com a ideia de que AI Safety diz respeito apenas ao comportamento final do sistema ou à mitigação de riscos para usuários abstratos. A segurança passa a ser entendida como um atributo de toda a cadeia sociotécnica, **desde a coleta e tratamento de dados até a implementação dos sistemas**. Nesse enquadramento, não é possível afirmar que um sistema de IA é “seguro” se sua produção depende da exploração sistemática de trabalho humano precarizado ou da exposição continuada de pessoas a danos psíquicos. A segurança, portanto, torna-se indissociável da dignidade do trabalho e da proteção daqueles que sustentam materialmente a chamada “automação inteligente”, ainda que permaneçam invisíveis nas narrativas dominantes sobre inovação.

O segundo argumento defendido pela Data Privacy Brasil propõe que o conceito de “AI Safety” deve abarcar **uma concepção ampliada de ecossistema informacional e democracia**, na qual a segurança não se limita a evitar falhas técnicas ou usos maliciosos isolados, mas envolve a proteção das condições comunicacionais que tornam possíveis processos democráticos legítimos. A partir das evidências produzidas pelo [“Observatório da IA nas Eleições” \(2025\)](#), desenvolvido em parceria com o Aláfia Lab, temos debatido como sistemas de IA generativa, mecanismos de recomendação e estratégias automatizadas de amplificação de conteúdo sintético podem produzir efeitos sistêmicos de desinformação, inundação informacional e manipulação do debate público, afetando diretamente a integridade de pleitos eleitorais e a confiança nas instituições democráticas.

Nessa perspectiva, “AI Safety” pode ser concebida como uma questão de ordem democrática comunicacional. O risco não reside apenas em conteúdos falsos isolados e na sua circulação, mas na capacidade de sistemas automatizados de distorcer ecossistemas informacionais inteiros, tornando-os hostis à deliberação pública, à pluralidade de vozes e à formação livre da vontade política. A segurança em IA exige mecanismos estruturais de contenção da produção e circulação massiva de desinformação, bem como instrumentos de transparência, responsabilização e auditoria dos sistemas que mediam o fluxo de informações em larga escala.

4. Perspectivas de participação da sociedade civil na Cúpula

As Cúpulas de IA não possuem procedimentos formalizados de participação da sociedade civil e possuem uma diferença enorme com os espaços de governança da Internet, como Internet Governance Forum da ONU (onde IA é um tema já discutido há anos).

As perspectivas de participação da sociedade civil são muito limitadas. As entidades civis precisam “cavar o próprio espaço” e articular contatos com think tanks e servidores públicos com contato direto com agentes estatais. Uma das formas de facilitação desse contato é a mediação com os Ministérios de Relações Exteriores.

A sociedade civil organizada tem feito um esforço coordenado para ampliação da participação civil na Cúpula. De um lado, [o CCG realizou uma parceria com a Global Network Initiative e conseguiu recursos junto à ISOC](#) para criar um projeto de participação multissetorial na Cúpula de Impacto da Índia (“Multistakeholder Approaches to AI Governance”)³. Do outro lado, redes como a [Aliança do Sul Global](#) têm criado grupos de trabalho para construção de consensos e publicação de cartas de posição para influenciar a redação de textos finais da Cúpula.

Há algumas estratégias que podem ser colocadas em marcha para ampliar a participação da sociedade civil na Cúpula.

A primeira diz respeito à institucionalização mínima de canais informais, mesmo na ausência de regras formais de participação. Se as Cúpulas não reconhecem oficialmente a sociedade civil como *stakeholder*, torna-se estratégico transformar práticas recorrentes em expectativas normativas. Isso pode ser feito, por exemplo, pela produção sistemática de papers, issue briefs e propostas de linguagem normativa que circulem antecipadamente entre delegações estatais e organizadores da Cúpula. Quando esses documentos passam a ser referenciados, citados ou incorporados em versões preliminares de textos finais, cria-se um precedente informal que, ao longo do tempo, aumenta o custo político da exclusão da sociedade civil.

Um segundo caminho envolve o fortalecimento da mediação diplomática como estratégia. A interlocução com Ministérios de Relações Exteriores pode ser aprofundada por meio da formação de coalizões estáveis entre organizações da sociedade civil e diplomatas sensíveis ao tema da governança democrática da IA. Isso inclui oferecer insumos técnicos, análises comparadas e propostas redacionais que facilitem o trabalho das delegações, especialmente de países do Sul Global com menor capacidade técnica

3 Rafael Zanatta, codiretor da Data Privacy Brasil, faz parte do conselho consultivo do projeto.

interna. Ao se tornarem fontes confiáveis de expertise, as entidades civis deixam de ser vistas apenas como demandantes de espaço e passam a ser reconhecidas como parceiras funcionais no processo diplomático, colocando o caráter multissetorial em prática no nível multilateral.

Um terceiro eixo diz respeito à produção coordenada de conhecimento estratégico orientado especificamente para o ciclo das Cúpulas. Além de cartas de posição e declarações conjuntas, a sociedade civil pode investir em análises comparativas entre versões sucessivas dos documentos finais, destacando publicamente quais contribuições foram incorporadas, quais foram diluídas e quais foram ignoradas. Esse tipo de rastreamento discursivo, quando bem documentado e amplamente divulgado, aumenta a visibilidade do papel da sociedade civil e cria mecanismos indiretos de accountability. Ao demonstrar empiricamente sua influência, a sociedade civil passa a disputar não apenas conteúdo, mas também legitimidade no processo⁴.

Outro elemento relevante é a articulação transversal entre regimes de governança, aproximando o debate das Cúpulas de IA de espaços já legitimados de participação multissetorial. Conectar agendas, princípios e práticas do IGF, do WSIS+20 e de fóruns regionais de governança da Internet às discussões da Cúpula de Impacto da Índia ajuda a evidenciar a anomalia institucional representada pela exclusão da sociedade civil, caso ela se confirme na Índia em fevereiro de 2026. Essa estratégia não busca importar mecanicamente o modelo do IGF, mas usar sua legitimidade acumulada como parâmetro normativo para questionar o desenho fechado das Cúpulas de IA, e preparar o campo para novos espaços como os AI Global Dialogues, no âmbito da ONU.

Além disso, como destacado pelo trabalho da Aliança do Sul Global, a efetiva participação da sociedade civil enfrenta problemas como (i) dificuldades para obtenção de visto pelo governo da Índia e (ii) viabilização de recursos para presença em Nova Delhi em fevereiro de 2026.

O trabalho da Aliança do Sul Global será estratégico para garantir que os membros indianos – em especial CCG e Aapti Institute - podem ser “canais de vocalização” dos interesses das 26 entidades que compõem a Aliança e possam aproximar tais entidades dos representantes governamentais da Índia responsáveis pela elaboração dos documentos (os pen holders, no jargão do ativismo).

Apesar das incertezas sobre a efetividade do trabalho da sociedade civil, o fato de ser uma Cúpula direcionada a assuntos de justiça social e realizada em um país do Sul Global (Índia) tem produzido um relevante efeito coletivo de pressão para que ONGs possam estar presentes e organizem um conjunto de eventos paralelos (*side events*) no entorno da Cúpula.

4 Essa é uma etapa que pode ser realizada em 2026 após a conclusão da Cúpula e após análise dos documentos finais produzidos pelo governo da Índia e por diplomatas envolvidos no processo.

5. Líderes dos grupos de trabalho

Capital Humano

Prof. TG Sitharam

O Prof. (Dr.) T G Sitharam assumiu a presidência do Conselho de Educação Técnica de Toda a Índia (AICTE), do Governo da Índia, em 21 de dezembro de 2022. Ele é um líder acadêmico visionário com mais de 35 anos de experiência moldando a educação técnica, capacitação, pesquisa e inovação. É o arquiteto de modelos escaláveis e integrados à tecnologia, e liderou as reformas da NEP 2020, parcerias com a indústria e projetos nacionais. Estabeleceu Centros de Excelência em IA, drones, cibersegurança e tecnologias emergentes, transformando instituições e impactando milhões de pessoas por meio de iniciativas digitais e de base. Com mais de quatro décadas de contribuições pioneiras em engenharia civil e geotécnica, desenvolvimento de infraestrutura e políticas de educação técnica, o Prof. T. G. Sitharam combina profundidade científica com visão de futuro para o desenvolvimento nacional. Ele está classificado entre os 2% melhores cientistas do mundo (Stanford) e é reconhecido como um líder raro que conecta a academia, o governo e a inovação comunitária.

Inclusão para Empoderamento Social

Mr Rajesh Aggarwal

O Sr. Rajesh Aggarwal ocupa atualmente o cargo de Secretário do Departamento de Empoderamento de Pessoas com Deficiência (DEPwD). Em diversas funções nos níveis federal e estadual, ele tem defendido o uso da tecnologia para aprimorar a governança, desempenhando um papel fundamental na criação e no aprimoramento do Digilocker, do Jeevan Praman e de outros importantes programas de governo eletrônico. O Sr. Aggarwal também possui uma trajetória notável na promoção da acessibilidade de serviços e infraestrutura públicos para pessoas com deficiência, incluindo a garantia de que sites governamentais e espaços públicos sejam inclusivos para todos.

IA Segura e Confiável

Prof. Balaraman Ravindran

O Prof. Balaraman Ravindran é o fundador e diretor da Escola Wadhwani de Ciência de Dados e IA, do Centro Robert Bosch para Ciência de Dados e IA e do Centro para IA Responsável (CeRAI) do IIT Madras. Com mais de três décadas de experiência em aprendizado de máquina e aprendizado por reforço, ele é uma das principais vozes da Índia em pesquisa de IA ética e responsável. Seu trabalho atual abrange aprendizado por reforço profundo, equidade algorítmica e governança de IA. Ele integra o grupo consultivo do governo indiano sobre IA e fez parte do comitê que elaborou o relatório FREE-AI para o Banco Central da Índia. Ele tem colaborado extensivamente com instituições de pesquisa globais e parceiros da indústria para promover sistemas de IA confiáveis e sensíveis ao contexto. O Prof. Ravindran é membro da Academia Nacional de Engenharia da Índia (INAE) e da Associação para o Avanço da Inteligência Artificial (AAAI) e é amplamente reconhecido por sua liderança no avanço da pesquisa e educação em IA na Índia.

Resiliência, Inovação e Eficiência

Mr Pankaj Agarwal

Shri Pankaj Agarwal, Secretário do Ministério da Energia do Governo da Índia, tem sido um arquiteto fundamental da transição energética da Índia, impulsionando reformas focadas em sustentabilidade, eficiência e inovação. Com vasta experiência em políticas públicas, infraestrutura e transformação digital, ele traz uma perspectiva única para o Grupo de Trabalho — conectando a eficiência da IA com a resiliência energética e o desenvolvimento nacional, e reforçando o compromisso da Índia com uma IA responsável e consciente dos recursos, que gere impacto sustentável.

Ciência

Prof. Abhay Karandikar

O Prof. Abhay Karandikar é atualmente Secretário do Departamento de Ciência e Tecnologia (DST). Antes de ingressar no DST em outubro de 2023, atuou como Diretor do IIT Kanpur de 18 de abril de 2018 a 25 de setembro de 2023 (cedido pelo IIT Bombay). Também foi Decano (Assuntos Docentes) e Chefe do Departamento de Engenharia Elétrica do IIT Bombay. Ele liderou um esforço nacional para a criação da Telecom Standards Development Society of India (TSDSI), o órgão de padronização de telecomunicações da Índia, com a participação de todas as partes interessadas. O Prof. Karandikar foi membro fundador e ex-presidente da TSDSI. Ele também foi membro (em tempo parcial) da Autoridade Reguladora de Telecomunicações da Índia (TRAI) de janeiro de 2018 a janeiro de 2021. O Prof. Karandikar possui diversas patentes concedidas e pendentes, contribuições para os padrões IEEE e 3GPP, capítulos em livros e um grande número de artigos publicados em periódicos e conferências internacionais. O Prof. Karandikar foi agraciado com a Medalha de Padrões da IEEE SA em dezembro de 2016, em Nova Jersey. Sua equipe também ganhou o prêmio Mozilla Open Innovation Challenge em março de 2017 por seu trabalho em banda larga rural e empoderamento digital na Índia rural.

Democratização de Recursos de IA

Mr Saurabh Garg

O Dr. Saurabh Garg ocupa atualmente o cargo de Secretário do Ministério de Estatística e Implementação de Programas (MoSPI) do Governo da Índia. Ele desempenhou um papel fundamental na definição do cenário da governança digital na Índia, principalmente como CEO da Autoridade de Identificação Única da Índia (UIDAI), onde impulsionou o programa Aadhaar e sua integração aos sistemas de assistência social. Ao longo de sua carreira, o Dr. Garg trabalhou extensivamente com financiamento de infraestrutura, transformação digital e reformas econômicas, contribuindo para iniciativas sobre Transferências Diretas de Benefícios, política de Investimento Estrangeiro Direto (IED) e inovação em governança nos níveis nacional e estadual.

IA para Crescimento Econômico e Bem Social

Ms. Debjani Ghosh

A Sra. Debjani Ghosh é uma Distinguished Fellow da NITI Aayog e a Arquiteta-Chefe do NITI Frontier Tech Hub, o pioneiro think tank da Índia dedicado a impulsionar a prontidão do país em tecnologias de ponta para o crescimento econômico acelerado e o desenvolvimento social. Com mais de 28 anos de experiência em liderança, incluindo sua atuação como a primeira mulher presidente da Nasscom e ex-diretora da Intel para o Sul da Ásia, ela esteve na vanguarda da ascensão da Índia como uma potência global em inovação. Na Nasscom, Debjani redefiniu o foco da organização em tecnologias de ponta, defendeu o ecossistema de tecnologia avançada da Índia e consolidou a posição do país como um parceiro confiável para inovação sustentável e centrada no ser humano. Sua visão da "Techade" e a estratégia "Inevitable India" inspiraram uma narrativa global sobre a liderança da Índia na transformação impulsionada pela tecnologia. Reconhecida entre as "Primeiras-Damas" da Índia pelo Presidente do país, ela foi destaque na lista "Líder de Tecnologia do Ano" da Vogue e na lista "40 acima de 40 - As Mulheres Mais Inspiradoras do Mundo" da FemaleOneZero.

Referências úteis

AI Impact Summit India

<https://impact.indiaai.gov.in/>

Aliança do Sul Global (Global South Alliance)

<https://globalsouthalliance.net/>

Declaração de São Luís (G20)

<https://www.gov.br/g20/pt-br/noticias/declaracao-de-sao-luis-recomendacoes-estrategicas-para-o-futuro-da-inteligencia-artificial>

Eventos preparatórios da Cúpula

<https://impact.indiaai.gov.in/home/pre-summit-events>

Grupos de Trabalho da Cúpula

<https://impact.indiaai.gov.in/working-groups>

Projeto “Enabling multistakeholder approaches to AI governance” do Global Network Initiative

<https://globalnetworkinitiative.org/enabling-multistakeholder-approaches-to-ai-governance/>

