



Plano Brasileiro de Inteligência Artificial (PBIA)

Relatório contendo os resultados da implementação e monitoramento do PBIA para o período de 2024/2025

Brasília – DF
2025

© Ministério da Ciência, Tecnologia e Inovação (MCTI)

Presidente da República

Luiz Inácio Lula da Silva

Ministra de Estado da Ciência, Tecnologia e Inovação

Luciana Santos

Secretário Executivo

Luis Manuel Rebelo Fernandes

Secretário de Ciência e Tecnologia para a Transformação Digital

Henrique de Oliveira Miguel

© Centro de Gestão e Estudos Estratégicos (CGEE)

Diretor-presidente

Fernando Cosme Rizzo Assunção

Diretores

Anderson Stevens Leonidas Gomes

Caetano Christophe Rosado Penna

Diretor de administração e finanças

Geraldo Nunes Sobrinho

Relatório anual das atividades de apoio à implementação e monitoramento PBIA - 2025

Plano Brasileiro de Inteligência Artificial (PBIA)

Grupo de Trabalho para Gestão do Plano Brasileiro de Inteligência Artificial (GT-PBIA)

Ministério de Ciência, Tecnologia e Inovação (Coordenação)

Titular: Luis Manuel Rebelo Fernandes

Suplente: Henrique de Oliveira Miguel

Casa Civil da Presidência da República

Titular: Rodrigo Rodrigues da Fonseca

Suplente: Sylvio César Koury Musolino Filho

Secretaria de Comunicação Social da Presidência da República

Titular: João Caldeira Brand Monteiro de Castro

Suplente: Renam Vinicius Carvalho Santos Brandão

Ministério da Fazenda

Titular: Igor Marchesini Ferreira

Suplente: Pablo Bello Arellano

Ministério da Gestão e da Inovação em Serviços Públicos

Titular: Miriam Barbuda Fernandes Chaves

Suplente: Thaciana Guimarães de Oliveira Cerqueira

Ministério do Desenvolvimento, Indústria, Comércio e Serviços

Titular: Cristiane Vianna Rauen

Suplente: Pollyana de Carvalho

Ministério da Educação

Titular: Fernando de Barros Filgueiras

Suplente: Ana Úngari dal Fabro

Ministério da Justiça e Segurança Pública

Titular: Ricardo de Lins Horta

Suplente: José Rocha de Carvalho Filho

Ministério das Comunicações

Titular: Alexandre Reis Siqueira Freire

Suplente: André Leandro Magalhães

Ministério das Relações Exteriores

Titular: Eugênio Vargas Garcia

Suplente: Guilherme Fitzgibbon Alves Pereira

Financiadora de Estudos e Projetos

Titular: Carlos A. Aragão C. Filho

Suplente: Elias Ramos de Souza

Banco Nacional de Desenvolvimento Econômico e Social

Titular: Marconi Edson Ferreira Viana

Suplente: Lucas Moura de Lucena

Conselho Nacional de Desenvolvimento Científico e Tecnológico

Titular: Geraldo Sorte

Suplente: Débora Peres Menezes

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

Titular: Gustavo Jardim Portella

Suplente: Regiane Kawakami

Empresa Brasileira de Pesquisa e Inovação Industrial

Titular: Igor Manhães Nazareth

Suplente: Thales Marçal Vieira Neto

Ministério da Saúde (MS)

Titular: Ana Estela Haddad

Suplente: Tiago Bahia Fontana

4

Grupo de Trabalho de Apoio Operacional

Secretário-executivo do Ministério da Ciência, Tecnologia e Inovação (Sexec/MCTI)

Luiz Manuel Rebelo Fernandes

Secretário da Secretaria de Ciência, Tecnologia e Inovação para Transformação Digital (Setad/MCTI)

Henrique de Oliveira Miguel

Equipe técnica do MCTI

Adriana Anunciatto Depieri

Amauri Casarin Junior

André Rafael Costa e Silva

Daniel Silva Boson

Eliana Cardoso Emediato de Azambuja

Guilherme de Paula Côrrea

Hamilton José Mendes da Silva

Rubens Caetano Barbosa de Souza

Equipe técnica do Centro de Gestão e Estudos Estratégicos (CGEE)

Caetano Christophe Rosado Penna (Diretor supervisor)

Caroline Nascimento Pereira (Líder de projeto)

Isabela Quadros Dantas Barros

Sumário

1. Introdução 6

Monitoramento e principais resultados em 2025..... 10

2. Ações Estruturantes 11

2.1. Eixo 1. Infraestrutura e Desenvolvimento de IA..... 13

2.1.1. Principais entregas do Eixo 1..... 13

2.1.2. Detalhamento das ações – Eixo 1..... 18

2.2. Eixo 2. Difusão, Formação e Capacitação..... 24

2.2.1. Principais entregas - Eixo 2..... 24

2.2.2. Detalhamento das ações - Eixo 2 28

2.3. Eixo 3. IA para Melhoria dos Serviços Públicos 32

2.3.1. Principais entregas - Eixo 3..... 32

2.3.2. Detalhamento das ações – Eixo 3..... 35

2.4. Eixo 4. IA para inovação empresarial..... 42

2.4.1. Principais entregas do Eixo 4..... 42

2.4.2. Detalhamento das ações – Eixo 4..... 46

2.5. Eixo 5. Apoio ao processo regulatório e de governança da IA..... 50

2.5.1. Principais entregas do Eixo 5..... 50

2.5.2. Detalhamento das ações – Eixo 5..... 51

3. Ações de Impacto Imediato..... 54

4. Novas ações a serem incorporadas ao PBIA 58

5. Atuação das instituições no PBIA – CNPq, Embrapii e FINEP 60

5.1. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)..... 60

5.1.1. Impactos resultantes das ações empreendidas pelo CNPq 62

5.2. Empresa Brasileira de Pesquisa e Inovação Industrial (Embrapii)..... 63

5.2.1. Impactos resultantes das ações empreendidas pela Embrapii 65

5.3. Financiadora de Estudos e Projetos (FINEP)..... 67

5.3.1. Impactos resultantes das ações empreendidas pela FINEP 71

Atividades de apoio à implementação do PBIA - 2025..... 73

Seminário Pilha de IA Nacional: Desafios para Autonomia Tecnológica e Soberania Digital 74

ANEXO 1. Relatório completo do Pilha de IA: Desafios para Autonomia Tecnológica e Soberania Digital..... 78

1. Introdução

O Plano Brasileiro de IA foi lançado oficialmente em julho de 2024, durante a V Conferência de Ciência, Tecnologia e Inovação (VCNCTI), em Brasília/DF, com investimentos na ordem de R\$ 23,03 bilhões de reais, entre investimentos públicos e privados, para 54 ações estruturantes nas áreas de infraestrutura e desenvolvimento da IA, formação, capacitação e difusão da IA, serviços públicos, inovação industrial e apoio na regulação e governança da IA. Além de ações de impacto em áreas diversas, como educação, saúde, gestão dos serviços públicos, agricultura, dentre outras.

Entre o lançamento na VCNCTI e as primeiras atividades de implementação, foram realizadas algumas ações institucionais, de forma a permitir a implantação do plano e garantir seu monitoramento e acompanhamento. A primeira atividade importante de 2025 foi a publicação da Resolução do CITDigital nº 2, de 8 de maio de 2025¹, que instituiu o Grupo de Trabalho para a gestão do Plano Brasileiro de Inteligência Artificial. Dentre as competências do GT estão:

- monitorar a execução do PBIA e, quando cabível, apresentar ao Comitê Executivo do CITDigital propostas de ajustes ao plano;
- organizar seus trabalhos em planos anuais, a serem apresentados ao Comitê Executivo do CITDigital;
- informar, quando provocado pelo coordenador do Comitê Executivo do CITDigital, sobre os seus trabalhos e o monitoramento do PBIA; e
- apresentar anualmente relatório de acompanhamento de execução do PBIA ao Comitê Executivo do CITDigital.

Em junho de 2025 ocorreu a incorporação do PBIA ao plano de investimentos do FNDCT, com incorporações de ações do PBIA a programas existentes do FNDCT e a criação do novo programa do FNDCT, Programa Brasileiro de Inteligência Artificial - *IA Brasil*, do Fundo Nacional de Desenvolvimento Científico e Tecnológico (FNDCT), com a missão de posicionar o Brasil como referência global em IA responsável e confiável, garantindo a soberania nacional, os direitos fundamentais, a proteção do trabalho e o desenvolvimento econômico sustentável.

Ainda em junho foi publicado o documento oficial do Plano Brasileiro de IA (PBIA)², contendo o diagnóstico da IA no Brasil, janelas de oportunidades para o Brasil, as premissas para uma IA para o Bem de Todos, o plano de ação do PBIA e a estrutura de governança e monitoramento do plano. O documento é uma extensão contendo maior

¹ <https://www.in.gov.br/en/web/dou/-/resolucao-citdigital-n-2-de-8-de-maio-de-2025-628599624>

² https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/plano-brasileiro-de-inteligencia-artificial-pbia-_vf.pdf

detalhamento do que foi apresentado na VCNCTI, de forma a dar maior clareza sobre onde o Brasil está posicionado e o que almeja alcançar nos próximos anos.

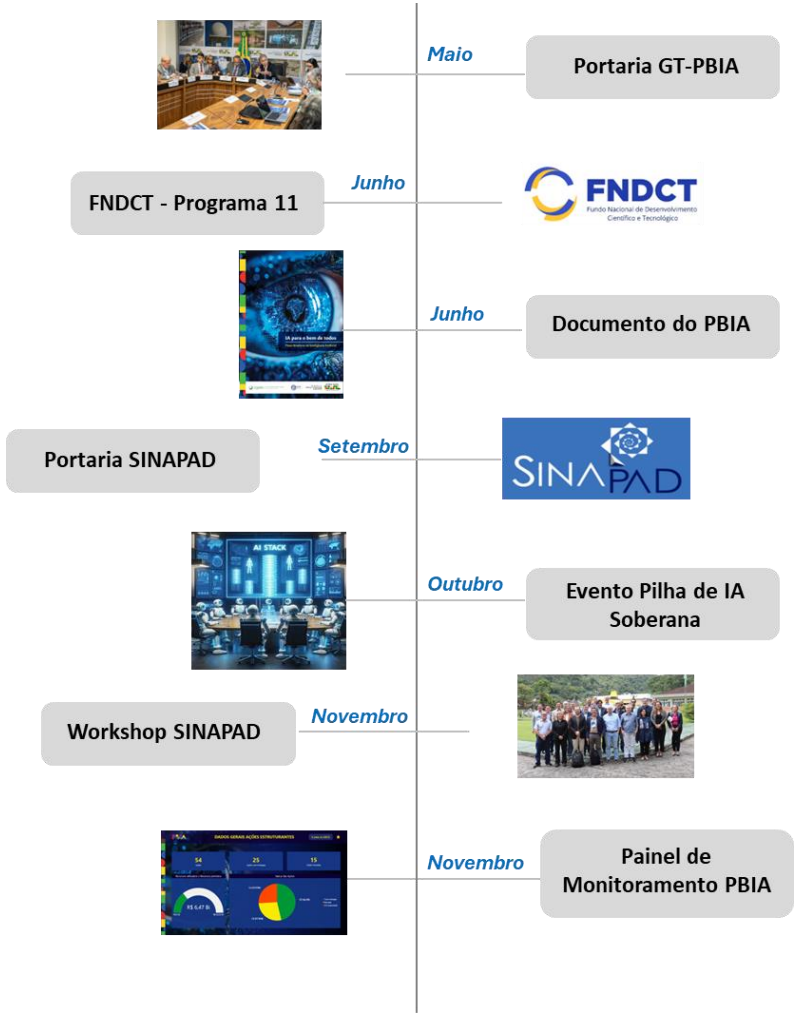


Figura 1. Marcos institucionais e eventos de apoio à implementação e monitoramento do PBIA

Em setembro foi instituída a nova regulamentação do Sistema Nacional de Processamento de Alto Desempenho (SINAPAD), estabelecida pela Portaria MCTI nº 9.445, de setembro de 2025, representando uma modernização significativa em relação ao modelo anterior (baseado fundamentalmente no Decreto de 2004 e na Portaria de 2006). O principal avanço é a transformação de um sistema focado apenas em gestão técnica centralizada para um sistema de governança estratégica, com regras claras de financiamento e um foco inédito em Inteligência Artificial. Dentre os principais avanços estão: criação dos CENAPAD-IA, com foco em inteligência artificial, permitindo que centros especializados em IA sejam integrados oficialmente à rede nacional, algo que a norma antiga não previa, focada em simulações numéricas e processamento de dados científicos gerais.

Em relação a governança, o novo SINAPAD instituiu um Conselho Diretor, com o LNCC mantendo um papel executivo importante, porém cabendo ao Conselho deliberar sobre a entrada de novos centros e para onde irão os novos investimentos. Outra mudança se refere à sustentabilidade financeira do sistema, com a possibilidade de cobrança pelo uso dos recursos computacionais, permitindo que os centros prestem serviços para o setor privado ou projetos específicos para custear as operações.

Sobre o compartilhamento dos recursos, a nova portaria estabelece que para ser um CENAPAD e receber fomento, o centro deve reservar no mínimo 25% de sua capacidade para uso aberto da comunidade científica/acadêmica externa, possibilitando que o investimento público no centro retorne como serviço à rede de pesquisa nacional. Por fim, também se estabelece critérios técnicos modernos, como conexão obrigatório às redes de alta velocidade da RNP, além da interoperabilidade de dados, viabilizando a Ciência Aberta.

Em outubro foi realizado o evento Seminário Pilha de IA Nacional: desafios para autonomia tecnológica e soberania digital, nos dias 01 e 02, em Brasília/DF. O evento foi coordenado pelo MCTI, com apoio do Centro de Gestão e Estudos Estratégicos (CGEE), Universidade Estadual de Campinas (Unicamp), Laboratório Nacional de Computação Científica (LNCC), Instituto Eldorado e SERPRO. O evento reuniu especialistas da academia, governo e setor produtivo para debater os caminhos do Brasil rumo a uma inteligência artificial mais justa, soberana e orientada à superação de desafios nacionais, conforme definido no PBIA.

Ao longo de dois dias, o evento buscou mapear desafios, identificar oportunidades e propor recomendações práticas para que o país avance na construção de uma pilha de IA nacional, capaz de garantir autonomia tecnológica, soberania digital e desenvolvimento sustentável. O evento contou com uma palestra do professor Edson Borin (Unicamp) e mais nove painéis com palestrantes de diversas instituições, que debateram temas como autonomia de infraestrutura tecnológica, atuação nas camadas superiores da Pilha de IA, desenvolvimento de IA para dispositivos móveis e embarcados, cooperações internacionais, SINAPAD, capacitação e formação técnica, colaborações estratégicas nacionais e articulação institucional para implementação das ações do PBIA. Na parte dois deste documento está apresentado um resumo com os principais resultados do seminário e o relatório final do evento.

Em novembro foi realizado o Workshop de Alinhamento Estratégico dos CENAPADs, no LNCC em Petrópolis/RJ, com a presença de representantes de CENAPADs existentes e de potenciais novos CENAPADs. O CGEE apoiou a realização do evento e a condução da oficina com os especialistas. Os resultados deste evento ainda não foram sistematizados e deverão ser entregues ao MCTI no final de dezembro de 2025.

Por fim, o CGEE, como instituição responsável por apoiar o monitoramento e implementação do PBIA, desenvolveu o Painel em Power BI do Plano, contendo

informações coletadas a partir do envio de formulários aos membros do GT do PBIA. O formulário contém campos para informar o status de implementação, entregas periódicas, data prevista para conclusão da ação, recursos financeiros previstos, recursos financeiros utilizados, fonte dos recursos, metas da ação, indicadores de monitoramento da ação, indicadores de impacto da ação e instituições parceiras envolvidas. O link para o formulário é: <https://forms.cgee.org.br/form/form-pbia>.

O painel apresenta as informações organizadas por tipo de ação (estruturantes, impacto e novas ações). E dentro de cada tipo, há subseções, como eixos para as ações estruturantes e setores para as ações de impacto. O painel está hospedado na página do MCTI e está aberto para toda a sociedade conferir os resultados do monitoramento. O link de acesso é: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/plano-brasileiro-de-inteligencia-artificial>.



Figura 2. Página inicial do Painel de Monitoramento do PBIA

As atividades supracitadas, principalmente os avanços institucionais, são fundamentais para a implementação e avanço do PBIA, o que permitirá maior atuação em 2026, do MCTI, com apoio do CGEE, no planejamento e realização das ações previstas no plano.

Monitoramento e principais resultados em 2025

1. Ações Estruturantes

O PBIA conta com um investimento robusto e um cronograma de ações estratégicas bem definido, sendo 54 ações estruturantes, divididas em cinco eixos. Os eixos do PBIA são:

- Eixo 1 - Infraestrutura e Desenvolvimento de IA
- Eixo 2 - Difusão, Formação e Capacitação
- Eixo 3 - IA para Melhoria dos Serviços Públicos
- Eixo 4 - IA para Inovação Empresarial
- Eixo 5 – Apoio ao Processo Regulatório e de Governança da IA

Do total das ações, 30 (55,6%) possuem entregas, 12 (22,2%) estão iniciadas e 13 ações (22,2%) estão em preparação, para monitoramento realizado em 30/11/2025 (Figura 3). Dentre os eixos, Difusão, Formação e Capacitação (Eixo 2), apresentou maior proporção de ações com entregas, 75% do total de 8 ações. As demais (2 ações) constam como iniciadas. O Eixo 3, IA para Melhoria dos Serviços Públicos, apresenta 58% de suas ações com entregas (11 ações), 26% em preparação e 16% iniciadas. O Eixo 4, IA para inovação Empresarial, apresenta 55% das ações com entregas, 22% como em preparação e 22% iniciadas. O Eixo 1, voltado para Infraestrutura e Desenvolvimento de IA, apresenta 46% de suas ações com entregas, 38,5% iniciadas e 15,5% em preparação.

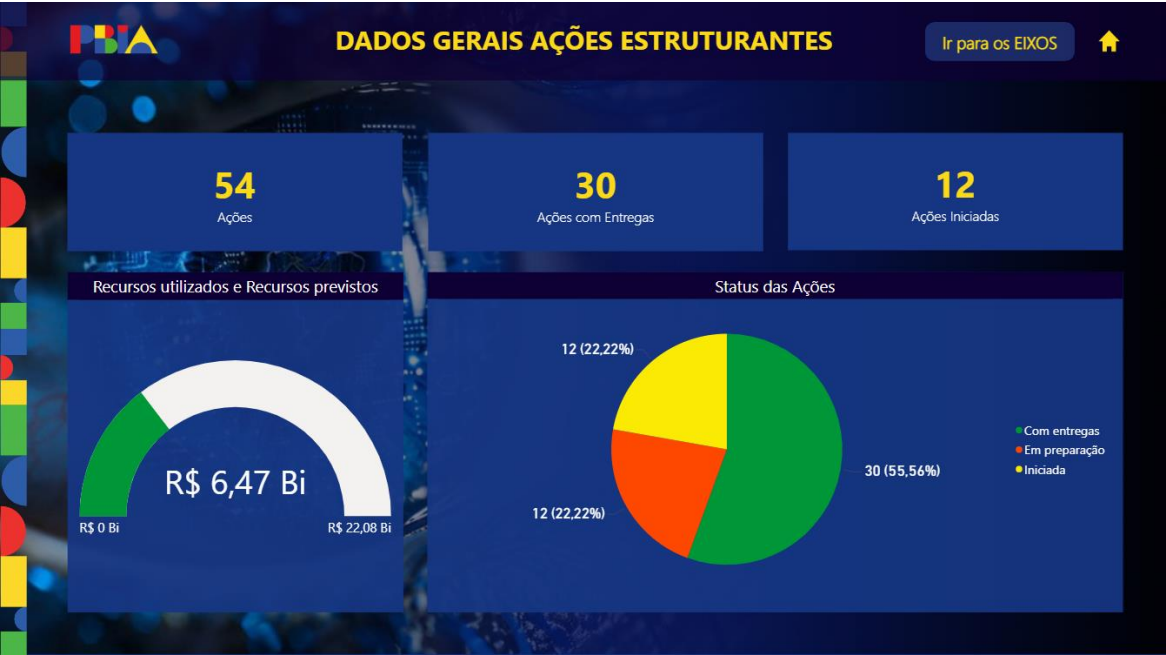


Figura 3. Dados gerais das ações estruturantes – PBIA – 30/11/2025
Fonte: Painel de Monitoramento PBIA - CGEE (2025)

Em relação aos recursos previstos para o PBIA, R\$ 23 bilhões, foram dispendidos até o momento, R\$, 6,47 bilhões, 28% do total. O Eixo 4 representou o maior montante, R\$ 5,98 bilhões dispendido no PBIA, 45% do total de R\$ 13,3 bilhões previstos no Eixo 4. O Eixo 1 apresentou o segundo maior montante, com R\$ 450 milhões dispendidos até o momento, representando cerca de 7,7% do total do eixo. As ações do Eixo 2 utilizaram R\$ 30,4 milhões, representando 2,6% do total previsto no eixo (R\$ 1,14 bilhões). O Eixo 3 utilizou R\$ 1,9 milhões do total previsto de R\$ 1,76 bilhões. Por fim, o Eixo 5 até o momento não possui recursos utilizados frente aos R\$ 101,25 milhões alocados para o eixo.

Importante destacar novamente a aprovação do Programa 11 do FNDCT, Programa IA Brasil, em junho de 2025, garantindo os recursos necessários para as ações do PBIA que possuem como previsão a utilização de recursos do FNDCT (cerca de 2/3 dos investimentos do plano). Apesar do PBIA ter ações financiadas com recursos de outros programas, a existência deste programa garante, em maior medida, os recursos necessários para as ações.

Eixo	Principais entregas
Eixo 1: Infraestrutura e Desenvolvimento de IA	- Ampliação da capacidade do supercomputador Santos Dumont. - Apoio a 04 Centros Temáticos de IA e edital com demanda de R\$ 755 milhões para 57 projetos temáticos. - Apoio a 808 projetos (bolsa de produtividade e universal) - 10 CPAs em IA - Apoio a 08 INCTs de IA em todas as regiões brasileiras.
Eixo 2: Difusão, Formação e Capacitação em IA	- 8.104 vagas em cursos de IA em 2024 totalizando 73.275 vagas no país. - Curso de IA na prática com 40 mil inscritos - Acordos de cooperação técnica com Microsoft, Google e Alura.
Eixo 3: IA para Melhoria dos Serviços Públicos	- 7 entregas principais: catalogação e interoperabilidade de dados, plataforma low-code com IA Generativa, Núcleo de IA de Governo (NIA), tecnologias de cibersegurança, qualificação de dados cadastrais, comunicação personalizada e difusão de informações
Eixo 4: IA para Inovação Empresarial	- Apoio a 596 projetos de IA até 7/2025 (R\$ 4,5 bilhões) (FINEP) - Apoio a 12 projetos desde o PBIA (R\$ 1,2 bilhões) (BNDES) - Financiamento de projeto de datacenter (BNDES)
Eixo 5: Apoio ao Processo Regulatório e de Governança da IA	- Guia de IA Generativa no serviço público - Lançamento do OBIA e desenvolvimento de indicadores em parceria com CGEE e MGI.

Quadro 1. Principais entregas – Visão geral PBIA

Fonte: CGEE (2025)

1.1. Eixo 1. Infraestrutura e Desenvolvimento de IA

O eixo 1, composto por 13 ações, tem como objetivo fomentar os investimentos essenciais em infraestrutura digital e computacional para sustentar o desenvolvimento e a implementação de sistemas de IA, com ênfase particular na promoção de sistemas de inteligência artificial sustentáveis. Abrange a expansão da capacidade computacional nacional, o aprimoramento das redes de comunicação e a estruturação de bases de dados robustas e acessíveis. Um aspecto central é o aproveitamento da vantagem competitiva do Brasil em sua matriz energética predominantemente limpa, visando a impulsionar o desenvolvimento sustentável da IA no País. Isso inclui a integração de fontes de energia renovável e a implementação de práticas de eficiência energética em toda a cadeia de desenvolvimento e operação de IA. Além disso, o eixo contempla o desenvolvimento de hardware especializado para IA e a criação de ambientes de teste e experimentação otimizados para eficiência energética. Busca-se também fomentar o desenvolvimento de modelos e ferramentas de IA adaptados ao contexto brasileiro, aproveitando as amplas bases de dados públicos disponíveis e priorizando soluções que maximizem a eficiência computacional e minimizem o impacto ambiental.

1.1.1. Principais entregas do Eixo 1

13

O eixo 1 apresentou avanços ao longo de 2025, como a atualização do supercomputador Santos Dumont no LNCC, com o recebimento de novos equipamentos em julho de 2025. Localizado em Petrópolis, Rio de Janeiro, o supercomputador já foi utilizado em milhares de projetos de pesquisa conduzidos por pesquisadores brasileiros, como a Dra. Ana Teresa Vasconcelos, que sequenciou o genoma da COVID-19 nele em 2020, e pode ser solicitado para uso por qualquer pesquisador ou instituição nacional.

Em 2024 o Santos Dumont estava em 89º lugar na lista dos supercomputadores mais poderosos do mundo. Em 2025 caiu para a posição 122º, entretanto é necessário ponderar os substanciais avanços e investimentos de outros países nesta área.

A atualização realizada baseia-se na arquitetura BullSequana XH3000, uma tecnologia escalável, projetada e fabricada em sua fábrica principal na França, que permitiu expandir em quatro vezes a capacidade de processamento do supercomputador, incluindo recursos adicionais dedicados à inteligência artificial.



Com isso, o SD passará a operar com 18,85 petaflops, ou quadrilhões de operações por segundo, o que representa um aumento de aproximadamente 575% em relação à especificação original de 2015. O supercomputador também integra um sistema patenteado de refrigeração líquida direta de alta eficiência (Direct Liquid Cooling – DLC), capaz de alcançar uma densidade cinco vezes maior do que sistemas tradicionais com refrigeração a ar, garantindo, assim, maior eficiência energética. Baseado na solução Enterprise da NVIDIA, o Santos Dumont agora conta com quatro CPUs Grace conectadas a quatro GPUs Hopper em cada nó acelerado nas blades BullSequana XH3515-H, tecnologia projetada para processar grandes volumes de dados em centros de dados dinâmicos, produzindo inteligência com máxima eficiência energética, essencial para trabalhos com IA. A CPU Grace oferece o dobro de desempenho por watt, o dobro da densidade de empacotamento e a maior largura de banda de memória em comparação com os servidores líderes atuais.

Ainda no âmbito da ação 1, a Petrobras lançou edital para aquisição de supercomputador, um HPC Acadêmico, no valor de R\$ 340 milhões, em que a UFRJ venceu o edital e o processo já foi encerrado.

Outra entrega, concernente à ação 2, é a nova regulamentação do Sistema Nacional de Processamento de Alto Desempenho (SINAPAD), estabelecida pela Portaria MCTI nº 9.445, de setembro de 2025, representando uma modernização significativa em relação ao modelo anterior (baseado fundamentalmente no Decreto de 2004 e na Portaria de 2006). O principal avanço é a transformação de um sistema focado apenas em gestão técnica centralizada para um sistema de governança estratégica, com regras claras de financiamento e um foco inédito em Inteligência Artificial. Dentre os principais avanços estão: criação dos CENAPAD-IA, permitindo que centros especializados em IA sejam integrados oficialmente à rede nacional, e a instituição de um Conselho Diretor, com o LNCC mantendo papel executivo importante, porém cabendo ao Conselho deliberar sobre a entrada de novos centros e os recursos a serem investidos. Também considera a sustentabilidade financeira do SINAPAD, com a possibilidade de cobrança pelo uso de recursos computacionais, permitindo que os centros prestem serviços para o setor privado ou projetos específicos para custear as operações.

A nova portaria também estabelece que para ser um CENAPAD e receber fomento, o centro deve reservar no mínimo 25% de sua capacidade para uso aberto da comunidade científica/acadêmica externa, possibilitando que o investimento público no centro retorne como serviço à rede de pesquisa nacional. Por fim, também se estabelece critérios técnicos modernos, como conexão obrigatória às redes de alta velocidade da RNP, além da interoperabilidade de dados, viabilizando a Ciência Aberta. Todos avanços vão permitir que a ação 2 seja implementada de forma a otimizar seus resultados, garantindo maior capacidade de processamento dos CENAPADs.

Outro destaque são os 08 Institutos Nacionais de Ciência e Tecnologia (INCTs) em Inteligência Artificial, que tiveram R\$ 98 milhões em recursos aprovados, com R\$ 25 milhões liberados em outubro de 2025. Os novos INCTs estão localizados no Pará (01), Pernambuco (03), Minas Gerais (02), São Paulo (01) e Rio Grande do Sul (01). Dos 08 INCTs, 06 pertencem a área do conhecimento de Ciências Exatas e da Terra, 01 às Engenharias e 01 Ciências Humanas.



Figura 4. INCTs-IA - Brasil

O Programa INCT é uma política estratégica do MCTI em parceria com o CNPq e as FAPs, reunindo redes multi-institucionais e interdisciplinares voltadas a temas complexos e prioritários para o país, com foco em excelência científica, cooperação internacional e aplicação em grandes desafios nacionais. O Quadro 2 apresenta detalhamento dos novos INCTs-IA, com a instituição e valor concedido.

INCT	Instituição	Área Predominante	Valor Total da Concessão (2025-2028)
Instituto Nacional de Inteligência Artificial para o Bem Social (IAPROBEM)	USP	Ciências Exatas e da Terra	R\$ 11.470.000,00
Computação Neural	UFPE	Ciências Exatas e da Terra	R\$ 12.441.000,00
Computação Quântica Aplicada	UFPE	Ciências Exatas e da Terra	R\$ 10.192.000,00
Inteligência Artificial e Sustentabilidade (INCT-IAS)	PUCR	Ciências Humanas	R\$ 11.255.000,00
Instituto Nacional de Ciência e Tecnologia para Engenharia de Software Baseada em e para Inteligência Artificial (INES.IA)	UFPE	Ciências Exatas e da Terra	R\$ 14.376.000,00
Instituto Nacional de Ciência e Tecnologia (INCT) em Inteligência Artificial Responsável para Linguística Computacional e Tratamento e Disseminação de Informação (TILD-IAR)	UFMG	Ciências Exatas e da Terra	R\$ 9.709.000,00
INCT IAmazônia: Instituto Nacional de Ciência e Tecnologia em Inteligência Artificial aplicada às cidades inteligentes e sustentáveis na Amazônia brasileira	UFPA	Engenharias	R\$ 13.075.000,00

INCT em Inteligência Artificial para Cibersegurança: Uma Abordagem Tecnológica e Social (INCT-IACiber)	UFMG	Ciências Exatas e da Terra	R\$ 10.303.000,00
TOTAL			R\$ 92.821.000,00

Quadro 2. INCTs em IA – instituição, área e valor da concessão entre 2025 e 2028
Fonte: CNPq (2025)

Ações com entregas

Do conjunto das 13 ações existentes no eixo 1, 06 ações já possuem entregas, como mostra a Tabela 1. A ação 1, *Supercomputador de IA*, atualização e aquisição, já possuem entregas, conforme citado acima, com a compra de equipamentos para a atualização do Santos Dumont e o lançamento do edital para aquisição do HPC Acadêmico da Petrobras. A atualização do Santos Dumont coloca esta ação como concluída, tendo como foco agora a aquisição do supercomputador, conforme descrito abaixo e no detalhamento da ação adiante.

Também possuem entregas a ação 2, *Ampliação da capacidade de processamento do CENAPAD*, que agora conta com uma nova portaria para regulamentar o SINAPAD, conforme detalhado acima e a realização de um workshop com os CENAPADs existentes e potenciais centros, de forma a entender a melhor estruturação para este sistema. Outra ação com entregas é a ação 4, *Pró-Infra IA*, com o apoio da FINEP a 4 Centros Temáticos de IA, além de um novo edital com demanda de R\$ 755 milhões para 57 projetos temáticos.

Ainda foi realizado do 1º Seminário da Pilha de IA Nacional, com especialistas relevantes no tema, discutindo possibilidades de modelos para desenvolvimento dos *stacks* no Brasil, no âmbito da ação 8, *Desenvolvimento de Pilha de Software para IA*. Resultado deste evento foi elaborado um relatório para apoiar a tomada de subsídios pelo MCTI. A ação 10, Fomento às atividades de P&D em IA também apresentou entregas, com 808 projetos aprovados e bolsas concedidas, além de uma nova chamada no tema a ser lançada até dezembro de 2025.

Por fim, a ação 11, *INCTs-IA*, apresentou apoio a 08 INCTs de IA nas regiões Norte, Nordeste, Sul e Sudeste, conforme mencionado acima.

Ação	Envolvidos	Principal Entrega	Recursos Utilizados
------	------------	-------------------	---------------------

Ação 1-A: Supercomputador de IA (atualização)	MCTI, Petrobras, FAPESP, LNCC, FINEP	LNCC: recebimento de novos equipamentos para ampliar capacidade do Santos Dumont. Petrobras: lançam. de edital para aquisição de supercomputador.	R\$ 148,8 milhões
Ação 1-B: Supercomputador de IA (aquisição)	MCTI, LNCC, FINEP	Realização de consulta (RFI) junto aos fabricantes. Publicação do edital de compra até dez/2025 (FINEP/LNCC). Aquisição no 1º trim/2026.	R\$ 100 milhões (até dez/2025)
Ação 2: Ampliação da capacidade proces. do CENAPAD	MCTI, LNCC	Publicação de nova portaria do SINAPAD e realização de workshop com CENAPADs.	-
Ação 4: Pró-Infra IA	FINEP, MCTI	FINEP apoiou 4 Centros Temáticos de IA com R\$ 52 milhões. Novo edital com demanda de R\$ 755 milhões para 57 projetos temáticos.	R\$ 52 milhões
Ação 8: Desenv. de Pilha de Software p/ IA	MCTI, FINEP, LNCC	1º Seminário da Pilha de IA Nacional e elaboração de relatório contendo subsídios para o desenvolvimento da ação.	-
Ação 10: Fomento às atividades de P&D em IA	CNPq, MCTI	808 projetos apoiados: Bolsas de Produtividade (296) e Universal (512). 10 CPAs em IA (R\$ 200 milhões com contrapartida). Lanç. de chamada de fomento às atividades de P&D em IA até 12/2025.	R\$ 125 milhões
Ação 11: INCTs-IA	MCTI, FINEP, CNPq, FAPs	CNPq: apoio a 8 INCTs de IA (meta original: 4), sendo 1 no RS, 1 em SP, 2 em MG, 1 em PA, e 3 em PE. Propostas selecionadas aguardando assinatura.	R\$ 98 milhões (aprovados)

Tabela 1. Ações com entregas – Eixo 1 – PBIA

Ações iniciadas

Dentre as ações iniciadas, ou seja, ações que estão em andamento, estão: ação 3, *Redes de conexão de alta velocidade*, com previsão de lançamento em 2026, a partir do projeto em formatação (e-Ciência para IA), com parceria da RNP e da Petrobras. A ação 5, *Parcerias internacionais para nós HPC e chips* também está em andamento, com elaboração da encomenda do *Reference Design de Hardware* do nó HPC com previsão de entrega até final de 2025 e contratação do consórcio que desenvolverá o desenho de referência em 2026.

A ação 6, *Parcerias Internacionais para Infraestrutura de P&D-IA*, está em andamento, com articulação e assinatura de acordo entre RNP e Redes Clara e Scalac. Outra ação em andamento é a 9, *IA baseada em dados nacionais (LLM)*, com encomendas de modelos de IA e lançamento de edital Mais Inovação Brasil em IA generativa. Por fim, a ação 7, *Pró-Infra IA Sustentável* está em fase de preparação do edital, com previsão de finalização até dezembro de 2025.

Ação	Envolvidos	Status Atual / Próximos Passos	Previsão
Ação 3: Redes de conexão de alta velocidade	RNP, Petrobras	e.Ciencia para IA – Projeto em formatação: RNP e Petrobras	Lançamento em 2026
Ação 5: Parcerias internacionais para nós HPC e chips	MCTI, LNCC	Encomenda do <i>Reference Design de Hardware</i> do nó HPC: contratação do consórcio, definição de etapas e monitoramento da execução	Em andamento
Ação 6: Parcerias internacionais para infraest. de P&D-IA	RNP, Clara, Scalac	Articulação e assinatura de acordo entre RNP/Redes Clara e Scalac	Em andamento
Ação 7: Pró-Infra IA Sustentável	MCTI, FINEP	Preparação do edital	Até dez/2025
Ação 9: IA baseada em dados nacionais (LLM)	MCTI, FINEP, USP, UFF	Encomendas de modelos de IA (USP/UFF/Unesp + Chile; Piauí/UFPI; LNCC c/ UE/SEGIB/Espanha) e lançam. de edital + Inovação BR em IA generativa	Em andamento

Tabela 2. Ações iniciadas – Eixo 1 – PBIA

Ações em preparação

Dentre as ações em preparação, ou seja, em discussão dos trâmites necessários, estão a ação 12, *Instituto Nacional de Informática*, com a encomenda para a o Porto Digital contendo a proposta do INI com foco em IA, com previsão de entrega da proposta para 2026. E a ação 13, *Fomento à P&D com América do Sul e África*, que está em processo de elaboração do edital pelo CNPq.

Ação	Envolvidos	Status Atual / Próximos Passos	Previsão
Ação 12: Instituto Nacional de Informática	MCTI, Porto Digital	Encomenda para criação do INI com foco em IA. Proposta em preparação.	Entrega em 2026
Ação 13: Fomento à P&D com América do Sul e África	CNPq, MCTI	Edital em elaboração pelo CNPq	Em elaboração

Tabela 3. Ações em preparação – Eixo 1 – PBIA

1.1.2. Detalhamento das ações – Eixo 1

Programa Nacional de Infraestrutura de IA

Ação 1A. Supercomputador de IA – Atualização

Meta: Atualização do supercomputador Santos Dumont no LNCC (5ª maior capacidade de processamento – TOP 500).

Entregas realizadas: LNCC recebeu novos equipamentos para ampliar a capacidade do Supercomputador Santos Dumont (nov/2024).

Petrobras – lançamento de edital para aquisição de supercomputador para pesquisa.

COM ENTREGAS



Fonte do recurso: FNDCT não-reembolsável, FAPESP e Petrobras
Instituições Envolvidas: MCTI, Petrobras, FAPESP, LNCC e FINEP

Ação 1B. Supercomputador de IA - Aquisição

Meta: Aquisição de supercomputador com alta capacidade de processamento, em cinco anos.

Entregas realizadas: Realização de consulta (RFI) junto aos fabricantes de supercomputadores localizados no país.

Entregas previstas até dez/2025: publicação de edital de compra até dezembro de 2025 (FINEP/LNCC).

COM ENTREGAS



Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MCTI, LNCC e FINEP

Ação 2. Ampliação da capacidade de processamento dos Centros Nacionais de Processamento de Alto Desempenho (Cenapads)

Meta: Ampliação da capacidade dos cinco Cenapads, distribuídos pelas cinco regiões do País, em dois anos.

Entregas: Publicação de nova Portaria relativa ao Sistema Nacional de Processamento de Alto Desempenho (SINAPAD).

Entregas previstas até dez/2025: Resolução do Comitê Gestor recredenciando os CENAPADs existentes e em operação.

COM ENTREGAS



Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MCTI, CENAPADs, LNCC e FINEP.

Ação 3. Redes de conexão de alta velocidade para supercomputação

Meta: Interligação de todos os centros de supercomputação brasileiros com redes de alta velocidade, até 2028.

Entregas previstas entre 2026 e 2028:

e-Ciência para IA – Projeto em formatação: RNP e Petrobras.



INICIADA

Fonte do recurso: FNDCT não-reembolsável e Petrobras

Instituições Envolvidas: MCTI, RNP e Petrobras

Ação 4. Pró-Infra IA

Meta: Apoio a 25 projetos em 5 anos.

Entregas realizadas: FINEP apoiou 04 Centros Temáticos de IA. Novo edital apresentou demanda global de R\$ 755 milhões para 57 projetos temáticos de IA.

Entregas previstas até dez/2025: Aprovação de anexo de referência e lançamento de chamada pública para projetos que atendam ação 4 e ação 7 do PBIA.



COM ENTREGAS

Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: FINEP e MCTI

20

Ação 5. Parcerias internacionais para desenvolvimento de nós de supercomputador e aceleradores de aprendizado de máquina

Meta: Estabelecimento de, ao menos, duas parcerias em 12 meses; desenvolvimento de aceleradores de processamento de máquina em 24 meses; e desenvolvimento de dois racks completos em 36 meses.

Entregas previstas até dez/2025:

Encomenda do projeto *Reference Design de Hardware* do nó HPC. Detalhamento: contratação do consórcio de instituições e empresas que desenvolverão o desenho de referência. Definição de etapas, fases e entregáveis intermediários e finais. Início de monitoramento da execução do projeto.



INICIADA

Fonte do recurso: FNDCT não-reembolsável, FNDCT reembolsável, setor privado e organismos internacionais
Instituições Envolvidas: MCTI e LNCC

Ação 6. Parcerias internacionais para compartilhamento de infraestrutura de P&D em IA

Meta: Ampliação significativa da capacidade de conexão da Rede Clara e da Rede Scalac em 24 meses; apoio a 30 projetos colaborativos que utilizem a infraestrutura brasileira de IA, até 2028.

Entregas previstas até dez/2025:
Articulação e assinatura de acordo entre RNP/Red Clara e Scalac

INICIADA



Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MCTI, Clara e Scalac

Programa de Sustentabilidade de Energias Renováveis para IA

Ação 7. Pró-Infra IA Sustentável

Meta: Apoio a 42 projetos em cinco anos.

Entregas previstas até dez/2025:
Lançamento do edital (até dezembro/2025)

INICIADA



Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MCTI e FINEP

Programa de Estrutura do Ecossistema de Dados e Software para IA

Ação 8. Desenvolvimento de Pilha de Software para IA

Meta: Desenvolvimento de uma pilha de software nacional completa para IA em 12 meses.

Entregas realizadas: 1º Seminário da Pilha de IA Nacional.

Entregas previstas até dez/2025: Encomenda ao LNCC para rede de ICT/empresas – pilha e fábrica de IA
Lançamento de edital Mais Inovação Brasil com linha para pilha de IA.

COM ENTREGAS



Fonte do recurso: FNDCT não-reembolsável e setor privado
Instituições Envolvidas: MCTI, LNCC e FINEP

Ação 9. IA baseada em dados nacionais (LLM em português)

Meta: ampliação da oferta de conjuntos de dados nacionais curados para treinamento; e construção de um modelo LLM robusto em termos de segurança para português em 12 meses.

Entregas previstas entre 2026 e 2028:

Encomendas de modelos de IA (USP/UFF/Unesp + Chile; Piauí/UFPI; consórcio LNCC e parceiros com UE/SEGIB/Espanha) (2026).
Lançamento de edital Mais Inovação Brasil em IA generativa (2026).

INICIADA



Fonte do recurso: FNDCT não-reembolsável, FNDCT reembolsável e setor privado
Instituições Envolvidas: MCTI, LNCC e FINEP

Programa de Pesquisa e Desenvolvimento em IA

Ação 10. Fomento às atividades de P&D em IA

Meta: metas em definição, específicas de cada iniciativa - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Entregas realizadas: Projeto de IA-CNPq com 808 projetos apoiados: Bolsas de Produtividade (296)/Universal (512). 10 CPAs – Centros de pesquisas científicas, tecnológicas e de inovação, aplicadas e orientadas à resolução de problemas.
Entregas previstas até dez/2025: Lançamento da chamada de fomento às atividades de P&D em IA.

COM ENTREGAS



Fonte do recurso: FNDCT não-reembolsável e CAPES
Instituições Envolvidas: MCTI e CNPq

Ação 11. INCT-IA

Meta: Estabelecimento de quatro centros temáticos interdisciplinares de estudos avançados em IA, até 2028.

Entregas realizadas: CPNq: apoio a 07 Institutos Nacionais de Ciência e Tecnologia de IA (PA, PE, MG, SP e RS).
Entregas previstas até dez/2025: Contratação dos projetos selecionados (edital INCT-2024)

COM ENTREGAS



Fonte do recurso: FNDCT não-reembolsável e FAPs
Instituições Envolvidas: MCTI, FINEP, CNPq e FAPs

Ação 12. Instituto Nacional de Informática com foco em IA

Meta: Múltiplas metas, incluindo o estabelecimento de cinco grupos de pesquisa em seis meses no âmbito do instituto; lançamento do Instituto em nove meses; e início das operações de apoio a startups em 12 meses.

Entregas previstas até dez/2025:
Aguardando proposta do Porto Digital.



EM PREPARAÇÃO

Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MCTI

Ação 13. Fomento a P&D em cooperação com América Latina, Caribe e África

Meta: Apoio a 100 projetos colaborativos em IA com países da América Latina, Caribe e África, até 2028.

Entregas previstas entre 2026 e 2028: Edital em elaboração com o CNPq.



EM PREPARAÇÃO

Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: CNPq e MCTI

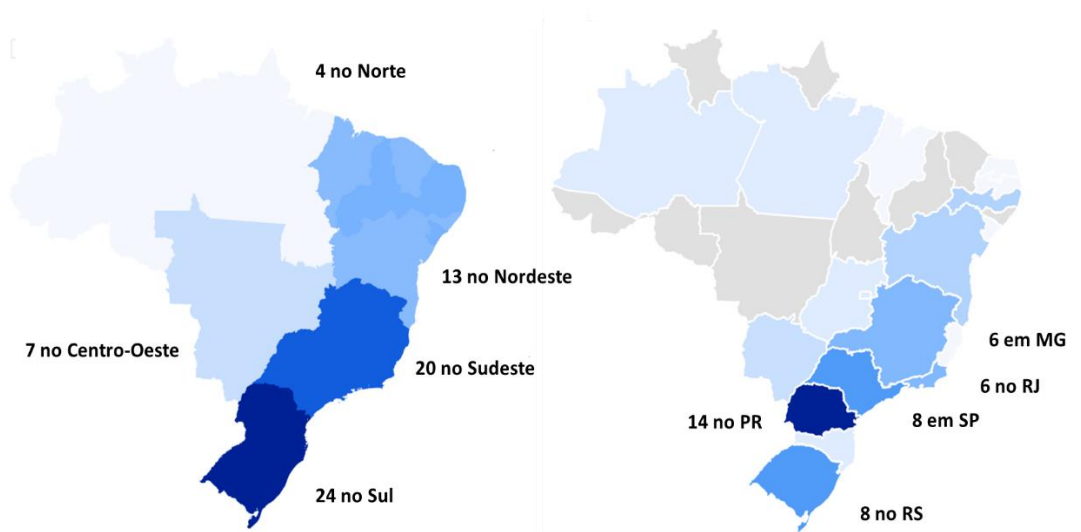
1.2. Eixo 2. Difusão, Formação e Capacitação

O eixo 2 é composto por 09 ações, que abrangem a formação e capacitação de profissionais em todos os níveis, desde a educação básica até a pós-graduação, bem como a qualificação e requalificação, visando a criar uma força de trabalho qualificada em IA. Busca aproveitar o capital humano brasileiro para impulsionar pesquisas de ponta e formar talentos em IA, suprimindo a demanda de capacitação, além de procurar reter talentos no País. Inclui também iniciativas de conscientização pública sobre os impactos e potenciais da IA, promovendo uma adoção informada e responsável da tecnologia, bem como a inclusão e a acessibilidade digital por meio da IA.

1.2.1. Principais entregas - Eixo 2

O tema da formação foi o que apresentou maior avanço neste eixo, com destaque para a ação 16, *IA na Graduação*, que previa a criação de cursos de graduação em IA e correlatos, com estímulo à criação de disciplinas optativas de programação, ciência de dados e IA, além de oferta de vagas no Fies. Até o final de novembro foram identificadas 73.275 vagas no país em cursos de IA, com 8.104 vagas criadas em 2024, superando a meta do PBIA, que previa 5 mil vagas ao ano. Dentre os cursos autorizados durante a vigência do PBIA estão UFMG, UFPB, UFPR e UEL.

De forma a captar a disponibilidade de cursos existentes, o CGEE está realizando, juntamente com o LNCC, o levantamento dos cursos de IA no Brasil³, graduação e pós-graduação. Foram reportados 70 cursos de IA, em 37 instituições distribuídas por 17 estados em todas as regiões brasileiras, conforme Figura 5.



³ O levantamento apresenta resultados parciais e a última atualização data de 30/09/2025.

Figura 5. Cursos de IA – Graduação e Pós-Graduação – Distribuição regional e estadual - Brasil – 2025
Fonte: CGEE e LNCC (2025)

Dentre os níveis de ensino, tem-se que dos cursos levantados até o momento, 39 são cursos de pós-graduação e 31 de graduação. E dentre os de pós, 23 cursos são de mestrado e doutorado acadêmicos e 16 de especialização (Figura 6).

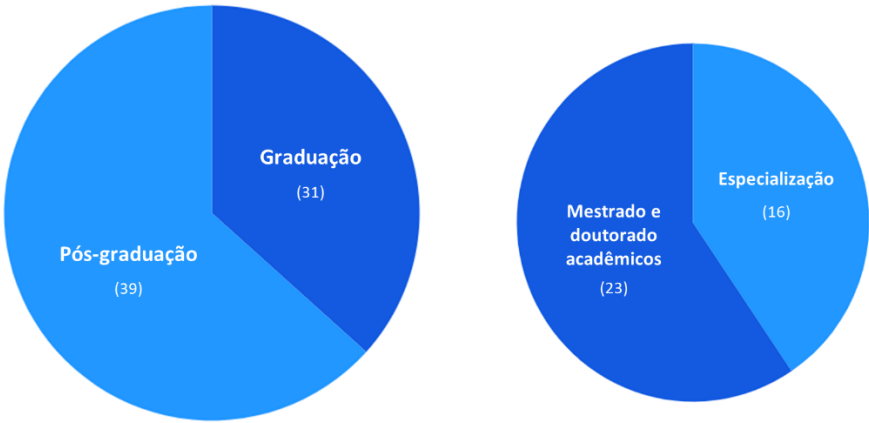


Figura 6. Cursos de IA – Graduação e Pós-Graduação – Brasil – 2025
Fonte: CGEE e LNCC (2025)

Dentre as áreas do conhecimento da CAPES, os cursos estão distribuídos por 8 grandes áreas e 16 áreas do conhecimento, conforme Figura 7, em que se observa o destaque da área da Ciência da Computação, seguida de Probabilidade e Estatística e Engenharias (Elétrica e Produção).

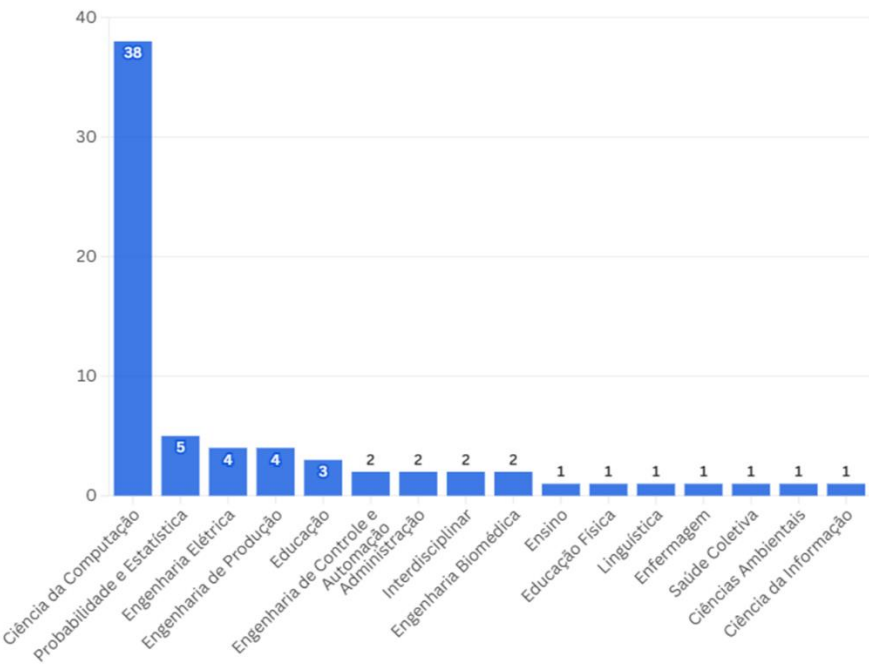


Figura 7. Áreas do conhecimento CAPES – Graduação e Pós-Graduação – Brasil – 2025
Fonte: CGEE e LNCC (2025)

Considerando apenas os bacharelados específicos em IA, há 20 cursos identificados na pesquisa, com 10 cursos iniciados em 2026, como se observa no Gráfico 1, distribuídos por quase todo o país, como se observa na Figura 8.

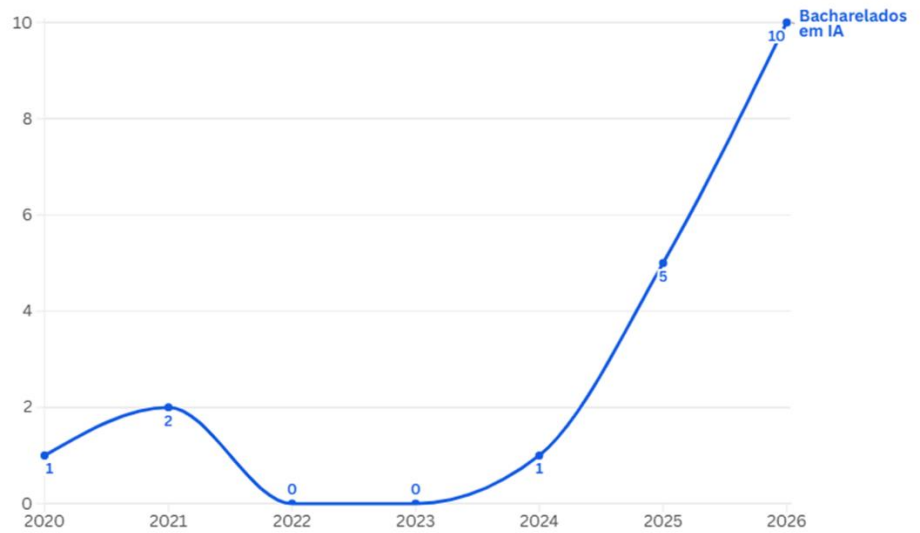


Gráfico 1. Bacharelado específico em IA
Fonte: CGEE e LNCC (2025)

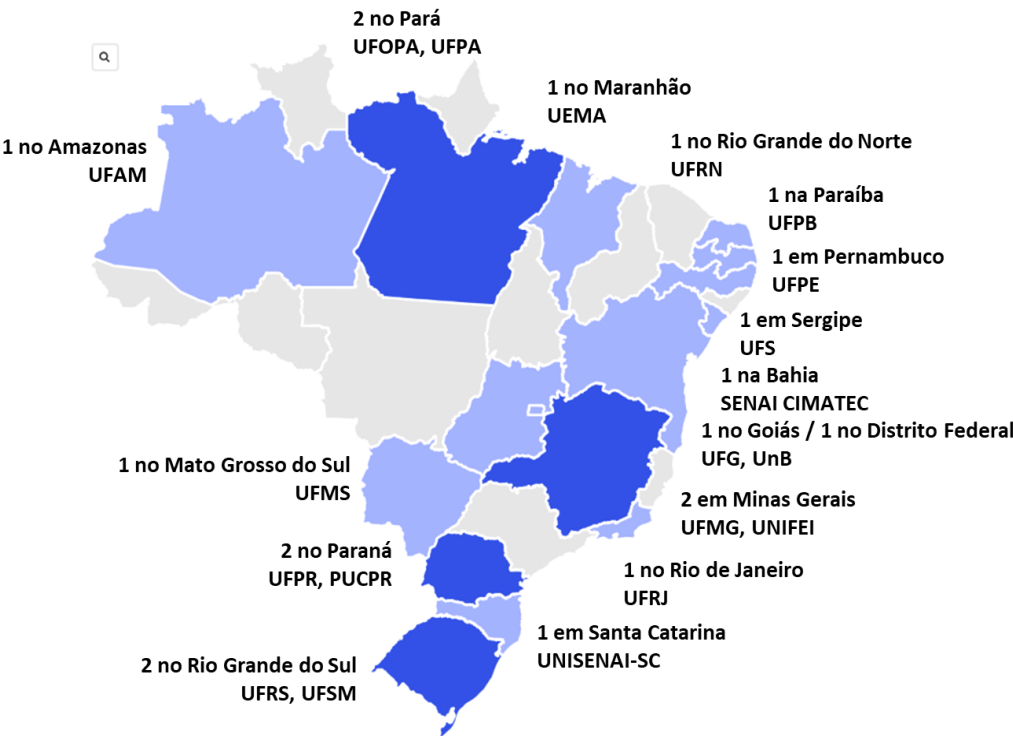


Figura 8. Bacharelado específico em IA
Fonte: CGEE e LNCC (2025)

Ações com entregas

Além da ação 16, dos cursos de IA, também há entregas na ação 14, *Difusão, divulgação e letramento em IA*, com a realização da 1º Olimpíada Nacional de IA (ONIA), com 716 mil estudantes inscritos e 4 representantes brasileiros na Olimpíada Internacional de IA, em Pequim, na China.

A ação 15, *Criação e ampliação de LIFEs*, também está em andamento com a criação de 27 novos laboratórios e ampliação de 20 laboratórios existentes. As ações 17 e 18, *Bolsas de IA para Graduação e Pós* e *Bolsas de Doutorado em IA no Exterior*, apresentaram dispêndio de cerca de R\$ 30 milhões em pagamento de bolsas (CAPES), com previsão de novas chamadas de bolsas de IA para 2026.

Na linha da capacitação, a ação 19, *Qualificação profissional em IA*, empreendida em grande medida pelo Sistema S com objetivo de ofertar cursos online de IA e cursos presenciais com estágios em empresas, está em implementação, com 40 mil inscritos em cursos do SEBRAE até novembro de 2025, considerando a meta do PBIA de 20 mil profissionais qualificados no 1º ano.

Ação	Envolvidos	Principal Entrega	Recursos Utilizados
Ação 14: Difusão, divulgação e letramento em IA	MCTI, LNCC, CNPq, MEC	1ª Olimpíada Nacional de IA (ONIA): 716 mil estudantes inscritos. 4 estudantes representaram o Brasil na Olimpíada Internacional (27% meninas premiadas).	Apoio institucional LNCC
Ação 15: Criação e ampliação de LIFEs	CAPES, MEC	Criação de 27 novos laboratórios e ampliação de 20 laboratórios existentes (até 2028).	Em apuração
Ação 16: IA na Graduação	MEC, CAPES	Cursos autorizados em UFMG, UFPB, UFPR e UEL. Total de 73.275 vagas no país (8.104 criadas em 2024). CENSO de cursos de graduação em andamento. Meta PBIA: 5 mil vagas/ano.	Em apuração
Ação 17: Bolsas de IA para Graduação e Pós	CAPES, CNPq, MEC	CAPES: Pagamento de bolsas de pós-graduação. CNPq: Lançamento de chamada de bolsas de IA para graduação e pós previsto para dez/2026.	R\$ 25,8 milhões (CAPES)
Ação 18: Bolsas de Doutorado em IA no Exterior	CAPES, CNPq, MEC	CAPES: Pagamento de bolsas de Doutorado no exterior. CNPq: Lançamento de chamada de bolsas de doutorado em IA no exterior previsto para dez/2026.	R\$ 4,4 milhões (CAPES)
Ação 19: Qualificação profissional em IA	Sistema S	SEBRAE: Curso de IA na prática com 40 mil inscritos (meta: 20 mil). Acordo de Cooperação Técnica com Microsoft, Google e Alura.	R\$ 80 mil

Tabela 4. Ações com entregas – Eixo 2 - PBIA

Ações em preparação

Dentre as ações em preparação estão a ação 20, *IA para Educação e Trabalho*, que atualmente está em fase de estruturação do programa, definição de parcerias e da plataforma para oferta dos cursos de capacitação. Para 2026 estão previstos o lançamento do portal de cursos de IA, estabelecimento de até 5 parcerias institucionais para capacitação/requalificação e capacitação de 2000 alunos por meio dos programas ofertados. Esta ação ganhou um novo componente, *Programa Nacional de Capacitação*, que está em formatação com previsão de R\$ 388 milhões.

Ação	Envolvidos	Status Atual / Próximos Passos	Previsão
Ação 20: IA para Educação e Trabalho	MEC, MCTI	Lançamento do portal de cursos em IA. Até 5 parcerias institucionais para capacitação/requalificação. 2.000 alunos capacitados por empresas parceiras.	1º sem/2026
Novo componente (Ação 20): Programa nacional de capacitação	MEC, MCTI	Iniciativa em formatação com valor previsto de R\$ 388 milhões.	Lançamento 1º sem/2026
Ação 21: Residência em TICs-IA	MCTI, FINEP	Chamada Pública para programas de formação em IA, visando capacitar até 60 mil desenvolvedores e usuários até 2028.	Até 2028

Tabela 5. Ações em preparação – Eixo 2 - PBIA

1.2.2. Detalhamento das ações - Eixo 2

Programa de difusão e divulgação da IA

Ação 14. Difusão, divulgação e literacia digital em IA

Meta: Criação de uma Olimpíada de IA em 12 meses (recorrência anual), com participação de escolas públicas de todas as regiões do País; aumento do percentual da população brasileira que declara ter um bom entendimento sobre IA para 85% (pesquisa de opinião Ipsos/AI Index Report) em dois anos.

Entregas realizadas: Realização da 1ª Olimpíada de IA (ONIA), com 716 mil estudantes inscritos na primeira etapa. 04 estudantes destaques representaram o Brasil na Olimpíada Internacional de IA em Pequim (08/2025).



Fonte do recurso: FNDCT não-reembolsável.
Instituições Envolvidas: MCTI, LNCC, CNPq e MEC.

COM ENTREGAS

Programa de Formação em IA

Ação 15. Laboratórios Interdisciplinares de Formação de Educadores (Lifes)

Meta: Criação de 27 novos Lifes; e ampliação de 20 Lifes existentes.

Entregas realizadas: Criação de 27 novos LIFEs e ampliação de 20 LIFEs existentes.



COM ENTREGAS

Fonte do recurso: CAPES
Instituições Envolvidas: CAPES e MEC.

Ação 16. IA na Graduação

Meta: Criação de pelo menos 5.000 vagas em cursos de IA em três anos; e disponibilização automática de 100% das vagas em cursos de ciência de dados e IA no Fies.

Entregas realizadas: Foram autorizados cursos de graduação em IA na UFMG, UFPB, UFPR e UEL. No total, a educação superior brasileira já conta com 73.275 vagas autorizadas nesses cursos, sendo 8.104



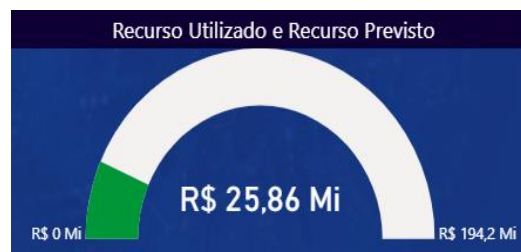
COM ENTREGAS

Fonte do recurso: CAPES
Instituições Envolvidas: MEC e CAPES

Ação 17. Bolsas de IA para Graduação e Pós-Graduação

Meta: Metas específicas de cada iniciativa em definição (Capes e CNPq).

Entregas realizadas: Pagamento de bolsas de pós-graduação (CAPES).



COM ENTREGAS

Fonte do recurso: CAPES e FNDCT não-reembolsável.
Instituições Envolvidas: CAPES, CNPq e MEC.

Ação 18. Bolsa de doutorado de IA no Exterior

Meta: Metas específicas de cada iniciativa em definição (Capes e CNPq).

Entregas realizadas: Pagamento de bolsas de Doutorado no Exterior (CAPES).



COM ENTREGAS

Fonte do recurso: CAPES e FNDCT não-reembolsável.
Instituições Envolvidas: CAPES, CNPq e MEC.

Programa de Capacitação, Qualificação e Requalificação em IA

Ação 19. Qualificação Profissional em IA

Meta: Qualificação de 20 mil profissionais no 1º ano, 30 mil no 2º ano e 50 no 3º ano; e implementação de 500 projetos-piloto de IA em empresas industriais em três anos.

Entregas realizadas: Curso de IA na prática para pequenos negócios (01/25) com 40 mil inscritos. Curso de IA na prática para gestores públicos.



COM ENTREGAS

Fonte do recurso: Sistema S e FNDCT não-reembolsável
Instituições Envolvidas: Sistema S

Ação 20. IA para Educação e Trabalho

Meta: Múltiplas metas, incluindo o estabelecimento de cinco grupos de pesquisa em seis meses no âmbito

Entregas previstas: Lançamento do Portal para disponibilização dos cursos de capacitação ofertados pelas empresas parceiras; Estabelecimento de até 5 parcerias institucionais para capacitação ou requalificação de profissionais em IA; Capacitação de 2000 alunos por meio dos programas ofertados pelas empresas parceiras; Programa Nacional de Capacitação em IA: iniciativa em formatação (valor previsto R\$ 388 milhões).



INICIADA

Fonte do recurso: Setor privado
Instituições Envolvidas: MEC e MCTI

Ação 21. Residência em TICs-IA

Meta: Múltiplas metas, incluindo o estabelecimento de cinco grupos de pesquisa em seis meses no âmbito

Entregas realizadas: Chamada pública para seleção de programas de formação e capacitação em IA, com foco em alcançar até 60 mil alunos desenvolvedores e usuários de IA até 2028.

**INICIADA**

Fonte do recurso: Setor privado e Projetos Prioritários da Lei de TICs (PPI)
Instituições Envolvidas: MCTI, FINEP e Softex.

1.3. Eixo 3. IA para Melhoria dos Serviços Públicos

Este eixo se concentra no desenvolvimento e implementação de soluções de IA para abordar gargalos específicos na administração pública e aprimorar a eficiência governamental, melhorando a qualidade dos serviços e processos. Envolve a identificação de áreas prioritárias para aplicação de IA, como saúde, educação, agricultura, meio ambiente, e segurança pública, e o desenvolvimento de projetos ainda em fase de ideação ou conceitualização. Também inclui a criação de diretrizes para a aquisição e uso ético de sistemas de IA pelo governo, e a utilização de IA para melhorar a tomada de decisões e a oferta de serviços públicos.

1.3.1. Principais entregas - Eixo 3

Ao longo de 2025, foi estruturado o Projeto IA Brasil, também chamado de INSPIRE, *Inteligência Artificial no Serviço Público com Inovação, Responsabilidade e Ética*, com foco na transformação digital do Estado brasileiro. A previsão de investimentos é de R\$ 390 milhões, sendo R\$ 100 milhões ainda em 2025 com encerramento em 2028. O projeto comporta um conjunto de ações que melhor captam as necessidades do eixo 3, em relação as ações inicialmente previstas no PBIA. As ações do projeto INSPIRE correspondem as ações 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 33 e 34 do PBIA. Dentre as novas ações estão:

- IA para catalogação e interoperabilidade de dados e serviços
- Plataforma *low-code* de soluções complexas com IA Generativa
- Núcleo de IA de Governo – NIA: desenvolvimento de soluções de IA em apoio a demais áreas do governo
- Núcleo de tecnologias voltadas para Segurança Cibernética no contexto das ameaças advindas do crescimento da IA
- IA para qualificação de dados cadastrais (gestão benefícios)
- IA para comunicação personalizada e melhoria de experiência dos usuários em programas sociais
- IA para personalização na difusão de notícias, informações de políticas públicas e serviços de Governo para gestores e população.

O objetivo do Projeto IA Brasil/INSPIRE é criar plataformas de IA para dar suporte e otimizar o desenvolvimento de soluções para políticas públicas, com base em dados qualificados, interoperáveis e protegidos, de modo a viabilizar a entrega de serviços públicos personalizados. Isto será feito por meio da catalogação e integração de diversas bases de dados do ecossistema de dados nacionais, como CadÚnico, saúde, educação, etc. O projeto também prevê o uso de IA para verificar a qualidade dos dados, assegurar sua interoperabilidade e integração entre as bases e aplicações e, também, desenvolver

chatbots e sistemas de recomendação de serviços de governo digital para o cidadão. De forma a estruturar o projeto, o CPQD está investindo na infraestrutura dedicada em suas instalações, com GPUs para o desenvolvimento de aplicações e treinamento de modelos de IA.

O Projeto IA Brasil/INSPIRE tem como envolvidos o Ministério da Gestão e da Inovação em Serviços Públicos (MGI) em parceria com o Centro de Pesquisa e Desenvolvimento em Telecomunicações (CPQD), a partir de uma encomenda do FNDCT, administrado pelo MCTI, com gestão pela FINEP.

Ações com entregas

Dentre as ações com entregas, estão as ações do Projeto INSPIRE, com 58 soluções de IA prospectadas e 16 priorizadas, 2 soluções de IA experimentadas, diagnóstico de adoção de IA nos órgãos públicos em 2024, realização de trilhas de capacitação em IA (12 mil servidores capacitados). Também foi catalogado mais de mil conjuntos de dados entre janeiro e julho de 2025, além do lançamento da Caixa Postal Gov.Br e contratação de plataforma de relacionamento com o cidadão. A ação 37, não inclusa no Projeto INSPIRE, também apresentou entregas, com a contratação de empresa de consultoria para elaboração do Termo de Referência da ação.

Ação	Envolvidos	Principal Entrega	Recursos Utilizados
Projeto IA Brasil (Ações 22-34)	MGI, MCTI, FINEP, CPQD	7 entregas principais: catalogação e interoperabilidade de dados, plataforma low-code com IA Generativa, Núcleo de IA de Governo (NIA), tecnologias de cibersegurança, qualificação de dados cadastrais, comunicação personalizada e difusão de informações.	R\$ 100 milhões (2025)
Ação 23: Prospecção e estruturação de projetos	MGI, MCTI, Serpro, Dataprev, FINEP	MGI: 58 soluções de IA prospectadas e 16 priorizadas.	R\$ 210 mil
Ação 24: Experimentação de projetos de IA	MGI, MCTI, FINEP, UnB, SERPRO, Dataprev, CPQD	MGI: 2 soluções de IA experimentadas.	R\$ 960 mil
Ação 25: Monitoramento do desenvolvimento e uso da IA	MGI, MCTI, FINEP, CPQD	MGI: Entregue autodiagnóstico de adoção de IA nos órgãos públicos do ano de 2024.	Parte de R\$ 4 milhões
Ação 26: Capacitação de servidores públicos	MGI, MCTI, FINEP, Enap, Serpro, CPQD	MGI: 4 trilhas de capacitação em IA realizadas (12 mil servidores capacitados) e 1 Cartilha de IA Generativa lançada.	Parte de R\$ 7,5 milhões

Ação 28: Catalogação e governança de dados	MGI, MCTI, FINEP, CPQD, CGU	MGI: Jan–jul/2025: 1.002 conjuntos de dados catalogados (meta EFGD: 2.000 até 2026).	Parte de R\$ 6 milhões
Ação 30: Personalização dos serviços públicos	MGI, SERPRO, Dataprev, CPQD, FINEP	MGI: Caixa Postal GOV.BR lançada e plataforma de relacionamento com cidadão contratada.	R\$ 700 mil
Ação 37: IA para gestão de pessoas	MGI, MCTI, FINEP, Enap	MGI: Consultoria com a Catálise para construção do TR.	Parte de R\$ 13 milhões

Tabela 6. Ações com entregas – Eixo 3 - PBIA

Ações iniciadas

Dentre as ações iniciadas estão a ação 27, Nuvem de Governo, que busca criar infraestrutura de nuvem privada ou comunitária gerida exclusivamente por órgãos ou empresas públicas. A ação está em implementação pelo SERPRO e Dataprev. A ação 40, Sistema de Previsão de Extremos Climáticos (SIPEC), está em andamento, com a estruturação do projeto e construção de oficinas para desenvolvimento do sistema.

Ação	Envolvidos	Status Atual / Próximos Passos	Previsão
Ação 27: Nuvem de Governo	SERPRO, Dataprev	Infraestrutura de nuvem soberana em implementação pelo SERPRO e Dataprev	Em implementação
Ação 40: SIPEC - Sistema de Previsão de Extremos Climáticos	MCTI, FINEP	Estruturação do projeto e construção de oficinas para desenvolvimento do sistema	Em andamento

Tabela 7. Ações iniciadas – Eixo 3 - PBIA

Ações em preparação

Dentre as ações em preparação, ou seja, ainda em estágio inicial estão a ação 35, IA para contratações públicas, ação 36, IA para gestão dos recursos da educação, ação 38, IA para Patrimônio da União, e ação 39, IA nas empresas estatais. Todas as ações estão em fase de elaboração de propostas, com exceção da ação 36, concernente ao MEC, que está em processo de revisão, bem como todas as demais ações referentes ao MEC.

Ação	Envolvidos	Status Atual / Próximos Passos	Previsão
Ação 35: IA para contratações públicas	MGI	Proposta em elaboração	Em elaboração
Ação 36: IA para gestão dos recursos da educação	MEC, FNDE	IA para simplificação e automatização da gestão e prestação de contas dos recursos do FNDE e Programa Dinheiro Direto na Escola (PDDE)	Em elaboração
Ação 38: IA para Patrimônio da União	MGI	Proposta em elaboração	Em elaboração
Ação 39: IA nas empresas estatais	MGI	Proposta em elaboração	Em elaboração

Tabela 8. Ações em preparação – Eixo 3 – PBIA

1.3.2. Detalhamento das ações – Eixo 3

Núcleo de IA do Governo Federal

Ação 22. Plataforma de Inteligência Artificial do Governo

Meta: Disponibilização da plataforma para desenvolvimento e operação em alta escala de projetos e modelos de IA para o governo, até 2026.

Entregas realizadas: IA para catalogação e interoperabilidade de dados e serviços; Núcleo de IA de Governo (NIA): desenvolvimento de soluções de IA em apoio a demais áreas do governo.

Entregas previstas: 3 soluções de IA generativa.

COM ENTREGAS



Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MGI, MCTI, FINEP, SERPRO, DATAPREV e CPQD

Ação 23. Prospeção e Estruturação de Projetos de IA

Meta: Disponibilização da plataforma para desenvolvimento e operação em alta escala de projetos e modelos de IA para o governo, até 2026.

Entregas realizadas: 58 soluções de IA prospectadas e 16 soluções de IA priorizadas e estruturadas; Núcleo de IA de Governo (NIA): desenvolvimento de soluções de IA em apoio a demais áreas do governo.

Entregas previstas: 4 soluções de IA prospectadas, priorizadas e estruturadas.

COM ENTREGAS



Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MGI, MCTI, SERPRO, Dataprev e FINEP

Ação 24. Experimentação de projetos de IA no governo

Meta: Disponibilização da plataforma para desenvolvimento e operação em alta escala de projetos e modelos de IA para o governo, até 2026.

Entregas realizadas: 2 soluções de IA experimentadas; Núcleo de IA de Governo (NIA); desenvolvimento de soluções de IA em apoio as demais áreas de governo.
Entregas previstas: 4 soluções de IA experimentadas.



Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MGI, MCTI, FINEP, UnB, SERPRO, Dataprev e CPQD

COM ENTREGAS

Ação 25. Monitoramento do Desenvolvimento e Uso da IA

Meta: Realização de levantamento anual sobre adoção da IA nos órgãos e entidades do governo federal.

Entregas realizadas: Autodiagnóstico de adoção de IA nos órgãos públicos do ano de 2024; Núcleo de IA de Governo (NIA); desenvolvimento de soluções de IA em apoio a demais áreas do governo.



Entregas previstas: Autodiagnóstico 2025.

Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MGI, MCTI, FINEP e CPQD

COM ENTREGAS

Ação 26. Capacitação de servidores públicos em IA

Meta: Capacitação de 115.000 servidores federais, até 2026 (20% do total)

Entregas realizadas: 4 trilhas de capacitação em IA realizadas (12 mil servidores capacitados) e 1 Cartilha de IA Generativa para o Serviço Público lançada; Núcleo de IA de Governo (NIA); desenvolvimento de soluções de IA em apoio a demais áreas de governo.
Entregas previstas: Lançamento de 5 trilhas de capacitação em IA; 1 cartilha de IA generativa (Guia de Prompt para servidores públicos).



Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MGI, MCTI, FINEP, ENAP, SERPRO e CPQD

COM ENTREGAS

Ação 27. Nuvem Soberana

Meta: Migração de dados classificados em grau de sigilo para ambiente de nuvem do governo, até 2027.



INICIADA

Fonte do recurso: SERPRO e Dataprev
Instituições Envolvidas: MGI, Serpro e Dataprev

Ação 28. Catalogação, Governança e Estratégia de Uso de Dados

Meta: Estabelecimento da política de governança de dados nos órgãos e entidades federais; elevação da maturidade de dois para três, até 2026; e catalogação de 2.000 conjuntos de dados do governo federal, até 2027.

Entregas realizadas: Jan-jul/202: 1.002 conjuntos de dados catalogados (meta EFGD: 2.000 até 2026); IA para catalogação e interoperabilidade de dados e serviços.

Entregas previstas: Publicação do Decreto de Governança de Dados.



COM ENTREGAS

Fonte do recurso: FNDCT não reembolsável
Instituições Envolvidas: MGI, MCTI, FINEP, CPQD e CGU

Ação 29. Integração e Réuso Estratégico de Dados

Meta: Economia de R\$ 6 bilhões com a utilização do Programa Conecta Gov.br, ao reduzir as exigências de documentos dos cidadãos, até 2026.

Entregas realizadas: IA para catalogação e interoperabilidade de dados e serviços.



COM ENTREGAS

Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MGI, CPQD, MCTI e FINEP

Ação 30. Personalização dos Serviços Públicos

Meta: Personalização da comunicação digital em 50 serviços públicos digitais, até 2026.

Entregas realizadas: Caixa Postal Gov.br lançada; Plataforma de relacionamento com cidadão contratada; IA para comunicação personalizada e melhoria de experiência dos usuários em programas sociais; IA para personalização na difusão de notícias, informações de políticas públicas e serviços de Governo para gestores e população.

COM ENTREGAS



Fonte do recurso: MGI e FNDCT não-reembolsável
Instituições Envolvidas: MGI, SERPRO, Dataprev, CPQD e FINEP

Ação 31. Privacidade e Segurança da Informação no Setor Público

Meta: Incremento no valor médio do índice de maturidade em privacidade e segurança da informação em dois décimos, até 2027.

Entregas realizadas: Reuniões preparatórias, planejamento para atendimento da ação, aguardando liberação de orçamento; Núcleo de tecnologias voltadas para Segurança Cibernética no contexto das ameaças advindas do crescimento da IA. Entregas previstas: Incremento de 0,05 no valor médio do índice de maturidade em privacidade e segurança da informação.

COM ENTREGAS



Fonte do recurso: FNDCT não-reembolsável.
Instituições Envolvidas: MGI, MCTI, FINEP e CPQD.

Ação 32. Infraestrutura para uso e aplicação de IA na educação

Meta: Estabelecimento de infraestrutura inicial a partir da organização, integração e operacionalização das bases de dados educacionais do MEC, até 2026.

Entregas realizadas: Em desenvolvimento em parceria com startup do MGI em. IA para catalogação e interoperabilidade de dados e serviços.

EM PREPARAÇÃO



Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MGI, MCTI, FINEP e CPQD

Programa de Soluções de IA para Serviços Públicos

Ação 33. Soluções de IA para o governo

Meta: Lançamento de duas chamadas anuais; e implementação de projetos estruturantes de IA em dez órgãos/ano.

Entregas realizadas: Priorização das soluções de IA, contratação da empresa Catalise para iniciação da etapa de enquadramento do problema (*design thinking*); 1. IA para catalogação e interoperabilidade de dados e serviços. 2. Plataforma *low-code* de soluções complexas com IA Generativa; 3. Núcleo de IA de Governo (NIA): desenvolvimento de soluções de IA em apoio a demais áreas do governo; 4. Núcleo de tecnologias voltadas para Segurança Cibernética no contexto das ameaças advindas do crescimento da IA; IA para qualificação de dados cadastrais (gestão benefícios); IA para comunicação personalizada e melhoria da experiência dos usuários em programas sociais.



Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MGI, MCTI, FINEP, ENAP e CPQD

COM ENTREGAS

Ação 34. Desenvolvimento de IA para Cibersegurança do Governo

Meta: Capacitação de 100% das equipes dos órgãos setoriais do Sistema de Administração dos Recursos de Tecnologia da Informação (Sisp), até 2026; e implantação de módulos de detecção e resposta a ameaças em 100% dos órgãos setoriais, até 2027.

Entregas realizadas: Reuniões preparatórias, planejamento para atendimento da ação, aguardando liberação de orçamento; Proposta em elaboração pelo GSI e MGI; Núcleo de tecnologias voltadas para Segurança Cibernética no contexto das ameaças advindas do crescimento da IA.



Fonte do recurso: FNDCT
Instituições Envolvidas: Não informado

INICIADA

Ação 35. Desenvolvimento de IA para aperfeiçoamento das contratações públicas

Meta: incorporação do uso de ferramentas de IA nas rotinas de contratação pública.

EM PREPARAÇÃO



Ação 36. IA para gestão dos recursos da educação

Meta: Desenvolvimento e implementação de solução para gestão de processos em até 24 meses.

Entregas realizadas: Em construção, estando em fase de testes e aplicação em processos do FNDE.

EM PREPARAÇÃO



Ação 37. Uso de IA para aperfeiçoar os processos de Gestão de Pessoas

Meta: Implementação de sete projetos de IA de alto impacto, relacionados à área de Gestão de Pessoas no serviço público, até 2028.

Entregas realizadas: Consultoria com a Catálise para construção do TR.

COM ENTREGAS



Ação 38. Uso de IA para aperfeiçoar os processos de patrimônio da União

Meta: Implementação de nove projetos de IA de alto impacto, relacionados à área de Gestão do Patrimônio da União, até 2028.

Entregas previstas: Proposta em elaboração.



EM PREPARAÇÃO

Fonte do recurso: FNDCT
Instituições Envolvidas: MGI

Ação 39. Apoio ao desenvolvimento de projetos de IA nas empresas estatais

Meta: Identificação de projetos de IA em curso nas empresas estatais, oportunidades para indução do ecossistema e desenvolvimento de soluções de IA no País.

Entregas previstas: Proposta em elaboração.



EM PREPARAÇÃO

Fonte do recurso: FNDCT
Instituições Envolvidas: MGI

Ação 40. SIPEC – Sistema inteligente de previsão de extremos climáticos

Meta: Estabelecimento de infraestrutura inicial a partir da organização, integração e operacionalização das bases de dados educacionais do MEC, até 2026.



INICIADA

Fonte do recurso: FNDCT
Instituições Envolvidas: MCTI e FINEP

1.4. Eixo 4. IA para inovação empresarial

Este eixo visa a promover o acesso amplo ao uso da IA no setor privado, incentivando a adoção de soluções de IA por empresas de todos os portes e setores. Busca apoiar, consolidar e expandir os setores que participam da cadeia produtiva da IA, incluindo as indústrias de Tecnologias da Informação e Comunicação (TIC) e microeletrônica. Simultaneamente, visa a aproveitar a IA para avançar a agenda da industrialização, posicionando o País como polo global de desenvolvimento e uso de IA.

1.4.1. Principais entregas do Eixo 4

Dentre os destaques do eixo estão a ampliação das unidades Embrapii, no âmbito da ação 42, *Apoio sistêmico à cadeia de valor de IA*, que tinha como metas o aumento de 20% ao ano no número de projetos de PD&I em IA, ampliação para 35 no número de Unidades Embrapii atuando em projetos de IA, estruturação da rede Embrapii de Competência em IA e a contratação de 5 projetos estruturantes em IA. Em relação ao número de projeto de PD&I em IA houve o crescimento de 116%, entre 2023 e 2024, e crescimento de 21%, de janeiro a setembro 2025, sobre o mesmo período de 2024. E há 23 Unidades Embrapii com IA nas competências técnica, além de 50 unidades que já contrataram projetos de IA, além de chamada para credenciamento de 3 novas Unidades para atuação em IA (proposta depende de recursos adicionais). A rede Embrapii de Competência em IA está em estruturação e a contratação de 5 projetos estruturantes em IA necessitam de recursos específicos para lançamento de chamada de projetos de alto impacto em IA.

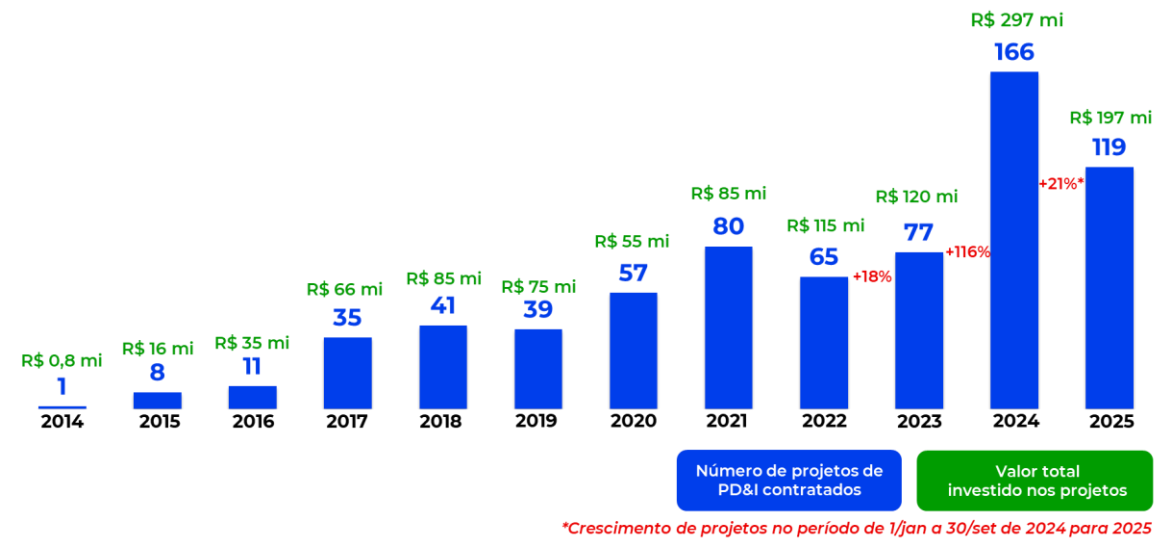


Gráfico 2. Evolução dos projetos da Embrapii em IA – 2014 – 2025
Fonte: Embrapii (2025)

O Gráfico 2 apresenta a evolução dos projetos da Embrapii em IA, entre 2014 e 2025, com destaque para o crescimento de 112,8% no número de projetos com IA, de 77 a 166 projetos, entre 2023 e 2024, e aumento de 143%, com valores de R\$ 122 milhões a R\$ 296 milhões, para o mesmo período.

A ação 48, *Soluções de IA para as missões da NIB*, também apresentou entregas expressivas, com apoio a 596 projetos por parte da FINEP, na ordem de R\$ 4,5 bilhões. Dentre os projetos de destaque estão:

Subvenção Direta (chamadas públicas) – R\$ 478 milhões, sendo destaques:

- Mais Inovação Brasil – Tecnologias Digitais (R\$ 137M e 15 projetos)
- Finep Startup IA (R\$ 64M e 41 projetos)
- Bioeconomia e Transformação Digital da Amazônia (R\$ 12,7M contratado e 8 projetos)
- Inovação para a Base Industrial de Defesa (R\$ 26M e 2 projetos)
- Mais Inovação – Cadeias Agroindustriais Sustentáveis (R\$ 35M e 3 projetos)
- Mais Inovação – Energias Renováveis (R\$ 16M e 2 projetos)
- Mais Inovação – Semicondutores (R\$ 12M e 01 projeto)
- Mais Inovação – Aviação Sustentável (R\$ 33M e 2 projetos)
- Mais Inovação Saúde (R\$ 5M e 1 projeto)
- Empresas Startups em Tecnologias Habilitadoras – Linha Temática V – Robótica - R\$ 19.110.678,79 em 10 projetos

Crédito (R\$ 2,214 bi)

- Direto: R\$ 1,7 bi (21 projetos)
- Descentralizado: R\$ 444M (150 projetos)

Como exemplo, o Mais Inovação Brasil – Tecnologias Digitais, apresentou os seguintes projetos aprovados:

Projetos aprovados – Tecnologias Digitais

Pré-treino e ajuste fino de LLM para interpretação e análise jurídica Descrição: pré-treinar e ajustar um LLM especializado no português e no domínio jurídico brasileiro. Utilizando dados gerais e jurídicos, como legislações e jurisprudência, o modelo será capaz de interpretar documentos jurídicos. Orçamento: R\$ 9,5 milhões Instituições envolvidas: Neuralmind, UFCA, UFG.	Arquitetura generativa auto personalizável Descrição: solução de IA generativa com LLMs para desenvolver projetos arquitetônicos autopersonalizáveis, traduzindo descrições textuais em opções alinhadas a normas técnicas e preferências subjetivas. Orçamento: R\$ 9,8 milhões Instituições envolvidas: Arqgen, USP, UNICAMP.	IA Generativa para cuidado personalizado em saúde integral Descrição: assistente de IA para atenção primária à saúde digital, ajudando a identificar pacientes prioritários, propor cuidados personalizados via mensagens e monitorar o engajamento para planejar ações futuras. Orçamento: R\$ 3,6 milhões Instituições envolvidas: Health Atendimento de Saúde Responsável e UFMG.
--	---	--

Projetos Aprovados – Cadeia do Agroindustrial

Drones para agricultura de precisão

Descrição: o projeto visa aprimorar a aeronave GD-25 para melhorar a agricultura de precisão, utilizando IA e análise de imagens. Inclui testes, certificação e desenvolvimento de plataforma para ortomosaicos e dados geolocalizados. O foco é reduzir recursos hídricos e emissões de carbono.

Orçamento: R\$ 5,4milhões (FNDCT não-reembolsável)

Instituições envolvidas: Global Drones

Plataforma digital de gestão, monitoramento para sustentabilidade

Descrição: o projeto desenvolve uma plataforma digital para monitorar e gerenciar emissões de GEE, otimizando operações agrícolas para menor impacto ambiental. Reduzir diesel, consumo e melhorar o manejo do solo, aumentando o sequestro de carbono.

Orçamento: R\$ 5,1 milhões (FNDCT não-reembolsável)

Instituições envolvidas: Bosh

Sensores com IA para agricultura de precisão

Descrição: o projeto SF Cortex AI desenvolve sensores com IA para aplicação precisa de insumos agrícolas, acopláveis a máquinas tradicionais. Inclui interface, telemetria, conexão online e controle inteligente para maior

Orçamento: R\$ 11,7 milhões (FNDCT não-reembolsável)

Instituições envolvidas: Savefarm

Exemplos de ações de crédito

Plano Hídrico e Indústria 4.0

Descrição: desenvolvimento de produtos e processos focados no conceito de Indústria 4.0 (IA, IoT, Gêmeos Digitais, Computação de Borda, Advanced Analytics, etc).

Orçamento: R\$ 150 milhões (Crédito)

Instituições envolvidas: São Martinho

Inovação e integração em saúde: arquiteturas interoperáveis e soluções assistida por IA

Descrição: desenvolvimento de soluções tecnológicas disruptivas para promover a interoperabilidade, a integração de dispositivos médicos e a aplicação

Orçamento: R\$ 97,4 milhões (Crédito)

Instituições envolvidas: Bosh

Ecossistema provedor de inteligência conversacional

Descrição: criação de um ecossistema provedor de inteligência conversacional direcionado pelo conceito de AI-First – utilizando IA e técnicas de machine learning com arquiteturas de nuvem.

Orçamento: R\$ 92,4 milhões (Crédito)

Instituições envolvidas: Blip

Ações com entregas

Além das duas ações supracitadas, também apresentaram entregas a ação 41, *Desenvolvimento de datacenters nacionais*, R\$ 180 milhões do BNDES, com o financiamento de um projeto de datacenter. A ação 44, *IA para MPes e MEIs*, com parceria do SEBRAE, também apresentou entregas, com estudos e desenvolvimento de soluções de IA para as micro e pequenas empresas. E a ação 45, *Recursos Humanos em Áreas Estratégicas – IA*, tem como previsão o lançamento de nova chamada RHAE-IA até dezembro de 2025.

Ação	Envolvidos	Principal Entrega	Recursos Utilizados
Ação 41: Desenvolvimento de datacenters nacionais	BNDES, FINEP	BNDES: Desde o lançamento do PBIA, o BNDES aprovou o financiamento de um projeto de datacenter	R\$ 180 milhões (BNDES)

Ação 42: Apoio à cadeia de valor de IA	FINEP, MCTI, Embrapii	Embrapii: Ampliação para 23 unidades em IA. Projetos com IA cresceram 112,8% (78 → 166 entre 2023/24). Valores aportados +143% (R\$ 122 mi → R\$ 296 mi)	R\$ 102,4 milhões
Ação 44: IA para MPEs e MEIs	BNDES, FINEP, SEBRAE	SEBRAE: Estudos e desenvolvimento de soluções de IA. Avatar 3D com IA generativa. Projeto com 9 microsserviços de IA para atendimento personalizado (10/25) e aprimoramento até 2028	Em apuração
Ação 45: Recursos Humanos em Áreas Estratégicas (RHAE)	MCTI, CNPq	Nova chamada RHAE – IA	R\$ 100 milhões (até 2027)
Ação 48: Soluções de IA para as missões da NIB	BNDES, FINEP	FINEP: Apoio a 596 projetos até 7/2025 (R\$ 4,5 bilhões). BNDES: Apoio a 12 projetos desde o PBIA (R\$ 1,2 bilhões)	R\$ 5,7 bilhões

Tabela 9. Ações com entregas – Eixo 4 – PBIA

Ações iniciadas

Dentre as ações iniciadas estão ação 43 e ação 47. A ação de *Apoio a startups de IA* no momento apresenta plano de trabalho e contrato de cooperação técnica em elaboração. E o *Centro Nacional de IA para a Indústria (CNIA4I)* está em fase de análise de proposta enviada pelo SENAI CIMATEC.

45

Ação	Envolvidos	Status Atual / Próximos Passos	Previsão
Ação 43: Apoio a startups de IA	SEBRAE, CEIA-UFG	SEBRAE: Elaboração de plano de trabalho e contrato de cooperação técnica com CEIA-UFG. Edital para seleção de 30 startups para programa de incubação (soluções inovadoras para pequenos negócios)	Até 12/2025 e até 2028
Ação 47: Centro Nacional de IA para a Indústria (CNIA4I)	SENAI CIMATEC, MCTI	Análise da proposta enviada pelo SENAI CIMATEC	Previsão de encomenda até 12/25

Tabela 10. Ações iniciadas – Eixo 4 – PBIA

Ações em preparação

Ação	Envolvidos	Status Atual / Próximos Passos	Previsão
------	------------	--------------------------------	----------

Ação 46: Retenção de talentos para inovação em IA	MCTI, CNPq	Lançamento da chamada de Retenção de talentos para inovação em IA	Até dez/2026
Ação 49: IA para aumento de produtividade de MPMEs e industriais	MCTI, SENAI	Tema sendo definido juntamente com ação 47. Realizado evento sobre Data Space no âmbito da Cooperação Brasil-Alemanha	Em definição

Tabela 11. Ações em preparação – Eixo 4 – PBIA

1.4.2. Detalhamento das ações – Eixo 4

Programa de Fomento à Cadeia de Valor da IA

Ação 41. Desenvolvimento de data centers nacionais

Meta: Metas específicas por iniciativa (BNDES e Finep

Entregas realizadas: BNDES: Desde o lançamento do PBIA o BNDES aprovou o financiamento a um projeto de data center no montante de R\$ 180 milhões.



COM ENTREGAS

Fonte do recurso: BNDES
Instituições Envolvidas: BNDES e FINEP

Ação 42. Apoio sistêmico à cadeia de valor da IA

Meta: Aumento de 20% a.a. do número de projetos de PD&I em IA; estruturação da Rede Embrapii de Competência em IA com 25 unidades e Centros de Competência, até 2026; ampliação de 15 para 35 no número de Unidades Embrapii habilitadas para projetos em IA, até 2026; e contratação de cinco projetos estruturantes em IA, até 2026.

Entregas realizadas: Ampliação para 23 unidades Embrapii em IA; projetos com IA cresceram 112,8% (78 para 166 entre 2023/24); valores aportados +143% (R\$ 122 mi para R\$ 296 mi, sendo R\$ 102,4 mi da Embrapii).



COM ENTREGAS

Fonte do recurso: FNDCT não-reembolsável.
Instituições Envolvidas: FINEP, MCTI e Embrapii

Ação 43. Apoio a startups de IA

Meta: Metas específicas por iniciativa (BNDES e Finep)

Entregas previstas até dez/2025:

SEBRAE: Plano de trabalho e contrato de cooperação técnica com o CEIA-UFG



INICIADA

Fonte do recurso: SEBRAE
Instituições Envolvidas: SEBRAE e CEIA/UFG

Ação 44. IA para MPes e MEIs

Meta: Programas de aceleração em IA em MPE de setores intensivos; execução de projetos-piloto de inclusão digital em três estados; e desenvolvimento e aplicação de estudos para aprimoramento no suporte aos MEI em 12 meses.

Entregas realizadas: Elaboração de estudos e desenvolvimento de soluções de IA para pequenos negócios; Avatar 3D com IA generativa para interação com MPes em eventos e agências SEBRAE. Projeto com 9 microsserviços de IA para atendimento personalizado (10/25) e aprimoramento do projeto (até 2028).



COM ENTREGAS

Fonte do recurso: SEBRAE e setor privado
Instituições Envolvidas: BNDES, FINEP e SEBRAE

Ação 45. Recursos Humanos em Áreas Estratégicas (RHAE) - IA

Meta: Concessão de 1.200 bolsas para pesquisadores, até 2026, focadas no desenvolvimento de soluções de IA em MPME.

Entregas previstas até dez/2025:

Lançamento de chamada de Recursos Humanos em Áreas Estratégicas (RHAE) em IA.



COM ENTREGAS

Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MCTI e CNPq

Ação 46. Retenção de talentos para inovação em IA

Meta: Apoio a 1.000 profissionais qualificados em IA em projetos de PD&I empresariais, até 2026, com retenção no País dos formados em IA.

Entregas previstas entre 2026 e 2028:
Lançamento da chamada de Retenção de Talentos para inovação em IA (2º sem/2026)



EM PREPARAÇÃO

Fonte do recurso: FNDCT não-reembolsável e setor privado
Instituições Envolvidas: MCTI e CNPq

Programa de IA para desafios da indústria brasileira

Ação 47. Centro Nacional de IA para a Indústria (CNIA4I)

Meta: Composição de redes de instituições em 12 meses; modelagem e início do programa de capacitação em 12 meses; e inauguração do CNIA4I em 18 meses.

Entregas realizadas: Proposta em estudo.
Entregas previstas entre 2026 e 2028:
Previsão de encomenda no 1º trim/2026



INICIADA

Fonte do recurso: FNDCT não-reembolsável e setor privado
Instituições Envolvidas: MCTI

Ação 48. Soluções de IA para missões da NIB

Meta: Financiamento de pelo menos 500 projetos de IA aplicados à indústria, até 2028.

Entregas realizadas: FINEP: apoio a 596 projetos até 7/2025 (R\$ 4,5 bilhões).
BNDES: apoio a 12 projetos desde o PBIA (R\$ 1,2 bilhões)



COM ENTREGAS

Fonte do recurso: BNDES e FINEP
Instituições Envolvidas: BNDES e FINEP

Ação 49. IA para aumento da produtividade de MPMEs industriais

Meta: Estruturação de base de dados e levantamento de requisitos do data space em seis meses; implantação da infraestrutura digital pública; e capacitação de instituições em um ano

Entregas previstas: em definição juntamente com a ação 47. Realização de evento sobre data space no âmbito da Cooperação Brasil-Alemanha.



EM PREPARAÇÃO

Fonte do recurso: Não informado
Instituições Envolvidas: MCTI e SENAI

1.5. Eixo 5. Apoio ao processo regulatório e de governança da IA

O Eixo 5 busca contribuir para a consolidação de um arcabouço de governança de IA no Brasil que promova a inovação, assegure o direito ao desenvolvimento, proteja os direitos humanos, a integridade da informação, os direitos autorais e os que lhe são conexos, o trabalho e os trabalhadores, e posicione o Brasil como referência em IA responsável e confiável.

1.5.1. Principais entregas do Eixo 5

O eixo 5 possui duas ações que já possuem entregas substantivas, como os *Guias Brasileiros de IA Responsável*, tendo sido lançado o Guia de IA Generativa no serviço público (SERPRO/MF/MGI), MJSP: ACT SEDIGI/MJSP-USP para pesquisas e MEC: versão preliminar do Referencial de Uso Responsável de IA na Educação (aguardando consulta pública) e workshop multissetorial.

Outra ação é o *Observatório Brasileiro de IA (OBIA)*, que tem como objetivo compilar, registrar e prover informações sobre os avanços das tecnologias habilitadoras da IA possibilitando análises sobre sua adoção e os seus principais impactos sobre a sociedade. O OBIA é uma iniciativa coordenada pelo NIC.br, com apoio do C4AI/USP, CGEE e Fundação SEADE.

Dentre outras informações, o OBIA provê indicadores brasileiros relacionados a IA, como uso e adoção da IA por setores como educação, empresas, governo e saúde; produção do conhecimento; e formação na área. A parceria com o CGEE vai na direção de desenvolver indicadores sobre produção científica e patentes em IA para o OBIA.



O OBIA também apresenta outros indicadores, realizados em parceria com o MGI e o CETIC.br, trazendo informações relevantes sobre a adoção da IA pela sociedade brasileira. O Observatório também apresenta estudos e outras informações relevantes sobre o tema da inteligência artificial.

Ações com entregas

Ação	Envolvidos	Principal Entrega	Recursos Utilizados
------	------------	-------------------	---------------------

Ação 50: Guias Brasileiros de IA Responsável	MGI, MJSP, USP, MEC, MMA, MDHC, SERPRO	SERPRO/MF/MGI: Guia de IA Generativa no serviço público lançado. MJSP: ACT SEDIGI/MJSP-USP para pesquisas. MEC: versão preliminar do Referencial de Uso Responsável de IA na Educação (aguardando consulta pública) e workshop multissetorial	Em apuração
Ação 52: Observatório Brasileiro de IA (OBIA)	CGI.br, NIC.br, MCTI, CGEE, C4AI e Fundação SEADE	NIC.br: Lançamento do OBIA em setembro/2024. Desenvolvimento de indicadores em parceria com CGEE e MGI. Realização de 2 seminários do OBIA. Disponível em: https://obia.nic.br .	Em apuração

Tabela 12. Ações com entregas – Eixo 5 – PBIA

Ações em preparação

As ações 51, *Centro Nacional de Transparência Algorítmica e IA Confiável*, ação 53, *Programa Política com Ciência – rede de apoio para governança*, e ação 54, *Participação do Brasil no debate internacional*, estão em elaboração, aguardando propostas para estruturação das ações.

Ação	Envolvidos	Status Atual / Próximos Passos	Previsão
Ação 51: Centro Nacional de Transparência Algorítmica e IA Confiável	MCTI	Proposta em elaboração	Em elaboração
Ação 53: Programa Política com Ciência – rede de apoio para governança	MCTI	Proposta em elaboração	Em elaboração
Ação 54: Participação do Brasil no debate internacional	MCTI, MRE	Proposta em elaboração	Em elaboração

Tabela 13. Ações em preparação – Eixo 5 – PBIA

1.5.2. Detalhamento das ações – Eixo 5

Programa de Aperfeiçoamento do Marco Regulatório para IA

Ação 50. Guias Brasileiros de IA Responsável

Meta: *Elaboração e publicação do Guia para IA Ética e Responsável em três meses; lançamento do Guia de IA para o Setor Público em 6 meses; e realização de workshops e eventos periódicos para divulgação dos guias.*

Entregas realizadas: SERPRO, MF e MGI: Lançamento de Guia de IA Generativa no serviço público. MJSP: ACT SEDIGI/MJSP-USP para pesquisas em IA. MEC: versão preliminar do Referencial de Uso Responsável de IA na Educação (aguardando consulta pública); realização de workshop multissetorial.

COM ENTREGAS

Fonte do recurso: Ação orçamentária
Instituições Envolvidas: MGI, MJSP, USP, MEC, MMA, MDHC e SERPRO

Ação 51. Centro Nacional de Transparência Algorítmica e IA Confiável

Meta: *Estabelecimento do Centro em até 120 dias; e publicação de estudo sobre riscos da IA e práticas de transparência algorítmica em seis meses após criação do Centro.*

Entregas previstas: Envio de proposta que estrutura um consórcio entre instituições (ICTs, INCTs, CPAs e CGI.br)

EM PREPARAÇÃO

Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MCTI

Programa de Apoio à Governança da IA**Ação 52. Observatório Brasileiro de IA (Obia)**

Meta: *Lançamento do Obia em 30 dias; e desenvolvimento contínuo de indicadores nas principais dimensões da IA.*

Entregas realizadas: Lançamento do OBIA; Desenvolvimento contínuo de indicadores em parceria com CGEE e MCGI; realização de 2 seminários do OBIA.

COM ENTREGAS

Fonte do recurso: NIB.br
Instituições Envolvidas: CGI.br, NIC.br, MCTI, CGEE, C4AI, Fundação Seade

Ação 53. Rede de apoio à governança da IA no Brasil

Meta: Estruturação da rede em 12 meses.

Entregas previstas até dez/2025: Publicação de portaria criando o Sistema Nacional de Tecnologia e Inovação em IA (SinIA); Estabelecimento ao CGEE para criação de núcleo de estudos avançados sobre soberania digital da IA, incluindo análises sobre aspectos estratégicos, técnicos e regulatórios; Estabelecimento de rede de ICTs para avançar na discussão de estratégia sobre impactos na população vulnerável.



Fonte do recurso: FNDCT não reembolsável
Instituições Envolvidas: MCTI

EM PREPARAÇÃO

Ação 54. Rede de apoio à participação do Brasil no debate internacional

Meta: Estruturação da rede em 12 meses.

Entregas previstas até dez/2025: Apoio ao estabelecimento da Rede Internacional de Colaboração em IA generativa soberana; Contratação de estudo abrangente sobre riscos, oportunidades e soluções seguras de compartilhamento de dados nacionais em cooperações internacionais para o desenvolvimento de modelos de IA generativa aberta.



Fonte do recurso: FNDCT não-reembolsável
Instituições Envolvidas: MCTI e MRE

2. Ações de Impacto Imediato

O PBIA possui 37 ações de impacto, distribuídas por 9 áreas: Agricultura e Pecuária; Defesa, Segurança Pública e Cibernética; Desenvolvimento Social; Educação; Gestão dos Serviços Públicos; Indústria, Comércio e Serviços; Meio Ambiente, Clima e Sustentabilidade; Saúde; e Trabalho e Emprego.

Dentre as ações de impacto, até o final de novembro, havia 1 ação concluída, a ação 9, Sistema Gestão Presente, da área de Educação. E 9 ações com entregas, nas áreas da Saúde (3), Educação (3), Gestão dos Serviços Públicos (2) e Defesa, Segurança Pública e Cibernética (1).



Figura 9. Dados gerais das ações de impacto – PBIA – 30/11/2025
Fonte: CGEE (2025)

Agricultura e Pecuária		
Ação	Envolvidos	Status Atual / Próximos Passos
Ação 08. Ater Digital - IA	Ematers, SNPA, Serpro, MGI, MCTI, MDA, MCOM.	Em preparação. Primeira etapa concluída
Ação 24. Cálculo do peso bovino por câmera 3D	Finep	Em acompanhamento
Ação 29. Embrapa – Rural GPT	Embrapa	Proposta em elaboração

Defesa, Segurança Pública e Cibernética		
Ação 37. ProtegIA para a segurança da população	MJSP, SESP, PNUD e PRF	SENASP/MJSP: Portaria de diretrizes nacionais para câmeras corporais. Edital selecionou 11 estados (R\$ 155 mi). Repasses do FNSP: R\$ 11,4 mi. 26 cursos com 917 policiais capacitados. PRF/MJSP: projeto de câmeras com testes e implantação inicial no RJ.
Desenvolvimento Social		
Ação 15. Acredita no primeiro passo IA	MDS	Em acompanhamento
Educação		
Ação 9. Sistema Gestão Presente	MEC, UFAL, Dataprev	Ação concluída. O SGP integra dados de 47,3 milhões de estudantes, centralizando informações acadêmicas (notas, frequência, matrícula, etc.). Permite monitoramento em tempo real, integração via API, transferências simplificadas e garante segurança e transparência dos dados.
Ação 10. Controle da qualidade das aquisições de alimentos para o PNAE	MEC e FNDE	Treinamento e gestão do sistema; Experimentação; Integração do sistema nas rotinas do FNDE
Ação 11. Sistema de predição e proteção de trajetória dos estudantes	MEC, NESS/UFAL, SEE-RO, SEE-MT, SEE-MG e SEE-MA	Piloto em fase de implementação, com formações, coleta de dados, diagnóstico na plataforma, intervenções em execução (até set/2025) e reavaliação em nov/2025.
Ação 12. Soluções adaptativas com IA generativa de avaliação formativa e diagnóstica para alfabetização e letramento	-	Em acompanhamento
Ação 13. Sistemas inteligentes de tutoria de matemática com IA generativa	-	Em acompanhamento
Ação 14. MelhorIA da aprendizagem e bem-estar dos estudantes	-	Em acompanhamento
Ação 33. Observatório de egressos da EPT	-	Em acompanhamento
Ação 34 SISSA - RFEPC	-	Em acompanhamento

Gestão dos Serviços Públicos		
Ação 16. Fiscaliza IA – Receita Federal do Brasil (RFB)	RFB e MF	Prova de conceito concluída (05/25) com resultados insatisfatórios, mas gerou aprendizados técnicos/metodológicos. Solução em redesenho, com novo ciclo de desenvolvimento e testes em andamento na Receita Federal.
Ação 17. Atende IA – Exterior	MRE	Chatbot ativo em 15 postos
Ação 35. Malha Fina IA	CGU	Em acompanhamento
Indústria, Comércio e Serviços		
Ação 26. IA para gestão jurídica	FINEP	Em acompanhamento
Ação 27. IA para comunicação corporativa	FINEP	Em acompanhamento
Meio Ambiente, Clima e Sustentabilidade		
Ação 25. IA para quantificação do estoque florestal do bioma Amazônico	-	Em acompanhamento
Ação 30. Projeto cidades verdes resilientes +IA	-	Em acompanhamento
Ação 31. Sistema de IA para o controle do desmatamento na cadeia da madeira da pecuária e da mineração e regularização ambiental de imóveis rurais	-	Em acompanhamento
Ação 32. Combate à biopirataria	-	Em acompanhamento
Saúde		
Ação 01: Prontuário falado no SUS	-	Em acompanhamento
Ação 02: IA para suporte a decisões de compras de medicamentos no SUS	MS, SECTICS, DAF, CGEAF	MS: Estruturação de piloto com 10 medicamentos e validação em 8 algoritmos com acurácia superior a 82% em modelos prévios
Ação 03: Otimização dos diagnósticos no SUS e	-	Em acompanhamento

predição prognóstica com IA		
Ação 04: IA em saúde bucal no SUS	-	Em acompanhamento
Ação 05: IA para detecção de anomalias nos procedimentos hospitalares e ambulatoriais no SUS	MS, DEMAS, CGDID, DRAC, SAES, DenaSUS	MS: Desenvolvimento de algoritmos de detecção de anomalias
Ação 06: IA para suporte à gestão de processos de judicialização no SUS	MS, DEMAS, CGDID, DJUD, SE	MS: Planejamento e prova de conceito, desenvolvimento da arquitetura de IA, protótipo funcional para extração de dados jurídicos e validação do protótipo
Ação 07: Idoso bem cuidado no SUS	-	Em acompanhamento
Ação 18: Desinfecção autônoma de ambientes	-	Em acompanhamento
Ação 19: IA e Big Data para tratamento de câncer	-	Em acompanhamento
Ação 20: IA Generativa para personalização no cuidado da saúde	-	Em acompanhamento
Ação 21: Prevenção de AVC e cardiopatia em clientes de saúde suplementar	-	Em acompanhamento
Ação 22: Retina sem Anomalias	-	Em acompanhamento
Ação 23: IA para otimização de reembolso de saúde	-	Em acompanhamento
Ação 28: IA nos hospitais da Ebserh/MEC	-	Em acompanhamento
Trabalho e Emprego		
Ação 36. Emprega + IA	-	Em acompanhamento

Tabela 14. Ações de impacto – status e entregas

3. Novas ações a serem incorporadas ao PBIA

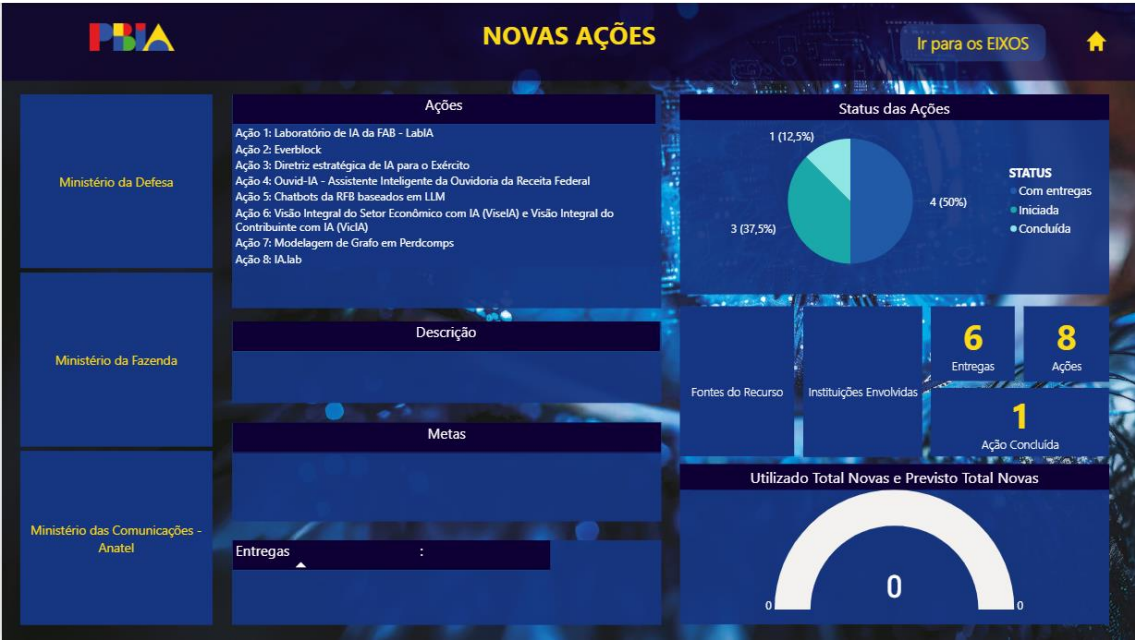


Figura 10. Novas ações a serem incorporadas ao PBIA

Ministério da Defesa	
Ação	Descrição
Laboratório de IA da FAB - LabIA	Criado para transformar operações da FAB por meio da IA. Suas aplicações incluem manutenção preditiva de aeronaves, simulação de cenários de combate, otimização logística, segurança cibernética e uso de LLMs para apoiar decisões e reduzir custos.
Everblock	Solução tecnológica brasileira para armazenamento seguro e imutável de dados, combinando blockchain pública e privada. Suas aplicações incluem registro oficial de documentos, rastreamento administrativo e logístico, defesa cibernética e validação de contratos inteligentes.
Diretriz estratégica de IA para o exército	Define orientações para uso da IA na Força Terrestre, visando apoiar decisões, aumentar a eficiência e promover pesquisa e inovação, com ênfase em controle humano, ética e segurança da informação.
Ministério da Fazenda	
OuvdIA – Assistente Inteligente da Ouvidoria da Receita Federal	Solução de IA para o SisOuvidor que usa modelos de linguagem natural para classificar manifestações, identificar urgências, gerar respostas e relatórios inteligentes com apoio de RAG, análise de sentimentos e mineração de dados.

Chatbots da RFB baseados em LLM	Desenvolvimento de chatbots para otimizar o atendimento ao contribuinte, ampliar o acesso à informação e reduzir a sobrecarga dos canais tradicionais.
Visão Integral do Setor Econômico com IA (ViselA) e Visão Integral do Contribuinte com IA (VicIA)	Solução de IA para elaboração automática de relatórios estratégicos Vise e Vic, para subsidiar planejamento e tomada de decisão.
Modelagem em Grafo de Perdcomps	IA para apuração de créditos e débitos de PIS e COFINS, identificando padrões, fraudes e pedidos corretos para agilizar pagamentos.
ANATEL	
IA.lab	Laboratório de IA da Anatel para formulação de estudos sobre IA no setor de telecomunicações, além de subsidiar decisões estratégicas e regulatórias. Apoiará a revisão do Regulamento Geral de Serviços de Telecomunicações

Tabela 15. Novas ações a serem incorporadas ao PBIA - descrição

4. Atuação das instituições no PBIA – CNPq, Embrapii e FINEP

4.1. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

O CNPq possui 08 ações no Plano Brasileiro de IA, que estão distribuídas por 3 eixos, conforme Quadro 3. O Eixo 1 possui 3 ações, com previsão de investimentos de R\$ 353 milhões entre 2025 e 2028. O Eixo 2, Difusão, Formação e Capacitação, possui 3 ações com investimentos de R\$ 350 milhões e o Eixo 4, IA para Inovação Empresarial, 2 ações, no montante de R\$ 220 milhões para o período de 2025 a 2028.

Eixo	Programa	Ação	2025-2028 (em Milhões)	2025 (em Milhões)	2026 (em Milhões)	2027 (em Milhões)	2028 (em Milhões)
EIXO 1 – Infraestrutura e Desenvolvimento de IA	1.4. Programa de Pesquisa e Desenvolvimento em IA	10. Fomento às atividades de P&D em IA	160,00	-	60,00	50,00	50,00
		11. INCTs – IA	93,00	25,00	25,00	23,00	20,00
		13. Fomento a P&D em cooperação com América Latina, Caribe e África	100,00	-	33,00	33,50	33,50
EIXO 2 – Difusão, Formação e Capacitação	2.1. Programa de Difusão e Divulgação da IA	14. Difusão, divulgação e literacia digital em IA	90,00	-	30,00	30,00	30,00
	2.2. Programa de Formação em IA	17. Bolsas de IA para Graduação e Pós-Graduação	180,00	-	60,00	60,00	60,00
		18. Bolsas de doutorado de IA no Exterior	80,00	-	27,00	27,00	26,00
EIXO 4 – IA para Inovação Empresarial	4.1. Programa de Fomento à Cadeia de Valor da IA	45. Recursos Humanos em Áreas Estratégicas (RHAe) – IA	100,00	10,00	45,00	45,00	-
		46. Retenção de talentos para inovação em IA	120,00	-	40,00	40,00	40,00
TOTAL CNPq			923,00	35,00	320,00	308,50	259,50

Quadro 3. Ações do CNPq no Plano Brasileiro de IA – 2025-2028

Fonte: CNPq e PBIA (2025)

A ação sob responsabilidade do CNPq com maior previsão de investimentos é a ação 17, *Bolsas de IA para Graduação e Pós-Graduação*, com R\$ 180 milhões ao longo do período. Em seguida vem a ação 10, *Fomento às atividades de P&D em IA*, com R\$ 160 milhões, e a ação 46, *Retenção de talentos para inovação em IA*, com aporte de R\$ 120 milhões. Importante mencionar o caráter estratégico destas ações, para a formação de mão-de-obra qualificada na área de IA, impulsionamento das atividades de P&D em IA no Brasil, por meio da produção científica e tecnológica em IA aplicada no Brasil, além da formação de nova geração de pesquisadores e o regresso de pesquisadores.

Entre 2024 e 2025 o CNPq investiu R\$ 162 milhões em infraestrutura e desenvolvimento da IA, sendo R\$ 25 milhões em INCTs – IA e R\$ 137 milhões em bolsas PQ/DT. E investiu R\$ 1 milhão em divulgação científica e R\$ 3,4 milhões em capacitação.



Gráfico 3. Investimentos CNPq – 2024-2025

Fonte: CNPq (2025)

Analisando período anterior, entre 2022 e 2025, é possível captar considerável volume de recursos investidos em projetos de Pesquisa e Desenvolvimento (P&D), Divulgação Científica e Capacitação de Recursos Humanos em temas relacionados a Inteligência Artificial. O CNPq investiu R\$ 259 milhões em cerca de 3.000 projetos na área no período mencionado.

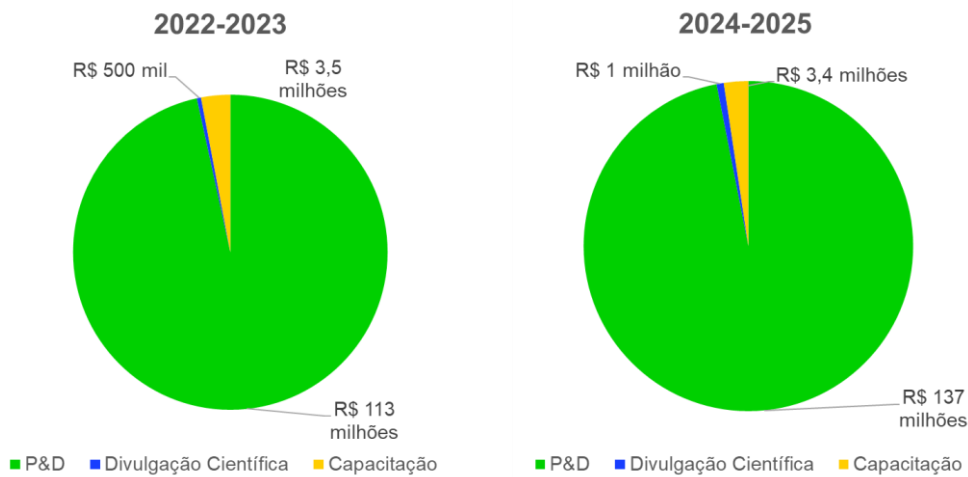


Gráfico 4. Investimentos CNPq em Projetos de P&D, Divulgação Científica e Capacitação de RH em IA – 2022-2025

Fonte: CNPq (2025)

Além disso, é importante mencionar novamente o volume de recursos investidos nos INCTs – IA, conforme Quadro 4. São 08 INCTs, com valor total de R\$ 93 milhões ao longo do período 2025 e 2028. Os INCTs possuem aporte médio de R\$ 11,6 milhões, chegando a R\$ 14,4 milhões concedidos para o INCT para Engenharia de Software Baseada em e para IA (INES.IA), e R\$ 13 milhões para o INCT Amazônia: Instituto Nacional de Ciência e Tecnologia em Inteligência Artificial aplicada às cidades inteligentes e sustentáveis na Amazônia brasileira.

Nome do INCT	Instituição	Área Predominante	Valor Total da Concessão (2025-2028)
Instituto Nacional de Inteligência Artificial para o Bem Social (IAPROBEM)	USP	Ciências Exatas e da Terra	R\$ 11.470.000,00
Computação Neural	UFPE	Ciências Exatas e da Terra	R\$ 12.441.000,00
Computação Quântica Aplicada	UFPE	Ciências Exatas e da Terra	R\$ 10.192.000,00
Inteligência Artificial e Sustentabilidade (INCT-IAS)	PUCR	Ciências Humanas	R\$ 11.255.000,00

Instituto Nacional de Ciência e Tecnologia para Engenharia de Software Baseada em e para Inteligência Artificial (INES.IA)	UFPE	Ciências Exatas e da Terra	R\$ 14.376.000,00
Instituto Nacional de Ciência e Tecnologia (INCT) em Inteligência Artificial Responsável para Linguística Computacional e Tratamento e Disseminação de Informação (TILD-IAR)	UFMG	Ciências Exatas e da Terra	R\$ 9.709.000,00
INCT IAmazônia: Instituto Nacional de Ciência e Tecnologia em Inteligência Artificial aplicada às cidades inteligentes e sustentáveis na Amazônia brasileira	UFPA	Engenharias	R\$ 13.075.000,00
INCT em Inteligência Artificial para Cibersegurança: Uma Abordagem Tecnológica e Social (INCT-IACiber)	UFMG	Ciências Exatas e da Terra	R\$ 10.303.000,00
TOTAL			R\$ 92.821.000,00

Quadro 4. INCTs-IA – IA

4.1.1. Impactos resultantes das ações empreendidas pelo CNPq

Com os investimentos realizados no período de 2022-2025, o CNPq despertou a vocação científica em ia de 962 estudantes do ensino médio e fundamental, por meio do programa de bolsas de iniciação científica júnior. No mesmo período foram apoiados 340 estudantes de graduação em temas relacionados à IA, pelo programa de bolsas de iniciação científica. No nível da pós-graduação o CNPq apoiou a formação de 773 estudantes de mestrado e 500 de doutorado, entre 2022 e 2025, no desenvolvimento e aplicação de soluções em IA para a resolução de problemas nas mais diversas áreas do conhecimento.

Em relação aos pesquisadores brasileiros residentes no exterior, um problema recorrentemente relatado como uma das causas da baixa disponibilidade de mão-de-obra qualificada no Brasil na área de IA, o CNPq promoveu o retorno, em 2024, de 96 pesquisadores para atuar em projetos de pesquisa relacionados à inteligência artificial, por meio do Programa Conhecimento Brasil. Ainda, o CNPq, por meio do Programa de Formação de Recursos Humanos em Áreas Estratégicas (RHAE), realizou chamadas no período de 2022 a 2025, que capacitaram 531 estudantes em 82 empresas para atuar no desenvolvimento de projetos em temas relacionadas à IA, notadamente nas áreas de ciência da computação, medicina, engenharia elétrica e agronomia.

Considerando os recursos previstos para os próximos anos, é possível afirmar que os resultados serão ainda maiores e robustos, permitindo ampliar o número e a qualidade da pesquisa em IA realizada no Brasil, assim como o número de pesquisadores formados anualmente, bem como a possibilidade de regresso de pesquisadores de alto nível que estão em instituições internacionais e podem retornar ao Brasil de modo a

contribuir para o avanço do P&D em IA no país, garantindo a soberania e o crescimento nacional.

4.2. Empresa Brasileira de Pesquisa e Inovação Industrial (Embrapii)

A Embrapii apresenta 01 ação no Plano Brasileiro de IA (PBIA), no Eixo 4, IA para Inovação Empresarial. A ação, Apoio sistêmico à cadeia de valor da IA, tem como objetivo fortalecer a capacidade produtiva nacional na cadeia de IA, por meio do aumento de projetos de PD&I em IA, aumento no número de Unidades Embrapii atuando em IA projetos de IA, estruturação de Rede Embrapii de Competência em IA e contratação de projetos estruturantes em IA.

a ampliação das unidades Embrapii, no âmbito da ação 42, Apoio sistêmico à cadeia de valor de IA, que tinha como metas o aumento de 20% ao ano no número de projetos de PD&I em IA, ampliar para 35 o número de Unidades Embrapii atuando em projeto de IA, estruturação da rede Embrapii de Competência em IA e contratação de 5 projetos estruturantes em IA. Em relação ao número de projeto de PD&I em IA houve o crescimento de 116% entre 2023 e 2024 e crescimento de 21% de janeiro a setembro 2025 sobre o mesmo período de 2024. O número de Unidades Embrapii com IA nas competências técnicas alcançou 23 e 50 unidades que já contrataram projetos de IA, além de chamada para credenciamento de 3 novas Unidades para atuação em IA (proposta depende de recursos adicionais). A rede Embrapii de Competência em IA está em estruturação e a contratação de 5 projetos estruturantes em IA necessitam de recursos específicos para lançar chamada de projetos de alto impacto em IA.

Analisando os investimentos gerais da Embrapii em PD&I entre 2014 e 2025, tem-se o total de 3.521 projetos contratados, com 1.139 ainda em execução. No universo de projetos contratados estão 2.392 empresas apoiadas, sendo 64% destas, micro e pequenas empresas. O total de recursos aportados no período em projetos de PD&I totalizam R\$ 6,7 bilhões alavancados, sendo 34,1% recursos oriundos da Embrapii, 16% das Unidades Embrapii e 49,9% das empresas apoiadas.

Ao analisar somente projetos de PD&I em IA para o mesmo período, 2014 a 2025, obtêm-se 699 projetos contratados, com 236 projetos ainda em execução. Estes projetos envolvem o apoio de 633 empresas, em que 69% são micro e pequenas empresas. O volume total de recursos alavancados por esses projetos soma R\$ 1,15 bilhão, com aporte de 49,2% das empresas, 15,1% das Unidades Embrapii e 35,6% da Embrapii. O Gráfico 5 apresenta a evolução no número de projetos e recursos investidos nos projetos de IA, com o pico de 166 projetos contratados em 2024. Os dados de 2025 se referem até 30 de setembro do mesmo ano, ou seja, estão subestimados, podendo apresentar maiores resultados até o final do ano. Observa-se considerável crescimento entre 2024 e 2025, aumento de 86,5% no número de projetos e aumento de 68% no valor total investido nos projetos.

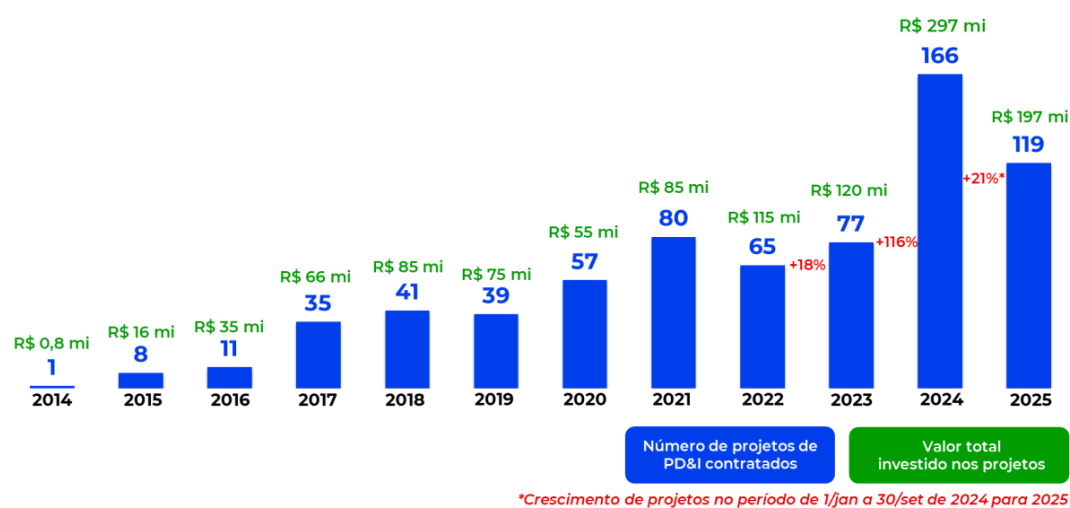


Gráfico 5. Evolução dos projetos Embrapii em IA – número e valor total investido – 2014 – 2025
Fonte: Embrapii (2025)

Em relação as Unidades Embrapii focadas em projeto de IA, atualmente há 50 unidades com projetos de IA. Em todo o Brasil há 91 Unidades Embrapii, sendo 46 no Sudeste, 17 no Nordeste, 17 no Sul, 6 no Centro-Oeste e 5 no Norte. As Unidades Embrapii com projetos em IA estão distribuídas conforme a Figura 11, com 22 unidades no Sudeste, a maioria em Minas Gerais (10), São Paulo (7), Rio de Janeiro (5), 14 unidades no Nordeste, Paraíba (4), Pernambuco (3), Ceará (3), Bahia (2), Alagoas (1) e Rio Grande do Norte (1). Na região Sul são 9 Unidades Embrapii, em Santa Catarina (4), Rio Grande do Sul (4) e Paraná (1). No Centro-Oeste são 3 Unidades, no Goiás, e no Norte, 2 Unidades no Amazonas.



Figura 11. Unidades Embrapii com projetos de IA – 2025
Fonte: Embrapii (2025)

Dentre as Unidades Embrapii com projetos de IA no período de 2022 a 2025, merecem destaque o Centro de Excelência em IA da Universidade Federal de Goiás (CEIA/UFG), o Centro de Engenharia Elétrica e Informática da Universidade Federal de Campinas Grande (CEEI-UFCG) e o Instituto Federal do Ceará (IFCE), como unidades destaque com maior número de projetos e valores aportados.

O CEIA/UFG teve 101 projetos de IA com 140 empresas ao longo dos anos 2022 a 2025, alavancando o montante de R\$ 189 milhões entre recursos Embrapii, CEIA e empresas. O CEEI-UFCG atuou em 66 projetos no período com 50 empresas, no valor total de R\$ 76 milhões. Por fim, o IFCE atuou em 36 projetos com 37 empresas no valor total de R\$ 34 milhões.

Analisando os projetos de destaque da Embrapii, o CEIA/UFG teve participação em dois projetos, o ZSCA, que desenvolve modelos de aprendizagem profundo para detecção de anomalias em imagens de colonoscopia, colaborando com a tomada de decisão do médico especialista para fins diagnósticos. Outro projeto de destaque do CEIA/UFG é o CILIA, que prevê o desenvolvimento de algoritmos para aprimorar a inspeção veicular em aplicativo que permite a realização de orçamentos para reparos em automóveis danificados em acidente com base em fotografias, sem a necessidade de um perito no local.

	Desenvolvimento de um modelo preditivo capaz de identificar riscos e falhas em linhas de gases industriais, emitindo alertas preventivos de vazamentos, rompimentos e desgastes para evitar desperdícios e acidentes.	UNIDADE EMBRAPII INDT
	Desenvolvimento de algoritmos para aprimorar a inspeção veicular em aplicativo que permite a realização de orçamentos para reparos em automóveis danificados em acidente com base em fotografias , sem a necessidade de um perito no local.	UNIDADE EMBRAPII CEIA - UFG
	Desenvolvimento de solução computacional em nuvem para otimizar o acompanhamento de pacientes diagnosticados com Hemofilia atendidos no hemocentro.	UNIDADE EMBRAPII SENAI CIMATEC
	Desenvolvimento de Dataset com medições de temperatura e pressão obtidas por sensores, junto à criação de algoritmo para correlacionar os dados com parâmetros dos poços de petróleo.	UNIDADE EMBRAPII ISI Sensoriamento
	Desenvolvimento de modelos de aprendizado profundo para detecção de anomalias em imagens de colonoscopia , colaborando com a tomada de decisão do médico especialista para fins diagnósticos.	UNIDADE EMBRAPII CEIA - UFG
	Desenvolvimento de soluções em Inteligência Artificial com técnicas de otimização para aumentar a eficiência e a sustentabilidade na indústria têxtil , por meio do planejamento otimizado a partir de parâmetros e dados.	UNIDADE EMBRAPII ISI Embarcados

Figura 12. Projetos de destaque – Embrapii – 2022-2025

Fonte: Embrapii

4.2.1. Impactos resultantes das ações empreendidas pela Embrapii

Os projetos e atividades da Embrapii, apresentados nesta seção, vem contribuindo de forma sólida e com grande impacto para a estruturação da cadeia de valor de IA no Brasil. Com a formação da infraestrutura regional para inovação em IA, com

a rede nacional de 50 Unidades Embrapii em todas as regiões, criando base institucional permanente e descentralizada para desenvolvimento tecnológico em IA.

Com a formação da Rede de Competência em IA e Centro de Competência especializado, estabelecendo capacidades duradouras que vão além de projetos pontuais, estruturando competências estratégicas no país. Ademais, as atividades apresentadas possuem um caráter de catalisador de soluções aplicadas, com portfólio de projetos conectando indústria e P&D para resolver desafios setoriais alinhados à Nova Indústria Brasil (NIB).

Como próximos passos da Embrapii de forma a intensificar suas atividades no PBIA, estão o lançamento de Chamada Pública para credenciamento do Centro de Competência em IA, na ordem de R\$ 60 milhões. Também está prevista a chamada pública para execução de projetos de alto impacto em IA, estimado em R\$ 50 milhões, e uma nova proposta, de chamada para credenciamento de 3 Unidades Embrapii para atuação em PD&I em áreas relacionadas a inteligência artificial, de R\$ 15 milhões. Para as próximas etapas a Embrapii aguarda deliberação sobre recursos adicionais para a execução das atividades.

4.3. Financiadora de Estudos e Projetos (FINEP)

A atuação da FINEP no Plano Brasileiro de IA se destaca pelo papel determinante no financiamento de projetos e ações do PBIA, por meio de recursos do Fundo Nacional de Desenvolvimento Científico e Tecnológico (FNDCT). Como mencionado na introdução, em junho de 2025 foi apresentado o Programa 11 – IA Brasil, do FNDCT, que tem como objetivo financiar de modo direcionado as ações voltadas para o desenvolvimento e adoção da IA no Brasil.

Entretanto, antes da existência do Programa 11, a FINEP já atuava no financiamento de projetos e pesquisas na área de inteligência artificial, como será detalhado nesta seção. Importante mencionar que a atuação da FINEP ocorre por diferentes instrumentos, como crédito e subvenção, ou seja, recursos reembolsáveis e recursos não reembolsáveis. A maior parte das ações do PBIA dependem de recursos reembolsáveis via FINEP e BNDES, ou seja, crédito. Por sua vez, o eixo 1, que busca desenvolver a infraestrutura de IA no país está baseada em recursos não reembolsáveis.

Segundo informações fornecidas pela FINEP, os investimentos no PBIA por parte da FINEP podem ser resumidos abaixo, em que se observa 79% dos recursos para o Eixo 4, na modalidade crédito (R\$ 2,2 bilhões), seguido do Eixo 3, 14%.

67

Eixo 1 – Infraestrutura e Desenvolvimento de IA

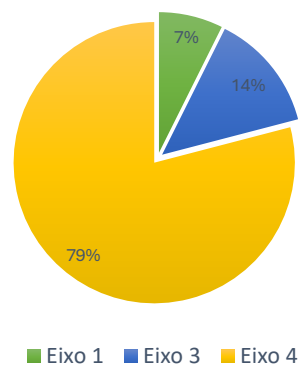
- › Liberados: R\$ 251 milhões
- › Aprovados: R\$ 45 milhões
- › Em análise: R\$ 675 milhões

Eixo 3 – IA para melhoria do serviço público

- › R\$ 462 milhões

Eixo 4 – IA para inovação empresarial (P&D)

- › Subvenção: R\$ 478 milhões
- › Crédito: R\$ 2,2 bilhões



Analisando os investimentos por eixo do PBIA, tem-se que no **Eixo 1, Infraestrutura e Desenvolvimento da IA** os seguintes projetos:

1. Centros Temáticos 2023:

- Demanda Total: R\$ 401.701.102,52 em 32 propostas
- Demanda Qualificada/Atendida: R\$ 181.723.162,08 em 14 propostas

2. Centros Temáticos 2024:

- Demanda Total: R\$ 660.988.053,70 em 50 propostas (em análise)

3. Outros investimentos

- Encomenda CENSIPAM - R\$ 14.483.768,53 (R\$ 7.453.166,17 liberados)
- Encomenda IA para Bolsa Família - R\$ 14,7 milhões (em análise)
- Projeto SisGAAz-2 (Marinha) - R\$ 74.999.402,68 (R\$ 49.999.402,68 liberados)
- Projeto Enxame de Drones (Exército) - R\$ 24.849.577,27 (R\$ 11.856.778,39 liberados)

No **Eixo 3, IA para Melhoria dos Serviços Públicos**, destacam-se os seguintes projetos:

Ação Transversal: Encomenda “Infraestrutura Nacional de Dados, Núcleo e Soluções de IA para o Serviço Público (R\$ 390M)

- Núcleo de IA do Governo Federal: Coordenação, padronização e capacitação para uso ético e responsável da IA na administração pública.
- Infraestrutura Nacional de Dados (IND): Governança, interoperabilidade e soberania de dados — incluindo a Nuvem Soberana.
- Soluções de IA para Serviços Públicos: Desenvolvimento e aplicação de soluções de IA para desafios estratégicos do setor público.

68

Subvenção Direta (Chamadas Públicas)

- Soluções IA Poder Público Rodadas 01, 02 e 03 (R\$ 72,3M e 24 projetos)

O projeto INSPIRE sob coordenação do MGI possui financiamento da FINEP, com objetivo de pesquisar e desenvolver tecnologias e aplicações de Inteligência Artificial no contexto do Eixo 3 do PBIA, “IA para melhoria do serviço público”. Este esforço visa aumentar as capacidades do governo federal na aplicação de IA de forma ética, segura e robusta, buscando a eficiência operacional na administração pública e a oferta de serviços digitais mais ágeis, eficazes e assertivos para o cidadão brasileiro.

As tecnologias e aplicações serão desenvolvidas para três pilares estratégicos:

- **Núcleo de IA do Governo Federal:** Visa a coordenar ações para transformar a administração pública por meio da IA ética e responsável. Isso inclui promover a colaboração, padronizar metodologias e recursos, e acelerar a transformação digital governamental, com suporte à definição de padrões para plataformas e capacitação de servidores públicos para projetos aplicados.

- **Infraestrutura Nacional de Dados (IND):** Conjunto de normas, políticas e ferramentas para governança, descoberta, interoperabilidade, acesso e proteção de dados. Inclui a Nuvem Soberana, que garante a segurança e soberania de dados críticos/sensíveis, armazenando-os em território nacional e assegurando privacidade. A meta é a migração de dados classificados em grau de sigilo para ambiente de nuvem do governo até 2027.
- **Programa de Soluções de IA para Serviços Públicos:** Fomenta o desenvolvimento e a implementação de soluções de IA para desafios estratégicos do setor público, promovendo a eficiência e a inovação nos serviços

Relativo à ação Soluções de IA para o poder público, foram realizadas 3 rodadas, com 24 projetos apoiados e R\$ 74,7 milhões em recursos de subvenção econômica. Os principais resultados e parceiros de cada rodada segue na Figura 13.



Figura 13. Rodadas Soluções de IA para o poder público



No **Eixo 4, IA para Inovação Empresarial**, destacam-se os seguintes projetos:

Subvenção Direta (chamadas Públicas) – R\$ 478 milhões, sendo destaques:

- Mais Inovação Brasil – Tecnologias Digitais (R\$ 137M e 15 projetos)
- Finep Startup IA (R\$ 64M e 41 projetos)
- Bioeconomia e Transformação Digital da Amazônia (R\$ 12,7M contratado e 8 projetos)
- Inovação para a Base Industrial de Defesa (R\$ 26M e 2 projetos)
- Mais Inovação – Cadeias Agroindustriais Sustentáveis (R\$ 35M e 3 projetos)
- Mais Inovação – Energias Renováveis (R\$ 16M e 2 projetos)
- Mais Inovação – Semicondutores (R\$ 12M e 01 projeto)
- Mais Inovação – Aviação Sustentável (R\$ 33M e 2 projetos)
- Mais Inovação Saúde (R\$ 5M e 1 projeto)
- Empresas Startups em Tecnologias Habilitadoras – Linha Temática V – Robótica - R\$ 19.110.678,79 em 10 projetos

Crédito (R\$ 2,214 bi)

- Direto: R\$ 1,7 bi (21 projetos)
- Descentralizado: R\$ 444M (150 projetos)

Projetos Aprovados – Tecnologias Digitais

Pré-Treino e Ajuste Fino de LLM para Interpretação e Análise Jurídica Pré-treinar e ajustar um LLM especializado no português e no domínio jurídico brasileiro. Utilizando dados gerais e jurídicos, como legislações e jurisprudências, o modelo será capaz de interpretar documentos jurídicos.  neuralmind^{ai} APOIO UFCA UNIVERSIDADE FEDERAL DO CARIRI UFPA UNIVERSIDADE FEDERAL DO PARÁ Apoio Finep R\$ 9,5 M Subvenção 0517/24	Arquitetura Generativa Auto Personalizável Solução de IA Generativa com LLMs para desenvolver projetos arquitetônicos autopersonalizáveis, traduzindo descrições textuais em opções alinhadas a normas técnicas e preferências subjetivas.  ARQGEN APOIO USP UNIVERSIDADE DE SÃO PAULO UNICAMP UNIVERSIDADE DE CAMPINAS Apoio Finep R\$ 9,8 M Subvenção 1654/24	IA Generativa para Cuidado Personalizado em Saúde Integral Assistente de IA para a Atenção Primária à Saúde Digital, ajudando a identificar pacientes prioritários, propor cuidados personalizados via mensagens e monitorar o engajamento para planejar ações futuras  HEALTH ATENDIMENTO DE SAÚDE RESPONSÁVEL LTDA APOIO FINEP Apoio Finep R\$ 3,6 M Subvenção 0501/24
---	--	--

Projetos Aprovados – Cadeia do Agroindustrial

Drones para agricultura de precisão O projeto visa aprimorar a aeronave GD-25 para melhorar a agricultura de precisão, utilizando IA e análise de imagens. Inclui testes, certificação e desenvolvimento de plataforma para ortomosaicos e dados geolocalizados. O foco é reduzir recursos hídricos e emissões de carbono.  Apoio Finep R\$5,4MM Não Reembolsável GLOBAL DRONES	Plataforma digital de gestão, monitoramento para sustentabilidade O projeto desenvolve uma plataforma digital para monitorar e gerenciar emissões de GEE, otimizando operações agrícolas para menor impacto ambiental. Reduzir diesel, consumo e melhorar o manejo do solo, aumentando o sequestro de carbono.  Apoio Finep R\$5,1MM Não Reembolsável BOSCH	Sensores com IA para agricultura de precisão O projeto SF Cortex AI desenvolve sensores com IA para aplicação precisa de insumos agrícolas, acopláveis a máquinas tradicionais. Inclui interface, telemetria, conexão online e controle inteligente para maior eficiência e sustentabilidade.  Apoio Finep R\$11,7MM Não Reembolsável SAVEFARM PODERIZAÇÃO SELETIVA
---	---	---

Exemplos de projetos aprovados (crédito)

Plano Hídrico e Indústria 4.0 Desenvolvimento de produtos e processos focados no conceito de Indústria 4.0 (Inteligência Artificial, IoT-Internet das Coisas, Gêmeos Digitais, Computação de Borda, Advanced Analytics, etc.).  Apoio Finep R\$ 150 M Crédito 0289/23 São Martinho	Inovação e Integração em Saúde: arquiteturas interoperáveis e soluções assistidas por inteligência artificial Desenvolvimento de soluções tecnológicas disruptivas para promover a interoperabilidade, a integração de dispositivos médicos e a aplicação de IA para otimizar a jornada de saúde.  Apoio Finep R\$ 97,4M Crédito 1838/24 MV PARTICIPAÇÕES	Ecosistema Provedor De Inteligência Conversacional Criação de um ecossistema provedor de inteligência conversacional direcionado pelo conceito de AI First – utilizando Inteligência Artificial e técnicas de Machine Learning com arquiteturas em nuvem.  Apoio Finep R\$ 92,4 M Crédito 1874/22 blip
--	---	---

4.3.1. Impactos resultantes das ações empreendidas pela FINEP

As ações que possuem apoio da FINEP são fundamentais para o desenvolvimento da cadeia de valor de IA no Brasil. A oferta de instrumentos de financiamento para que os projetos de IA se concretizem, especialmente aqueles que demandam investimento ou risco tecnológico elevado, promovem o desenvolvimento das tecnologias habilitadoras e suas aplicações. Também permite o apoio para que a pesquisa aplicada e infraestrutura de IA (hardware, HPC, datacenters) sejam estruturadas e difundidas.

As ações também promovem estímulo de mercado para soluções de IA, tanto em empresas como no setor público, criando demanda local. E, promovem a especialização nacional e talentos, na medida em que há apoio à contratação de profissionais especializados e capacitação no âmbito dos projetos de P&D.

Também possuem grande impacto para o desenvolvimento do ecossistema de P&D de IA no Brasil, pois os recursos permitem investir em infraestrutura mais moderna, permitindo aos pesquisadores atuarem em modelos mais complexos e treinarem redes neurais mais pesadas, resolvendo problemas exigentes de IA e realizarem pesquisa aplicada. Ademais, fomenta o desenvolvimento e estabelecimento de laboratórios nacionais com capacidade de competir internacionalmente.

As atividades da FINEP também fomentam seleções públicas específicas para apoio a iniciativas de IA, orientando os investimentos, definindo os problemas a serem priorizados e estimulando aspectos prioritários, como redução de desigualdades regionais, parceria entre empresas e ICTs e contratação de mestres e doutores. Por fim, permite a indução de demanda, visibilidade, articulação institucional e previsibilidade de recursos, que é fundamental para projetos de P&D.

Atividades de apoio à implementação do PBIA - 2025

Seminário Pilha de IA Nacional: Desafios para Autonomia Tecnológica e Soberania Digital

O Seminário Nacional sobre a Pilha de IA reuniu governo, academia, centros de pesquisa e setor privado para discutir os caminhos para construir uma **Pilha de Inteligência Artificial soberana**, abrangendo desde hardware e infraestrutura até modelos, aplicações, governança e formação profissional. O encontro evidenciou oportunidades, gargalos estruturais e encaminhamentos para políticas públicas, financiamento, cooperação e governança de dados.

Principais Temas, Desafios e Encaminhamentos

Tema 1 – Infraestrutura Tecnológica e Soberania em Hardware

O debate destacou que a base da pilha de IA depende de **HPC, servidores, GPUs/TPUs**, sistemas de interconexão, armazenamento e frameworks de execução. Países líderes possuem domínio sobre essa camada.

Gargalos e desafios

- Forte **dependência de hardware estrangeiro**, especialmente GPUs NVIDIA.
- Licenciamento proprietário que limita soberania sobre software e drivers.
- Baixa disponibilidade de serviços nacionais de HPC e altos custos de importação.
- Fragmentação das infraestruturas existentes, dificultando o compartilhamento.

Encaminhamentos

- Desenvolver política nacional de soberania digital integrada.
- Incentivar design de chips, *packaging* e modelos *fabless* (arquitetura RISC-V).
- Criar centros interconectados de HPC com governança compartilhada.
- Investir em tecnologias *open source* críticas (sistemas de arquivo paralelos, *runtimes*, *frameworks*).

Tema 2 – Camadas Superiores: Software, Frameworks e Modelos

O Brasil possui potencial para atuar com protagonismo nas camadas de maior valor agregado: bibliotecas intermediárias, modelos pré-treinados, MLOps, governança e aplicações setoriais.

Gargalos e desafios

- Baixa cultura de open source e pouca produção nacional de frameworks.
- Escassez de startups deep tech e pouca interoperabilidade entre soluções.
- Falta de padrões nacionais de integração e governança de modelos.
- Complexidade crescente da observabilidade e auditoria dos modelos.

Encaminhamentos

- Criar **plataforma nacional compartilhada** com templates e stacks reutilizáveis.
- Desenvolver **SLMs (small language models)** brasileiros com foco no português e multimodalidade.
- Fortalecer cultura open source e apoio a startups.
- Estimular reposição científica e automação via MLOps e AutoML.

Tema 3 – IA Embarcada e Dispositivos Móveis

Os debates reforçaram que muitas aplicações brasileiras exigem inferência na borda, devido à conectividade limitada e necessidade de decisões em tempo real.

Gargalos e desafios

- Limitada disponibilidade de dados de qualidade para treinar modelos embarcados.
- Restrições severas de energia, memória e processamento nos dispositivos.
- Infraestrutura nacional ainda fragmentada para pesquisa aplicada.
- Falta de domínio em técnicas de compressão, quantização e TinyML.

Encaminhamentos

- Fortalecer pesquisa aplicada em compressão de modelos, TinyML e visão computacional.
- Desenvolver hardware nacional para aplicações embarcadas.
- Integrar esforços acadêmicos, industriais e de governo em programas cooperativos.
- Priorizar aplicações estratégicas (saúde, agricultura, óleo e gás, automação).

75

Tema 4 – Ciência Aberta, Open Source e Cooperação Internacional

O painel abordou oportunidades e riscos da ciência aberta, governança de dados e internacionalização da pesquisa.

Gargalos e desafios

- Resistência institucional e desconhecimento de boas práticas de dados.
- Falta de infraestrutura nacional para repositórios interoperáveis.
- Assimetria nas parcerias internacionais, que podem gerar dependência tecnológica.
- Riscos de segurança e soberania no uso de plataformas não nacionais.

Encaminhamentos

- Criar **política nacional de ciência aberta** com padrões FAIR e segurança by design.
- Estruturar rede nacional de governança e certificação de dados científicos.
- Apoiar projetos open source estratégicos e consórcios regionais de dados.

- Fortalecer cooperação Sul-Sul e parcerias focadas em transferência tecnológica.

Tema 5 – Governança Nacional: PBIA, SINAPAD e SINIA

O Brasil avança na constituição de um ecossistema federado de HPC, IA e dados, integrando CENAPADs, LNCC e RNP.

Gargalos e desafios

- Sustentabilidade financeira e necessidade de orçamento recorrente.
- Desuniformidade técnica entre centros e baixa articulação com startups.
- Carência de programas e métricas nacionais de desempenho e impacto.

Encaminhamentos

- Implementar governança unificada entre MCTI, LNCC, RNP e CENAPADs.
- Concluir a interconexão nacional de HPC com rede dedicada de 100–200 Gbps.
- Criar fundo permanente para atualização tecnológica.
- Abrir infraestrutura pública para P&D empresarial e startups.

Tema 6 – Capacitação e Formação Profissional

A formação é central para a soberania, abrangendo desde fundamentos de computação até IA avançada.

Gargalos e desafios

- Déficit de formação profunda em computação, matemática e sistemas.
- Evasão de talentos e migração para o exterior.
- Formação excessivamente voltada ao uso de ferramentas, não aos fundamentos.
- Baixa oferta de estágios e vivências em HPC e IA aplicada.

Encaminhamentos

- Criar política nacional de formação em IA articulada com MEC, CAPES e CNPq.
- Expandir cursos de graduação e pós-graduação e incentivar mobilidade acadêmica.
- Implantar programas de treinamento em HPC, clusters, paralelismo, CUDA, SYCL etc.
- Criar cursos de verão e centros nacionais de formação aplicada.

Tema 7 – Alianças Estratégicas e Colaborações Nacionais

Diversas iniciativas isoladas existem no Brasil; o desafio é conectá-las.

Gargalos e desafios

- Baixa coordenação entre universidades, ICTs e empresas.

- Dependência de fornecedores estrangeiros para hardware crítico.
- Falta de programas nacionais de cooperação continuada.

Encaminhamentos

- Consolidar redes nacionais e regionais (ex.: consórcios estaduais, SCALAC).
- Estimular acordos de transferência tecnológica como o caso Paraná–Índia.
- Criar ecossistemas de inovação integrados entre HPC, dados e IA.

3. Conclusão Geral

O seminário revela que o Brasil possui **capacidade científica, base industrial e ativos institucionais** suficientes para liderar uma estratégia nacional de IA soberana. Contudo, isso exige:

- **Planejamento de longo prazo**, coordenação federativa e continuidade política;
- **Investimentos estruturantes** em hardware, software e dados;
- **Formação intensiva de talentos** com forte base técnico-científica;
- **Infraestrutura aberta, federada e segura**, baseada em padrões interoperáveis;
- **Cooperação estratégica nacional e internacional** com foco em transferência de tecnologia.

O conjunto de encaminhamentos desenha um caminho para transformar o potencial brasileiro em capacidade tecnológica real, reduzindo dependências e fortalecendo a soberania digital.

ANEXO 1. Relatório completo do Pilha de IA: Desafios para Autonomia Tecnológica e Soberania Digital

O evento foi idealizado pelo Ministério da Ciência, Tecnologia e Inovação (MCTI), com o apoio do Centro de Gestão e Estudos Estratégicos (CGEE), do Instituto de Inteligência Artificial do Laboratório Nacional de Computação Científica (IIA-LNCC), do Instituto de Pesquisas ELDORADO, do Instituto de Computação da UNICAMP (IC/UNICAMP) e do Serviço Federal de Processamento de Dados (SERPRO).

Comitê Técnico

André Rafael (MCTI)
Guilherme Corrêa (MCTI)
Carla Osthoff (LNCC)
Edson Borin (IC/UNICAMP)
Mateus Pierre (Instituto de Pesquisas ELDORADO)

Comitê Organizador

Caroline Nascimento Pereira (CGEE)
Isabela Quadros Dantas Barros (CGEE)
Felipe Paronetto Guines (Instituto de Pesquisas ELDORADO)
Marina Leme Bonfanti (Instituto de Pesquisas ELDORADO)

Apoio

Centro de Gestão e Estudos Estratégicos (CGEE)
Instituto Eldorado
Laboratório Nacional de Computação Científica (LNCC)
Serviço de Processamento (SERPRO)
Universidade Estadual de Campinas (UNICAMP)

Elaboração do relatório final

Centro de Gestão e Estudos Estratégicos com o apoio da Unicamp, LNCC e Instituto Eldorado.

Sumário

Introdução	80
Pilha de IA Brasileira: Oportunidades e Barreiras	81
Painel 1 – Caminhos para Sustentar a IA com Autonomia de Infraestrutura Tecnológica Nacional	88
Painel 2 - Atuação do Brasil nas Camadas Superiores da Pilha de IA	98
Painel 3: Como o Brasil está posicionado no desenvolvimento de IA para dispositivos móveis e embarcados	104
Painel 4: Ciência aberta, open source e cooperações internacionais: limites e potenciais	111
Painel 5: Avanços do Plano Brasileiro de Inteligência Artificial (PBIA) – Construindo o novo Sinapad e o Sistema Nacional de Tecnologia e Inovação em IA (SinIA)	121
Painel 6: Capacitação e Formação Técnica para a Construção de uma Pilha de IA Soberana	130
Painel 7: Conexões que sustentam a IA: Onde estão as alianças/colaborações estratégicas nacionais para a construção de uma Pilha de IA soberana?	138
Painel 8: Como retroalimentar o Plano Brasileiro de IA, conectando ciência, indústria e governo no financiamento da soberania digital brasileira em IA?	153
Painel 9: Reflexão sobre Desafios, Lacunas e Oportunidades para a Pilha de IA Nacional.....	167
ANEXO 1. Propostas consolidadas a partir dos subsídios do Seminário Nacional da Pilha de IA	174
ANEXO 2. Programação do Seminário.....	177
ANEXO 3. Contribuições do Youtube	181
ANEXO 4. Contribuição via Strateegia.....	183
ANEXO 5. Respostas às perguntas levantadas por painel	187

Introdução

O Seminário “Pilha de IA Nacional: Desafios para Autonomia Tecnológica e Soberania Digital”, que ocorreu nos dias 1 e 2 de outubro de 2025, em Brasília, teve como objetivo debater os desafios, oportunidades e estratégias para o fortalecimento da autonomia tecnológica e soberania digital do Brasil no campo da Inteligência Artificial. O evento reuniu representantes do governo, academia, indústria e centros de pesquisa, promovendo discussões técnicas e institucionais sobre todos os componentes da pilha de IA, desde infraestrutura de *hardware* até aplicações avançadas, com foco em políticas públicas, capacitação, financiamento e cooperação nacional e internacional. O conteúdo do documento reflete as principais ideias, consensos e divergências identificadas ao longo dos painéis, propondo caminhos concretos para o desenvolvimento sustentável da Pilha de IA Nacional.

Ao longo do seminário, foram apresentados principais desafios técnicos, institucionais e estratégicos para a construção de uma pilha de IA soberana no Brasil. São abordados temas como infraestrutura computacional, domínio de componentes de *software*, capacitação de recursos humanos, modelos de financiamento, ciência aberta, cooperação internacional, e articulação entre academia, governo e setor produtivo. O relatório propõe recomendações e próximos passos para orientar políticas públicas, pesquisas e investimentos, visando consolidar a autonomia tecnológica nacional em IA.

Também foram apresentadas as experiências nacionais e internacionais que ilustram a relevância de se construir uma Pilha de IA Soberana, capaz de reduzir a dependência tecnológica e fomentar a produção de conhecimento e inovação local. As trocas entre os participantes apontaram caminhos concretos para o fortalecimento das infraestruturas abertas, o incentivo à formação de recursos humanos qualificados e a promoção de projetos cooperativos entre academia, governo e setor produtivo.

O documento do seminário apresenta as apresentações realizadas pelos palestrantes em seus respectivos painéis, com uma síntese ao final de cada painel com as principais propostas e recomendações, além de uma análise SWOT para cada painel. Também apresenta as contribuições advindas do *youtube*, da plataforma *strategie* e contribuições de especialistas participantes do evento.

Posto isso, este relatório reúne as principais ideias, reflexões e encaminhamentos discutidos durante o evento, oferecendo uma síntese das contribuições que emergiram das mesas temáticas e dos debates coletivos. Mais do que registrar as discussões realizadas, o documento busca servir como subsídio estratégico para a formulação de políticas e ações voltadas ao avanço de uma Inteligência Artificial soberana, inclusiva e sustentável no Brasil.

Pilha de IA Brasileira: Oportunidades e Barreiras

Professor Edson Borin (UNICAMP)

A pilha de inteligência artificial (IA) brasileira é o conjunto organizado de dados, *hardware* e tecnologias de *software* que operam de forma integrada para possibilitar o funcionamento de sistemas de IA. Ela envolve desde a infraestrutura física de processamento até as aplicações finais, incluindo sistemas operacionais, bibliotecas, *frameworks* e modelos de aprendizado de máquina. O objetivo é estabelecer uma base técnica comum para os debates sobre a construção de uma pilha de IA soberana no Brasil e levantar questões estratégicas sobre seu desenvolvimento e consolidação.

Investir em IA significa explorar o valor estratégico dos dados que o país já possui. Fontes como a Agência Nacional do Petróleo (ANP), o Sistema Único de Saúde (SUS), a Embrapa, o IBGE, o Portal Dados Gov.br, o INEP, o CNPq, a Receita Federal, o Banco Central (com as transações via PIX) e outras instituições produzem continuamente informações que descrevem o perfil do Brasil — quem somos, como vivemos e o que produzimos. Esses dados representam um ativo de valor inestimável, capaz de transformar o país se forem utilizados de forma estratégica. Sua exploração pode melhorar os serviços públicos em escala sem precedentes, fortalecer a economia e ampliar a competitividade internacional.

O Brasil possui um mercado interno robusto, ampla capacidade científica, instituições consolidadas de pesquisa e ensino, uma base industrial diversificada e um ambiente favorável à inovação e ao empreendedorismo. Para transformar o potencial dos dados em valor real, são necessários três componentes fundamentais: *softwares* de IA capazes de treinar e operar modelos complexos, infraestrutura computacional de alto desempenho e recursos humanos altamente qualificados para utilizar essas ferramentas e extrair conhecimento dos dados disponíveis.

Esses elementos, quando integrados, podem gerar avanços expressivos em múltiplos setores. Na saúde, a IA pode contribuir para diagnósticos mais precisos e para a otimização do SUS; na educação, pode permitir o ensino personalizado, adaptado às particularidades regionais; no meio ambiente, pode auxiliar na previsão de eventos climáticos e no monitoramento ambiental; e, na governança pública, pode promover maior transparência e ampliar a participação social.

A pilha de IA é composta por camadas interdependentes, nas quais cada componente desempenha uma função específica. A base é o *hardware*, sobre o qual operam os sistemas operacionais e *runtimes* que controlam o processamento e a execução de programas. Acima, encontram-se as bibliotecas e *frameworks*, como o PyTorch, que sustentam o desenvolvimento dos modelos de aprendizado de máquina, os quais, por sua vez, dão origem às aplicações práticas, como o ChatSUS. Essa estrutura modular

permite a substituição gradual de componentes sem comprometer o funcionamento do sistema, o que possibilita um processo de nacionalização progressivo da tecnologia (Figura 1).

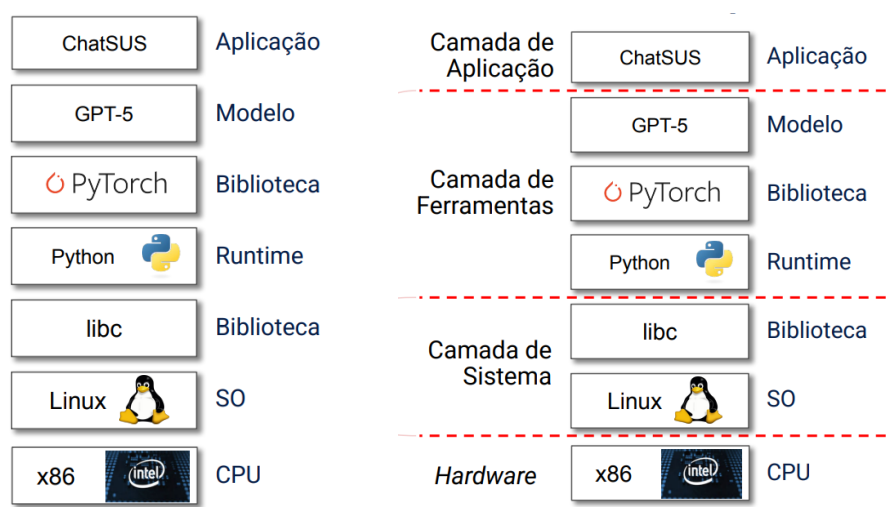


Figura 1. Exemplo de pilha de IA

Existem diferentes tipos de pilha, dependendo do contexto e da aplicação. A pilha de HPC (*High Performance Computing*) é voltada ao treinamento de grandes modelos e requer o uso de supercomputadores e sistemas de interconexão de alto desempenho. A Figura 2 apresenta um problema complexo, pois agora o *hardware* não é apenas um computador, é um sistema computacional com múltiplos computadores, sistemas de interconexão de alto desempenho. Acima tem uma camada de *software* de sistema para permitir que possa ser utilizado, com sistema operacional, contêineres de HPC, *software* de gerência de *job* e escalonamento do *cluster*, *software* para armazenamento de dados, *software* para desenvolvimento de aplicações, compiladores e bibliotecas. Numa camada posterior tem as camadas de ferramentas mais especializadas para domínio da IA, como *frameworks* de IA e *frameworks* de armazenamento e curadoria de dados e serviços de inferência e, por fim, a cada de aplicação, como o *chatSUS*.

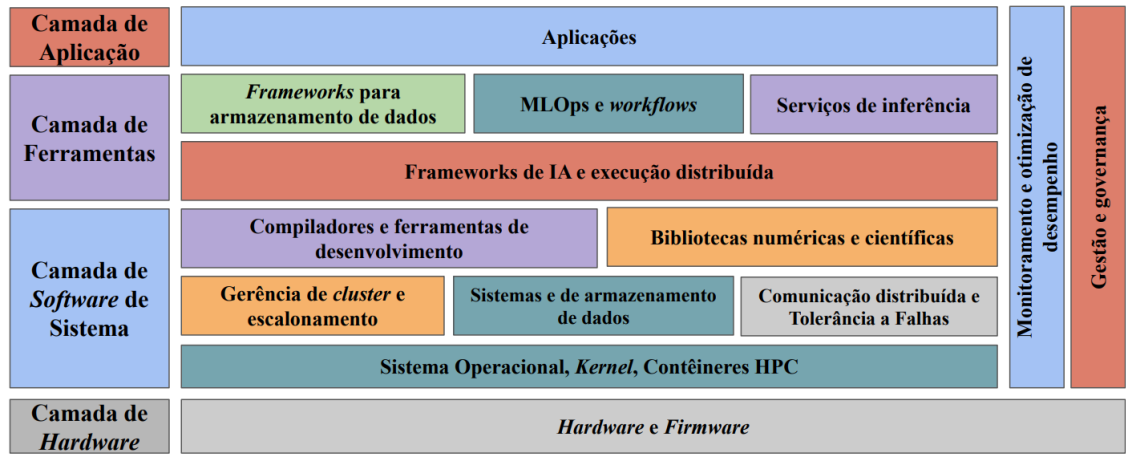


Figura 2. Pilha de HPC para IA

Nesse ambiente, o foco é maximizar a capacidade de processamento, ainda que o consumo de energia seja elevado. Já a pilha embarcada é compacta e otimizada para dispositivos móveis, priorizando a eficiência energética e o baixo consumo de recursos. Ambas são essenciais e complementares: enquanto a primeira garante o desenvolvimento e o treinamento de modelos, a segunda permite sua aplicação em larga escala.

Para trazer perspectiva para o problema, têm-se os diferentes cenários em que as pilhas vivem. Da esquerda para a direita, os tipos de sistemas computacionais onde se verá os sistemas de IA sendo executados. No começo o sistema de HPC, depois nuvem computacional, depois os pequenos *clusters* de HPC que estão distribuídos pelas universidades brasileiras, seguido de dispositivos de *desktops* e laptops, smartphones e dispositivos embarcados. O pré-treinamento de grandes modelos exige infraestrutura de HPC ou nuvem computacional, pois muitos desses sistemas de nuvem rodam em grandes *datacenters* que são praticamente sistemas de HPC. (Figura 3).

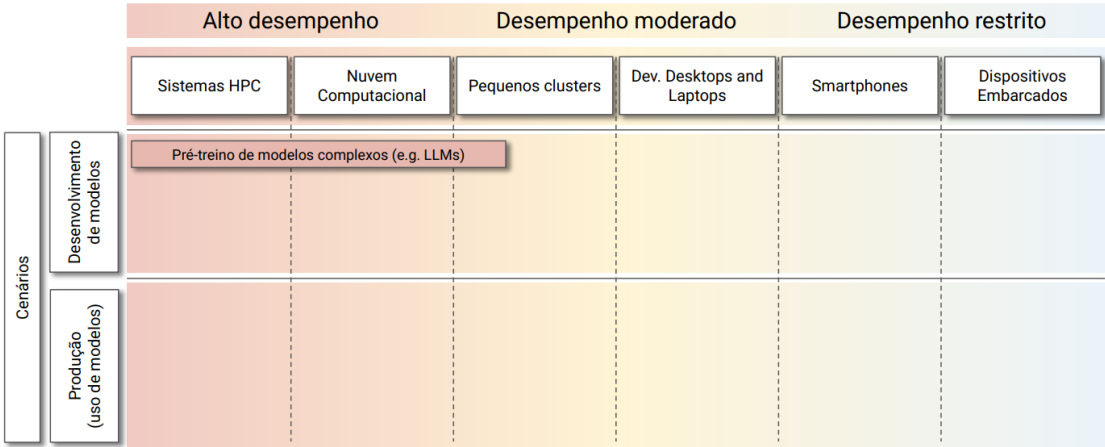


Figura 3. Desenvolvimento de modelos de IA em sistemas computacionais de alto desempenho

Por sua vez, o refino, processo de especialização de modelos para áreas específicas, como saúde ou direito, pode ser realizado em *clusters* menores. Esse cenário não é tão barato, mas pode ser feito em *clusters* menores ou nuvem computacional e não precisa de um super sistema de HPC para realizar o refino (Figura 4).

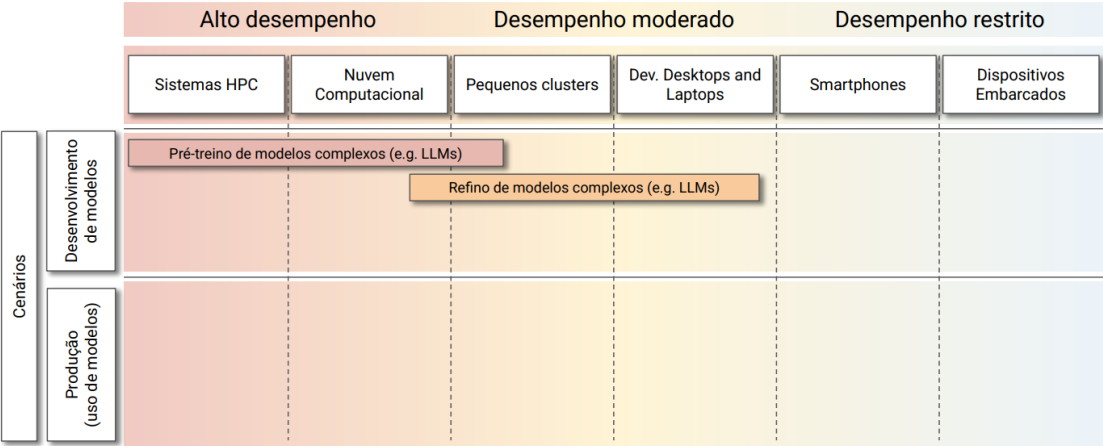


Figura 4. Desenvolvimento de modelos de IA em sistemas computacionais com desempenho moderado

Já modelos de média complexidade, baseados em técnicas clássicas de aprendizado de máquina, continuam relevantes para diversos problemas e podem ser treinados em computadores convencionais. Também chamado de modelos clássicos, como as árvores de decisão e suporte vector machines, que são modelos que existem há muito tempo e eles servem para resolver grande parte dos problemas que temos hoje em dia, ou seja, muito dos problemas existentes não precisa de um GPT para resolver. E é possível treinar esses modelos em *desktops* ou em *clusters* pequenos, sem necessidade de GPU (Figura 5).

84

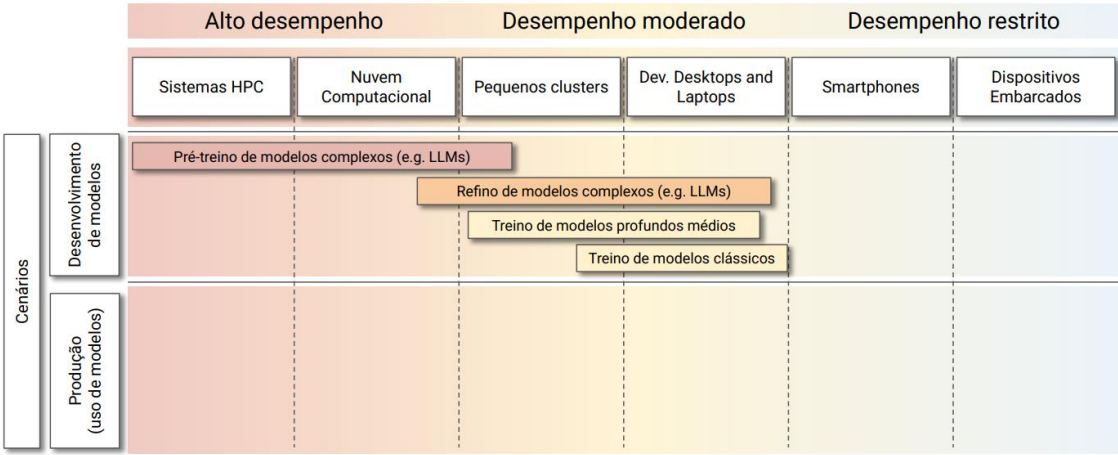


Figura 5. Desenvolvimento de modelos de IA mais simples (p.ex. modelos clássicos) em sistemas computacionais com desempenho moderado

No estágio final, de produção, os modelos são disponibilizados em nuvem ou em dispositivos móveis, dependendo da complexidade e do custo computacional (Figura 6). Aqui há dois ambientes contrastantes, do lado direito tem os cenários onde vai se executar esses modelos dentro de *smartphones* e dispositivos embarcados (modelos compactos com baixo uso de energia), e do lado esquerdo teremos uma gama grande de

modelos que vão rodar no que se chama de nuvem computacional. Porque quando se tem um modelo de *machine learning* que o custo dele não é trivial, ou seja, não é possível rodar no celular, utiliza-se um servidor em datacenter, por meio do que se chama de modelo computacional de nuvem para acessar. Exemplo: quando abre o ChatGPT no celular, não se está rodando o modelo ali, a mensagem é enviada pela internet até um servidor e é executada em um *data center* em uma GPU altamente potente, sendo o *backbone* de oferta de serviços de IA no mundo inteiro.

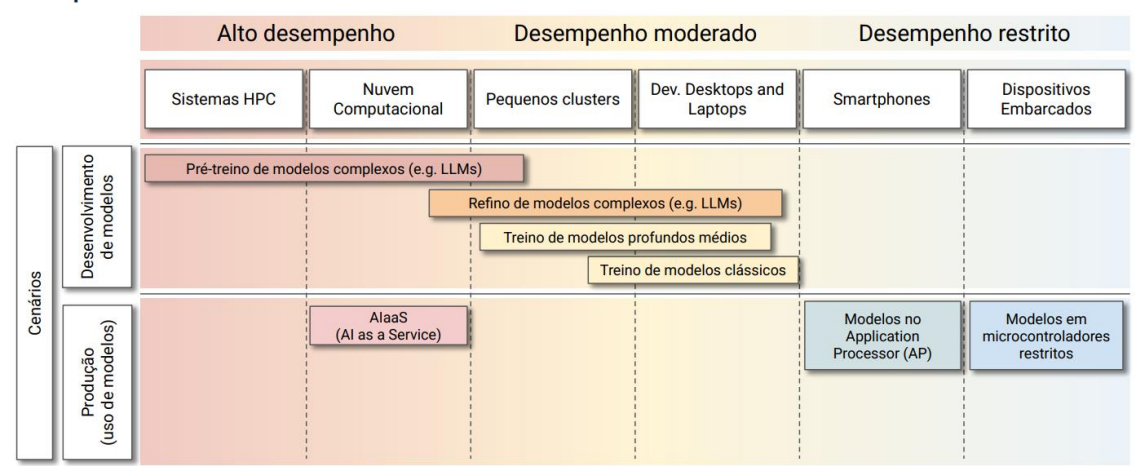


Figura 6. Cenário de produção, onde os modelos realizam inferência em smartphones, dispositivos embarcados e na nuvem computacional (AlaaS)

Não se pode deixar de fora também os problemas complexos de ciência e engenharia que não vão precisar ser servidos no celular, nem na nuvem. Aqui está se falando de previsão climática, detecção de cobertura vegetal do país, previsão de safra, ou seja, problemas específicos que serão executados dentro de sistemas de HPC por um conjunto restrito de pessoas e pesquisadores.

O objetivo estratégico da pilha de IA brasileira é extrair valor dos dados nacionais com segurança, reduzir a dependência tecnológica externa e fortalecer a soberania digital do país. Para isso, é fundamental investir em infraestrutura computacional nacional, no domínio dos componentes de *software*, no desenvolvimento de modelos de IA treinados com dados brasileiros, na formação e retenção de talentos especializados e na criação de modelos de financiamento sustentáveis.

Entre as principais barreiras estão a escassez de serviços competitivos de HPC para IA, a burocracia e os custos de importação de equipamentos, a desvalorização cambial e a dificuldade de acesso a essas infraestruturas por parte de startups e pesquisadores. Ainda assim, há oportunidades promissoras, como o avanço do Plano Brasileiro de Inteligência Artificial (PBIA), a criação do programa REDATA e a modernização da rede SINAPAD (Figura 7), que pode evoluir para uma nuvem nacional integrada de HPC, com compartilhamento seguro de dados e recursos.

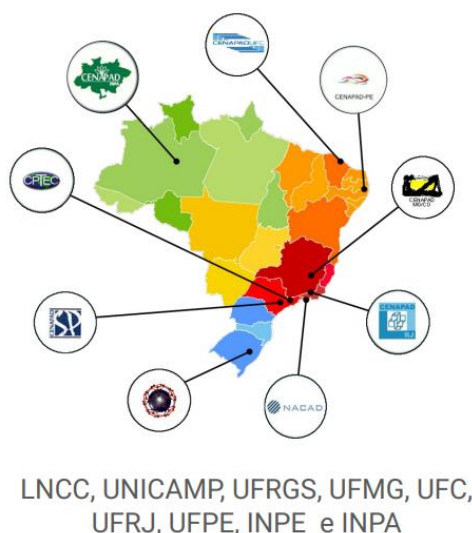


Figura 7. Distribuição de infraestrutura computacional no Brasil

No campo do *software*, os desafios decorrem da diversidade e da complexidade dos componentes da pilha, além da necessidade de profissionais altamente capacitados em sistemas e otimização. O foco inicial pode ser o treinamento de modelos e o desenvolvimento de camadas de abstração que permitam uma nacionalização gradual dos componentes. A capacidade científica das universidades e institutos de pesquisa brasileiros oferecem uma base sólida para esse avanço.

A formação e retenção de talentos é outro ponto crítico. O país precisa de profissionais com conhecimento profundo em IA e sistemas computacionais, não apenas de programas de letramento. Além disso, é fundamental criar mecanismos de incentivo ao empreendedorismo tecnológico e à inovação acadêmica, oferecendo condições para que jovens talentos possam experimentar, errar e desenvolver soluções de impacto.

Em termos de financiamento, o Brasil dispõe de mecanismos robustos, como FINEP, FNDCT, PPI, Embrapii e Lei de Informática. No entanto, é necessário aprimorar a coordenação entre governo, fundações de amparo à pesquisa, universidades e setor produtivo, garantindo um fluxo contínuo de investimento e apoio a *startups*, pequenas e médias empresas. É igualmente importante assegurar que esses instrumentos de fomento sejam complementares e sustentáveis a longo prazo.

O Brasil reúne condições únicas para se tornar referência em inteligência artificial soberana: possui vastos repositórios de dados, um mercado interno amplo e diversificado, uma comunidade científica consolidada e um ecossistema de inovação em expansão. O desafio, contudo, é transformar esse potencial em ação coordenada e duradoura. Para isso, é imprescindível expandir a infraestrutura computacional nacional, dominar o *software* e as tecnologias que compõem a pilha de IA, formar e reter profissionais qualificados e adotar modelos de financiamento estáveis e de longo prazo.

Com visão estratégica e continuidade institucional, o país tem condições de construir uma Pilha de IA Nacional que garanta autonomia tecnológica, promova inovação e fortaleça a soberania digital do Brasil.

Painel 1 – Caminhos para Sustentar a IA com Autonomia de Infraestrutura Tecnológica Nacional

Este painel explora as camadas mais profundas da pilha de IA, aquelas que lidam com a infraestrutura de *hardware*, como *clusters* de alto desempenho (HPC), GPUs, TPUs e *frameworks* de execução. A discussão abordará o quanto o Brasil precisa avançar para prover a base computacional necessária para o desenvolvimento e a operação de soluções avançadas de IA.

Moderação: Renata Mielli (*Coordenadora do Comitê Gestor da Internet no Brasil e assessora especial da Ministra da Ciência, Tecnologia e Inovação - MCTI*)

Marcelo Pinheiro

Diretor e fundador da Versatus HPC

Há uma questão de política internacional que orienta a estrutura de investimentos na pilha de *software* e compreender esse contexto é fundamental. Até a década de 1990, o mundo estava inserido em outra configuração geopolítica, e esse tipo de equipamento, o *hardware*, era extremamente proprietário. Havia grande diversidade de *hardwares*, com vários fabricantes, sobretudo nos Estados Unidos, que desenvolveram a base da atual era digital, e cada fabricante possuía seu próprio sistema operacional. Os processadores eram muito distintos entre si e os códigos eram criados para arquiteturas específicas.

Com o avanço da globalização, houve a exigência de transferência de propriedade intelectual para os países que recebiam as fábricas, como a China e outras nações asiáticas. Paralelamente, surgiu o movimento *open source*, no início da década de 1990, com destaque para o Linux, projeto de código aberto mais conhecido.

Atualmente, um supercomputador ou uma infraestrutura de alto desempenho é amplamente definida por *software*. E as tecnologias utilizadas são, em grande medida, derivadas das tecnologias de computadores pessoais (PCs). A propriedade intelectual presente nos supercomputadores foi desenvolvida inicialmente pela indústria de PCs, seguida pela indústria de games, com as GPUs, que se tornou um setor gigantesco. Atualmente, a indústria de dispositivos móveis detém grande parte da propriedade intelectual e financia a estrutura de pesquisa e desenvolvimento (P&D).

Atualmente, um equipamento desse tipo é composto por uma rede de múltiplos PCs (servidores) e tem suas origens no desenvolvimento dos consoles de videogame. O processador mais poderoso da NVIDIA, por exemplo, utiliza uma arquitetura ARM, a mesma presente em smartphones, em conjunto com GPUs, tudo isso financiado por uma indústria de larga escala. E fazendo novamente o resgate histórico, antes dos anos 1990, o P&D era majoritariamente financiado pelo Estado, hoje, esse papel é desempenhado pelas empresas.

Nos últimos dez anos, nota-se uma mudança significativa na dinâmica da globalização, com empresas como a NVIDIA produzindo *hardwares* padronizados e amplamente utilizados, mas com *softwares* proprietários. Por outro lado, empresas como o Google, que oferecem *softwares* abertos, mas operam com *hardwares* exclusivos, como suas TPUs, que rodam apenas em sua infraestrutura de nuvem.

No que se refere ao *software*, há uma grande oportunidade para o Brasil: o uso de tecnologias *open source*. Por serem de código aberto, elas podem ser apropriadas e desenvolvidas localmente, desde que haja investimento em grupos de pesquisa, empresas e universidades capazes de compreender e aprimorar essas tecnologias. Porém, o simples acesso ao código não é suficiente, é preciso domínio técnico e estratégico. E, embora o *software* seja, em certa medida, mais acessível que o *hardware*, também apresenta desafios significativos. Considerando a pilha apresentada pelo professor Borin, observa-se que o sistema de armazenamento, equivalente, em um PC, ao disco rígido, é um ponto central para a soberania tecnológica. Esses sistemas, conhecidos como sistemas de arquivos paralelos, são extremamente complexos, e alguns dos melhores projetos existentes são de código aberto. Por isso, é essencial que o país desenvolva competência e domínio sobre essas tecnologias.

Luca Belli

Professor de Direito - FGV Rio de Janeiro

89

O conceito de pilha (*stack*) ainda não é amplamente difundido para identificar as camadas de um sistema. A Índia, com seu projeto de transformação digital, adotou o termo em sua iniciativa denominada *Índia Stack*, e, atualmente, a Europa segue caminho semelhante, desenvolvendo o *EuroStack*.

O conceito de *stack* refere-se a sistemas digitais compostos por múltiplos elementos e dimensões. Além dos aspectos tecnológicos, é fundamental considerar dimensões como cibersegurança, energia e gestão de riscos, que são essenciais para a robustez e sustentabilidade desses sistemas. Também é importante destacar a conectividade, frequentemente negligenciada, mas que representa uma vulnerabilidade sistêmica significativa.

Atualmente, o Brasil depende fortemente da importação de GPUs, TPUs e outros componentes de infraestrutura de ponta, produzidos por um número restrito de empresas, em sua maioria sediadas nos Estados Unidos. E é pouco realista que o país consiga, no curto prazo, fabricar GPUs capazes de competir com as da NVIDIA. No entanto, é possível avançar em outras frentes, como o desenvolvimento de semicondutores mais básicos e o aproveitamento de tecnologias *open source*, especialmente aquelas baseadas em arquiteturas abertas, como o RISC-V.

Esse avanço pode ocorrer não necessariamente pela produção física, mas pelo design, adotando o modelo *fabless*, o mesmo utilizado pela NVIDIA, que projeta seus chips, enquanto a fabricação é realizada pela TSMC, em Taiwan, responsável por aproximadamente 90% da produção mundial de semicondutores de alto desempenho.

O Brasil poderia investir no desenvolvimento de uma indústria de *design* de *chips*, especialmente de componentes mais simples, e também em áreas como *packaging* e empilhamento (encapsulamento) de *chips*, com foco no curto prazo. Em médio e longo prazo, o país poderia buscar ascender na cadeia de valor, o que exige estabilidade institucional, política industrial consistente e formação de recursos humanos qualificados. O PBIA (Plano Brasileiro de Inteligência Artificial) propõe ações de capacitação, mas é necessário ir além: é fundamental desenvolver *software* nacional.

O aproveitamento pleno de GPUs da NVIDIA, por exemplo, depende de *softwares* proprietários, como o CUDA e o AI Enterprise, ambos desenvolvidos pela empresa. Assim, mesmo após a compra do *hardware*, é necessário adquirir licenças contínuas para utilização de suas funcionalidades mais avançadas, o que reforça a dependência tecnológica. Outro ponto essencial é conectar as diversas dimensões da pilha. Já se mencionou a importância de federar sistemas e até mesmo de aproveitar estruturas de federação já existentes, que ainda não são plenamente exploradas. Também se destacou o papel do *humanware*, ou seja, das pessoas. Por fim, é importante discutir a questão dos dados.

Atualmente, a população brasileira, incluindo estudantes e servidores da administração pública, tem contribuído, de forma direta ou indireta, para o treinamento de modelos de empresas estrangeiras. Enquanto prevalecerem políticas como o *zero rating* aplicado à Meta AI, que permite o uso gratuito de aplicativos patrocinados, essa situação persistirá. Hoje, cerca de 78% dos brasileiros têm acesso gratuito e contínuo aos serviços da Meta, o que representa uma vantagem competitiva significativa para a empresa.

A *OpenAI* percebeu que, sem um modelo semelhante de aplicativos patrocinados, não conseguiria competir com a Meta AI. Outras empresas, como a *Perplexity*, também adotaram essa estratégia. Entretanto, permitir esse tipo de prática, que é, inclusive, ilegal segundo o artigo 9º do Marco Civil da Internet, representa um grave equívoco. Permitir que empresas como o *Google* ofereçam, por exemplo, o *Gemini* gratuitamente para estudantes, apenas para treinar seus modelos, significa ceder dados estratégicos de forma gratuita.

Desta forma, não adianta desenvolver uma pilha nacional de IA se, ao mesmo tempo, os dados brasileiros continuam sendo entregues a grandes corporações estrangeiras. Esse tipo de exploração de dados deve ser tributado e regulado. Desde 2019, a China reconhece os dados como um fator de produção, orientando suas políticas industriais e produtivas com base nesse princípio. E a Índia, por sua vez, adotou a neutralidade de rede, proibindo práticas que ainda persistem no Brasil.

Leandro Rosa dos Santos

Vice-presidente e Diretor de estratégia e inovação - Positivo Tecnologia

Para ser eficaz, a soberania digital deve proporcionar três resultados essenciais e interdependentes:

- **Soberania de Dados:** consiste na conformidade com todas as leis e regulamentações aplicáveis à privacidade, segurança e residência dos dados.
- **Soberania Operacional:** refere-se à garantia de que a infraestrutura crítica associada a aplicações intensivas em dados permaneça sempre ativa, acessível e sob controle nacional.
- **Soberania Digital:** diz respeito ao nível de controle que uma organização ou Estado exerce sobre seus ativos digitais — incluindo dados, *softwares*, conteúdos e infraestrutura tecnológica.

Desafios para a Construção de uma Nuvem Soberana de Dados e Inteligência Artificial

- **Soberania de Dados:** As arquiteturas de nuvem pública oferecidas por empresas multinacionais estão sujeitas às legislações de seus países de origem, como o *Foreign Intelligence Surveillance Act* (FISA) e o *Cloud Act*, que permitem ao governo dos Estados Unidos acessar dados armazenados por empresas sediadas em território norte-americano, independentemente da localização física dos dados. Desafios semelhantes são observados em parcerias com empresas chinesas e de outros países, especialmente em contextos relacionados à segurança nacional.
- **Soberania Operacional:** Exige a existência de uma infraestrutura robusta, atualizada, redundante, escalável e resiliente. No entanto, tecnologias fornecidas por empresas estrangeiras têm se tornado objeto de restrições comerciais e políticas, limitando o acesso de determinados países a componentes estratégicos.
- **Soberania Digital:** Implica a necessidade de controle efetivo — tanto operacional quanto tecnológico — sobre dados, *softwares*, conteúdos digitais e infraestrutura crítica, de modo a reduzir dependências externas e garantir autonomia tecnológica.

Serviços	 Naver	 SIAM.AI	 AI Factories		 Brasil IA
Sovereignty AI Applications	Consultant, Development (Energy, Health, Etc), Deployment	Consultant, Development (Health, SmatCity, Etc), Deployment (WIPRO)	Consultant, Development (Energy, Health, Etc), Deployment	Use Cases in Health, Transport, Energy and Administration	Ecosistema NACIONAL de Empresas
Sovereignty Data	Consultant, Storage, Building, Management	Consultant, Storage, Building, Management	Consultant, Storage, Building, Management	Consultant, Storage, Building, Management	Consultant, Storage, Building, Management
Sovereignty AI Models	Open Source LLM, Enterprise LLM (HyperCLOV X) Tunning, Evaluation, LLM and AI Development	Customization of LLMs for local language, Enterprise Models, (SiamGPT) Customization	Development and Inference	Open Source LLM, Enterprise LLM (Mistral), Tunning, Evaluation, LLM and AI Development	Open Source LLM, Enterprise LLM, Tunning, Evaluation, Development
Sovereignty Data and AI Cloud Infrastructure	Consultant, Optimization, Data and LLMOps, Services	Consultant, Optimization, Data and LLMOps, Services	Consultant, Optimization, Data and LLMOps, Services	Consultant, Optimization, Data and LLMOps, Services	Consultant, Optimization, Data and LLMOps, Services
Sovereignty Datacenter	Bare Metal, Establishment, Operation Management	Establishment, Operation Management	15 AI Factories all over EUROPE (100,000 advanced AI processors)	Bare Metal, Establishment, Operation Management	Bare Metal, Establishment, Operation Management

Figura 8. Exemplos de países com iniciativas de nuvem e IA soberana

Diversos países têm implementado políticas e programas voltados à soberania digital, estruturando ecossistemas nacionais de nuvem e inteligência artificial (Figura 8). Essas iniciativas compartilham um princípio comum: o foco na contratação pública estratégica, orientada para um ecossistema de parceiros comprometidos com o interesse do cidadão.

Nesses modelos, o debate não se limita à tecnologia, mas à transformação concreta dos serviços públicos e à melhoria da qualidade de vida da população — em áreas como saúde, segurança, educação e administração pública. A partir da definição dos desafios prioritários, são desenvolvidas soluções que envolvem o uso de supercomputadores, dados massivos e sistemas de inferência voltados ao benefício coletivo e à soberania nacional.

A cooperação é um elemento central desses projetos, e o *open source* surge como um instrumento estratégico para promover a colaboração entre governos, academia, empresas e sociedade civil.

Elementos Comuns aos Projetos de Nuvem Soberana

- **Soberania de Dados:** Estruturação e armazenamento de dados de acordo com seu nível de criticidade, priorizando soluções *on-premises* e provedores nacionais de nuvem, com o uso restrito de nuvens estrangeiras apenas para dados não sensíveis e não críticos.
- **Soberania Operacional:** Investimentos em redundância e resiliência para dados e processamento em toda a cadeia — desde infraestruturas locais até nuvens nacionais e internacionais — garantindo continuidade operacional.

- **Soberania Digital:** Formação de um ecossistema nacional de empresas voltado à gestão da infraestrutura, supervisão dos serviços em nuvem, desenvolvimento e manutenção de sistemas de inteligência artificial, e à criação de competências técnicas em diversas especialidades relacionadas.

Além das camadas da pilha tecnológica apresentadas anteriormente, o funcionamento de um *datacenter* requer diversos elementos complementares, muitos deles baseados em soluções *open source*, como apresentado em Figura 9:

- sistemas de supervisão e orquestração,
- mecanismos de convergência de perfis e serviços de nuvem,
- camadas de gerenciamento e interoperabilidade.

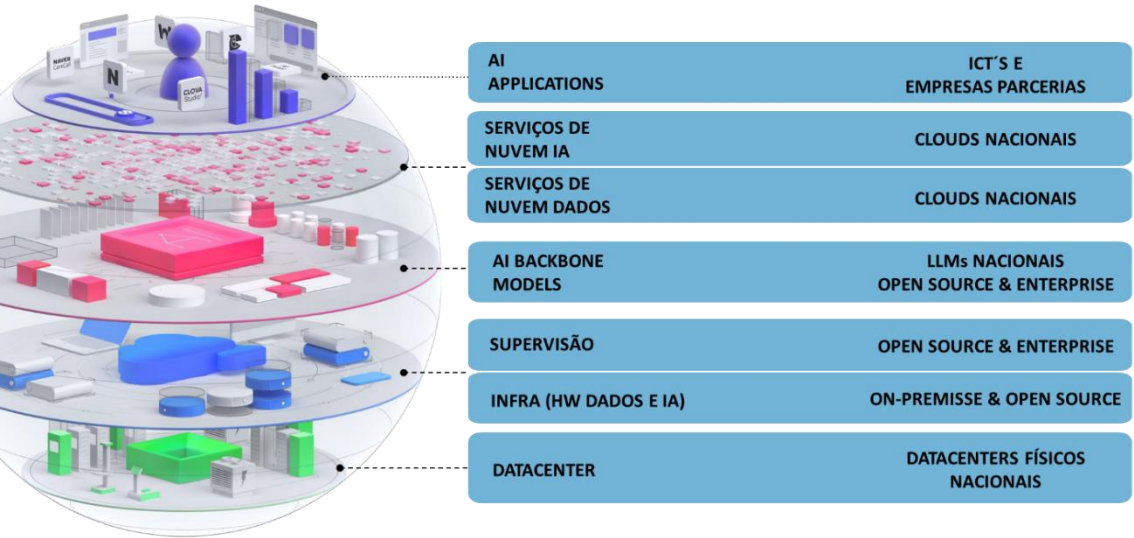


Figura 9. Camadas da soberania digital

A partir desses elementos, é possível desenhar arquiteturas de nuvem soberana adaptadas às necessidades nacionais. O Brasil, em particular, já dispõe de competências técnicas e institucionais relevantes, que podem ser mobilizadas para o desenvolvimento de soluções próprias e sustentáveis no campo da soberania digital.

Guido Araújo

Professor do Instituto de Computação - Unicamp

É fundamental revisitar o passado para compreender a atuação e resultados do Brasil em alguns setores e, de modo a inspirar a atuação no desenvolvimento da pilha de IA e identificar em quais áreas o país pode alcançar maior grau de soberania. Um exemplo emblemático se refere a Embraer e a indústria aeronáutica nacional, fruto de um esforço contínuo e de longo prazo. Ainda assim, o país não é plenamente soberano nesse setor,

uma vez que componentes estratégicos, como motores e sistemas eletrônicos, são fabricados no exterior. Esse caso demonstra que é possível gerar valor e desenvolver uma indústria robusta sem produzir internamente todos os elementos da cadeia produtiva.

Outro exemplo é a Petrobras, cuja tecnologia empregada na exploração do pré-sal foi desenvolvida em colaboração com parceiros internacionais. Isso evidencia que a construção de uma indústria nacional sólida pode ocorrer por meio de alianças estratégicas de longo prazo, com parceiros tecnológicos consolidados, capazes de contribuir para a agregação de valor e o fortalecimento do ecossistema local. Em síntese, o Brasil conseguiu desenvolver setores industriais relevantes, mesmo sem possuir soberania completa em todos os elos da cadeia.

Com isso, o Brasil poderia novamente realizar colaborações internacionais e desenvolver gradualmente as camadas da pilha, baseado em um planejamento de longo prazo, assim como fez a China. Em 2015 o gigante asiático reconheceu a importância estratégica do *hardware* como base da soberania digital. Por meio da *Chinese Academy of Sciences*, o país uniu-se à *RISC-V International*, uma iniciativa voltada ao desenvolvimento de arquiteturas de *hardware* aberto. Essa decisão refletiu uma visão de longo prazo e a compreensão de que o movimento *open source* seria determinante para o futuro tecnológico, percepção que a China já tinha desde aproximadamente 2010.

O resultado dessa estratégia é visível hoje: o país desenvolveu suas próprias GPUs e componentes de alto desempenho. Embora não sejam as mais eficientes do mundo, a China não busca competir diretamente na corrida da inteligência artificial generativa, como fazem empresas ocidentais voltadas a modelos como o ChatGPT, mas sim aplicar a IA em setores industriais estratégicos, por meio da aplicação da IA na robótica, veículos elétricos e automação. Em outras palavras, a China incorporou a IA em áreas em que possuía competitividade, utilizando-a como vetor de inovação produtiva.

A Europa, por sua vez, iniciou esse debate relativamente tarde, em 2015, quando iniciaram as discussões sobre o *hardware stack* e os riscos associados à dependência tecnológica. Foram necessários seis anos para que o Commissariado Europeu estruturasse o *European Chips Act*, uma política voltada ao fortalecimento da indústria de semicondutores no continente, com iniciativas relevantes, incluindo o surgimento de empresas dedicadas ao desenvolvimento de chips abertos, como a *OpenChip*.

As duas experiências apresentadas evidenciam que para alcançar a soberania tecnológica se faz necessário o compromisso estratégico do Estado, investimentos de longo prazo, muitas vezes ao longo de décadas, e obtenção de resultados graduais e intermediários. Este processo requer resiliência institucional, planejamento contínuo e clareza de propósito quanto aos objetivos nacionais. Ademais, a consolidação de um ecossistema soberano depende da indução e fortalecimento das capacidades locais, estimulando o avanço da indústria e da pesquisa aplicada. Assim, é essencial que o

Estado atue como articulador de políticas públicas e financiador de inovação, mantendo o foco no desenvolvimento tecnológico autônomo.

Paralelamente, a atuação das universidades é fundamental para o alcance da soberania, pois constituem a base intelectual e científica nacional, são responsáveis pela formação de capital humano qualificado e pela geração de conhecimento. No entanto, a manutenção efetiva da soberania é realizada por empresas e laboratórios estatais, que transformam o conhecimento em capacidade tecnológica e produtiva.

Em suma, para o professor Guido, o caminho para a soberania tecnológica pode ser resumido em três atos paralelos:

- *Software Stack*: o desenvolvimento de camadas de *software* é essencial, e o modelo *open source* representa um facilitador estratégico para a autonomia tecnológica.
- *Servidores*: é possível alcançar independência tecnológica no desenvolvimento e produção de servidores nacionais, reduzindo a dependência de fornecedores estrangeiros.
- *Aceleradores*: a ênfase deve ser colocada na inferência, em vez do treinamento de larga escala, explorando oportunidades de inovação em aplicações de IA mais específicas e otimizadas.

Carlos Rodrigo

Coordenador do Centro de Excelência de Ciências de Dados e Inteligência Artificial do Serpro

O SERPRO tem realizado investimentos significativos em GPUs com o objetivo de reduzir o tempo de treinamento de modelos de inteligência artificial. Essa estratégia tem levado à ampliação e modernização contínua do parque tecnológico, de modo a atender à crescente demanda por processamento e armazenamento de dados. Atualmente, a principal preocupação da instituição é diversificar fornecedores e fortalecer a soberania tecnológica. Nesse sentido, o SERPRO vem buscando alternativas às GPUs da NVIDIA, incorporando TPUs da *Google* e avaliando soluções provenientes da China, que têm apresentado desempenho comparável às tecnologias já em uso.

O objetivo de longo prazo é ampliar a oferta de *hardware* nacional, o que requer uma articulação sólida entre Instituições de Ciência e Tecnologia (ICTs), universidades, *startups* e governo. Essa cooperação é essencial para a criação de um ecossistema inovador e sustentável, capaz de reduzir a dependência de tecnologias estrangeiras e promover o desenvolvimento local.

Além disso, é importante destacar que nem todas as demandas computacionais exigem alta capacidade de processamento. Em muitos casos, o uso eficiente dos recursos

disponíveis depende diretamente da qualificação dos profissionais envolvidos, em que equipes bem capacitadas são capazes de otimizar o uso da infraestrutura existente, reduzindo custos e aumentando a eficiência operacional.

Por fim, enfatiza-se a necessidade de articulação e cooperação entre todos os atores do ecossistema tecnológico nacional. Como ressaltado, não é viável, técnica, econômica ou institucionalmente, desenvolver soluções de forma isolada. Somente por meio de uma ação coordenada entre os diversos entes públicos e privados será possível alcançar autonomia tecnológica e soberania digital de forma sustentável.

Considerações finais

O painel 1 reuniu especialistas de diferentes setores para debater os caminhos necessários para o Brasil alcançar autonomia em infraestrutura tecnológica voltada à inteligência artificial. O foco principal esteve nas camadas mais profundas da pilha de IA, abrangendo *hardware*, sistemas operacionais e as primeiras camadas de *software*, elementos essenciais para definir supercomputadores e infraestruturas de alto desempenho. Desta forma, será possível ter soberania de dados, operacional e digital.

Dentre as principais propostas deste painel, destacam-se:

- Políticas públicas e governança, criar uma política nacional de soberania digital e infraestrutura de IA, integrando iniciativas de *hardware*, *software*, dados e formação; estabelecer mecanismos de governança federada entre governo, universidades e empresas para definir prioridades tecnológicas e de investimento; e fortalecer o marco legal de dados e cibersegurança, incluindo regulação sobre uso e tributação de dados por empresas estrangeiras.
- Desenvolvimento tecnológico, investir em *hardware* nacional, priorizando áreas de design de chips, *packaging* e servidores com modelo *fabless*; promover o uso estratégico de tecnologias *open source*, incentivando o domínio e a contribuição brasileira em projetos globais; e estimular a produção de sistemas de arquivos paralelos e *frameworks* de execução de IA como áreas críticas de soberania.
- Formação e capacitação, expandir programas de formação em computação de alto desempenho, engenharia de chips e IA aplicada; incentivar a criação de laboratórios públicos-privados focados em P&D de *hardware* e *software* soberanos; e promover políticas de retenção de talentos e atração de pesquisadores em tecnologias críticas.
- Cooperação e ecossistema, estimular parcerias entre ICTs, startups e empresas nacionais, com apoio financeiro e institucional do Estado; criar centros nacionais de computação de alto desempenho (HPC) interconectados, com governança compartilhada; e incentivar a interoperabilidade e federação de sistemas públicos, reduzindo a fragmentação política.

O painel reforçou que a busca pela autonomia tecnológica exige esforços coordenados, investimentos estratégicos e políticas públicas voltadas ao fortalecimento da infraestrutura nacional, com vistas à soberania digital e ao protagonismo brasileiro no cenário internacional de inteligência artificial

Matriz SWOT – Painel 1

Forças	Fraquezas
• Capacidade científica instalada em universidades e ICTs. • Experiência de empresas públicas como o SERPRO na modernização e uso de IA. • Existência de programas nacionais (PBIA, MCTI) que podem articular políticas setoriais. • Crescente maturidade em tecnologias <i>open source</i> e potencial de cooperação internacional. • Possibilidade de financiamento por meio de bancos públicos (BNDES).	• Forte dependência de <i>hardware</i> e <i>software</i> estrangeiros (NVIDIA, Google, AWS). • Baixo investimento continuado em P&D de semicondutores e infraestrutura nacional. • Fragmentação entre governo, academia e setor privado. • Escassez de mão de obra especializada em design de chips, HPC e cibersegurança.
Oportunidades	Ameaças
• Aproveitar o movimento global por autonomia digital (ex.: <i>Chips Act</i>, <i>India Stack</i>, <i>EuroStack</i>). • Participar de ecossistemas open source internacionais (ex.: <i>RISC-V</i>). • Incentivar <i>startups</i> nacionais em <i>hardware</i> e IA aplicada. • Desenvolver infraestrutura nacional de nuvem soberana e HPC público.	• Concentração global da indústria de semicondutores em poucos países (EUA, Taiwan, China). • Restrições geopolíticas e sanções tecnológicas. Captura de dados brasileiros por big techs sob políticas de zero rating. • Falta de continuidade política e orçamentária em projetos de longo prazo.

Figura 10. Matriz SWOT - Painel 1

Painel 2 - Atuação do Brasil nas Camadas Superiores da Pilha de IA

Este painel discute como o Brasil pode atuar com protagonismo nas camadas superiores da Pilha de IA, como bibliotecas, *frameworks*, modelos pré-treinados, plataformas de desenvolvimento, MLOps e ferramentas de aplicação. O objetivo é refletir os riscos da dependência tecnológica e explorar caminhos para a soberania digital na construção de uma base sólida para o futuro da IA no Brasil.

Moderador: Caetano Penna (*Diretor no Centro de Gestão e Estudos Estratégicos - CGEE*)

Cleilton Rocha

Líder em ciência de dados e inteligência artificial - Instituto Atlântico

Refletir sobre o desafio de posicionar o Brasil nas camadas superiores da pilha de IA evidencia que, para desenvolver soluções de ponta, é fundamental cuidar de três elementos centrais: código, dados e modelos inteligentes. Há uma necessidade clara de reproduzir e entregar valor ao usuário final, que se encontra na extremidade desse ecossistema.

Outro desafio relevante nas camadas superiores é a construção de serviços. Atualmente, o país apresenta uma baixa cultura de sistemas *open source* e possui poucas *startups* de alta densidade tecnológica, o que impacta a interoperabilidade e gera alta granularidade no desenvolvimento de soluções de IA voltadas às áreas temáticas que utilizam tecnologia nacional. Entretanto, o Brasil possui grande potencial para superar esse cenário, uma vez que exporta mão de obra altamente qualificada para países desenvolvidos, sendo amplamente reconhecido pela excelência na entrega de serviços de *software* para o mercado internacional.

Ao considerar os sistemas inteligentes e a *stack* de MLOps, é imprescindível garantir que os modelos sejam auditáveis e observáveis, bem como assegurar a devida governança dos dados, permitindo a auditoria de resultados e a mitigação de problemas como *data drift* e *model drift*. Durante a pandemia, por exemplo, muitos modelos inteligentes não apresentaram bom desempenho justamente pela ausência de ferramentas robustas de observabilidade em produção.

Para gerar valor aos setores estratégicos da economia e facilitar o desenvolvimento de soluções na ponta, é essencial que o governo atue como catalisador e disseminador de uma cultura baseada em sistemas *open source*, além de ser indutor e fomentador de *startups* de alta densidade tecnológica. Essa atuação permitiria a tropicalização de serviços e o fortalecimento do ecossistema nacional de inovação. Um exemplo recente foi o investimento de uma grande empresa global, a *Data Bricks*, em uma startup brasileira, fato que demonstra o potencial de atração de investimentos internacionais quando há serviços locais de qualidade e colaboração efetiva.

Além disso, o país pode avançar com o desenvolvimento de uma plataforma compartilhada, composta por repositórios públicos com *templates* de *stacks*, capazes de atender múltiplos usuários e diversas demandas nacionais. Essa plataforma deveria ser acompanhada da criação de padrões comuns de integração e colaboração, fortalecendo a articulação entre governo, ICTs, universidades e setor privado.

Ao promover um ecossistema mais colaborativo e investir na automação de MLOps e na adoção de AutoML, torna-se possível democratizar o acesso à IA, permitindo que pesquisadores ainda não especializados na área desenvolvam suas pesquisas e alcancem resultados com maior eficiência. Essa abordagem também favorece a reprodutibilidade científica, possibilitando, por exemplo, que uma pesquisa desenvolvida na UFRN seja facilmente reproduzida na UNICAMP.

Por fim, ao disponibilizar uma plataforma como serviço (PaaS) para entrega de soluções de IA, o Brasil poderá oferecer ferramentas de criação de sistemas inteligentes com interfaces *low code*, ampliando a democratização da IA e fortalecendo a capacidade inovadora de diversos setores da sociedade.

Igor Régis

Especialista em tecnologia da informação - Banco do Brasil

O Banco do Brasil vem desenvolvendo uma camada tecnológica que viabiliza o acesso a uma *stack* de IA por meio da criação de uma plataforma habilitadora de tecnologia. Essa iniciativa está alinhada à estratégia da instituição de fortalecer a integração entre tecnologia, negócios e experiência do cliente.

A estrutura organizacional foi desenhada em camadas distintas, sendo a primeira voltada à entrega de objetivos de negócio e à geração de valor para o consumidor que utiliza os serviços bancários. Outra camada é composta pelas linhas de plataforma, formadas por grupos de equipes responsáveis por criar abstrações tecnológicas com o uso de *frameworks*, bibliotecas *open source* e soluções contratadas ou adquiridas pela instituição. Todo esse ecossistema é orquestrado para facilitar o trabalho dos cientistas de dados e dos usuários finais, permitindo que concentrem esforços na criação de soluções, sem a necessidade de lidar com complexidades técnicas intermediárias.

Durante a apresentação, foi destacada a Plataforma A2B2 do Banco do Brasil, que tem como foco principal a oferta de capacidades de IA e *analytics* em modelo de autosserviço para o desenvolvimento de soluções e sistemas de negócio. Essa plataforma está estruturada em plataforma e tecnologia, provendo os recursos necessários para o desenvolvimento e integração de soluções baseadas em IA e análise de dados; experiência e processos, buscando oferecer um ambiente intuitivo, completo e eficiente para os desenvolvedores, promovendo uma experiência aprimorada no uso da plataforma; capacitação e comunidade, com o objetivo de capacitar os colaboradores

para o uso efetivo da Plataforma A2B2, fomentando uma comunidade colaborativa voltada à evolução contínua e à inovação; e inovação e P&D, para a exploração de novas tecnologias, à geração de conhecimento avançado e à incorporação de inovações à plataforma, fortalecendo o ecossistema tecnológico do Banco do Brasil.

Karine Ferreira

Pesquisadora e Docente – Instituto Nacional de Pesquisas Espaciais (INPE)

Foi apresentada a forma como o Instituto Nacional de Pesquisas Espaciais (INPE) está estruturando e avançando na consolidação de sua pilha de IA. Já existem, inclusive, pilhas de IA implementadas em algumas instituições federais, como o próprio INPE, aplicadas em áreas estratégicas para o Governo Federal. No caso desta instituição, destacou-se a pilha de IA aplicada ao monitoramento ambiental e à gestão sustentável de recursos naturais. O INPE detém um amplo conjunto de dados de observação da Terra, composto por imagens e informações meteorológicas, e está desenvolvendo um ambiente de ciência de dados que permitirá aos usuários utilizarem esses dados de forma interativa e integrada.

A instituição armazena grande volume de dados produzidos por satélites de diversos países, além dos satélites brasileiros, totalizando aproximadamente 2 petabytes de informações referentes a todo o território nacional. Para processar esse volume de dados, o INPE conta com um centro de dados em constante aprimoramento, equipado com um supercomputador de 1,6 petaflops e uma infraestrutura física de 1.000 m² de área útil, além de sistemas de suporte, como climatização, energia e segurança, continuamente atualizados.

Dentre as principais iniciativas em andamento, destaca-se o BDC-Lab (Ambiente Computacional de Ciência de Dados Geoespaciais), uma plataforma de ciência de dados voltada ao processamento e análise de big data utilizando métodos de IA, diretamente na infraestrutura computacional do INPE, sem necessidade de download dos dados. Dentre os usuários desse ambiente estão instituições como a Embrapa e o IBGE, que conseguem realizar o processamento diretamente nos servidores do INPE.

A palestrante ressaltou que o INPE detém soberania nacional em observação da Terra, ao consolidar um ecossistema de big data próprio, com centro de processamento de dados e uma pilha de ciência de dados capaz de tratar e analisar as informações coletadas de forma independente. Embora o instituto mantenha autonomia tecnológica e não dependa de plataformas estrangeiras, ele também mantém interoperabilidade e diálogo com outras plataformas internacionais, promovendo cooperação científica e tecnológica sem comprometer a soberania nacional.

Norberto Alves Ferreira

Gerente de Solução IoT e IA (CPQD)

O CPQD tem se dedicado ao desenvolvimento e à estruturação de uma pilha de IA, que pode ser observada na proposta abaixo (Figura 11), sobre como compreende esse conceito e em quais camadas de *software* o Brasil deve concentrar seus investimentos, considerando uma arquitetura composta por seis camadas.



Figura 11. Pilha de IA – Proposta CPQD

O CPQD tem atuado de forma limitada no desenvolvimento de *frameworks* básicos, uma vez que entende que há poucas oportunidades de inovação nessa camada. As ações mais relevantes, nesse caso, seriam de acompanhamento tecnológico e de desenvolvimento pontual. As melhores perspectivas de avanço estão nas camadas superiores da pilha, especialmente nas bibliotecas intermediárias, onde é possível incorporar elementos da cultura e da diversidade brasileiras, gerando diferenciações competitivas em áreas e nichos específicos.

Na camada de modelos pré-treinados, o foco tem sido o uso e a adaptação de modelos já existentes, especialmente os modelos médios (SLMs - Small Language Models), que apresentam potencial para soluções nacionais ao incorporarem o idioma português e suas especificidades linguísticas. Prevê-se, em breve, a integração de recursos multimodais a esses modelos, o que poderá aumentar a eficiência e a aplicabilidade da IA em diversos setores.

Em relação às ferramentas de MLOps e Governança, o CPQD destaca o avanço nos *frameworks* de agentes e orquestração, com trabalhos intensivos na criação de agentes setoriais. A instituição acredita que essa será uma frente estratégica para o avanço do país no desenvolvimento de aplicações de IA.

Por fim, na camada de aplicações e soluções verticais, o Brasil já apresenta progressos significativos, com o desenvolvimento de *chatbots*, portais inteligentes, soluções de predição agrícola, diagnóstico médico, entre outras iniciativas. Essas soluções demonstram a capacidade nacional de aplicar IA de forma adaptada à realidade brasileira, explorando oportunidades locais e setoriais.

De maneira geral, o CPQD ressalta que todas as camadas da pilha são relevantes e que o desenvolvimento deve ocorrer de forma colaborativa e integrada, especialmente diante dos recursos limitados do Programa Brasileiro de Inteligência Artificial (PBIA) frente aos investimentos internacionais. Assim, o avanço sustentável e coordenado em múltiplas camadas é essencial para fortalecer a autonomia e a soberania digital do país.

Samuel Xavier de Souza

Pesquisador e docente – Universidade Federal do Rio Grande do Norte (UFRN)

A UFRN, por meio do Núcleo de Processamento de Alto Desempenho (NPAD/UFRN), apresentou sua visão sobre a pilha de *software* e manifestou a expectativa de integrar o Sistema Nacional de Processamento de Alto Desempenho (SINAPAD). A instituição já presta serviços à comunidade acadêmica e tem como objetivo ampliar o atendimento a um número cada vez maior de usuários.

O foco principal do NPAD é tornar a computação uma ferramenta cotidiana e acessível à população, o que tem se refletido no aumento significativo de publicações científicas na área de IA, elaboradas por pesquisadores de diversas áreas do conhecimento, como biologia, física, telecomunicações, entre outras, e não apenas por profissionais da ciência da computação. Tal avanço foi possível porque o NPAD estruturou uma equipe especializada para oferecer suporte técnico e científico a pesquisadores de diferentes campos, além de contar com infraestrutura própria, adquiridas com recursos da Finep.

Também se ressaltou a importância da formação de profissionais qualificados para o uso e o desenvolvimento de aplicações nas camadas superiores da pilha de *software*, destacando que a capacidade humana é um fator crítico de sucesso nesse processo. Foi citado, como exemplo, o caso do *DeepSeek*, que se destacou ao alcançar resultados expressivos em *software* utilizando uma fração do *hardware* empregado pela *OpenAI*.

O representante do NPAD enfatizou, por fim, que é fundamental investir na capacitação de pessoas aptas a apoiar usuários com demandas de supercomputação e Inteligência Artificial, fortalecendo a autossuficiência nacional e o uso estratégico da infraestrutura computacional avançada disponível no país.

Considerações finais

No painel 2 observa-se a convergência em torno de um mesmo objetivo: fortalecer a capacidade do Brasil em IA, com ênfase em soberania tecnológica, infraestrutura computacional, formação de pessoas e integração institucional.

Dentre as principais propostas deste painel se destacam:

- Estruturação da pilha nacional de IA composta por múltiplas camadas, *frameworks*, bibliotecas, modelos e aplicações; foco no desenvolvimento de

camadas superiores, onde o país pode gerar maior valor agregado, adaptando soluções ao contexto e à cultura nacionais; incentivo à tropicalização tecnológica, incorporando o idioma português, diversidade regional e soluções específicas a setores estratégicos; e desenvolvimento de modelos médios (SLMs) e integração da multimodalidade (texto, imagem, som e vídeo) para atender demandas nacionais.

- Criação de plataforma nacional compartilhada, com repositórios públicos de *templates* e *stacks open source*, promovendo interoperabilidade e colaboração entre governo, ICTs, universidade e setor privado.
- Governança, ética e soberania digital, com garantia de governança e auditabilidade dos modelos de IA e dados utilizados; implementação de ferramentas de observabilidade e controle de desempenho, reduzindo riscos de *data drift* e *model drift*; autonomia tecnológica nacional; e atuação do governo como catalisador, fomentando startups de alta densidade e cultura *open source*.
- Formação, capacitação e cultura de inovação, com investimento em formação de profissionais qualificados para atuar nas diversas camadas da pilha de IA (dados, modelos e aplicações); criação de comunidades de prática e capacitação para democratizar o uso da IA; incentivo à pesquisa interdisciplinar; e promoção de ecossistemas colaborativas, integrando universidades, ICTs, governo e setor privado.

Matriz SWOT – Painel 2

103

Forças	Fraquezas
• Existência de instituições de referência no Brasil. • Experiências concretas com plataformas de IA (A2B2, BDC-Lab, NPAD). Grande volume de dados nacionais estratégicos (observação da Terra, meteorologia, finanças). • Competência técnica reconhecida internacionalmente e exportação de mão de obra qualificada. • Engajamento de múltiplos setores (público, privado e acadêmico).	• Fragmentação de iniciativas e ausência de integração nacional da pilha de IA. • Escassez de financiamento e recursos contínuos para P&D. • Baixa cultura de <i>open source</i> e colaboração tecnológica. • Deficiência de profissionais em MLOps, governança e segurança de IA. • Dependência parcial de tecnologias e infraestruturas estrangeiras.
Oportunidades	Ameaças
• Crescimento global do mercado de IA e demanda por soluções éticas e soberanas. • Potencial de atrair investimentos internacionais para soluções locais de qualidade. • Cooperação internacional sem perda de soberania (exemplo: interoperabilidade INPE). • Expansão da IA para novos campos científicos e sociais. • Avanço de tecnologias <i>open source</i> e acessíveis (AutoML, <i>low code</i>, SLMs).	• Domínio das Big Techs e risco de dependência tecnológica. • Barreiras geopolíticas e restrições de acesso a hardware e modelos. • Concorrência internacional em aceleração exponencial. • Instabilidade econômica e orçamentária em programas nacionais de P&D. • Falta de coordenação política e continuidade institucional em políticas de IA.

Figura 12. Matriz SWOT – Painel 2

Painel 3: Como o Brasil está posicionado no desenvolvimento de IA para dispositivos móveis e embarcados

Este painel discute o papel estratégico dos dispositivos móveis e embarcados na construção de uma Pilha de IA eficiente, acessível e soberana. A partir de uma perspectiva técnico-científica, serão debatidos os avanços em pesquisa e desenvolvimento no Brasil voltados à execução de modelos de IA na borda, em smartphones, sensores, microcontroladores e outros dispositivos com recursos limitados.

Moderação: Eduardo Matos de Brito Júnior (*Especialista da Embrapii*)

Ariston Brandão

Gerente Sênior de desenvolvimento de IA embarcada - SIDI

A experiência do SIDI no desenvolvimento de uma solução de Inteligência Artificial voltada à indústria sucroalcooleira trouxe aprendizados valiosos sobre os desafios práticos da implementação de IA em ambientes reais. Inicialmente, a equipe identificou que os dados disponíveis não apresentavam qualidade satisfatória, comprometendo a eficiência do modelo. A partir dessa constatação, iniciou-se um trabalho de aprimoramento e otimização dos dados, buscando aumentar a precisão e reduzir o tempo de inferência, que até então era excessivamente elevado. Com o avanço das capacitações técnicas e metodológicas, o grupo conseguiu diminuir significativamente o tempo de inferência, demonstrando a importância de investir em formação especializada e boas práticas de engenharia de IA.

Outro ponto crítico identificado foi o gargalo de conectividade, que limitava o desempenho da solução. Essa limitação evidenciou a necessidade de migrar a aplicação para um ambiente embarcado, reduzindo a dependência de comunicação constante com servidores remotos. No entanto, durante a implementação, constatou-se que o mundo dos sistemas embarcados apresenta desafios próprios, o poder computacional mais restrito, a memória é limitada e o consumo energético se torna uma variável crucial a ser otimizada. Tais diferenças exigem uma abordagem técnica e estratégica distinta da tradicional IA em ambiente de alto desempenho (AI box). Ao transitar para dispositivos mais acessíveis, como celulares, que possuem maior capacidade de processamento, mas ainda apresentam limitações para a execução de modelos de linguagem de grande porte (LLMs), torna-se evidente a importância de planejar a inferência na borda de maneira pragmática.

Esse planejamento deve considerar infraestrutura de conectividade, capacitação técnica e parcerias estratégicas. A cooperação entre governo, iniciativa privada e instituições de pesquisa é essencial para consolidar um ecossistema sustentável de desenvolvimento e

aplicação de IA embarcada no país. Além disso, é fundamental fomentar colaborações internacionais, que ampliem o intercâmbio de conhecimento, experiências e tecnologias.

Mateus Pierre

Diretor de pesquisa e desenvolvimento – Instituto Eldorado

Nos debates recentes sobre soberania em Inteligência Artificial (IA), tem se destacado a importância de compreender que não é necessário desenvolver internamente todas as camadas da pilha tecnológica. O foco deve estar, sobretudo, nas camadas em que há uma ampla oferta de soluções abertas (*open source*), permitindo que o país domine essas tecnologias, fortaleça sua autonomia digital e firme sua soberania tecnológica. Ao mesmo tempo, é essencial identificar quais camadas demandam desenvolvimento nacional próprio, voltado às necessidades, características e prioridades estratégicas do Brasil. Essa capacidade de analisar criticamente a pilha tecnológica e definir os pontos de investimento prioritários é fundamental para orientar políticas públicas, alocação de recursos e estratégias de inovação.

Mesmo no campo dos sistemas embarcados, observa-se que o desenvolvimento de modelos e aplicações ocorre nas camadas superiores da pilha, reforçando a importância de investir também em *software* e algoritmos, sem perder de vista o ecossistema de *hardware*. Nesta camada, o país tem avançado em iniciativas estratégicas, como a adoção e fortalecimento da arquitetura RISC-V, considerada aberta e promissora para a soberania tecnológica. Há também o reconhecimento da importância de estimular o desenvolvimento de circuitos integrados e chips voltados à IA, o que ampliaria a capacidade nacional de projetar e fabricar componentes críticos.

Ao refletir sobre os desafios e potencialidades do desenvolvimento soberano em IA, é fundamental considerar as especificidades nacionais e as oportunidades de aplicação em setores estratégicos. O setor público brasileiro, por exemplo, dispõe de uma riqueza de dados sem precedentes, como os provenientes do Sistema Único de Saúde (SUS), o que possibilita o desenvolvimento de modelos voltados à melhoria dos serviços de saúde e à inclusão de inteligência em áreas remotas do país. Além disso, o desenvolvimento de soluções personalizadas e embarcadas pode representar um caminho alternativo e promissor para o Brasil, especialmente em setores onde a infraestrutura de conectividade é limitada ou onde há demanda por autonomia local.

Assim como ocorreu com o celular, em que o *software* passou a ter papel mais determinante que o *hardware*, e como vem acontecendo no setor automotivo, observa-se uma transformação estrutural na forma de conceber tecnologias. Nessa transição, o valor passa a estar no domínio das camadas lógicas e inteligentes, mais do que nos componentes físicos. E de forma a viabilizar essa trajetória, é essencial mobilizar instrumentos de fomento e políticas públicas que incentivem o desenvolvimento tecnológico nacional. Mecanismos como a Lei de Informática (Lei de TICs), o Programa

de Apoio ao Desenvolvimento Tecnológico da Indústria de Semicondutores (PADIS) e outros programas de incentivo à pesquisa e inovação são fundamentais para estimular o desenvolvimento de modelos, algoritmos e tecnologias de IA, tanto na nuvem quanto na borda.

Fernando dos Santos

Especialista em Visão Computacional – Venturus

Observando o cenário global, há uma tendência crescente na concepção de modelos de aprendizado de máquina cada vez mais complexos, que exigem grandes volumes de dados, tanto em quantidade quanto em variedade. O treinamento desses modelos demanda elevado poder computacional, o que, conseqüentemente, implica alto consumo de energia. Essa exigência está diretamente relacionada à busca por maior desempenho e precisão nos resultados.

Entretanto, quando se trata de dispositivos móveis e sistemas embarcados, o contexto é distinto. Nesses casos, espera-se que os modelos operem com menor capacidade computacional, menor dependência de servidores de borda e tempos de resposta reduzidos, garantindo assim eficiência operacional para as empresas e melhor experiência para os usuários.

Diante desse desafio, surge a questão: como trazer a qualidade dos modelos complexos para dispositivos móveis e embarcados? Atualmente, a literatura científica e o mercado oferecem alternativas voltadas à compressão de modelos, que permitem reduzir a complexidade de arquiteturas robustas sem comprometer, em muitos casos, o desempenho. Estudos demonstram que, em determinadas situações, a compressão pode resultar em níveis de performance equivalentes aos modelos originais.

Outro aspecto relevante é a otimização de modelos, na qual se observa que abordagens mais simples podem, por vezes, superar em eficiência as estratégias excessivamente complexas. Esse ponto é crucial, pois o domínio da otimização de modelos amplia a capacidade técnica dos profissionais e incentiva o uso consciente da complexidade, evitando o emprego de soluções sofisticadas sem o devido entendimento de sua aplicabilidade.

Além disso, modelos otimizados e redes neurais mais enxutas tendem a ser mais interpretáveis, característica especialmente relevante em áreas sensíveis, como a saúde, onde a explicabilidade é um fator determinante.

Com base nesse contexto, é possível identificar diversas oportunidades, sobretudo no campo dos sistemas embarcados. Entre elas, destacam-se:

- Desenvolvimento de *hardware* nacional e de robôs com dispositivos móveis embarcados;

- Visão computacional aplicável a setores como o de óleo e gás, por exemplo, na detecção de incidentes em plataformas;
- Avanço em tecnologias de voz e linguagem, com foco no português brasileiro;
- Criação de agentes com RAG (*Retrieval-Augmented Generation*) capazes de realizar processamento de texto de forma local, sem depender exclusivamente da nuvem, funcionando também em dispositivos móveis.

Essas iniciativas representam não apenas oportunidades de mercado, mas também frentes promissoras de pesquisa científica, nas quais o desenvolvimento de modelos menores e eficientes pode gerar alto valor tecnológico e estratégico.

Leandro Villas (UNICAMP)

Professor Unicamp

Antes de abordar o tema da Inteligência Artificial (IA) na borda, é importante compreender o modelo tradicional de IA na nuvem, ilustrado na Figura 13. Tipicamente, os dispositivos dependiam de uma infraestrutura baseada em nuvem para transmitir dados, onde eram realizadas as análises e inferências, cujos resultados eram posteriormente enviados de volta aos dispositivos.

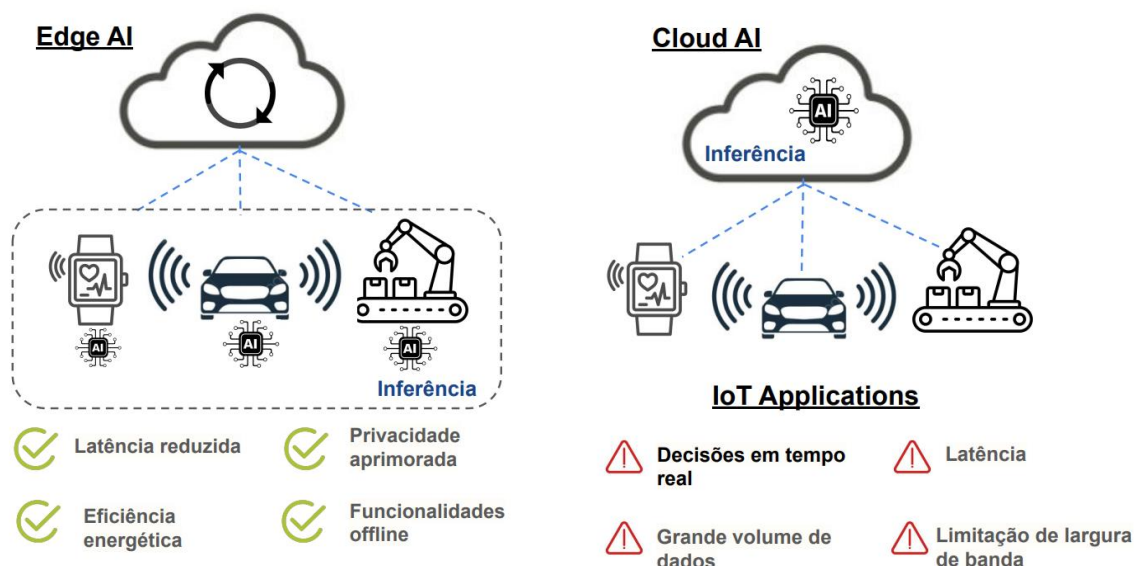


Figura 13. IA na borda x IA na nuvem

Entretanto, esse modelo computacional centralizado passou a apresentar limitações à medida que as aplicações de Internet das Coisas (IoT) passaram a demandar decisões em tempo real. O grande volume de dados gerados, aliado à latência e à restrição de largura de banda, torna a computação em nuvem inviável para diversos casos de uso.

Nesse contexto, surgiu o paradigma da IA na borda (Edge AI), que consiste em levar a inteligência tradicionalmente hospedada na nuvem para os próprios dispositivos. Dessa forma, as inferências passam a ser realizadas localmente, sem a necessidade de interação constante com servidores remotos.

A principal vantagem desse modelo é a eliminação da necessidade de transmissão contínua de dados para a nuvem, o que proporciona latência reduzida, maior privacidade, uma vez que os dados permanecem nos dispositivos, e funcionalidades offline, aspecto especialmente relevante no contexto brasileiro, diante dos desafios de conectividade em diversas regiões. Em suma, a inteligência artificial na borda representa a capacidade de executar algoritmos inteligentes próximos à origem dos dados, garantindo rapidez, privacidade e autonomia, sem depender permanentemente da nuvem.

Dois conceitos complementares concretizam essa visão:

- *TinyML*: corresponde a um conjunto de práticas, técnicas e ferramentas que permitem a execução de modelos de IA em dispositivos com restrições computacionais e energéticas, como microcontroladores e sensores.
- *Federated Learning* (Aprendizado Federado): treinamento distribuído de modelos diretamente nos dispositivos locais, mantendo os dados em segurança e trocando apenas os parâmetros atualizados com o servidor central. Esse servidor agrega as atualizações e retorna um modelo global aprimorado, sem que os dados precisem sair dos dispositivos.

108

Para tornar os modelos de IA viáveis nesse ambiente, é necessário enfrentar diversos desafios técnicos, entre eles:

- *Conexão intermitente*: a instabilidade da conectividade exige arquiteturas resilientes e autônomas, capazes de operar de forma independente da nuvem.
- *Baixa latência*: as inferências locais devem ocorrer rapidamente e com mínima dependência da rede, especialmente em aplicações críticas.
- *Consumo energético restrito*: muitos dispositivos embarcados dependem de baterias ou fontes limitadas de energia, demandando soluções de IA com alta eficiência energética.
- *Memória e processamento limitados*: microcontroladores e sensores geralmente possuem menos de 1 MB de RAM e capacidade de processamento reduzida, o que impõe desafios significativos à execução de modelos complexos.

Dentre as soluções disponíveis para viabilizar o uso de IA na borda, destacam-se:

- *Quantização de modelos*: reduz a precisão numérica dos pesos e ativações, diminuindo o uso de memória e o custo operacional, além de acelerar a inferência, ainda que possa haver pequena perda de acurácia.

- *Pruning (poda de modelos)*: remove pesos ou conexões menos relevantes das redes neurais, reduzindo o tamanho do modelo, acelerando o processamento e diminuindo o consumo de memória, também com possível perda leve de precisão.
- *Destilação do conhecimento (Knowledge Distillation)*: transfere o conhecimento de um modelo grande (teacher) para um modelo menor (student), resultando em arquiteturas mais leves e rápidas, com redução de custo computacional e mínima perda de desempenho.

No contexto brasileiro, há avanços importantes nesse campo. O RIAC (Rede de Inovação em Aplicações Críticas) tem desenvolvido soluções de redução de modelos, *pruning* e destilação do conhecimento, contribuindo para o aprimoramento de arquiteturas de IA aplicáveis à borda. Além disso, existem *frameworks* adaptados para ambientes embarcados, voltados à otimização e portabilidade, bem como bibliotecas ultraleves para microcontroladores que dispensam o uso de sistemas operacionais. Essas soluções resultam em menor consumo de memória e energia, sendo amplamente utilizadas em projetos de TinyML.

Considerações finais

O Brasil tem competências emergentes e ecossistemas institucionais consolidados que o posicionam de forma promissora no campo da IA embarcada. Entretanto, há gargalos estruturais, especialmente em dados, *hardware* e integração de esforços, que limitam a escala e a soberania tecnológica. O fortalecimento de políticas públicas, a formação de talentos e a articulação entre governo, academia e indústria são fatores críticos para consolidar a presença nacional nesse setor estratégico.

Dentre as propostas de destaque, estão:

- Fortalecimento do ecossistema nacional de IA embarcada, envolvendo governo, academia e indústria em arranjos cooperativos. Centros de P&D e startups como agentes de experimentação aplicada e fomento contínuo e previsível por meio de políticas públicas (Lei de TICs, PADIS, Programas Embrapii, Finep, etc).
- Desenvolvimento de competências técnicas e engenharia de IA, abrangendo formação em engenharia de dados, engenharia de modelos e compressão de redes neurais, com foco em eficiência e explicabilidade, e desenvolvimento de *software* embarcado e otimização de *hardware*, com ênfase em restrições de energia, memória e processamento. Essas competências são fundamentais para reduzir o tempo de inferência, aumentar a confiabilidade e viabilizar aplicações de IA em ambientes reais.
- Soberania tecnológica e domínio estratégico da pilha de IA, priorizando o domínio das camadas abertas (open source) da pilha de IA, desenvolvimento nacional de algoritmos, *frameworks* e arquiteturas abertas (como RISC-V), fortalecimento da

capacidade de projetar e fabricar chips e circuitos integrados para aplicações em IA, e o uso inteligente de dados públicos e estratégicos, como os do SUS, para aplicações de alto impacto social.

- Inovação em IA na borda e computação eficiente, impulsionada por desafios de conectividade e necessidade de decisões em tempo real, otimização energética e de recursos computacionais, e segurança e privacidade de dados e inferências locais. Dentre as tecnologias-chave estão: TinyML, aprendizado federado e compressão, quantização, *pruning* e desligação do conhecimento.
- Aplicações estratégicas e oportunidades prioritárias considerando o potencial da IA embarcada, como saúde-pública, agronegócio e indústria sucroalcooleira, óleo e gás, educação e serviços públicos digitais e robótica móvel e sistemas autônomos.
- Cooperação internacional e intercâmbio de conhecimento, indispensável para acelerar o desenvolvimento nacional de IA, com a participação em consórcios globais de open source e padronização tecnológica, projetos bilaterais e multilaterais para intercâmbio de conhecimento e formação de redes de pesquisa internacionais voltadas à IA embarcada e à soberania tecnologia.

Matriz SWOT - Painel 3

Forças	Fraquezas
• Base crescente de instituições de pesquisa e inovação (SIDI, Eldorado, Venturus, Unicamp, RIAC). • Capacitação técnica em crescimento e boas práticas de engenharia de IA. • Experiência prática em IA embarcada e otimização de modelos. • Potencial de dados públicos (ex: SUS) para aplicações sociais e estratégicas. • Apoio de políticas públicas e programas de incentivo (Lei de TICs, PADIS).	• Baixa qualidade e disponibilidade de dados nacionais para treinamento de IA. • Infraestrutura limitada de conectividade e hardware no país. • Dependência de tecnologias estrangeiras nas camadas mais profundas da pilha tecnológica. • Escassez de profissionais especializados em IA na borda e engenharia de hardware. • Fragmentação do ecossistema e ausência de coordenação nacional integrada. • Desafios para otimizar os modelos.
Oportunidades	Ameaças
• Desenvolvimento de hardware nacional e chips dedicados à IA. • A latência, alto volume de dados e a restrição de banda tornam a nuvem inviável para aplicações que exigem respostas em tempo real, como IoT. • Aplicações estratégicas em setores como saúde, agricultura, óleo e gás. • Avanço em IA explicável, TinyML e aprendizado federado. • Cooperação internacional e parcerias público-privadas. • Expansão do uso de IA offline em áreas com baixa conectividade. • Arquitetura RISC-V e o desenvolvimento de chips nacionais são estratégicos para a autonomia tecnológica.	• Concorrência internacional com grandes players dominando a cadeia global. • Falta de investimento contínuo e descontinuidade de políticas públicas. • Vazamento de dados e riscos de segurança em sistemas conectados. • Barreiras regulatórias e dificuldades na adoção de padrões abertos. • Obsolescência tecnológica rápida e risco de dependência de plataformas externas. • Desafios da IA na borda: limitações de memória, processamento e consumo energético em dispositivos embarcados. • Modelos de IA cada vez mais complexos, exigindo grandes volumes de dados e alto poder computacional.

Figura 14. Matriz SWOT - Painel 3

Painel 4: Ciência aberta, open source e cooperações internacionais: limites e potenciais

Este painel discute os desafios e oportunidades relacionados à adoção de práticas de ciência aberta, uso de tecnologias *open source* e parcerias internacionais na construção da Pilha de IA nacional. O objetivo é refletir como essas abordagens podem acelerar o desenvolvimento tecnológico, ampliar a colaboração e reduzir dependências, ao mesmo tempo em que se avaliam os riscos associados à soberania, segurança, governança e sustentabilidade dos projetos.

Moderação: Adriana Thomé (*Coordenadora Geral de Cooperação Multilateral, MCTI*)

Bianca Amaro

Coordenadora de Governança de Dados – Ministério de Ciência, Tecnologia e Inovação (MCTI)

A ciência aberta tornou-se um tema recorrente, especialmente no âmbito da gestão. Contudo, ainda existem diversas formulações, muitas delas bem estruturadas, sobre como ela deveria ser implementada a partir de seus elementos constituintes. A Figura 15 abaixo ilustra alguns dos pilares que compõem o conceito de ciência aberta.



Figura 15. Elementos da ciência aberta

Ciência aberta significa o compartilhamento de todos os processos da pesquisa científica, tratando-se de um conceito internacional, não restrito ao contexto brasileiro, que orienta ações em instâncias de cooperação global. E organismos como a OCDE possuem diretrizes específicas que exigem do Brasil iniciativas voltadas à promoção da ciência aberta. Ademais, discussões em fóruns multilaterais, como os BRICS e o G20, também contemplam ações relacionadas a esse tema. Porém a implementação da ciência aberta encontra alguns desafios, porém grandes oportunidades, descritas abaixo:

- Disseminar a cultura e a prática da ciência aberta, superando a resistência existente em diversos segmentos da comunidade científica.
- Criar e expandir infraestruturas adequadas para sustentar a ciência aberta, promovendo sua disseminação em todo o território nacional.
- Capacitar pesquisadores nas práticas de ciência aberta, uma vez que, em geral, ainda há desconhecimento sobre o tratamento adequado dos dados de pesquisa.
- Revisar os critérios de avaliação científica, incorporando as práticas de ciência aberta como fatores positivos e incentivadores do depósito de dados em repositórios públicos.
- Ampliar a cooperação nacional, regional e internacional, por meio do intercâmbio de dados e resultados de pesquisa.
- Organizar os dados de pesquisa nacionais, atualmente em grande parte armazenados em computadores pessoais, muitos deles oriundos de projetos financiados com recursos públicos.
- Fornecer dados confiáveis para aplicações em inteligência artificial (IA), especialmente em modelos de linguagem (LLMs), de modo que essas tecnologias tragam benefícios concretos à sociedade.
- Estabelecer uma governança de dados aberta e soberana, garantindo segurança, interoperabilidade e controle nacional sobre as informações.

Soberania e Governança de Dados

A soberania de dados pode ser compreendida como a capacidade máxima de um Estado exercer poder social, de forma que suas normas e decisões prevaleçam sobre as de grupos intermediários, como famílias, escolas ou empresas. No contexto das relações internacionais, o conceito está associado à igualdade entre Estados e à independência na tomada de decisões. Em termos práticos, a soberania de dados implica que um país deve ter autoridade sobre os dados gerados, coletados, armazenados e processados em seu território. Isso inclui a capacidade de regular, proteger, acessar e utilizar essas informações de acordo com suas leis, políticas e interesses nacionais.

Para lidar com esse vasto universo informacional, é essencial estabelecer uma governança de dados voltada à ciência soberana, equilibrando as necessidades de cooperação e compartilhamento com as de proteção e segurança. Além disso, a arquitetura de dados deve contemplar infraestruturas organizadas e interoperáveis entre

A governança de dados deve também abranger informações relacionadas à ciência, tecnologia e inovação (CT&I), alinhadas à transformação digital do MCTI. Isso inclui a interoperabilidade entre a Lei de TICs, a Lei do Bem, o Conecta.gov.br e o Cadastro Único de ICTs. Adicionalmente, é necessária a reclassificação das bases de dados

inventariadas pelo MCTI, visando à requalificação do Plano de Dados Abertos (PDA) e à consolidação da Rede Nacional de Indicadores Estaduais de CT&I.

Por fim, os dados resultantes das atividades científicas e tecnológicas brasileiras devem garantir segurança, interoperabilidade e integridade, de modo a criar um ambiente institucional propício à tomada de decisões estratégicas sobre a governança de dados de CT&I no Brasil, com base na definição de padrões de interoperabilidade, conjuntos mínimos de dados e regras claras de integração entre sistemas.

Danielle Ayres

Diretora de Segurança da Informação no Gabinete de Segurança Institucional (GSI)

O recorte apresentado no Seminário da Pilha de IA é central para o Gabinete de Segurança Institucional (GSI). Ao considerar o ecossistema decorrente do desenvolvimento da inteligência artificial (IA) voltada ao Estado, dois pontos se destacam como fundamentais para o GSI. O primeiro é a segurança como elemento central em todo o processo. Não há como construir bancos de dados, modelos de linguagem (LLMs) ou outras infraestruturas tecnológicas sem que estejam intrinsecamente conectadas à proteção dos sistemas e dos dados gerados.

O segundo elemento é o aspecto multidimensional da IA, de forma que suas soluções não serão unilaterais, exigindo articulação entre diferentes órgãos e instituições, como o MCTI, GSI, MRE, dentre outros. Além disso, é imprescindível valorizar as capacidades das instituições federais de ensino e pesquisa, que devem atuar como motores desse processo de desenvolvimento e inovação.

Com base nessas premissas, foram definidos quatro eixos centrais para orientar ações voltadas à tríade soberania, segurança e inteligência artificial:

- *Segurança desde a concepção (security by design) e por padrão (security by default)*: toda ferramenta, componente ou etapa da arquitetura de IA deve incorporar, desde sua concepção, mecanismos de segurança estruturais e automáticos, garantindo resiliência e confiabilidade desde o início do desenvolvimento.
- *Sistemas auditáveis*: é necessário assegurar que os sistemas sejam passíveis de auditoria, de modo que a capacidade técnica e regulatória se torne um modelo de referência. Esse processo não deve ser visto como um entrave ou uma ameaça, mas como uma garantia de confiança e transparência no desenvolvimento tecnológico.
- *Cooperação estratégica*: para ampliar a segurança, eficiência e qualidade de cada elemento da cadeia de IA, é essencial promover diálogos e parcerias em níveis internacional, nacional e regional. Essas cooperações devem estar alinhadas às demandas estratégicas do Brasil, tanto sob o ponto de vista tecnológico quanto

protetivo. Na América Latina, em especial, tais parcerias podem favorecer a criação de um grande repositório regional de dados, posicionando o Brasil como um protagonista nesse processo.

- *Governança e certificação de dados*: é fundamental estabelecer mecanismos de governança voltados à proteção e certificação dos dados, definindo padrões de qualidade e segurança. Essa iniciativa já vem sendo desenvolvida em conjunto com outros órgãos públicos, buscando padronizar processos e garantir integridade e soberania sobre as informações nacionais.

Por fim, é importante priorizar a cooperação com parceiros do Sul Global e da América Latina, sem excluir a colaboração com países tecnologicamente mais avançados. Contudo, tais parcerias devem ser construídas de forma estratégica, focadas na criação de padrões e arquiteturas de sistemas que atendam às demandas e interesses convergentes do Brasil e da região.

Guilherme Fitzgibbon

Assessor Especial – Ministério das Relações Exteriores (MRE)

O Ministério das Relações Exteriores (MRE) coordena o Programa Diplomacia da Inovação, que contempla o tema das tecnologias digitais e tem como objetivo posicionar o Brasil como um ator relevante nas discussões internacionais sobre ciência, tecnologia e inovação (CT&I). A iniciativa abrange diversas vertentes, incluindo ações de cooperação internacional, promoção da imagem do Brasil no exterior e articulação em instâncias multilaterais voltadas à negociação de temas estratégicos para o desenvolvimento científico e tecnológico.

No campo da cooperação internacional, destaca-se o esforço de conectar ecossistemas brasileiros de inovação com parceiros estrangeiros. A abertura e a transparência são requisitos fundamentais para que essas colaborações sejam eficazes, uma vez que a cooperação exige tanto a criação de soluções conjuntas quanto a adaptação (“tropicalização”) de soluções desenvolvidas em outros contextos. Esse processo depende, contudo, da disposição dos parceiros internacionais em atuar em áreas de interesse mútuo, respeitando as prioridades e capacidades nacionais.

Considerando o objetivo do Brasil em construir uma “pilha de inteligência artificial” completa, do *hardware* ao *software*, abrangendo desde semicondutores até aplicações de alto nível, o MRE possui formas de colaboração neste processo, seja estimulando parcerias tecnológicas confiáveis, seja pela oferta de soluções completas, reforçando a soberania tecnológica nacional. Para isso, é essencial estabelecer parcerias que contribuam para o desenvolvimento de capacidades locais. O MRE atua como elo de conexão entre diferentes órgãos governamentais e universidades brasileiras com parceiros estratégicos internacionais previamente mapeados por suas missões diplomáticas no exterior.

Um exemplo recente é a parceria entre o Itamaraty e a Malásia voltada ao fomento do setor brasileiro de semicondutores. Embora a Malásia não seja um parceiro tradicional da ciência brasileira, identificou-se uma complementaridade significativa entre os dois países, especialmente no campo da formação e transferência de conhecimento técnico, o que pode impulsionar a indústria nacional de semicondutores. Espera-se que acordos concretos sejam assinados em breve, consolidando essa cooperação.

Além da promoção de parcerias, o MRE também exerce papel central na negociação e difusão da visão brasileira sobre CT&I em fóruns internacionais. Um exemplo notável é a declaração no BRICS sobre Inteligência Artificial, na qual o Brasil, por meio do MRE e em colaboração com o MCTI, defende princípios de código aberto e inovação aberta, alinhados às discussões conduzidas no âmbito do G20.

No que se refere à ciência aberta, é necessário cautela no contexto internacional. Embora o Brasil defenda a abertura e o compartilhamento de dados científicos, nem todos os países compartilham dessa perspectiva. Alguns buscam vincular o Brasil a modelos de desenvolvimento científico dependentes, baseados em relações comerciais e de consumo de soluções estrangeiras, e não em colaborações equitativas.

Dessa forma, é fundamental que o Brasil afirme, nos fóruns internacionais, seu modelo próprio de ciência aberta, baseado em infraestruturas públicas, instituições de pesquisa consolidadas e cooperação internacional equilibrada. Esse modelo não apenas reforça a imagem do país como referência para outras nações, mas também cria oportunidades para atrair talentos e fomentar intercâmbios científicos. Posto isso, o Brasil tem interesse em ampliar suas parcerias internacionais em CT&I, mas é imprescindível que essas relações sejam estabelecidas com regras claras e mutuamente benéficas, de modo a preservar a soberania nacional e fortalecer a autonomia científica e tecnológica do país.

Jorge Edison

Pesquisador Sênior - Fundação Araucária

O professor Jorge apresentou o programa para implantação da Rede de Computação de Alto Desempenho (HPC) no estado do Paraná, desenvolvida em parceria com a Índia. A iniciativa parte do reconhecimento de que, sem o domínio da infraestrutura basilar, o *hardware*, falar em soberania tecnológica torna-se um discurso meramente retórico. Para alcançar um grau efetivo de independência na área de tecnologia da informação, é indispensável dominar essa camada fundamental da cadeia tecnológica.

Com esse propósito, o projeto da Fundação Araucária, em cooperação com instituições da Índia, prevê a construção de uma rede de supercomputadores (HPCs) com tecnologia inicialmente indiana, mas com transferência progressiva de conhecimento e produção para uma empresa nacional. Espera-se que todos os equipamentos sejam montados no Brasil, promovendo a internalização de competências e capacidades industriais.

Embora o objetivo possa parecer de longo prazo, já foi firmado um acordo com o *Centre for Development of Advanced Computing (C-DAC)*, instituição pública de pesquisa da Índia responsável pelo desenvolvimento de supercomputadores, atuante na transferência de tecnologia para empresas nacionais indianas. No projeto com o Governo do Paraná, haverá a transferência de tecnologia para empresas brasileiras, possibilitando a formação de uma base produtiva local em computação de alto desempenho.

Essa cooperação é de relevância estratégica, considerando que a dependência externa em *hardware* representa altos custos econômicos e tecnológicos para o Brasil. As limitações não se restringem apenas ao acesso e aos custos das tecnologias estrangeiras, mas também às restrições de uso impostas por empresas internacionais. Em diversos casos, as licenças de uso de equipamentos importados proíbem determinadas pesquisas ou até mesmo a execução de certos algoritmos. Há inclusive situações em que empresas norte-americanas restringem o acesso a instalações de HPC a pessoas de determinadas nacionalidades, o que reforça a vulnerabilidade de depender de sistemas controlados por atores externos. Diante desse cenário, a Fundação Araucária busca superar essas limitações por meio da construção e montagem dos equipamentos em território nacional, garantindo maior autonomia e segurança tecnológica.

Importante mencionar que o histórico de dependência tecnológica brasileira é antigo. Já na década de 1980, os Estados Unidos restringiram a venda de um supercomputador a uma universidade brasileira. Na mesma época, a Índia enfrentou situação semelhante, fato que serviu como alerta estratégico para o país asiático, levando-o a criar o C-DAC e a desenvolver seu próprio programa de tecnologia da informação e computação avançada. Em 1989, os indianos produziram o primeiro supercomputador soberano e, atualmente, contam com 47 supercomputadores interconectados em uma estrutura denominada *National Supercomputing Mission*, cuja governança tecnológica apresenta elementos que o Brasil pretende adotar como referência.

Com isso, o Paraná será o primeiro estado brasileiro a implantar uma rede de 9 supercomputadores interligados por uma rede de alta velocidade, distribuídos entre as 7 universidades estaduais e 1 instituto de pesquisa, utilizando predominantemente tecnologia indiana. O projeto prevê, a médio prazo, a fabricação integral dos equipamentos no Brasil, consolidando o retorno do país à área de computação de alto desempenho. Em suma, o projeto representa um marco de cooperação Sul-Sul, demonstrando que a integração entre países em desenvolvimento pode gerar resultados concretos em inovação, transferência de tecnologia e soberania digital.

Prof. Philippe Olivier Navaux

Professor do Instituto de Informática da Universidade Federal do Rio Grande do Sul (UFRGS) e Presidente do Conselho do SCALAC

O professor iniciou sua fala trazendo a relação simbiótica entre computação de alto desempenho (HPC) e IA, uma vez que não é possível desenvolver IA sem alto poder de processamento computacional, sendo, portanto, o HPC um elemento essencial para o avanço das tecnologias de IA. Dentro deste cenário, o Brasil apresenta relativa relevância, figurando na 8ª posição, com 9 supercomputadores ativos, segundo o *ranking* das TOP500 máquinas de HPC do mundo (dados de 2023), o que demonstra uma presença relevante no cenário global de infraestrutura computacional (Figura 16. Presença do Brasil em número de Sistemas de HPC

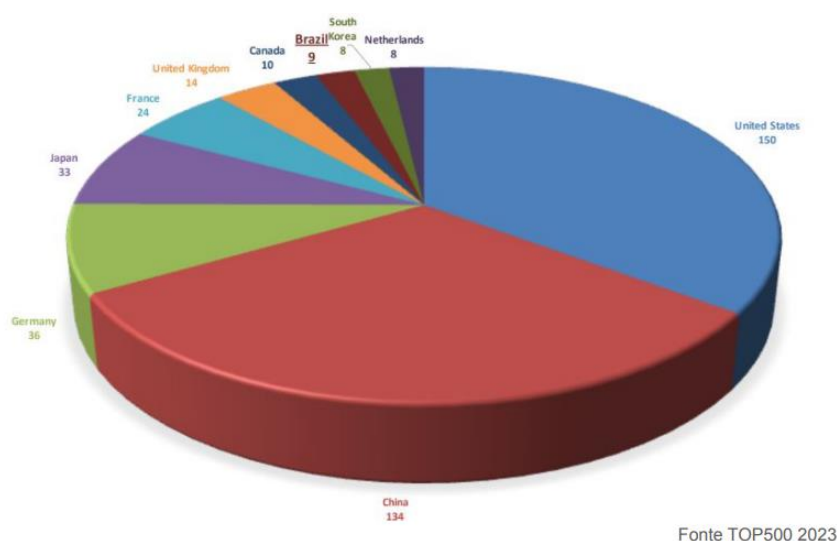


Figura 16. Presença do Brasil em número de Sistemas de HPC

Além da infraestrutura tecnológica, o Brasil apresenta capacidade significativa de formação de recursos humanos altamente qualificados. Atualmente, há 85 programas de pós-graduação voltados à área (Figura 17), formando aproximadamente 400 doutores por ano. Segundo dados do GeoCapes (2024), foram distribuídas 909 bolsas de doutorado e 875 bolsas de mestrado, evidenciando o compromisso do país com o fortalecimento da pesquisa científica e tecnológica.



Figura 17. Número de programas e formação de alunos de pós-graduação

Também é importante mencionar que se encontra em desenvolvimento, em parceria com o Laboratório Nacional de Computação Científica (LNCC), um observatório dedicado ao mapeamento de grupos e iniciativas de pesquisa em Inteligência Artificial no Brasil. O objetivo é identificar competências, promover sinergias e apoiar políticas públicas voltadas à consolidação da IA como área estratégica para o país. Alguns números iniciais apontam que há 78 grupos de IA no Brasil e 12 centros de HPC (>70TFLOP) (Figura 18).

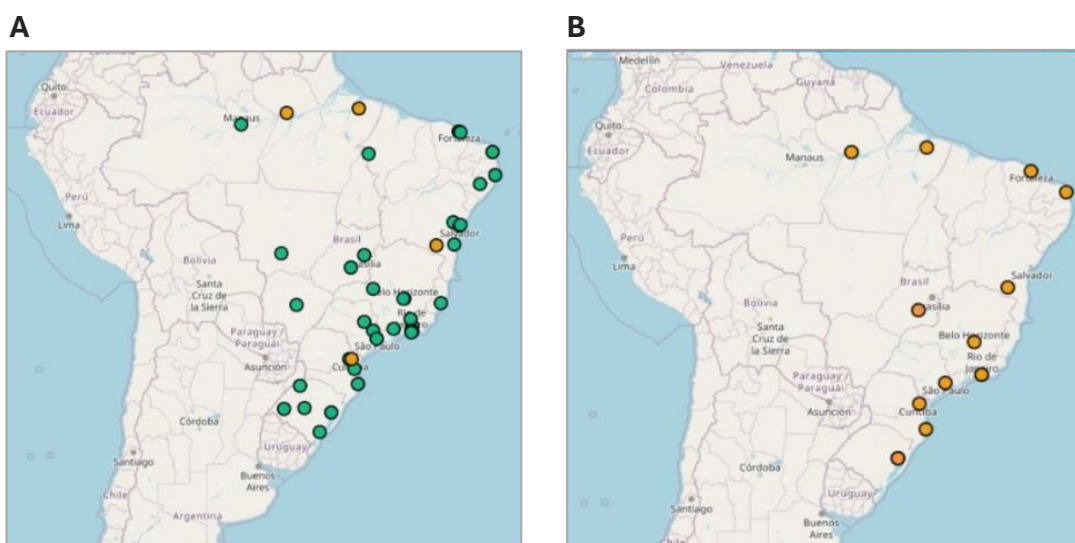


Figura 18. Grupos de IA e Centros de HPC no Brasil

No contexto latino-americano, destaca-se o papel do Sistema de Computação Avançada da América Latina e Caribe (SCALAC), que pode ser considerado o equivalente regional do Sistema Nacional de Processamento de Alto Desempenho (SINAPAD) no Brasil. O SCALAC representa uma iniciativa de cooperação entre diversos países da América Latina, amparada por Memorandos de Entendimento (MOUs) firmados entre instituições de pesquisa da região. No âmbito do SCALAC, está em curso um levantamento regional

de máquinas e infraestruturas de HPC, cujo resultado preliminar indica que a maioria dos supercomputadores latino-americanos se encontra no Brasil, sendo grande parte pertencente a Petrobras ou financiada por ela.

Do ponto de vista estratégico, o SCALAC tem como proposta expandir sua atuação em parceria com o LNCC, promovendo a integração entre o SINAPAD e a rede latino-americana de HPC, de modo a ampliar a cooperação científica e tecnológica na região. Essa integração continental tem sido acompanhada atentamente, e esforços concretos vêm sendo realizados para fortalecer os laços técnicos e institucionais entre os países da América Latina no campo da computação de alto desempenho e inteligência artificial.

Por fim, é importante destacar o Índice Latino-Americano de Inteligência Artificial (ILIA), elaborado pelo Chile, que busca mensurar e comparar o desenvolvimento da IA entre os países latino-americanos.

Primeiramente, a ciência não tem fronteiras e está por trás de tudo que se está falando no seminário. Ele começa com a simbiose entre HPC e IA, ou seja, não se faz IA sem poder de processamento, para tal se precisa de HPC.

Considerações finais

O painel 4 do foi dedicado à discussão dos desafios e oportunidades relacionados à adoção de práticas de ciência aberta, uso de tecnologias *open source* e parcerias internacionais na construção da pilha de IA nacional. O objetivo central foi refletir sobre como essas abordagens podem acelerar o desenvolvimento tecnológico, ampliar a colaboração e reduzir dependências, ao mesmo tempo em que se avaliam os riscos associados à soberania, segurança, governança e sustentabilidade dos projetos.

- Consolidação de uma Política Nacional de Ciência Aberta, alinhada às diretrizes da OCDE, BRICS e G20, articulando abertura, segurança e soberania de dados científicos; e estabelecimento de metas nacionais para adoção progressiva de práticas de ciência aberta por universidades, institutos de pesquisa e órgãos públicos de fomento; e estimular a formação de consórcio latino-americano de governança de dados soberanos, com padrões regionais de interoperabilidade e segurança.
- Infraestrutura e capacitação, com a criação de rede nacional de infraestrutura para ciência aberta, integrando repositório de dados, supercomputadores e serviços de nuvem soberana; investimento na formação de pesquisadores e gestores de dados científicos, com ênfase em interoperabilidade, FAIR data (*findable, accessible, interoperable e reusable*) e cibersegurança.
- Avaliação e métricas de ciência aberta relacionadas ao compartilhamento de dados e produção de *software* livre nos critérios de avaliação de pesquisadores e instituições; e reconhecimento de boas práticas de abertura e reuso de dados públicos financiados com recursos nacionais.

- Governança e interoperabilidade de dados, estruturando um sistema nacional de governança de dados científicos, integrando os diversos instrumentos como Lei de TICs, Lei do Bem, Conecta.br, CadÚnico, etc.; estabelecimento de padrões de interoperabilidade, conjunto mínimos de dados e regras de integração entre sistemas públicos; criação de protocolos de certificação e auditoria de repositórios e bancos de dados científicos nacionais, e adoção de princípios de segurança desde a concepção (*security by design*) em toda a cadeia da pilha de IA.
- Ecossistema *open source* com fomento da adoção de tecnologias abertas como pilar estratégico de soberania tecnológica; financiamento de projetos colaborativos *open source* com foco em infraestrutura, *frameworks* de IA e repositório de dados científicos.

Por fim, os temas da segurança, soberania e ciência aberta convergem em uma visão comum: construir um ecossistema científico e tecnológico nacional robusto, ético e estratégico, capaz de proteger seus dados, gerar inovação e projetar o Brasil como um protagonista global na transformação digital e na inteligência artificial.

Matriz SWOT – Painel 4

Forças	Fraquezas
• Ampla rede de universidades, instituições de fomento e instituições públicas de pesquisa (Capes, CNPq, LNCC, etc.). • Capacidade científica consolidada, com centenas de doutores formados anualmente. • Tradição em diplomacia científica e experiência do MRE em cooperação internacional. • Participação ativa em fóruns internacionais (BRICS, G20, OCDE) e defesa da ciência aberta. • Projetos de cooperação Sul–Sul em estruturação (Índia, Malásia e Governo do Paraná)	• Baixa coordenação entre órgãos governamentais e fragmentação de políticas de CT&I. • Dependência de hardware e infraestrutura estrangeira, especialmente em HPC e IA. • Falta de financiamento contínuo e sustentável para P&D. • Ausência de padrões nacionais unificados de interoperabilidade e governança de dados. • Dificuldade em transformar conhecimento científico em inovação aplicada e industrial.
Oportunidades	Ameaças
• Expansão da rede de HPC e IA nacional e regional (SCALAC, SINAPAD, LNCC). • Fortalecimento da diplomacia da inovação e acesso a novos mercados tecnológicos. • Adoção da ciência aberta como vetor de transparência e cooperação internacional. • Desenvolvimento de ecossistemas regionais de inovação e integração latino-americana. • Transferência de tecnologia e aprendizado com países do Sul Global (Índia, Malásia).	• Risco de dependência tecnológica de potências estrangeiras e perda de soberania de dados. • Barreiras comerciais e restrições tecnológicas impostas por empresas e governos estrangeiros. • Vazamento ou apropriação indevida de dados científicos e estratégicos. • Desigualdades estruturais entre países parceiros, limitando o equilíbrio das cooperações. • Falta de marcos regulatórios claros sobre IA, segurança cibernética e governança digital.

Figura 19. Matriz SWOT – Painel 4

Painel 5: Avanços do Plano Brasileiro de Inteligência Artificial (PBIA) – Construindo o novo Sinapad e o Sistema Nacional de Tecnologia e Inovação em IA (SinIA)

Este painel trata dos avanços do Plano Brasileiro de Inteligência Artificial, bem como das iniciativas que estão sendo construídas para fortalecer a governança do ecossistema de inteligência artificial brasileiro. O objetivo é dialogar com a comunidade científica e empresarial nacional sobre perspectivas dos arranjos de governança previstos, buscando formas de aperfeiçoá-los.

Moderação: Washington Luís Ribeiro (*Coordenador geral de informação científica e tecnológica do Ibict*)

Hugo Valadares

Diretor do Departamento de Ciência, Tecnologia e Inovação Digital – Ministério de Ciência, Tecnologia e Inovação (MCTI)

É essencial que o Brasil desenvolva competências próprias na criação de grandes modelos de linguagem (LLMs), e o Ministério da Ciência, Tecnologia e Inovação (MCTI) tem apoiado iniciativas voltadas a esse objetivo. No entanto, esse não deve ser o foco exclusivo, especialmente diante da limitação de recursos financeiros e de infraestrutura existentes no país. É necessário equilibrar esforços entre recursos humanos, dados e infraestrutura, priorizando áreas em que o Brasil possa obter maior impacto estratégico e sustentável. Nesse sentido, o Programa Brasileiro de Inteligência Artificial (PBIA) contempla essas três dimensões fundamentais, além de outros objetivos estruturantes, como a promoção do P&D nacional e o fomento à cooperação internacional em inteligência artificial (IA).

A cooperação científica é um elemento central desse processo e vem evoluindo de forma significativa. No passado, era comum práticas isoladas de pesquisa, que se mostraram ineficientes e insustentáveis, sobretudo em relação ao uso de *hardware*, dado o alto custo de implantação e manutenção de infraestrutura computacional. E no desenvolvimento de *software*, a colaboração entre laboratórios e instituições torna-se ainda mais crucial para acelerar a inovação e otimizar recursos. Dado esse contexto, destaca-se a importância da Portaria nº 5.156/2004 do Sistema Nacional de Processamento de Alto Desempenho (SINAPAD), que instituiu o sistema com os seguintes objetivos:

- Prestar serviços de computação de alto desempenho sob demanda a universidades, institutos de pesquisa e outras instituições públicas e privadas;
- Apoiar o desenvolvimento de produtos e aplicações em processamento de alto desempenho (PAD);
- Fomentar e apoiar a formação de pessoal especializado na área;

- Transferir conhecimento e tecnologia em PAD;
- Difundir a cultura e ampliar o uso do processamento de alto desempenho no país.

O documento passou por revisão e, em setembro de 2025, foi publicada uma nova portaria do SINAPAD⁴, com aperfeiçoamentos importantes. Dentre estes, destaca-se o artigo que estabelece que os Centros Nacionais de Processamento de Alto Desempenho (CENAPADs) que venham a receber recursos do MCTI passarão a integrar formalmente o SINAPAD. A nova norma também institui uma estrutura de governança simplificada, composta pelo Laboratório Nacional de Computação Científica (LNCC), pela Secretaria de Transformação Digital (SETAD/MCTI), pela Rede Nacional de Ensino e Pesquisa (RNP) e por um representante (*rotativo*) dos CENAPADs. Além disso, determina que cada CENAPAD deve disponibilizar ao SINAPAD pelo menos 25% de sua capacidade computacional total, fortalecendo o princípio de compartilhamento e otimização de recursos.

Essas medidas abrem caminho para a realização de eventos, comissões técnicas e grupos de trabalho, voltados à produção e disseminação de informações estratégicas sobre HPC e IA, estimulando sinergias entre governo, academia e setor produtivo. E o próximo passo é aproximar os CENAPADs de *startups* e empresas nacionais, de modo a facilitar o acesso a recursos de HPC e reduzir a dependência de serviços estrangeiros. Essa integração contribuirá para o fortalecimento do ecossistema nacional de inovação, incentivando o uso de infraestrutura pública de alto desempenho para o desenvolvimento de soluções tecnológicas competitivas e soberanas.

Fábio Borges

Diretor e professor do Laboratório Nacional de Computação Científica (LNCC)

O Laboratório Nacional de Computação Científica (LNCC) é uma das principais instituições brasileiras dedicadas à pesquisa e ao desenvolvimento de inovação em computação de alto desempenho, desempenhando papel estratégico na promoção da supercomputação no Brasil. A instituição mantém um programa de pós-graduação de excelência, avaliado com nota máxima pela CAPES, que contribui diretamente para a formação de recursos humanos altamente qualificados na área.

O LNCC foi responsável pela coordenação do Sistema Nacional de Processamento de Alto Desempenho (SINAPAD) por mais de 20 anos, desempenhando papel central na sua consolidação. Em 2024, a gestão do SINAPAD foi transferida para o Ministério da Ciência, Tecnologia e Inovação (MCTI), o que demonstra o reconhecimento do papel estratégico do sistema e o fortalecimento da governança nacional da infraestrutura de HPC.

⁴ Portaria n° 9.445/2025. Disponível em: <https://www.in.gov.br/en/web/dou/-/portaria-mcti-n-9.445-de-26-de-setembro-de-2025-659322619>

O LNCC tem atuado de forma ativa no fortalecimento do ecossistema nacional de supercomputação, sendo responsável pela distribuição de aproximadamente R\$ 50 milhões em investimentos destinados aos Centros Nacionais de Processamento de Alto Desempenho (CENAPADs). Dentre os projetos atualmente executados no supercomputador Santos Dumont, localizado no LNCC, destacam-se as pesquisas na área de química, setor de grande relevância industrial e científica para o Brasil (Figura 20).

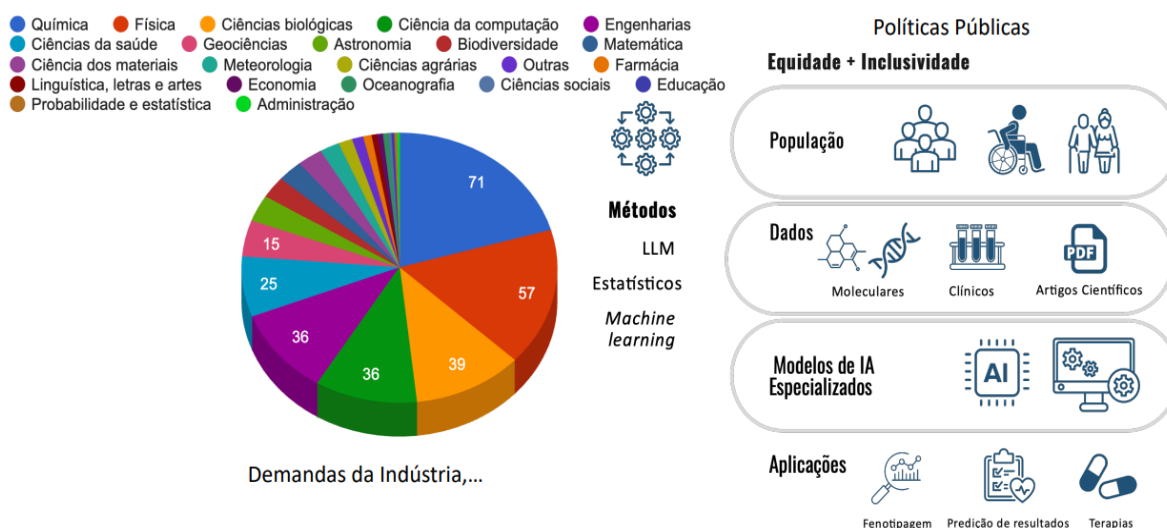


Figura 20. Projetos no Santos Dumont

Quando se analisa a estrutura da computação de alto desempenho, observa-se que as camadas superiores estão frequentemente associadas a aplicações de interesse do setor privado, dependentes de transações comerciais e financeiras, enquanto as camadas inferiores, que sustentam o funcionamento de todo o sistema, dependem de subsídios públicos, tanto no Brasil quanto em outros países. Essas camadas de base, embora possuam menor valor econômico agregado, são essenciais para o desenvolvimento científico e tecnológico e, portanto, requerem financiamento governamental contínuo.

O modelo teórico de Camadas de Interconexão de Sistemas Abertos (*Open Systems Interconnection – OSI*) (Figura 21), foi concebido para padronizar a comunicação entre sistemas computacionais. No entanto, sua implementação prática foi demorada, o que levou à consolidação do *Transmission Control Protocol / Internet Protocol (TCP/IP)*, que, apesar de ter se tornado o padrão global da Internet, não previa mecanismos nativos de segurança. Essa limitação resultou, posteriormente, na criação de camadas virtuais adicionais voltadas à proteção e criptografia de dados, um ponto crítico também observado na pilha de *software* de inteligência artificial (IA), que demanda arquiteturas seguras e eficientes desde os níveis mais básicos.

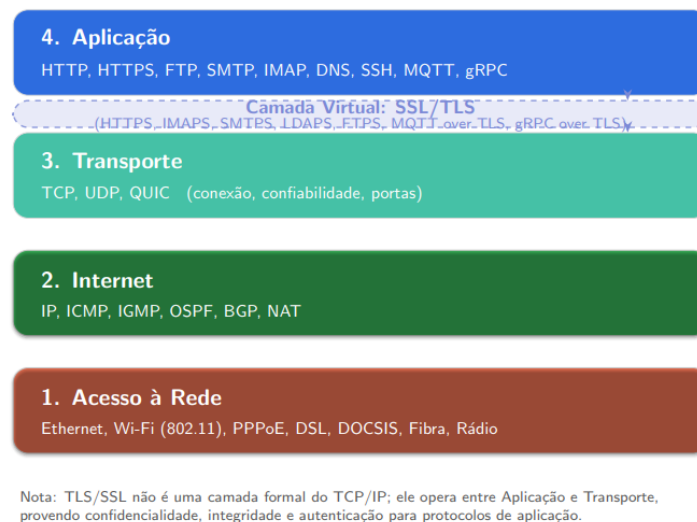


Figura 21. Modelo Teórico – Camadas de Interconexão de Sistemas Abertos (OSI)

O LNCC também se destaca pela busca constante de eficiência energética e computacional. A instituição realiza análises detalhadas de custo e desempenho, calculando o número exato de processadores necessários para resolver diferentes tipos de problemas, com o objetivo de otimizar o uso dos recursos de HPC. Essa preocupação com a eficiência é crucial, especialmente em dispositivos amplamente utilizados, como celulares e sistemas embarcados. Por isso, é fundamental que os algoritmos utilizados nas camadas mais baixas da computação sejam otimizados e energeticamente eficientes, evitando o uso de “algoritmos gulosos” que elevam o consumo energético e reduzem a sustentabilidade dos sistemas.

124

José Ferreira

Assessor acadêmico e científico da (RNP)

Foi apresentado um breve histórico da Rede Nacional de Ensino e Pesquisa (RNP), destacando o crescimento da demanda por novos serviços voltados à pesquisa científica e alta computação. Desde sua criação, a RNP tem se empenhado em apoiar e atender a comunidade acadêmica e científica brasileira, oferecendo soluções inovadoras de conectividade, armazenamento e processamento de dados.

Com o avanço da IA, a RNP passou a disponibilizar serviços de computação em nuvem e ambientes de contêineres para pesquisadores, ampliando suas capacidades tecnológicas. Atualmente, desenvolve um projeto na área de 5G/6G, no qual a IA desempenha papel central, e disponibiliza infraestruturas baseadas em GPU, possibilitando o acesso de grupos de pesquisa a recursos computacionais de alto desempenho. Desta forma, a RNP representa o elo essencial de conectividade entre os

supercomputadores nacionais, desempenhando papel estratégico na proposta de integração submetida ao PBIA, que tem como objetivos principais:

- *Oferta de recursos distribuídos de armazenamento e processamento baseados em GPU, com foco em:*
 - Agilizar o atendimento aos usuários;
 - Garantir o acesso de projetos de pequeno porte;
 - Impulsionar a pesquisa e a capacitação em IA.
- *Interconexão de infraestruturas de alto desempenho:*
 - Integrar o supercomputador Santos Dumont, os CENAPADs existentes e a nova infraestrutura distribuída, por meio de uma rede de alto desempenho dedicada e segregada da rede de produção;
 - Permitir a rápida movimentação de dados entre centros de forma segura e eficiente;
 - Garantir baixas latências, possibilitando processamento paralelo;
 - Construir uma infraestrutura federada de processamento para IA.
- *Definição de um modelo de governança eficiente:*
 - Estabelecer políticas e diretrizes de uso dos recursos;
 - Capacitar e apoiar os usuários no uso das infraestruturas;
 - Assegurar um ambiente seguro e em conformidade com as normas regulatórias.
- *Sustentabilidade e evolução da infraestrutura:*
 - Implementar um modelo de sustentabilidade de longo prazo, com definição clara sobre responsabilidades financeiras e operacionais após a conclusão do projeto;
 - Garantir a continuidade e a expansão das capacidades computacionais ao longo do tempo.

Em termos de processamento, a RNP, devido à sua ampla integração com a comunidade científica, pretende atender principalmente à camada 3, voltada a modelos de IA de pequeno porte. Esses modelos se caracterizam por:

- Possuírem menos de centenas de bilhões de parâmetros;
- Utilizarem conjuntos de dados de tamanho limitado (inferiores a dezenas de terabytes);

- Demandarem curtos períodos de processamento (poucos dias);
- Exigirem agilidade no atendimento das solicitações dos pesquisadores.

A governança dessa infraestrutura de nuvens baseadas em GPU está sendo estruturada pela RNP, que já possui parcerias com centros de dados localizados em São Paulo, Fortaleza e Brasília. Nesses locais será hospedada a infraestrutura de armazenamento e processamento, todos interligados à rede de alta velocidade da RNP, garantindo disponibilidade e eficiência no acesso aos recursos computacionais (Figura 22).

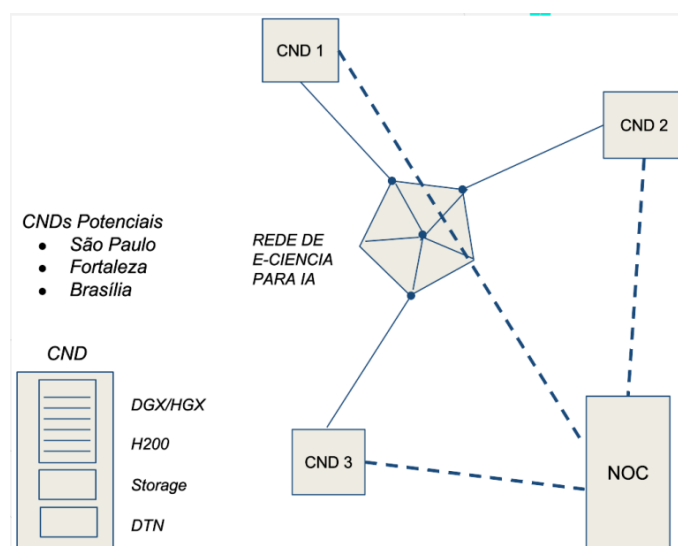


Figura 22. Nuvem de GPUs – RNP

O projeto prevê a implantação de uma rede dedicada de alta capacidade, com enlaces entre 100 e 200 Gbps, interligando os 9 CENAPADs e permitindo a expansão da conectividade para novos parceiros (Figura 23). Essa estrutura permitirá transferência rápida e segura de grandes volumes de dados, elemento essencial para o treinamento e execução de modelos de IA.

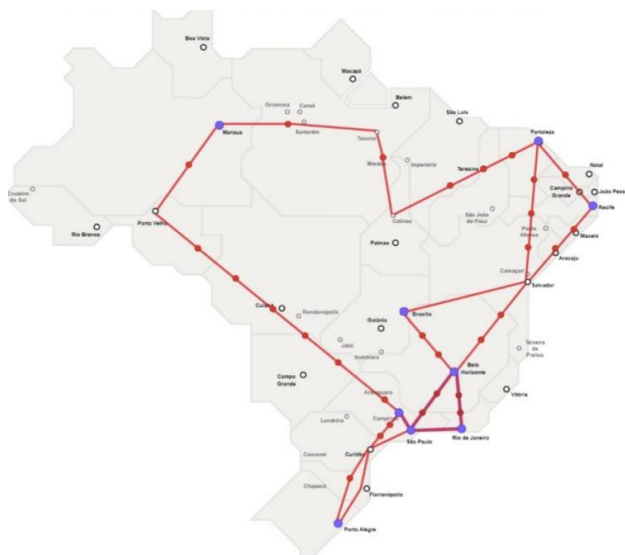


Figura 23. Rede de Alto Desempenho para IA

A agilidade na movimentação de dados entre centros de processamento é um fator determinante para o uso eficiente dos recursos. Por exemplo, caso um centro (A) não disponha de capacidade suficiente para o treinamento de um modelo, os dados poderão ser rapidamente transferidos para outro centro (B), otimizando o uso dos recursos disponíveis em toda a rede. Ademais, com uma rede de alta velocidade e baixa latência, a RNP poderá realizar transferências rápidas e seguras. A instituição possui expertise consolidada em *Data Transfer Nodes (DTNs)*, utilizados em projetos complexos com grandes empresas e centros de pesquisa, como a Petrobras, para viabilizar transferências massivas de dados científicos.

Em relação a estrutura de governança, esta deve ser revisada periodicamente, com monitoramento contínuo do uso dos recursos, visando otimizar o desempenho e evitar subutilização. Além disso, é essencial identificar gargalos operacionais e propor ajustes técnicos e gerenciais. Outro tema prioritário é a capacitação de usuários, em que a RNP planeja oferecer programas de treinamento e cursos especializados, em parceria com pesquisadores e instituições acadêmicas, para maximizar o aproveitamento da infraestrutura e estimular o desenvolvimento de novas aplicações em IA.

Por fim, a sustentabilidade financeira e operacional da infraestrutura deve ser objeto de discussão permanente, considerando o alto custo de manutenção e a necessidade de planejamento para atualização tecnológica contínua. O desafio é não apenas preservar a infraestrutura existente, mas também assegurar sua evolução no longo prazo, em consonância com as demandas crescentes da comunidade científica e tecnológica brasileira.

Considerações finais

O painel 5 do Seminário de Pilha de IA abordou os avanços do Plano Brasileiro de Inteligência Artificial (PBIA), com foco na construção do novo SINAPAD e do Sistema Nacional de Tecnologia e Inovação em IA (SinIA). O objetivo foi fortalecer a governança do ecossistema nacional de IA, promovendo diálogo entre a comunidade científica e empresarial para aprimorar esses arranjos de governança. O conjunto das palestras trouxe as seguintes propostas:

- Consolidação de governança integrada com gestão compartilhada entre LNCC, SINAPAD, RNP e MCTI, com foco em eficiência transparência; criação de comitês técnicos interinstitucionais permanentes para deliberações sobre investimentos, interoperabilidade e normas de segurança.
- Infraestrutura soberana, concluir a interconexão dos 9 CENAPADs por meio da rede de alta capacidade da RNP (100-200Gbps), garantindo baixa latência, segurança e redundância; ampliar a rede federada para incluir novos polos regionais de supercomputação em universidades e parques tecnológicos; e integrar o SINAPAD ao SINIA, promovendo interoperabilidade plena entre recursos de HPC, armazenamento, IA e *big data*.
- Sustentabilidade a longo prazo, com previsão de orçamento recorrente, atualização tecnológica e eficiência energética; criação de fundo permanente de manutenção e atualização, financiado por recursos e contrapartidas privadas.
- Expansão de programas de pós-graduação em HPC e IA aplicada, articulando universidades e CENAPADS como polos de formação e pesquisa; estabelecimento de parcerias com RNP, LNCC, etc, para ofertar de cursos sobre uso de supercomputadores, gestão de dados e desenvolvimento de algoritmos eficientes.
- Certificação para usuários de HPC e IA reconhecendo competências em programação paralela, otimização de código e gestão de dados; criação de comunidade aberta de usuários SINAPAD/SINIA com fóruns, *hackathons* e eventos técnicos voltados à troca de experiência e colaboração científica.
- Infraestrutura como serviço, permitindo que pequenas empresas acessem recursos computacionais sob demanda, com custos subsidiados; promoção de editais para viabilizar o uso do SINAPAD por startups com foco em P&D de IA aplicada à indústria, saúde e agricultura.
- Desenvolvimento de sistema nacional de indicadores de HPC e IA, avaliando desempenho, eficiência energética, impacto científico e geração de inovação; relatórios anuais de desempenho do SINAPAD/SINIA, assegurando transparência e prestação de contas.

A integração entre LNCC, SINAPAD e RNP constitui um marco no fortalecimento da infraestrutura científica nacional. A consolidação de uma rede federada de HPC e IA,

sustentada por governança compartilhada, conectividade avançada e políticas de uso colaborativo, posiciona o Brasil para reduzir vulnerabilidades externas e ampliar sua autonomia tecnológica. O sucesso dessa estratégia dependerá da continuidade do apoio institucional do MCTI, da sustentabilidade financeira e da formação constante de especialistas, garantindo que a infraestrutura digital científica nacional se torne um ativo estratégico de soberania e inovação.

Matriz SWOT – Painei 5

Forças	Fraquezas
- Infraestrutura consolidada com o LNCC e o supercomputador Santos Dumont. - Existência do SINAPAD como estrutura nacional de coordenação e compartilhamento de recursos de HPC. - RNP como rede de alta conectividade com abrangência nacional e expertise em transferência de dados (DTNs). - Governança formalizada com participação de MCTI, LNCC, RNP e CENAPADs. - Capacidade científica e de formação de recursos humanos altamente qualificados.	- Financiamento instável e dependência de repasses públicos anuais. - Fragmentação entre instituições e baixa integração operacional em alguns centros. - Capacidade computacional ainda limitada em relação a países líderes. - Carência de políticas contínuas de atualização tecnológica e de reposição de equipamentos. - Falta de modelos sustentáveis de manutenção e operação no longo prazo.
Oportunidades	Ameaças
- Expansão da rede federada de HPC e IA interconectada via RNP. - Parcerias estratégicas no âmbito do PBIA e com países do Sul Global (Índia, Malásia, América Latina). - Inclusão de startups e empresas nacionais no uso da infraestrutura HPC e IA. - Fortalecimento da diplomacia científica e inserção internacional do Brasil em HPC e IA. - Implementação de uma governança transparente e de políticas de uso compartilhado.	- Dependência tecnológica externa em hardware e GPUs - Crescimento do custo de energia e manutenção de data centers. - Risco de centralização excessiva de recursos em poucos centros. - Vulnerabilidades cibernéticas e riscos à segurança de dados sensíveis. - Falta de continuidade política e orçamentária em projetos de longo prazo.

Figura 24. Matriz SWOT – Painei 5

Painel 6: Capacitação e Formação Técnica para a Construção de uma Pilha de IA Soberana

Este painel discute como preparar e reter os profissionais para atuar de forma especializada nas camadas da pilha de IA identificadas como promissoras para a soberania nacional. O objetivo é identificar lacunas atuais na formação profissional, estratégias de qualificação técnica e o papel das universidades, centros de pesquisa e empresas na construção de uma força de trabalho preparada para sustentar o ecossistema de IA no Brasil.

Moderação: Guilherme de Paula Corrêa (*Coordenador Geral de Tecnologias Digitais - MCTI*)

Alba Melo

Professora do Departamento de Ciência da Computação da Universidade de Brasília (UnB)

A apresentação focou nas iniciativas da Universidade de Brasília (UnB) em termos de formação de recursos humanos na área de inteligência artificial e sobre como deve ser o perfil de um profissional para trabalhar nesta área. Dentre outras iniciativas, está a criação de um bacharelado em Inteligência Artificial (IA), com início previsto para 2026. Além desse curso, estão sendo planejadas novas formações, especialmente em nível de pós-graduação, com linhas de pesquisa voltadas à IA, dentro dos cursos de pós-graduação, o acadêmico, atualmente nota 5, e o profissional, nota 4.

O perfil do profissional em Inteligência Artificial deve contemplar, em primeiro lugar, uma sólida formação nos fundamentos de computação, microeletrônica, matemática e estatística. E é essencial compreender o funcionamento de um sistema computacional e seus componentes, como CPU, hierarquia de memória, dispositivos de entrada e saída, bem como distinguir o que é uma GPU (e o que não é), além de compreender conceitos de *clusters* de computadores e interconexões entre sistemas. Desta forma, a professora Alba enfatiza que um profissional de IA deva ir além de apenas treinar modelos e conhecer ferramentas.

Também é indispensável o domínio de programação paralela, incluindo tecnologias como OpenMP (*multicore*), MPI, sockets (para múltiplas máquinas), CUDA, OpenCL (padrão aberto) e SYCL (para placas gráficas). Ademais, o profissional deve possuir conhecimento aprofundado em Inteligência Artificial e aprendizado de Máquina, compreendendo técnicas como *k-means* (kMN), SVM, redes neurais e redes neurais profundas, visto que há pessoas que treinam modelos sem o devido entendimento teórico sobre redes neurais. Por fim, é necessário o conhecimento em *big data* e *data analytics*.

Entre as competências esperadas dos profissionais de IA, destacam-se:

- *Análise e seleção de técnicas*: capacidade de compreender problemas complexos e selecionar a técnica de IA mais adequada para solucioná-los, e até mesmo desenvolver novas abordagens quando necessário;
- *Implementação e otimização*: habilidade para implementar as técnicas escolhidas de forma eficiente, aplicando a devida parametrização e otimização;
- *Dimensionamento de infraestrutura*: competência para projetar o sistema computacional adequado, avaliando a complexidade algorítmica (tempo e memória) a fim de garantir desempenho e escalabilidade;
- *Adaptação e aprendizado contínuo*: disposição para acompanhar e adotar proativamente novas técnicas e ferramentas que surgem constantemente na área.

Atualmente, observa-se que muitos profissionais carecem de fundamentos essenciais, limitando-se ao uso de ferramentas sem adequada parametrização, o que frequentemente resulta em análises e dados inconsistentes. Para ampliar o número de profissionais qualificados em IA, propõe-se a implementação de uma política nacional de formação em Inteligência Artificial, aliada a cursos e capacitações específicas, com fomento governamental em larga escala. Essa política deveria contemplar o treinamento de estudantes de ciência da computação e engenharia em infraestruturas de HPC (Computação de Alto Desempenho), nos moldes do Programa CI Expert, com oferta de cursos, estágios em empresas nacionais e experiências em centros de supercomputação, tanto no Brasil quanto no exterior.

Outra frente importante seria oferecer formação em fundamentos de IA para estudantes de áreas STEM, garantindo que compreendam conceitos como redes neurais e saibam aplicar as ferramentas atuais de maneira fundamentada. Essa capacitação poderia incluir estágios em centros de supercomputação e o estímulo à pós-graduação *stricto sensu* (mestrado e doutorado) em IA, com períodos de pesquisa no exterior e acesso a laboratórios equipados com GPUs no Brasil. O objetivo é formar mestres e doutores capazes de propor novas técnicas e estratégias de IA, contribuindo para o desenvolvimento de uma IA soberana.

Contudo, a formação de profissionais, por si só, não é suficiente. É necessário estimular a indústria nacional a contratar esses profissionais altamente qualificados. Além disso, é fundamental divulgar a Inteligência Artificial e a computação de alto desempenho (HPC) entre estudantes do ensino médio, por meio de cursos de curta duração e visitas a universidades. Por fim, deve-se fomentar colaborações internacionais, especialmente com países da América Latina e do Sul Global, promovendo tanto o envio quanto o recebimento de estudantes, e criando uma Federação de Supercomputadores para uso compartilhado entre os países envolvidos.

Evânio Antônio de Araújo

Secretário de Gestão da Informação, Inovação e Avaliação de Políticas Educacionais do Ministério da Educação (MEC)

O Ministério da Educação (MEC) está presente em praticamente todos os eixos do Plano Brasileiro de Inteligência Artificial (PBIA). No que se refere, em especial, ao eixo de Capacitação e Formação, destaca-se a preocupação não apenas com o uso da Inteligência Artificial, mas também com a compreensão de seus fundamentos e implicações, abrangendo todos os níveis de ensino, e não somente o ensino superior.

O MEC atua por meio de três frentes principais de trabalho, todas fortemente alinhadas ao objetivo de fomentar o aumento das matrículas. A primeira frente envolve a promoção de cursos de Inteligência Artificial, impulsionada pelo interesse crescente das universidades e dos próprios estudantes, o que reduz a necessidade de uma atuação direta e intensiva do MEC na promoção dessas formações.

A segunda frente está relacionada à pós-graduação em IA, conduzida em articulação com a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). O objetivo é fomentar e catalisar programas de mestrado e doutorado voltados à área, por meio de bolsas específicas, incluindo o doutorado sanduíche, que possibilita vivências internacionais, tanto acadêmicas quanto empresariais. Essa iniciativa busca fortalecer a formação de profissionais com as competências necessárias para atuar no setor de Inteligência Artificial.

A terceira frente concentra-se em estimular e apoiar iniciativas institucionais voltadas à criação de centros de pesquisa e ao fortalecimento de ações já em curso nas universidades. O MEC tem buscado contribuir para aprofundar a sinergia entre instituições de ensino, reconhecendo que muitas dessas iniciativas já têm ocorrido de forma autônoma. Entretanto, há um potencial significativo a ser explorado por meio da institucionalização dessas ações, permitindo que o debate e a cooperação aconteçam com maior regularidade, continuidade e dinamismo.

Por fim, o MEC também tem se debruçado sobre a construção de uma infraestrutura nacional de dados da educação, reconhecendo as dificuldades existentes quanto à compartilhamento, integração e uso desses dados. Além disso, está em discussão a criação de um *sandbox* regulatório destinado a estimular o desenvolvimento e a experimentação de soluções baseadas em Inteligência Artificial no ambiente educacional.

Anderson Soares

Professor do Instituto de Informática da Universidade Federal de Goiás (UFG) e Coordenador do Centro de Excelência de IA (CEIA-UFG)

Atualmente vive-se um momento singular de oportunidades no campo da IA em todo o mundo, com grande visibilidade e valorização da IA, representando uma oportunidade estratégica de geração de valor, sendo fundamental compreender que a cadeia de desenvolvimento é integrada, englobando as etapas laboratorial, de produtização e de aplicação prática. Por isso, é essencial saber extrair valor de cada fase desse processo. O Centro de Excelência em Inteligência Artificial (CEIA) vem aproveitando esse momento favorável e atualmente tem mais de 1.000 pesquisadores envolvidos em mais de 70 projetos, desenvolvidos em parceria com startups, empresas e outras instituições.

A Universidade Federal de Goiás (UFG) foi pioneira ao criar, em 2019, o curso de bacharelado em Inteligência Artificial, cujo projeto pedagógico do curso foi concebido com a premissa de oferecer menos conteúdo teórico fragmentado e mais foco em projetos práticos. Atualmente, o curso segue para a sua terceira turma de formandos, apresentando uma taxa de evasão de apenas 3%, um índice considerado muito baixo para cursos da área de exatas, o que demonstra o sucesso da proposta curricular. Ademais, a universidade tem conseguido articular de forma eficaz a integração entre os projetos acadêmicos e o setor produtivo, fortalecendo a formação prática e aplicada dos estudantes.

No que se refere à retenção de talentos, a instituição tem obtido êxito ao oferecer projetos de pesquisa e desenvolvimento desafiadores e com remuneração atrativa, o que faz com que muitos estudantes reconsiderem propostas de trabalho no exterior, com remuneração em dólar ou euro, para permanecer na universidade e continuar se desenvolvendo academicamente.

Luciana Gomes

Coordenadora de Propriedade Intelectual, Negociação e Prospecção de Parcerias (CNPq)

O Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) possui um amplo portfólio de instrumentos de fomento à pesquisa brasileira. No contexto deste painel, foi apresentado o conjunto de ferramentas de apoio ao Ministério da Ciência, Tecnologia e Inovação (MCTI), com foco no Plano Brasileiro de Inteligência Artificial (PBIA). Essas iniciativas abrangem tanto o apoio à pilha tecnológica da IA quanto o fomento à capacitação e à fixação de talentos e pesquisadores na área.

Atualmente, o CNPq mantém 91 mil bolsas de pesquisa ativas, conforme dados atualizados em seu painel de fomento. Esse conjunto inclui bolsas de vocação tecnológica e outras modalidades voltadas a diferentes níveis e estágios da pesquisa científica, tecnológica e de inovação, sendo, portanto, capaz de atender a toda a cadeia de formação e produção de conhecimento. Adicionalmente, o CNPq apoia 21 mil bolsas voltadas ao letramento digital, no âmbito do programa Mais Ciência na Escola, destinadas a estudantes do 6º ano do ensino fundamental até o ensino médio. Essa

iniciativa integra o conjunto de ações de suporte ao PBIA, ao promover a formação de competências digitais desde a educação básica.

No contexto da vocação para a pesquisa, o CNPq também oferece programas de bolsas de iniciação científica e tecnológica, programas institucionais regulares e bolsas de mestrado e doutorado, compondo um ecossistema de formação contínua de pesquisadores. Destaca-se, nesse âmbito, o Programa de Mestrado e Doutorado Acadêmico para Inovação, que utiliza o mesmo modelo de concessão de bolsas dos programas de pós-graduação institucionais, mas com um diferencial importante: trata-se de projetos de interesse do setor empresarial, que promovem a aproximação entre estudantes e empresas, fortalecendo a integração entre academia e mercado.

Outra modalidade relevante é a das bolsas de fixação de mestres e doutores, que contam com um robusto conjunto de instrumentos de apoio, entre eles o Programa RHAE - Recursos Humanos em Áreas Estratégicas. Esse programa contempla não apenas a fixação de mestres e doutores, mas também o apoio a micro, pequenas e médias empresas (MPMEs), permitindo inclusive que o coordenador do projeto seja bolsista. Além disso, o CNPq mantém parcerias com fundações de desenvolvimento científico e tecnológico regionais, ampliando as possibilidades de fomento, e com o setor privado, por meio de programas que abrangem desde estudantes do ensino médio até mestres e doutores, com foco na empregabilidade e na formação de capital humano altamente qualificado.

Por fim, o portfólio de cooperação internacional do CNPq confere maior alcance e profundidade às ações de capacitação e intercâmbio de talentos, favorecendo tanto a formação avançada de pesquisadores brasileiros quanto a absorção desses profissionais em ambientes empresariais e tecnológicos, fortalecendo a presença do Brasil no cenário global da pesquisa em Inteligência Artificial.

Wagner Meira

Professor do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais (UFMG)

O professor destacou as metas de formação do Plano Brasileiro de Inteligência Artificial (PBIA), que preveem a oferta de 5.000 vagas em cursos de IA. Ressaltou que as metas são, de fato, ambiciosas, e que devem ser assim, porém observou que a chamada “pilha de IA” talvez não possa ser considerada uma pilha propriamente dita, mas algo muito mais dinâmico e volátil. Ao citar dados do *Hugging Face*, mencionou que é disponibilizado um novo modelo de linguagem (LLM) por dia na plataforma, e derivados de LLMs são publicados a cada três minutos, o que evidencia a velocidade exponencial da evolução da área. Diante desse cenário, questionou: “*Enquanto isso, ainda estamos discutindo se faremos uma LLM brasileira.*” Segundo ele, a IA tem avançado de forma tão acelerada

que as LLMs já se tornaram praticamente *commodities*, e o país deveria estar discutindo como alcançar minimamente esse patamar tecnológico e produtivo.

Apontou que a parte mais complexa da IA talvez não sejam os modelos em si, mas sim o campo da IA responsável, que é muito mais desafiador do que a chamada “IA não responsável”, referindo-se àquela que se tornou amplamente conhecida após críticas e debates públicos. Explicou que, quando se fala em IA responsável, a complexidade dobra de tamanho, pois ela abrange não apenas os aspectos técnicos, mas também éticos, sociais e culturais, especialmente em um país onde o áudio do WhatsApp ainda é instrumento de inclusão digital. Segundo ele, trata-se de uma realidade preocupante, pois parte significativa da população brasileira corre o risco de ser cultural e educacionalmente marginalizada, por não possuir condições de compreender o impacto das transformações tecnológicas em curso.

Ao abordar o tema da soberania tecnológica, reconheceu sua importância e concordou que os dados são um ativo estratégico fundamental. Contudo, alertou que a atenção excessiva ao debate sobre dados pode estar desviando o foco de um problema igualmente crítico: a formação e a retenção dos estudantes. Citou como exemplo a Universidade Federal de Minas Gerais (UFMG), onde aproximadamente 30% dos alunos recém-formados já estão atuando de forma remota para o exterior, configurando uma diáspora intelectual sem precedentes. A seu ver, esse fenômeno somente poderia ser revertido por meio de fatores macroeconômicos excepcionais, como uma forte variação cambial ou até mesmo eventos globais extremos, a exemplo de uma guerra mundial.

Ao discutir a formação em IA, destacou que todos os bacharelados atualmente existentes, como ciência da computação, engenharia da computação, engenharia de *software*, ciência de dados e inteligência artificial, são plenamente pertinentes, pois já são *AI-enabled*, isto é, formam profissionais que dominam competências aplicáveis à área e que já estão sendo contratados pelo mercado. Lembrou que o Brasil forma cerca de 20 mil bacharéis e mais de 80 mil tecnólogos por ano nessas áreas, defendendo que a atenção deve se voltar também a esses novos formandos e graduandos, e não apenas à criação de novos cursos.

Ressaltou ainda que é indispensável manter a qualidade e a solidez da formação, condição essencial para resistir à volatilidade da evolução tecnológica da IA e considera positiva a criação de cursos específicos de IA, pois eles contribuem para oxigenar e modernizar os cursos já existentes. Na UFMG, há atualmente cinco cursos de graduação na área de computação, com entrada anual de cerca de 340 estudantes, e o aspecto mais enriquecedor tem sido a interação entre os alunos de diferentes cursos e disciplinas, o que favorece o intercâmbio de experiências e conhecimentos. Destacou ainda a importância da atuação de centros como o CEIA/UFMG e da UFMG, observando que ainda persiste no mercado a percepção de que os estudantes se formam com pouca maturidade profissional. Para enfrentar esse desafio, defendeu estratégias mais

proativas, que estimulem a resolução de problemas reais e o trabalho com dados concretos desde a graduação, prática que o CEIA tem realizado com êxito.

Por fim, sugeriu que essa experiência poderia ser ampliada para outras instituições, destacando que a mobilidade acadêmica entre universidades brasileiras ainda é muito limitada. Propôs, que o CEIA crie um curso de verão, destinado a atrair estudantes de todo o país, permitindo o intercâmbio de conhecimentos e o acesso a centros de excelência, como o CEIA. Enfatizou, assim, a necessidade de maior integração entre os cursos da área de computação, com acesso ampliado a desafios e problemas reais, como forma de fortalecer a formação e o ecossistema de inovação em IA no Brasil.

Otávio Ciribelli

Gerente Geral de Tecnologia da Informação da Petrobras

A Petrobras dispõe de uma capacidade computacional robusta, com mais de 15 supercomputadores em operação, destacando-se o Pegaso, que por 5 anos foi o maior e mais ecoeficiente supercomputador da América Latina. Atualmente, a capacidade total instalada é de 235 PFLOPS (projeção de 300 PFLOPS para 2026), com 6 supercomputadores listados nos *rankings* internacionais Top500 e Green500. No momento, 3% da capacidade dessas máquinas é dedicada a aplicações em IA e aprendizado de máquina (*machine learning*), com a meta de atingir 13% em 2026.

Desde 2019, a Petrobras investe em supercomputadores para serem utilizados por Instituições de Ciência e Tecnologia (ICTs) em todo o Brasil. Além de prover capacidade computacional para a academia brasileira, os investimentos nesses centros de HPC acadêmico permitem fomentar o uso de novas arquiteturas computacionais e soluções inovadoras. A empresa também desenvolveu sua própria LLM, baseada em dados nacionais que representam três décadas de conhecimento acumulado em engenharia estratégica. O modelo foi integralmente treinado e operado em supercomputadores nacionais, Santos Dumont (LNCC), Tatu, Gaia e Pegasus, em um ambiente totalmente controlado pela Petrobras e executado localmente, sem utilização de APIs externas.

No eixo de capacitação e desenvolvimento de competências, a Petrobras já capacitou cerca de 2.000 profissionais em trilhas de ciência de dados e *analytics*, além de proporcionar acesso ao hub de modelos de IA a aproximadamente 9.000 colaboradores. Externamente, destaca-se o Programa Autonomia e Renda – Petrobras Tecnologia, desenvolvido em parceria com o SENAI, que tem como objetivo capacitar 420 jovens em 7 estados brasileiros em temas estratégicos de Tecnologia da Informação e Comunicação (TIC), incluindo robotização, IA, computação em nuvem e desenvolvimento *low code*.

Considerações finais

O Brasil possui fundamentos sólidos e ativos institucionais relevantes para desenvolver estratégias de capacitação em IA robusta e soberana, incluindo infraestrutura de supercomputação, programas de fomento consolidados e universidades engajadas. No entanto, enfrenta desafios estruturais, como a fragmentação institucional, a evasão de talentos e a necessidade de formação em IA ética e responsável. A consolidação de uma política nacional de formação em IA, articulada entre educação, pesquisa e setor produtivo, é essencial para maximizar as oportunidades e mitigar as ameaças. Essa política deve combinar investimento em infraestrutura HPC, modernização curricular, parcerias internacionais e incentivos à inovação, garantindo a formação contínua e de qualidade de profissionais capazes de sustentar o desenvolvimento tecnológico e soberano do país.

Matriz SWOT – Painel 6

Forças	Fraquezas
- Infraestrutura nacional robusta de HPC, com +15 supercomputadores (Pegasus, Santos Dumont, Gaia etc.) - Universidades pioneiras (UnB, UFG, UFMG) com cursos e metodologias inovadoras baseadas em projetos. - Programas consolidados de fomento (CNPq, CAPES, RHAE, Acadêmico para Inovação). - Parcerias estratégicas com o setor produtivo (Petrobras, SENAI, CEIA, ICTs). - Centros de excelência como o CEIA, com +1.000 pesquisadores e 70 projetos ativos.	- Defasagem de formação em fundamentos técnicos (computação, matemática, estatística, microeletrônica). - Fragmentação institucional e baixa integração entre universidades, centros de pesquisa e empresas. - Formação incipiente em IA responsável (ética, regulação e impactos sociais). - Fuga de talentos para o exterior e baixa retenção de profissionais qualificados. - Falta de sistematização e coordenação das ações de capacitação do PBIA.
Oportunidades	Ameaças
- Expansão da educação em IA em todos os níveis (fundamental, médio e superior). - Criação de uma Política Nacional de Formação em IA, articulando MEC, MCTI, CNPq e setor privado. - Fortalecimento da cooperação internacional (América Latina e Sul Global) e criação de uma Federação de Supercomputadores. - Sandboxes regulatórios e inovação segura em ambientes educacionais e industriais. - Alta demanda global e nacional por profissionais de IA e dados.	- Avanço acelerado da IA e rápida obsolescência de currículos e tecnologias. - Dependência de tecnologias e APIs estrangeiras, com risco à soberania tecnológica. - Desigualdade digital e educacional que limita o acesso à formação em IA. - Baixa absorção de talentos pela indústria nacional e risco de “desemprego qualificado”. - Restrições orçamentárias e instabilidade fiscal que ameaçam programas de fomento e bolsas.

Figura 25. Matriz SWOT – Painel 6

Painel 7: Conexões que sustentam a IA: Onde estão as alianças/colaborações estratégicas nacionais para a construção de uma Pilha de IA soberana?

Este painel debate as colaborações nacionais e internacionais que estão contribuindo para a construção da pilha de IA. A partir de exemplos concretos de cooperação são analisadas articulações entre academia, centros de pesquisas e governo. O objetivo é identificar os arranjos já existentes e as oportunidades futuras de colaboração que possam fortalecer a autonomia tecnológica do Brasil em todas as camadas da pilha, da infraestrutura ao nível de aplicação.

Moderadora: Beatriz Vasconcelos (*Secretária Adjunta da Secretaria de Articulação e Monitoramento da Casa Civil*)

Carla Osthoff

Tecnologista sênior e Professora do Laboratório Nacional de Computação Científica (LNCC)

O CENAPAD integra o Sistema Nacional de Processamento de Alto Desempenho (SINAPAD). Dentre os centros que compõem o SINAPAD, o LNCC é o único vinculado diretamente ao Ministério da Ciência, Tecnologia e Inovação (MCTI). Conforme apresentado pelo diretor do LNCC, o LNCC abriga o supercomputador Santos Dumont, financiado pela Petrobrás e o mais potente do SINAPAD. E está como o segundo mais potente da América Latina na lista Top 500, sendo o primeiro colocado entre os acadêmicos da região e o número 1 na categoria Green. O Santos Dumont é disponibilizado gratuitamente à comunidade de pesquisa, um aspecto de grande relevância, apoiando atualmente cerca de 170 projetos de pesquisa, distribuídos em 12 estados brasileiros e atendendo cerca de 2 mil usuários ativos. Desde 2016, as áreas de pesquisa contempladas abrangem praticamente todos os campos científicos que demandam computação de alto desempenho, conforme Figura 26.

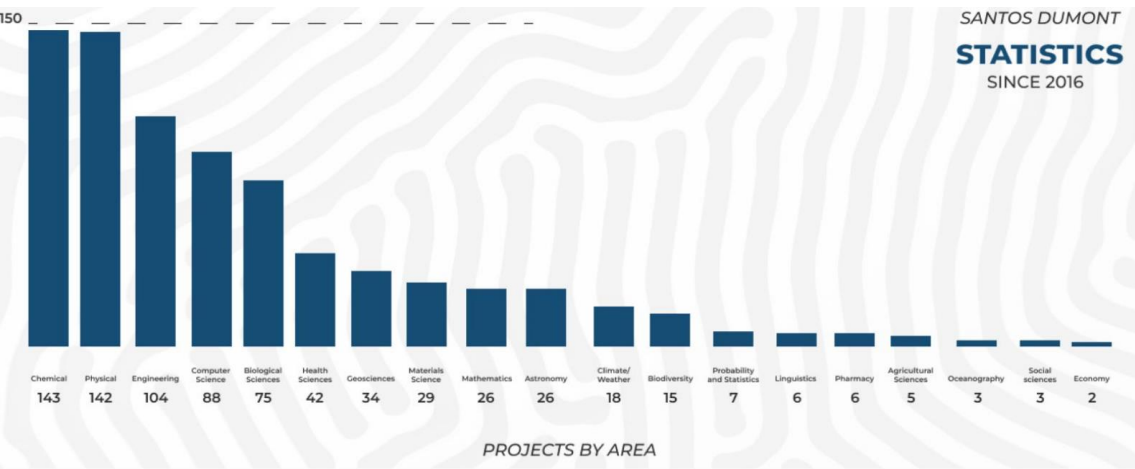


Figura 26. Áreas de pesquisa dos projetos realizados com apoio do supercomputador Santos Dumont desde 2016

Nos últimos anos, temos observado um grande crescimento no uso de aplicações de IA, que vêm se inserindo em praticamente todas as áreas de pesquisa. A IA generativa, em especial, tem se expandido de forma significativa, alcançando diversos campos do conhecimento. Esse movimento é corroborado pelo relatório [The Brazilian Landscape of Science, Technology and Innovation in Artificial Intelligence](#) (CGEE, 2025), que evidencia o aumento do número de pesquisas voltadas à área de inteligência artificial em diferentes setores e domínios científicos (Figura 27).

The Brazilian Landscape of Science, Technology, and Innovation in Artificial Intelligence

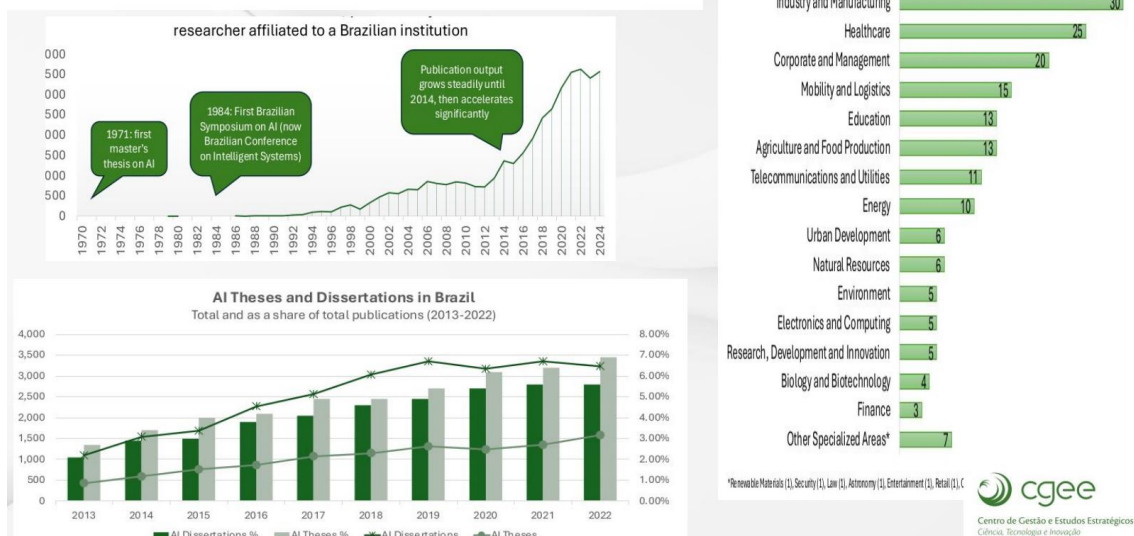


Figura 27. Pesquisa brasileira em inteligência artificial

Atualmente, o supercomputador Santos Dumont abriga uma ampla variedade de aplicações voltadas à inteligência artificial, e não apenas as aplicações tradicionais de computação de alto desempenho (HPC), como aquelas voltadas à modelagem de clima, tempo e outras áreas que demandam grande capacidade computacional. Carla apresentou a pilha de *software* para IA no contexto de HPC, mencionada anteriormente pelo professor Edson Borin. Essa pilha compreende todas as camadas de *software* presentes no supercomputador, que possibilitam a execução das aplicações, e a mesma estrutura é utilizada na infraestrutura atualmente em operação no LNCC. Um aspecto relevante destacado por Carla é que todas as camadas dessa pilha de *software* são baseadas em código aberto (*open source*). No entanto, as distribuições utilizadas são customizadas por cada fabricante, o que é necessário para garantir suporte técnico contínuo (24x7), segurança operacional e controle total sobre o sistema — fatores que explicam por que essas distribuições acabam sendo fechadas.

A proposta apresentada aqui consiste na criação de um consórcio voltado à substituição gradual dessas camadas, que são abertas, porém de distribuição fechada, por

distribuições totalmente abertas. Para isso, é essencial a implementação de um sistema experimental de menor porte, um mini supercomputador de pesquisa, que permita realizar experimentos e modificações nas camadas de *software* sem interferir nas aplicações em produção executadas no Santos Dumont. Esse sistema paralelo deverá replicar as camadas do supercomputador em operação, permitindo substituir progressivamente cada camada por sua versão aberta equivalente, até que se alcance um desempenho comparável ao ambiente de produção. Essa é a ideia central do projeto, conforme a Figura 28.

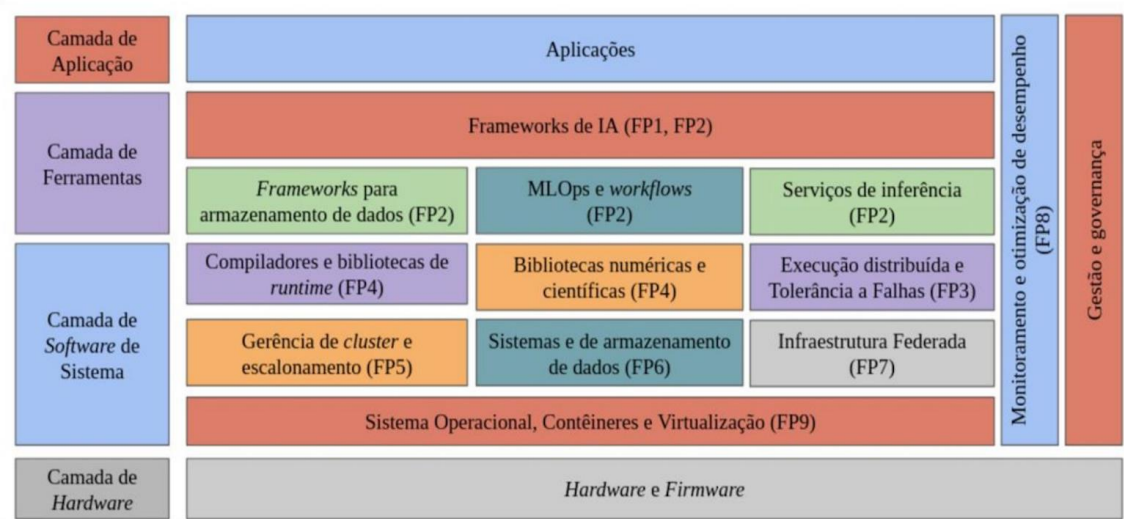


Figura 28. Pilha de *software* para IA no contexto de HPC: possui ‘distribuição proprietária’ baseada em open source

Há um mapeamento das instituições que se destacam em excelência de pesquisa em cada uma das camadas da pilha de *software* e que podem atuar no consórcio proposto. As instituições envolvidas são: LNCC, UNICAMP, UFRGS, UFRN, UnB, UFF, UFSC e CEFET/RJ.

- **Sistema de Armazenamento de Dados (FP6):** Há cerca de dez anos, quando o primeiro supercomputador foi instalado, estabeleceu-se uma parceria com a UFRGS, uma das instituições de referência na área. O grupo liderado pelo professor Philippe Navaux tem colaborado desde então no estudo e aprimoramento de sistemas de arquivos paralelos. Apesar do avanço obtido ao longo dessa década, as possibilidades de evolução ainda são limitadas pelo ambiente de produção.
- **Gerência de Cluster e Escalonamento (FP5):** O mesmo grupo também se destaca na área de gerência de *clusters* e escalonamento. Sob a liderança do professor Arthur Lorenzon, têm sido desenvolvidos estudos e publicações relevantes nessa temática. No entanto, o trabalho permanece limitado devido à impossibilidade de realizar alterações diretas no ambiente de produção. Além disso, há uma colaboração com o grupo do CEFET/RJ, liderado por Diego

Gavardo, que atua em projetos de grids computacionais e desenvolvimento de soluções voltadas à gestão e distribuição de recursos em ambientes HPC, também enfrentando restrições similares.

- **Frameworks de Inteligência Artificial (FP1 e FP2):** Em parceria com o grupo da UNICAMP/HIAAC, coordenado pelo professor Edson Borin, são conduzidos estudos sobre *frameworks* de IA voltados ao treinamento de modelos fundacionais para séries temporais e imagens, bem como ao desenvolvimento de modelos complexos em sistemas HPC. O grupo também atua na criação de modelos de aprendizado profundo voltados à segmentação e predição de atributos sísmicos, além de modelos para classificação e detecção de anomalias em séries temporais.
- **Compiladores e Bibliotecas de Runtime (FP4):** Os professores Guido Araújo e Hervé Yviquel, do Instituto de Computação da UNICAMP, lideram pesquisas recentes em compiladores voltados para IA, como TVM (Apache) e GLOW (Meta). As atividades incluem geração de código e otimizações para a arquitetura RISC-V, além de larga experiência em aceleração de aplicações em GPU. O grupo também desenvolve técnicas de otimização para convoluções, *benchmarks* de desempenho e parcerias com empresas como a Petrobras, visando à execução eficiente de modelos de aprendizado de máquina em *clusters*.
- **Execução Distribuída e Tolerância a Falhas (FP3):** O grupo dos professores Guido Araújo e Sandro Rigo (UNICAMP) possui experiência consolidada em *runtimes* e otimização de aplicações de IA e HPC em *clusters* de grande escala com GPUs. Contribuem para projetos de *software open source* estratégicos como LLVM, OpenMP e Spack, fortalecendo as pilhas de *software* de supercomputadores. Mantêm ainda colaborações com empresas de referência (Petrobras, AMD) e centros internacionais de excelência (ANL, BSC, INSA/França). O grupo do professor Samuel Xavier, da UFRN, também integra essa frente, com experiência em balanceamento dinâmico de carga, tolerância a falhas e escalonamento adaptativo em sistemas distribuídos. O grupo desenvolve bibliotecas resilientes com foco em confiabilidade de execução, contando com a infraestrutura do NPAD/UFRN para testes e validação de protótipos em *clusters* e supercomputadores.
- **Sistema Operacional, Contêineres e Virtualização (FP9):** Sob a liderança do professor Samuel Xavier (UFRN), com experiência em sistemas operacionais e arquiteturas tolerantes a falhas, com pesquisa em kernels experimentais e execução segura. O grupo possui competência técnica para o desenvolvimento de um kernel em Rust compatível com Linux, buscando aliar segurança, privacidade e interoperabilidade. Essa atividade requer uma infraestrutura e máquina independente, isolada da máquina de produção.
- **Monitoramento e Otimização de Desempenho (FP8):** O grupo da UFRN, também sob coordenação do professor Samuel Xavier, realiza pesquisas em

monitoramento e otimização de desempenho que abrangem diversas camadas da pilha. As iniciativas incluem instrumentação de baixo *overhead*, avaliação reprodutível de desempenho e consumo energético, modelagem analítica para identificação de gargalos em CPU, GPU e memória, além de estratégias de autoajuste (*auto-tuning*) e visualização de escalabilidade.

- **Gestão e Governança:** A segurança foi identificada como uma camada transversal a todo o sistema. Nesse contexto, há colaboração com o Instituto Militar de Engenharia (IME) e diálogo com o Instituto Eldorado, representado por Cibele Coscione e Matheus Pierre, para o desenvolvimento de soluções integradas de segurança e governança.
- **Infraestrutura Federada (FP7):** A camada de infraestrutura em nuvem é particularmente sensível. Conforme apresentado anteriormente, há interesse do governo de que o projeto de IA seja apoiado por diversas máquinas do sistema SINAPAD, que deverão operar de forma colaborativa. No âmbito do PBIA, busca-se fomentar também a cooperação com países da América Latina. O LNCC mantém um memorando de entendimento com a SCALAC, associação latino-americana de computação de alto desempenho, para o desenvolvimento de uma rede federada de compartilhamento de recursos e pesquisa. Além disso, o LNCC integra o BRICS IA com HPC, em diálogo com países do grupo, e estamos conversando sobre a criação de uma nuvem federada conjunta. Participam ainda a UnB (professora Alba Melo), a UFF (professora Lúcia Drummond) e a UFSC (professor Márcio Castro), que possuem ampla experiência em projetos internacionais de *cloud computing*. A RNP, representada por José Ferreira Rezende, atua como elo de integração entre universidades e instituições de pesquisa, garantindo conectividade e segurança. É fundamental que o projeto se mantenha aberto e interoperável, assegurando compatibilidade entre todas as comunidades envolvidas.

O projeto proposto envolve aproximadamente 200 profissionais acadêmicos e baseia-se na substituição gradual das camadas proprietárias por equivalentes de código aberto. Carla ressaltou que a iniciativa não contempla integralmente todas as necessidades da pilha de IA, motivo pelo qual foi realizado, em parceria com a Sociedade Brasileira de Computação (SBC), um levantamento de grupos de pesquisa em HPC em 75 instituições brasileiras (Figura 29). O objetivo é identificar os projetos em andamento em cada camada da pilha de IA, gerando subsídios para órgãos de fomento e apoio à formulação de editais voltados ao fortalecimento da pesquisa nacional em HPC e IA.

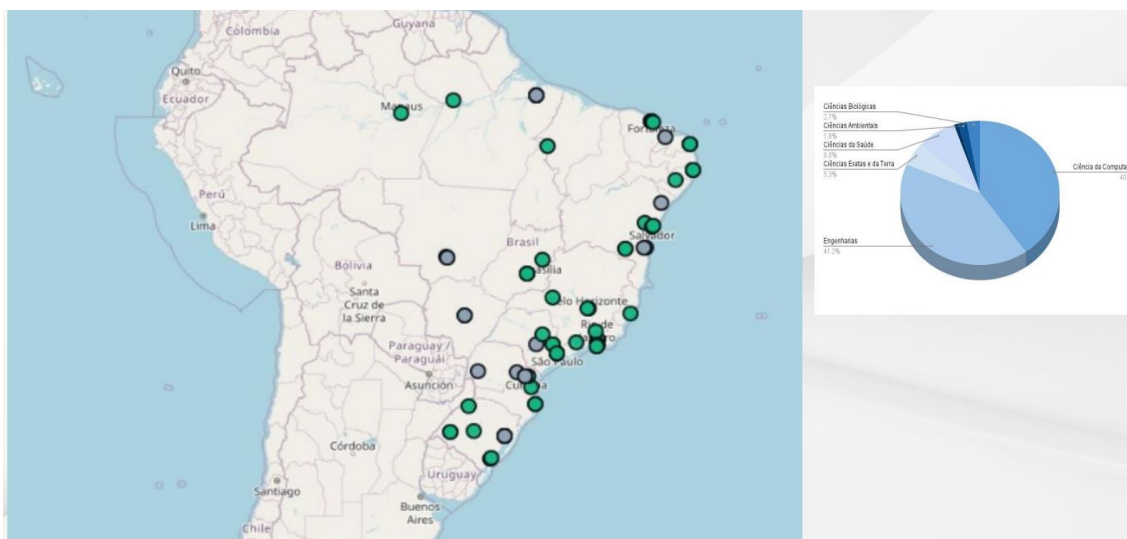


Figura 29. Levantamento de grupos de pesquisa em HPC (SBC)

Por fim, o universo da IA é muito mais amplo, abrangendo além da computação de alto desempenho, sistemas embarcados e diversas outras aplicações que demandam abordagens complementares. Durante as reuniões e debates realizados, ficou o entendimento de que os países que atualmente dominam essas tecnologias iniciaram seus esforços há mais de uma década.

Deste modo, é fundamental que o Brasil estabeleça um *roadmap*, uma estratégia, com metas e prazos bem definidos, que permita orientar o desenvolvimento tecnológico de forma consistente ao longo do tempo. Mesmo no caso desse projeto da pilha de HPC, foi pensado em um horizonte de três anos, mas há consenso de que a trajetória necessária é mais longa e complexa. Nesse contexto, destaca-se a necessidade de uma estratégia maior, que defina objetivos e prioridades claras. Para tal, algumas questões devem ser endereçadas, como: deve-se priorizar a infraestrutura de HPC ou o desenvolvimento de produtos embarcados? Quais tecnologias o país pretende dominar, e em que prazos? O que queremos dominar, em quanto tempo e em qual ponto? Essas são perguntas muito complexas e precisam ser debatidas coletivamente.

Deybson de S. Cipriano

Presidente da Confederação ASSESPRO

A ASSESPRO é a primeira entidade associativa empresarial de tecnologia do país, fundada em 1976, com atuação na defesa do setor produtivo de tecnologia no Brasil. À época da fundação, o foco estava no processamento de dados, o que se reflete na própria sigla da instituição. A ASSESPRO reúne mais de três mil empresas associadas, das quais entre 90% e 95% são micro, pequenas e médias empresas (MPMEs), presentes em todas

as unidades da federação. A missão principal da entidade é manter a atividade econômica e o fortalecimento dessas empresas.

Entre as ações já discutidas no âmbito do associativismo, destaca-se a necessidade de uma governança colaborativa em inteligência artificial. O PBIA já representa um passo importante nessa direção, mas é preciso ampliar e fortalecer essa governança. E o setor produtivo deve ser visto como uma ferramenta e um instrumento que está em inteira disposição. A ASSESPRO conta com cases de sucesso de suas empresas associadas tanto no Brasil quanto no exterior.

Um dos principais desafios identificados é a capacitação e retenção de talentos, uma dificuldade global, mas particularmente acentuada no Brasil devido à concorrência internacional, à valorização de moedas estrangeiras e às condições diferenciadas de trabalho. Embora o trabalho remoto tenha se popularizado após a pandemia, essa prática já é realidade no setor tecnológico há mais de duas décadas.

Em 1998, Deybson integrou uma comissão internacional de inteligência artificial, quando os temas centrais eram *big data*, interfaces amigáveis e superprocessadores. Na época, acreditava-se que a IA alcançaria um nível comparável ao da inteligência humana, que não se concretizou, levando a um período de esfriamento do interesse global. Entretanto, países emergentes que mantiveram investimentos contínuos em tecnologia conseguiram resultados expressivos e conseguiram reter mão de obra qualificada. No caso brasileiro, a dimensão continental do país cria tanto oportunidades quanto desafios. Embora o Brasil possa ser visto como um “paraíso profissional” para talentos estrangeiros de algumas áreas, muitos profissionais brasileiros ainda buscam oportunidades fora do país. Por isso, é essencial criar linhas de crédito e incentivos fiscais específicos para a IA, pauta já defendida pela ASSESPRO em diversos fóruns e pleitos.

Entretanto, o investimento em tecnologia de ponta é caro, especialmente nas camadas mais avançadas das pilhas de IA, e o setor produtivo sofre com a escassez de crédito e incentivos adequados. Outra questão crucial é evitar entraves na regulamentação, pois um ambiente regulatório rígido pode gerar dependência tecnológica e reduzir a competitividade nacional. O momento atual é propício para o desenvolvimento de tecnologias e *frameworks* próprios, a fim de garantir soberania e independência tecnológica.

A conexão entre academia e empresas é também um pilar essencial. A ASSESPRO mantém parcerias com universidades de referência, como UNICAMP e UFPE, aproximando as demandas do mercado das pesquisas acadêmicas. Essa interação é fundamental para alinhar o desenvolvimento tecnológico às necessidades reais da sociedade. Também é importante garantir infraestrutura computacional compartilhada de IA, permitindo que instituições e empresas com menor capacidade de investimento tenham acesso a recursos de alto desempenho, fomentando competitividade e inovação.

Dentre as contribuições da ASSESPRO para o fortalecimento da inteligência artificial estão a disseminação do conhecimento, oferta de informações atualizadas sobre as melhores práticas e tendências corporativas, promoção do debate em eventos e fóruns, representação do setor junto a órgãos governamentais e reguladores para favorecer um ambiente de desenvolvimento ético e inovador no Brasil, e fomento ao *networking*, conectando empresas e *startups* que buscam soluções em IA com aquelas que já são referência na área.

Por fim, cabe uma provocação: “O que queremos ser nessa corrida — os *Flintstones* ou os *Jetsons*?” A reflexão aponta para o papel que o Brasil deve assumir nos próximos anos. Ademais, destacou que o PBIA será um instrumento importante para conectar os diversos atores e o setor produtivo. A ASSESPRO mantém-se à disposição para parcerias público-privadas e investimentos, reforçando que, diante da timidez dos investimentos públicos, o setor privado está pronto para cooperar, seja por meio de acordos técnicos ou institucionais. Esse é o caminho para superar discussões fragmentadas e avançar de forma conjunta, pois, como ressaltado, a academia é fundamental, sem ciência, não há progresso.

Kiko Reis

Diretor da Magalu Cloud

A Magalu Cloud está em operação desde 2021, com atuação em duas grandes regiões do Brasil, Sudeste e Nordeste, totalizando 5 data centers, com capacidade de processamento de cerca de 10 mil cores por mês, com centenas de milhares de cores ativos e centenas de GPUs em funcionamento. A estrutura apresentada atende a aproximadamente 50% da demanda de nuvem pública utilizada pelo Magazine Luiza, que está entre os 10 maiores usuários de nuvem pública do país. E, desde 2023, cerca de mil empresas externas ao Magalu passaram a utilizar seus serviços.

O sistema operacional é, em essência, um grande integrador da pilha, conectando o *hardware*, o *software* básico e as diversas aplicações que se desenvolvem sobre essa base. Entretanto, no Brasil o lado da oferta apresenta maior destaque quando se fala de pilha, considerando a oferta de *data centers* e a capacidade de estímulo à criação e aumento de competitividade do setor. Em outros termos, o olhar está voltado da base para o topo, analisando as necessidades de infraestrutura, qual *software* básico será utilizado e como atrair profissionais experientes para montar uma pilha competente. Entretanto, observa-se a falta de demanda brasileira, visto que os grandes sistemas de inteligência artificial do mundo estão hoje nas mãos de 10 grandes empresas globais, em que 8 delas comprem mais de 1 gigawatt de capacidade computacional, sendo que nenhuma delas é brasileira. E é importante mencionar que nos anos 2000, durante a transição para a internet, existiam centenas de empresas brasileiras criando produtos voltados para o público nacional, desenvolvendo soluções próprias, porém esse

processo foi descontinuado. Desta forma, se quisermos desenvolver uma pilha competente, é necessário resolver o problema da demanda.

Grandes plataformas como Instagram, TikTok, ChatGPT e Google são altamente consumidas pela sociedade brasileira, mas apresentam baixo envolvimento de profissionais brasileiros no desenvolvimento destas plataformas. Apesar do alto potencial do mercado consumidor, canaliza-se pouco a demanda para o desenvolvimento interno. Como exemplo, a OpenAI oferece um produto B2C, voltado diretamente para o usuário, que ao obter respostas da IA generativa, arrasta toda uma pilha de cima a baixo. A partir desta aplicação sugeriram fábricas de GPUs, empresas de *chips* e *data centers* multigigawatt nos Estados Unidos, que não configuram o local mais adequado para esse tipo de infraestrutura. Porém, como havia um produto de grande sucesso puxando a demanda, toda a pilha se reorganizou para sustentá-lo.

Quando a Magalu Cloud começou a ser construída, foram prospectadas parcerias com diversos institutos brasileiros para rodar projetos e promover integração. No entanto, perceberam que, se não for possível alinhar a demanda, corre-se o risco de construir uma pilha que não alcance o impacto desejado. A Magalu Cloud quer ser protagonista e o Brasil tem condições, dada as condições atuais, como o fato de ser uma das maiores economias do mundo. Mas, para isso, é necessário adotar uma estratégia diferente do que se fez no passado, focado em muitas apostas pequenas e isoladas. No momento, se faz necessário apostas estruturantes, capazes de fomentar grandes agregadores de demanda, porque são eles que determinarão o sucesso de uma pilha tecnológica.

Em suma, o problema do Brasil não é falta de recursos, mas sim problemas na demanda, porque faltam aplicações sendo construídas no Brasil. Como exemplo, uma tarefa simples, como indicar o endereço do SERPRO no convite do seminário, utiliza-se o pin do Google Maps, e não de uma *startup*. Desta forma, o verdadeiro problema é agregar demanda: se ela existir, os recursos e as soluções virão naturalmente, porque há muita gente capacitada e disposta a construir soluções

Renan Gaya

Diretor do Departamento de Infraestrutura de Dados Públicos (MGI)

O olhar sobre a demanda é, de fato, muito importante, mas Renan acredita que o natural no mercado é se guiar por qualidade e preço. No setor de tecnologia, há uma concentração muito definida de qualidade com preços competitivos, e essa combinação cria uma inércia natural: mesmo quem trabalha diretamente com o tema tende a utilizar aquilo que oferece melhor qualidade pelo menor custo. Por isso, a soberania não virá por inércia. Já estamos atrasados em algumas camadas dessas pilhas, e isso significa que a soberania precisa ser uma política pública, porque ela tem custo. Produtos soberanos podem ser mais caros, e por isso talvez não sejam naturalmente aceitos pelo mercado. O Brasil é, sim, um grande demandante de tecnologia, não apenas pelo governo, mas pela

própria população. Somos um país altamente digitalizado e estamos entre os principais consumidores de internet do mundo, no top 3 a 5 em tempo de uso. Temos uma demanda enorme, mas que segue a inércia da qualidade e do baixo custo, o que reforça a necessidade de políticas públicas que sustentem a soberania digital. Há muitos anos já existem iniciativas relacionadas à soberania tecnológica, mas vivemos hoje um momento geopolítico novo, em que riscos antes distantes começam a se aproximar. Situações como as sanções internacionais, os conflitos no Oriente Médio ou as tensões comerciais dos Estados Unidos mostram que as tecnologias também são instrumentos de poder geopolítico. O que antes parecia um cenário hipotético agora começa a se concretizar.

Nesse contexto, governança e articulação tornam-se elementos fundamentais. Precisamos fazer o “*match*” entre demanda e oferta. As duas existem no país, mas ainda não se conectam plenamente. Às vezes, a oferta nacional não é tão competitiva ou não oferece todas as funcionalidades das grandes empresas globais, mas ela já existe, e é preciso fazer essa ponte funcionar. E esse alinhamento deve ocorrer em múltiplas camadas, entre academia, pesquisa, mercado e governo. Nos últimos cinco anos, o Brasil avançou em instrumentos que permitem ao Estado usar seu poder de compra para fomentar inovação, como o Marco Legal das Startups, a nova Lei de Licitações e a Lei de Governo Digital. Ademais, Serpro e Dataprev têm papel estratégico nesse processo de digitalização governamental.

Apesar dos avanços normativos, esses instrumentos ainda são subutilizados. Em dez anos foram realizadas menos de 100 encomendas tecnológicas, frente a mais de 20 mil contratações por ano. Ou seja, as ferramentas existem, mas seu potencial permanece pouco explorado. A soberania tecnológica, portanto, depende de decisão política e investimento deliberado. O Brasil possui laboratórios e hubs de inovação distribuídos pelo território e destaca-se na América Latina pelo ecossistema de startups. No entanto, o universo das govtechs ainda é reduzido, com 500 mapeadas. A limitação se deve à escassez de ações práticas e estímulos consistentes. Para tal, é necessário ampliar conexões, criar condições reais para o encontro entre oferta e demanda e investir em governança para consolidar essas pontes.

Há, contudo, avanços substantivos nos últimos anos. O MCTI, FINEP, MGI e ENAP têm atuado de forma integrada para identificar desafios reais da administração pública e lançar editais voltados a startups nacionais. Com recursos do FNDCT, são 25 projetos em andamento, com entregas previstas ainda neste ano, em áreas como saúde, agricultura e meio ambiente. O esforço é relevante, mas o potencial de expansão é expressivo. É preciso fortalecer o uso dos instrumentos existentes para conectar as etapas da cadeia de inovação. Em 2024 o governo planejou investir cerca de R\$ 15 bilhões em infraestrutura e serviços de *software*, o que demonstra a centralidade do setor.

Pesquisa recente com órgãos públicos que utilizam inteligência artificial mostrou que o principal recurso usado para o desenvolvimento de soluções é a equipe interna de

servidores, seguida por parcerias com universidades e, por último, pela iniciativa privada. Isso evidencia a base técnica robusta do setor público, especialmente em instituições como a Receita Federal, o Ministério da Fazenda, o Banco Central e o Tesouro Nacional. As universidades atuam como parceiras fundamentais, e as compras públicas, mesmo não voltadas à inovação, já estimulam o desenvolvimento tecnológico nacional.

Empresas públicas como Serpro e Dataprev desempenham papel estratégico na soberania tecnológica. Muitos sistemas governamentais são complexos e específicos, como o gov.br, com 170 milhões de usuários, o CadÚnico, com 90 milhões, e a folha de pagamento do INSS, que atende 40 milhões de pessoas por mês. Esses sistemas demandam soluções sob medida, reforçando a necessidade de empresas públicas estáveis e capacitadas, já que licitações periódicas podem comprometer a continuidade de serviços críticos. Essas empresas têm promovido o retorno do conhecimento e da operação tecnológica ao país, adotando soluções *open source* e infraestrutura própria.

Além disso, há crescente integração entre universidades, *startups* e empresas públicas, em articulação com MCTI e FINEP, para fortalecer o ecossistema de inovação. Ainda há amplo espaço para aprimorar o uso dos instrumentos disponíveis e avançar nas camadas mais profundas da pilha tecnológica, do *hardware* às aplicações finais. As compras públicas configuram um canal estratégico dessa integração. Contudo, é necessário intensificar o uso dos mecanismos existentes para que impactem efetivamente o desenvolvimento da inteligência artificial e consolidem a soberania tecnológica nacional.

Teresa Ludermir

Professora (UFPE) e Diretora do Centro de Pesquisa Aplicada em IA para Segurança Cibernética

A professora Teresa Ludermir destaca que uma das conexões mais relevantes para o desenvolvimento da inteligência artificial é a integração com a academia. Segundo ela, não é possível discutir soberania tecnológica sem reconhecer o papel central da ciência de ponta, produzida por pesquisadores altamente qualificados, com sólida formação e dedicação contínua à pesquisa científica. Em sua primeira intervenção, Teresa enfatiza que a soberania requer conexões com o setor privado, o governo e a sociedade civil, mas que sem ciência não há soberania. Em toda a cadeia, da infraestrutura às aplicações, as universidades e centros de pesquisa exercem papel essencial, pois são nelas que se realiza o avanço científico que sustenta o progresso tecnológico.

A professora reforça que alcançar autonomia em inteligência artificial depende de uma coordenação multissetorial, que envolva governo, academia e sociedade civil na construção de uma pilha nacional de IA. Esse processo, contudo, exige financiamento contínuo e planejamento de longo prazo. Sem uma visão estruturada para os próximos 10 ou 15 anos, afirma, “estaremos apenas improvisando”. Investimentos pontuais e descontínuos não produzem resultados duradouros. A formação de um pesquisador leva

décadas de estudo e prática, e, ao atingir maturidade, esse profissional passa a formar novas gerações, multiplicando o conhecimento acumulado.

As universidades são, portanto, o núcleo da produção científica de excelência no Brasil. Mesmo diante de crises e restrições orçamentárias, a ciência nacional resistiu, sustentada pelo empenho de seus pesquisadores, que seguiram atuando, muitas vezes em condições adversas. A contribuição da academia é decisiva para o desenvolvimento de uma pilha nacional de IA, em que a ciência é a base da soberania tecnológica, e nela é preciso investir para garantir um avanço autônomo e sustentável.

O mapeamento da pilha de IA no Brasil (2025) mostra avanços, mas também limitações. Teresa reconhece que, se conduzido com responsabilidade, o país tem potencial para escalar seus modelos de IA. Contudo, ressalta que os recursos financeiros do PBIA destinados à academia têm sido insuficientes. Segundo ela, esses recursos não estão chegando efetivamente às universidades e grupos de pesquisa. Coordenadora do primeiro INCT de Inteligência Artificial Responsável do país, Teresa lembra que o projeto não foi financiado pelo PBIA, pois já havia sido aprovado anteriormente.

Embora o edital dos INCTs tenha priorizado a área de IA e aprovado mais sete centros, houve cortes orçamentários significativos, dificultando o repasse de verbas para infraestrutura e equipamentos, especialmente nas regiões Norte e Nordeste. Ainda assim, ela considera o PBIA uma iniciativa notável, afirmando ter sido “uma das maiores alegrias de sua vida acadêmica nos últimos anos” ver o Brasil instituir um plano nacional de inteligência artificial. Se bem executado, o PBIA poderá impulsionar fortemente o desenvolvimento científico e tecnológico do país.

Teresa observa que grande parte dos pesquisadores brasileiros possui experiência internacional, o que permite aprender com modelos bem-sucedidos de outros países e adaptá-los à realidade nacional. Defende uma gestão tripartite, que envolva governo, academia e sociedade civil, garantindo infraestrutura acessível e equitativa e promovendo uma inteligência artificial responsável. Como exemplo, cita o Mila, no Canadá, onde há forte integração entre governo e pesquisa acadêmica, além de cátedras de IA que incentivam jovens pesquisadores a permanecerem na área. Segundo ela, o Brasil precisa de iniciativas semelhantes para assegurar a renovação das novas gerações e a continuidade da soberania tecnológica (Figura 30).

País	Institutos/ Iniciativas	Experiências e Características Principais
França	INRIA/3IA	Hub nacional com centros temáticos; modelo de cooperação público-privada; definição de metas de médio e longo prazo (5 a 10 anos).
Alemanha	DFKI	Parceria público-privada permanente como “ponte” entre universidades e indústria; articulação em rede com Fraunhofer, Max Planck e universidades.
Canadá	CIFAR/Mila/ Vector/Amii	Estrutura baseada em cátedras de IA e infraestrutura de computação dedicada (PAICE).

Estados Unidos	CHIPS/NSF/NIST	Investimento abrangente (<i>end-to-end</i>), da fabricação de chips à definição de padrões; criação de institutos temáticos.
China	—	Planejamento de ciclo completo (" <i>full-stack</i> "); <i>datacenters</i> em rede; estímulo a campeões nacionais; definição de metas por fase.

Figura 30. Lições internacionais para adaptação ao Brasil

Para garantir a governança e soberania dos dados, a professora propõe a criação de um Conselho Nacional de Inteligência Artificial, que vem sendo chamado de SIN-IA (ou CONAPIA, sugestão própria). O órgão teria funções de governança e execução, estruturado sob o modelo tripartite, com participação do governo, da academia e da sociedade civil. Entre suas atribuições centrais estariam a soberania de rede e de dados, equilibrando o acesso amplo à informação com a proteção de dados estratégicos, fundamentais para a autonomia nacional. O conselho também seria responsável por articular os grandes desafios científicos e tecnológicos do país, assegurando que as decisões estratégicas contemplem a dimensão acadêmica e científica. A academia deve estar no centro dessas discussões, contribuindo com pesquisa, formação e conhecimento que sustentem o avanço tecnológico nacional.

PROPOSTA - CONAPIA (Governança e execução)

- Conselho tripartite: Governo (MCTI/MEC/MDIC/ANPD), Universidades/ICTs/FAPs/EMBRAPII, Indústria/Sociedade.
- Compute.br: fila única nacional com SLA/transparência e quotas regionais.
- Data Trusts BR: trusts setoriais (saúde, justiça, clima, educação) com LGPD/FAIR/red teaming.
- Model Zoo PT-BR: fundo anual para LLMs e multimodal + avaliações públicas e governança de pesos/datasets.
- Grand Challenges: missões anuais com pilotos → escala (SUS, agro carbono-zero, ciberdefesa, educação continuada).

Figura 31. Sugestão de Conselho de Governança para IA

A professora considera que o decreto do MCTI provavelmente trará mais detalhes sobre essa estrutura, mas ressalta que suas propostas partem de modelos já consolidados, como o CGI.br, que atua de forma bem-sucedida na governança multissetorial da internet. Assim, defende a adoção de um modelo colaborativo semelhante para a governança da inteligência artificial, com interação equilibrada entre governo, academia e sociedade civil.

Outro ponto enfatizado é a necessidade de inclusão regional, com a criação de *hubs* regionais, especialmente no Norte e no Nordeste, voltados à formação de pessoas, hospedagem de nós de dados e fortalecimento de capacidades locais. Ressalta, contudo, que os investimentos nessas regiões ainda são insuficientes, e que as ações precisam ganhar maior escala e efetividade. E acredita que as universidades estão preparadas para contribuir, tanto em infraestrutura quanto em pesquisa e formação de talentos. Por fim, sugere a criação de uma empresa pública nacional de inteligência artificial, com papel gestor e operacional, comparável em importância à Petrobras, mas voltada à IA. Essa instituição teria visão abrangente da pilha tecnológica, coordenando esforços desde o *hardware* até as aplicações, consolidando assim uma estratégia nacional robusta para o setor.

Considerações finais

O Painel 7 destacou a importância das alianças e colaborações estratégicas nacionais para a construção de uma pilha de inteligência artificial soberana no Brasil, evidenciando a necessidade de coordenação entre governo, academia, centros de pesquisa e setor produtivo. As discussões revelaram avanços concretos, como a consolidação da infraestrutura de HPC e o protagonismo do LNCC com o supercomputador Santos Dumont, além do engajamento de instituições de excelência que já colaboram em diferentes camadas da pilha de *software*. Também ficou evidente o papel central da academia na produção científica e na formação de talentos, essencial para sustentar a soberania tecnológica do país.

151

Dentre as principais propostas se destacam:

- Governança nacional de IA (tripartite), com a instituição de um Conselho Nacional de IA, de natureza executiva e deliberativa; criação de câmaras técnicas permanentes; e inspiração no modelo CGI.br como referência de processo decisório e transparência.
- Consórcio *open source* para a pilha de HPC/IA, com financiamento do mini-Santo Dumont para substituição gradual das camadas com distribuição fechada por distribuições 100% abertas; formalização de consórcio entre instituições; e *roadmap* técnico por frentes (apresentação da Carla Osthoff).
- Compute.br e infraestrutura federada por meio de operacionalizada de fila única nacional e federação (SINAPAD, nuvens acadêmicas e provedores nacionais via RNP); marketplace público de recursos (CPU/GPU e armazenamento) com preços de referência e crédito de uso para MPMEs, govtechs e grupos emergentes.
- Política de compras públicas pro-inovação, com ampliação do uso de encomenda tecnológica, criação de catálogo de soluções e lotes para MPMEs;

inclusão de critérios de avaliação: desempenho, segurança, abertura (portabilidade, padrões) e pegada de carbono.

- Capital humano e retenção de talentos, por meio de cátedras de IA (modelo Mila); bolsas de longa duração e *visiting-engineers* indústria-academia; crédito, incentivos fiscais para IA (P&D e aceleração de MPMEs); e trilhas nacionais de engenharia de plataformas de IA.
- Hubs regionais e inclusão com implantação de hubs no Norte e Nordeste com nós de dados, GPU compartilhadas e programas de formação e reserva de quotas regionais na fila Compute.br.

De modo geral, o painel reforçou que a soberania em IA depende de planejamento de longo prazo, investimentos contínuos e coordenação multissetorial. O consenso entre os participantes foi o de que o Brasil tem potencial e competência técnica para se tornar protagonista, desde que atue de forma articulada, estratégica e sustentada ao longo do tempo.

Matriz SWOT – Painel 7

Forças	Fraquezas
• Infraestrutura nacional de HPC consolidada (LNCC/Santos Dumont). • Base científica forte e ampla rede de universidades. • Uso de software aberto em todas as camadas da pilha. • Empresas públicas estruturadas (Serpro, Dataprev). • Instrumentos de fomento ativos (PBIA, FNDCT, BNDES, Finep). • Grandes volumes de dados públicos com potencial de uso. • Parcerias academia-empresa e integração internacional em HPC.	• Fragmentação e falta de coordenação entre iniciativas. • Demanda nacional desagregada e pouco alinhamento de mercado. • Recursos insuficientes e lentidão no repasse à academia. • Dependência de distribuições fechadas e suporte externo. • Escassez e evasão de talentos qualificados. • Uso limitado de instrumentos de fomento e compras públicas. • Falta de roadmap e prioridades definidas.
Oportunidades	Ameaças
• Criação de consórcio nacional de software aberto. • Mini supercomputador para pesquisa e testes. • Expansão de nuvem federada nacional e latino-americana. • Missões estratégicas de IA em saúde, agricultura e energia. • Governança tripartite e inspirada no CGI.br. • Fomento regional e hubs de inovação no Norte e Nordeste. • Formação massiva e políticas de capacitação em IA.	• Domínio global das big techs e lock-in tecnológico. • Riscos geopolíticos e restrições de exportação. • Vazamento e dependência externa de dados. • Descontinuidade de políticas e orçamentos. • Fuga de cérebros e competição internacional. • Limitações de infraestrutura e energia. • Permanecer apenas como consumidor de tecnologia.

Figura 32. Matriz SWOT - Painel 7

Painel 8: Como retroalimentar o Plano Brasileiro de IA, conectando ciência, indústria e governo no financiamento da soberania digital brasileira em IA?

Este painel explora como os instrumentos de fomento e financiamento podem ser aprimorados ou criados para apoiar o desenvolvimento de tecnologias nas camadas da pilha de IA identificadas como necessárias para a soberania digital brasileira. A discussão buscará identificar lacunas nos mecanismos atuais, propor novas abordagens e refletir sobre o papel das políticas públicas, da articulação interinstitucional e da colaboração com o setor privado para garantir investimentos sustentáveis e estratégicos.

Moderador: Hugo Valadares (*Diretor do Departamento de Ciência, Tecnologia e Inovação Digital do MCTI*)

Ana Estela Haddad

Secretária de Informação e Saúde Digital do Ministério da Saúde

A ideia é entender o SUS como um estudo de caso para os desafios relacionados à soberania digital. Pela sua complexidade e capilaridade nacional, o SUS oferece um modelo tangível de como articular governo, ciência e indústria para financiar e construir uma inteligência artificial soberana. O SUS é responsável por cerca de 2,8 bilhões de atendimentos por ano, o que significa uma quantidade imensa de dados produzidos continuamente. O PBIA é um marco fundamental, um roteiro estratégico que aponta um caminho para o país. Mas, como toda estratégia viva, ele precisa ser nutrido e constantemente reavaliado, conectado com a realidade e as potencialidades nacionais, ajustando-se ao longo do tempo.

A secretária afirma que nosso desafio é retroalimentar esse plano, criando um ciclo virtuoso que una a genialidade da ciência, a força da indústria e o papel indutor do Estado, garantindo soberania digital em inteligência artificial. O SUS não é apenas o maior sistema público de saúde universal do mundo, ele é uma fonte inestimável de dados e um laboratório vivo de inovação em larga escala. A partir dessa perspectiva, a da saúde, que toca cada cidadão, podemos desenhar um modelo de inteligência artificial impulsionado no Brasil.

Considerando o papel do governo como articulador de missões e financiador estratégico, é possível definir missões do SUS que orientem o desenvolvimento da IA no país. Por exemplo:

- IA para redução da mortalidade materno-infantil na Amazônia Legal, utilizando modelos preditivos para identificar gestantes de risco e otimizar a alocação de recursos;
- IA para vigilância epidemiológica em tempo real, capaz de prever surtos de dengue, zika e Chikungunya a partir de dados climáticos, sociais e de saúde;

- IA para personalização de tratamentos oncológicos no SUS, com algoritmos brasileiros treinados com dados da nossa população, para apoiar decisões clínicas e ampliar o acesso a terapias de ponta.

Hoje já contamos com uma infraestrutura de dados espaciais que permite o uso de camadas múltiplas de informação (*multilayers*), possibilitando cruzamentos e análises complexas e modelos de predição mais precisos. Propõe também fomento em cascata, direcionando recursos do BNDES e da FINEP para projetos alinhados a essas missões, além da expansão de modelos como a Embrapii para a área da saúde. Os dados devem ser tratados como um ativo soberano, com repositórios anonimizados, seguros e de alta qualidade, com protocolos éticos e de privacidade rigorosos. Esses dados são o pré-sal da nossa era digital e devem servir para treinar algoritmos brasileiros, garantindo soluções de IA que reflitam a diversidade genética e epidemiológica da nossa população.

Como exemplo, há uma iniciativa do Ministério da Saúde em andamento, iniciada em sua primeira gestão no MS, quando atuou com formação e educação permanente em saúde, onde foi iniciado o trabalho com telessaúde em parceria com universidades. Por meio do projeto foi desenvolvido núcleos de telessaúde em diversas instituições, entre elas a UFMG, cujo núcleo no Hospital das Clínicas se especializou em segunda opinião e laudos de eletrocardiogramas. Atualmente, esse núcleo acumula 11 milhões de laudos emitidos para a população e estão investindo na criação de um centro de IA e predição. Dentre os resultados estão trabalhos publicados e a realização de predições em cardiologia com base em dados da população brasileira. Isso demonstra o potencial de redes de pesquisa orientadas, de formação de talentos e de integração entre ciência e serviço público.

Também se destaca o papel da indústria como vetor de desenvolvimento e a necessidade de fortalecer compras públicas voltadas à inovação, *sandboxes* regulatórios e mecanismos de certificação de algoritmos. O ideal é criar um ciclo virtuoso, em que a ciência responde com pesquisa de ponta e formação de pessoas, a indústria transforma esse conhecimento em soluções escaláveis, e o SUS adota essas soluções, melhorando o atendimento à população e gerando novos dados que retroalimentam a pesquisa. Com isso, é possível ter a capacidade de desenvolver e controlar as tecnologias estratégicas para o futuro, garantindo que a IA à disposição dos brasileiros seja pensada, treinada e aplicada por brasileiros e voltada a resolver os desafios sociais e estruturais tipicamente brasileiros.

Os gargalos, porém, ainda são muitos e de naturezas diversas, como a relação público-privada, que precisa de maior maturidade. No campo da saúde e no SUS, há avanços consideráveis da Secretaria de Informação e Saúde Digital, preparando o terreno para avançar de forma segura. Atualmente o MS está migrando os sistemas para uma nuvem de governo, e nesse processo não pode haver interrupções. Na saúde, em especial, não há margem para falhas, e um erro em um sistema de suporte à vida é inaceitável. De forma a avançar nesse sentido, optou-se por uma arquitetura de interoperabilidade, e já

temos a Rede Nacional de Dados em Saúde (RNDS) instituída por decreto presidencial. Ela está se refletindo na disseminação de dados em várias frentes:

- para o cidadão, por meio do Meu SUS Digital;
- para o profissional de saúde, pela plataforma SUS Digital Profissional;
- e, para o gestor, pelo SUS Digital Gestor, que reúne e agrega informações para tomada de decisão em políticas públicas.

O avanço tem sido gradual, a passos firmes e largos. Há aplicações de inteligência artificial em desenvolvimento, e muito mais virá pela frente. O lançamento do PBIA e o projeto de lei para regulamentação da IA, em tramitação no Congresso, e eventos como este nos mostram que o Brasil está trilhando o caminho certo. O percurso é complexo, mas a oportunidade é única. Se governo, academia e indústria trabalharem juntos, será possível construir uma inteligência artificial não apenas inteligente, mas sábia, justa e profundamente brasileira.

João Paulo Papa

Docente do Departamento de Computação da UNESP-Bauru

O professor João possui especialidade na ciência básica, em especial na matemática da IA e trabalha há um ano no Programa Estratégico de Ciência da Computação e IA da FAPESP. Devido à essa experiência, observa que os desafios em São Paulo são muito semelhantes aos discutidos no evento, como a necessidade de ter um programa específico para IA, posto que, muitos programas de pesquisa já utilizam IA em suas linhas de atuação. A discussão na FAPESP, e que pode ser trazida para o cenário nacional, é a necessidade de identificar competências em diferentes áreas, por meio do mapeamento de grupos de pesquisa que demandam IA para acelerar suas próprias áreas de atuação e aqueles que desenvolvem a IA básica, para os quais, os investimentos precisam ser perenes.

Hoje, o fomento concentra-se em grandes empresas e aplicações, com recursos insuficientes para a pesquisa fundamental. Aplicações e parcerias produtivas são valiosas, mas não substituem a base científica, onde o conhecimento nasce. Se almejamos um modelo de linguagem próprio (à semelhança do *DeepSeek*), precisamos articular grupos de IA de base e equipes de infraestrutura. Ecossistemas de excelência unem empresas, universidades e governo de modo estruturado, com iniciativas aplicadas em curso, por exemplo, a melhoria do portal do Ministério da Saúde, com investimento relevante. Contudo, a burocracia universitária frequentemente inviabiliza parcerias com empresas, exigindo mecanismos para agilizar cooperação e contratação. Muitas vezes, as empresas interrompem projetos porque não conseguem integrar as soluções desenvolvidas ou não têm clareza sobre o que realmente desejam aplicar com IA.

Por fim, urge desburocratizar as relações entre universidades e setor público. A visão de que docentes atuam quase por filantropia, sem poder receber por contribuições específicas, é anacrônica. Regras devem evoluir para permitir remuneração e incentivos adequados pela colaboração com projetos públicos, sob critérios de transparência e conflito de interesses. Com isso, muitos pesquisadores migram para parcerias privadas porque lá há bolsas e reconhecimento direto, e o setor público precisa criar mecanismos ágeis e justos para valorizar esse trabalho e, assim, ampliar o impacto social da ciência.

Rafael Moreira

Fundador e CEO da Bertha Capital

A IA é um tema multifacetado, envolvendo recursos humanos, academia, demanda de mercado, soberania de dados, arcabouço legislativo e, além disso, a complexa integração entre governo e setor produtivo. Montar um programa que consiga articular todas essas dimensões, especialmente no uso de aplicações de IA dentro do governo e no campo das compras públicas, é um desafio enorme.

A Bertha Capital atua como investidores de startups e buscam fortalecer o ecossistema, gerenciando cerca de 1 bilhão de reais em ativo, com participações em 54 startups em diferentes estágios de desenvolvimento, com aportes variando de R\$ 1 milhão a R\$ 40 milhões. Como resultado, 2.500 empregos diretos gerados e maior conexão entre academia e mercado, com muitos professores se tornando sócios ou conselheiros das startups apoiadas pela Bertha, permitindo transformar pesquisa científica em inovação aplicada. O modelo também busca trazer corporações para investir com a empresa, criando um ambiente colaborativo entre empresas estabelecidas e empreendedores emergentes. O objetivo é fortalecer a soberania tecnológica do Brasil, investindo em empresas de base tecnológica e formando empreendedores comprometidos com o desenvolvimento nacional.

Com o apoio do Ministério da Ciência, Tecnologia e Inovação (MCTI), será anunciado o primeiro Fundo de Investimentos em Participações (FIP) voltado exclusivamente para *startups* de inteligência artificial, com recursos provenientes da Lei de Informática. O fundo terá capital inicial de R\$ 100 milhões, destinado exclusivamente a *startups* em estágio inicial. Durante a criação desse fundo, foi realizada uma análise setorial, avaliando o panorama de investimentos em IA no Brasil e no exterior, que movimenta valores na casa de centenas de bilhões de dólares, e as projeções para 2027 indicam um volume de US\$ 394 bilhões (Figura 33).

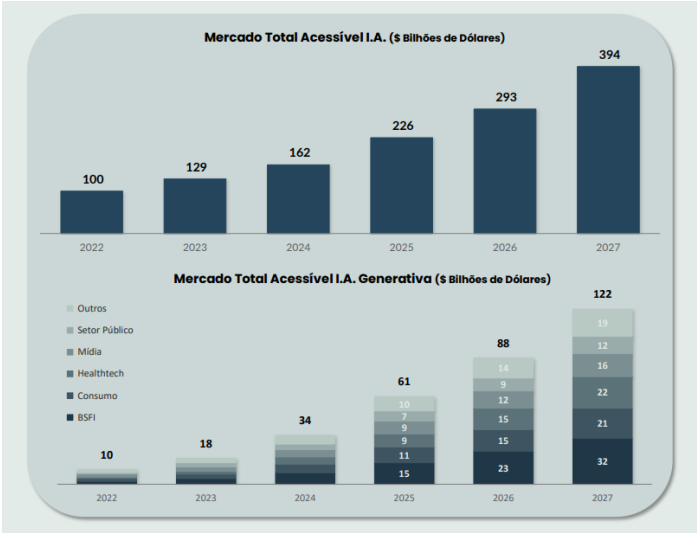


Figura 33. Mercado global de inteligência artificial: crescimentos e projeções (2022-2027)

Observa-se um crescimento muito intenso no uso da IA, especialmente da IA generativa, com projeções indicando que setor público, saúde, mídia, consumo e soluções financeiras, serão os principais demandantes até 2027, liderando a adoção e o investimento em IA nos próximos anos. Os dados globais mostram uma taxa média de crescimento anual entre 31% e 32% na demanda por tecnologias de IA, o que evidencia a velocidade e a escala dessa transformação. Quando se olha para onde está a produção dos modelos de larga escala, percebe-se dois movimentos distintos: geração e o da adoção de IA. O Brasil ainda vive um grande *trade-off* entre produzir e adotar essas tecnologias. Em alguns casos, o Brasil gera modelos próprios, mas em muitos outros está apenas adotando soluções estrangeiras.

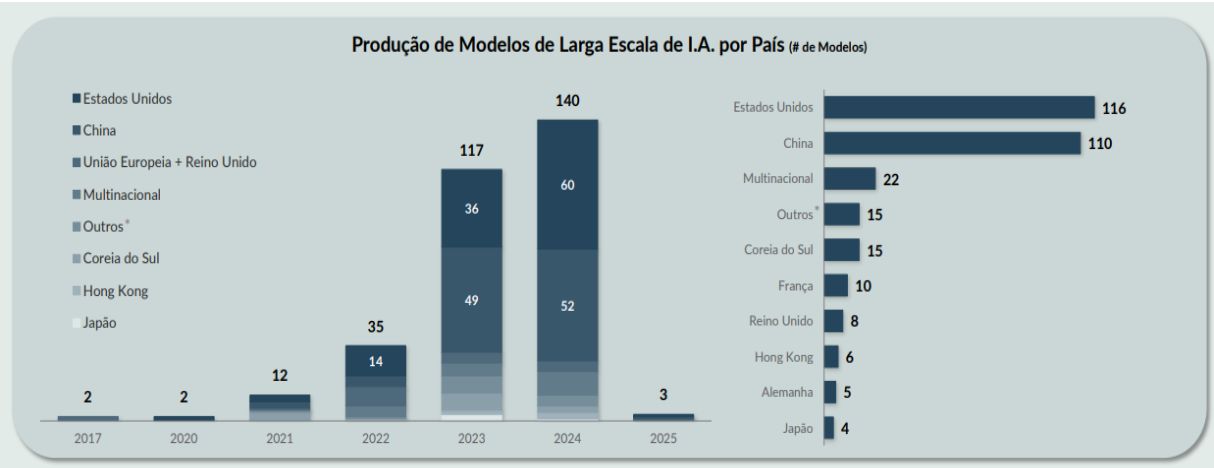


Figura 34. Produção de modelos de larga escala de inteligência artificial por país (2017-2025)

Globalmente há uma corrida global por investimentos em IA, com predominância dos Estados Unidos e da China, seguidos, ainda que a certa distância, pela União Europeia. Os Estados Unidos contam com aproximadamente 116 modelos LLM desenvolvidos e a

China cerca de 110 (Figura 34). Há ainda 22 modelos produzidos por empresas multinacionais, que desenvolvem suas próprias soluções de IA, e depois disso observamos uma curva de Pareto, em que o número de produtores cai rapidamente. Por sua vez, a América Latina apresenta números mais modestos, cerca de 0,1% da produção mundial. Entretanto, há movimento importantes na região, com o início de novos investimentos e projetos estruturantes, com possibilidade de intensificação nos próximos anos, com as novas iniciativas amadurecendo. Ainda assim, os Estados Unidos e a China continuam liderando amplamente (Figura 35).

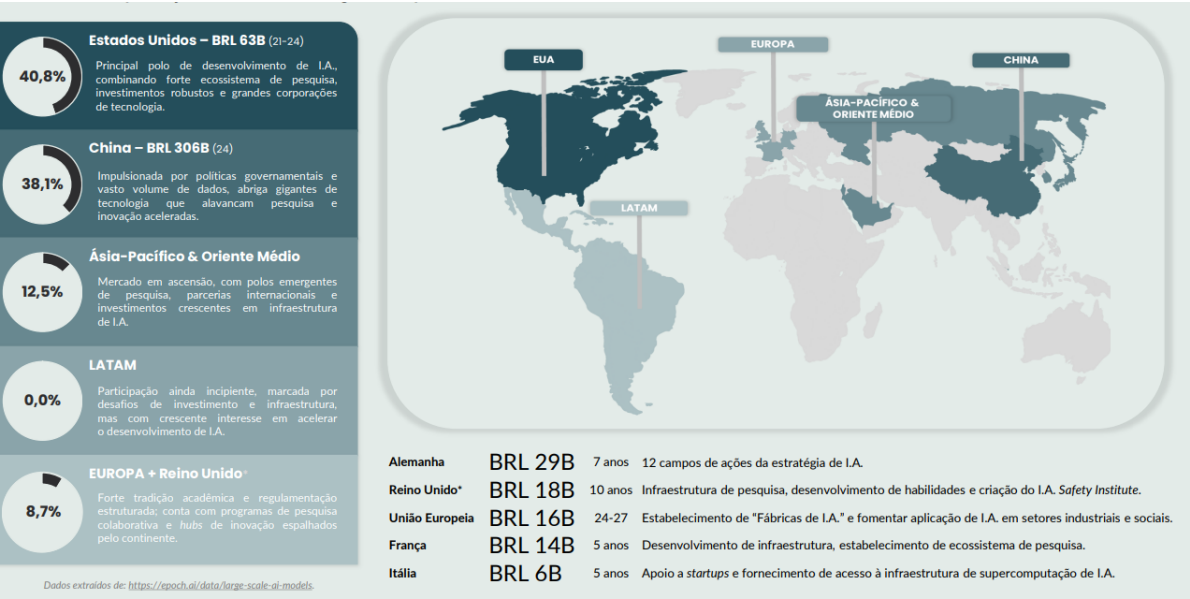


Figura 35. Investimentos públicos em modelos de IA de larga escala

O financiamento global em IA, especialmente sob a ótica da indústria de venture capital, está acelerando significativamente seus aportes. Os cinco maiores investimentos realizados em 2024 foram direcionados a empresas como *Databricks*, *OpenAI* e *Anthropic*, com valores entre 4 e 10 bilhões de dólares cada, com financiamento global de US\$ 100,4 bilhões, sendo os maiores aportes para empresas de infraestrutura de IA. Ademais, há registro de uma *startup* com três meses de existência alcançando um *valuation* de 2 bilhões de dólares (Figura 36).



Figura 36. Financiamento global em empresas e startups de IA (2020-2024)

A América Latina apresenta crescentes aportes de *venture capital* voltados à IA, principalmente a partir de 2024, quando os fundos de investimento começam a se mobilizar com maior força na região, e novos anúncios de aportes têm sido feitos recentemente. Dentre os destaques, está a *Blip*, que recebeu o maior investimento em 2024. A empresa atua com soluções de IA voltadas para atendimento via WhatsApp. Outro caso relevante é o da *Teachy*, que captou aproximadamente 7 milhões de dólares para o desenvolvimento de tecnologias de IA aplicadas à educação (Figura 37)

159

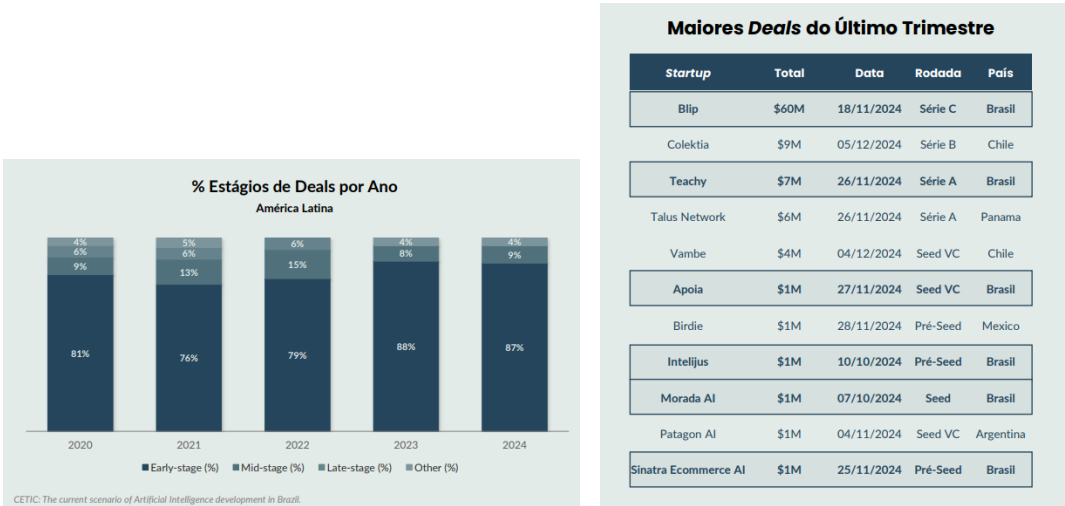


Figura 37. Investimentos em inteligência artificial na América Latina (2020-2024)

Em suma, nos últimos anos houve uma corrida dos países em investimento na base tecnológica. A Europa, em relação aos Estados Unidos, menor crescimento do PIB (redução de 25%) nos últimos dez anos, e a IA abriu uma oportunidade de, no médio prazo, conseguir retomar o controle tecnológico e estruturar nova base industrial a partir

de tecnologia. As técnicas de IA existem há um longo tempo, mas vem sendo potencializadas pela capacidade de processamento massivo e por toda a evolução tecnológica associada. E, assim como aconteceu com outras tecnologias, determinados grupos e países passaram a dominar e controlar os avanços.

Deste modo, a IA tem o potencial de mudar o patamar. Os desenvolvimentos tem ocorrido em maior medida no B2C, mas muitos países estão preocupados com o desafio de estruturar e acessar dados no contexto B2B, dentro das corporações e cadeias produtivas. Com isso emergem questionamentos importantes, como o acesso aos dados, pois se não há dados disponíveis para alimentar os modelos, há um problema estrutural.

Luiza Sidônio

Banco Nacional de Desenvolvimento Econômico Social (BNDES)

O papel do BNDES é fundamental para o contexto do financiamento da IA, que vem atuando em diversas linhas para o tema, como na Nova Indústria Brasil (NIB), que inclui grupos de trabalho dedicados à IA e busca coordenar as ações entre diferentes frentes. O PBIA e a NIB estão conectados, e a NIB, por sua vez, é estruturada a partir de missões. O apoio financeiro se materializa por meio do Plano Mais Produção, que surgiu inspirado na ideia de criar um “plano safra” para a indústria brasileira, uma política de fomento que tenha continuidade ano após ano, sem os soluços e descontinuidades que historicamente marcam as políticas industriais no país. A proposta busca unir diferentes atores de financiamento, como Caixa Econômica Federal, BNDES, BNB, Banco do Brasil, Embrapii e Finep, formando um arranjo integrado de apoio ao desenvolvimento industrial e tecnológico, com recursos consideráveis, da ordem de R\$ 611 bilhões (Figura 38).

160



Figura 38. Nova Indústria Brasil e Plano Mais Produção

Analisando o potencial da IA, a tecnologia possui grande oportunidade de impactar diferentes setores e segmentos da economia, além da capacidade em gerar transformações profundas também do ponto de vista sustentável, especialmente porque o Brasil possui uma das matrizes energéticas mais verdes do mundo, podendo ser utilizada na gestão de recursos naturais, no monitoramento de desmatamento, na saúde pública e em diversas outras áreas. Na agricultura temos um potencial extraordinário, e a IA pode ser usada para aumentar a produtividade, promover a sustentabilidade e estimular a inclusão social no campo.

De forma resumida, o ecossistema de IA pode ser categorizado em quatro setores: temos os desenvolvedores de modelos fundacionais, como *OpenAI* e *DeepSeek*, e no Brasil, o Maricata e o Amazônia; na parte de *hardware* especializado, como *chips* e *hardwares* embarcados, temos no mundo a *NVIDIA* e, no Brasil, empresas como a *Positivo*; na parte de integradores, plataformas e desenvolvedores, estão aqueles que conectam a tecnologia de IA aos clientes finais, como as *fintechs* e as **govtechs**; e, por fim, a infraestrutura, composta por *cloud*, *data centers* e plataformas de inferência (Figura 39).



Figura 39. Fomento do BNDES em IA categorizado em grupos

A parte de infraestrutura e integradores recebe maior apoio pelo adensamento empresarial no Brasil nessas áreas e as baixas barreiras de entrada. Ademais, tem-se o diferencial de energia verde em *data centers*, como o Redata, com energia mais sustentável, e um setor de TI desenvolvido, com especializações setoriais e soluções em IoT. O BNDES, junto a Finep, no programa Mais Inovação, oferece uma taxa TR para apoiar projetos de inovação e digitalização, e tem atuado por meio do mercado de capitais, com FIPs para apoio a *startups* e empresas inovadoras, e um FIP específico para IA (Figura 40).



Figura 40. Atuação do BNDES em IA categorizado em grupos – Infraestrutura e integradores

Além disso, tem sido realizadas chamadas públicas, como a que ocorreu recentemente para centros de P&D, com grande participação de empresas. Dentro do Plano Mais Inovação, o BNDES pode apoiar investimentos em plantas pioneiras, projetos de P&D e planos de investimento completos das empresas. O programa foi lançado recentemente, com foco em equipamentos 4.0, atuando de forma direta e indireta, financiando projetos de R\$ 10 milhões para as regiões Norte e Nordeste e de R\$ 20 milhões para o restante do país. Também atuam com agentes financeiros parceiros que operam linhas do BNDES, e também possuem o programa de digitalização no Plano Brasil Mais Produtivo, voltado para MPMEs. Por fim, o FUST também oferece condições especiais para *data centers*. De forma ampla, o montante do BNDES em aprovações relacionadas à IA totalizam R\$ 1,38 bilhão, divididos em 7 FIPs (Figura 41).

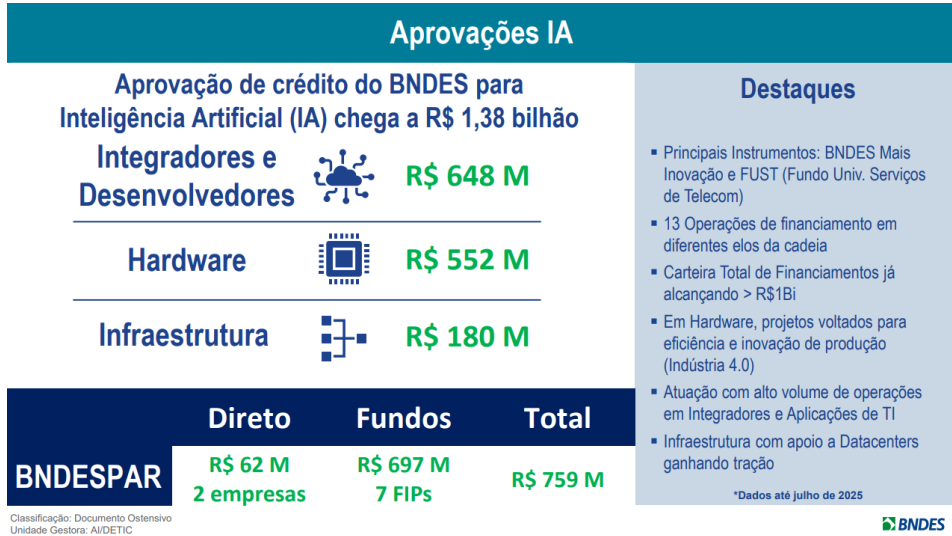


Figura 41. Aprovação de crédito do BNDES para inteligência artificial

Em suma, a IA apresenta um mundo de possibilidades e dentro das pilhas existem diversas camadas e possibilidades. Por enquanto, o Brasil ainda tem muito a refletir, desde a estruturação até o desenho do caminho que deseja seguir, afinal, o mundo está diante de uma mudança de paradigma tecnológico, com possibilidade de transformação profunda que traz consigo novas oportunidades. Apesar dos recursos disponíveis, eles são finitos, então é fundamental definir com clareza onde serão aplicados.

Joana Souza de Meirelles

Superintendente da área de saúde e transformação digital - FINEP

O tema é complexo, tanto em relação à decisão de soberania quanto ao reconhecimento de que se trata de uma escolha política, com implicações financeiras significativas. Esse contexto reforça a necessidade de coordenação e continuidade. Os investimentos em inteligência artificial devem ser permanentes, atravessar diferentes governos e manter-se estáveis diante de mudanças políticas para alcançar resultados sustentáveis. Além disso, os montantes aplicados precisam ser compatíveis com as metas estabelecidas. Ainda há grande discrepância entre os recursos brasileiros e os investimentos realizados pelas principais potências mundiais. Apesar dessa diferença, observa-se avanço significativo, impulsionado por iniciativas como o programa Inspire, resultado da articulação entre o MGI, MCTI, Casa Civil, CPQD e Finep, que mobilizou R\$ 390 milhões para fortalecer a agenda de IA no governo.

É essencial estabelecer uma estratégia clara, com prioridades definidas. O PBIA apresenta eixos estruturados com fontes de recursos e objetivos bem delineados, alinhando-se à Nova Indústria Brasil (NIB) e à política industrial orientada por missões. No campo do financiamento, a principal fonte de recursos da Finep é o FNDCT, que foi recentemente descontingenciado, garantindo fluxo contínuo de recursos não reembolsáveis para ICTs, universidades e projetos empresariais de maior risco tecnológico. A aprovação do PL 847 ampliou o acesso da Finep ao estoque do FNDCT, permitindo movimentar mais de R\$ 22 bilhões em crédito e fortalecer a integração com a academia.

Embora o país ainda esteja distante dos níveis de investimento das grandes potências, trata-se de um avanço expressivo, especialmente nos aportes voltados à infraestrutura, pesquisa básica e inovação. A Finep tem buscado tornar os investimentos mais assertivos, aumentando as chances de sucesso em projetos complexos. O progresso em inovação depende da integração entre ciência, academia e setor produtivo. Por isso, os editais de subvenção têm estimulado arranjos colaborativos em que as empresas proponentes devem incluir ICTs como parceiras. No programa Conecta, dentro da linha de crédito, há incentivo à contratação de ICTs, previsto nos indicadores de avaliação dos instrumentos.

A Finep também implementou um bônus de 1,5 ponto percentual na taxa de crédito para empresas que envolvam ICTs em seus projetos, exigindo investimento mínimo de R\$ 4 milhões na contratação de uma ICT, 5% do valor total do projeto ou a inclusão de mestres e doutores em pelo menos 20% da equipe executora. Essa medida reforça o compromisso da instituição em promover integração entre os diferentes atores do ecossistema de inovação. As novas linhas de crédito, baseadas em TR + 2,5%, têm gerado maior demanda, oferecendo custos mais baixos e prazos adequados para projetos de risco tecnológico elevado.

Outra inovação relevante é a adoção de editais em fluxo contínuo, que conferem flexibilidade e maior tempo para a formação de parcerias, além de permitir a reavaliação de propostas anteriormente indeferidas. A Finep prepara também um novo edital voltado a tecnologias digitais e pilhas de *software*, com ênfase em subvenção econômica e participação de ICTs como parceiras estratégicas.

Persistem, entretanto, desafios clássicos de financiamento, como a dificuldade de acesso ao crédito por pequenas empresas, a falta de garantias e a escassez de recursos não reembolsáveis. Para enfrentá-los, a Finep tem ampliado a atuação de agentes descentralizados, o número de editais de subvenção e o uso de Fundos de Investimento em Participações (FIPs) voltados a empresas de base tecnológica. A integração entre instrumentos de apoio é considerada essencial. A Finep tem desenvolvido editais conjuntos com o BNDES, fortalecendo parcerias e promovendo convergência entre crédito e subvenção, assegurando complementaridade entre os mecanismos de fomento à inovação.

164

Considerações finais

O painel 8 discutiu como fortalecer o Plano Brasileiro de Inteligência Artificial (PBIA) por meio da integração entre governo, ciência, indústria e instrumentos de financiamento, buscando consolidar a soberania digital do Brasil. As falas destacaram que essa construção exige planejamento de longo prazo, coordenação interinstitucional e continuidade de políticas públicas, com investimentos sustentáveis e estratégicos. O SUS foi apresentado como um exemplo concreto de articulação entre dados, inovação e serviço público, demonstrando o potencial de gerar soluções em IA voltadas às necessidades nacionais. Também foi enfatizada a importância de fortalecer a base científica e tecnológica, investir em infraestrutura e reduzir a burocracia que limita a colaboração entre universidades, empresas e órgãos públicos.

Dentre as principais propostas, destacam-se:

- Governança e alinhamento estratégico, por meio de estabelecimento de mecanismos permanentes de coordenação interinstitucional entre MCTI, FINEP, BNDES, Casa Civil e ministérios setoriais para o acompanhamento e atualização

contínua do PBIA; e garantia de integração com a NIB, de forma a assegurar estratégias de inovação, industrialização e soberania digital.

- Uso de SUS como modelo de soberania digital, servindo como estudo de caso para o desenvolvimento de IA soberana, dada a sua escala, diversidade de dados e capilaridade nacional; tratar os dados do SUS como ativos soberanos; e criar repositórios federados de dados em saúde, anonimizados e interoperáveis.
- Financiamento e instrumentos de fomento, com editais contínuos e recorrentes, com valores compatíveis aos desafios tecnológicos e maior flexibilidade para formação de parcerias; promover sinergia entre crédito e subvenção com editais conjuntos Finep-BNDES; expandir o uso de Fundos de Investimentos em Participações (FIPs) e de linhas de crédito específicas para IA, com condições especiais para projetos de maior risco tecnológico.
- Compras públicas inovadoras e *sandboxes* regulatórios, com maior estímulo ao desenvolvimento de soluções nacionais em IA para o setor público. E garantir que a IA desenvolvida no país seja treinada, operada e aplicada por brasileiros, garantindo a diversidade.
- Integração entre ciência, indústria e governo, com estímulo a arranjos colaborativos entre indústrias, empresas e Estado, priorizando transferência de tecnologia e formação de talentos; simplificar e desburocratizar as relações entre universidades e setor público, permitindo remuneração transparente a docentes e pesquisadores que contribuam em projetos públicos; fortalecer programas de inovação aberta, como o Conecta Finep, para cooperação efetiva entre empresas e ICTs.
- FIP-IA, voltado para startups de IA, com recursos de fundos existentes; fortalecer programas de incubação e aceleração de base científica, conectando pesquisadores e empreendedores; e estimular a criação de startups de IA orientadas por missões públicas, conectadas ao PBIA e às metas nacionais de soberania tecnológica.

O painel concluiu que retroalimentar o PBIA passa por construir uma estrutura coordenada, perene e colaborativa, capaz de unir ciência, setor produtivo e Estado em torno de uma visão comum de soberania tecnológica.

Matriz SWOT – Painel 8

Forças	Fraquezas
• PBIA consolidado como marco estratégico nacional. • Grande volume de dados públicos, especialmente do SUS. • Infraestrutura de dados e interoperabilidade em expansão (RNDS). • Capacidade científica sólida e redes acadêmicas qualificadas. • Instituições financeiras estruturadas (BNDES, Finep, Embrapii). • Programas de fomento ativos e articulação entre ministérios. • Setor de startups e capital de risco em crescimento.	• Falta de coordenação e continuidade entre políticas. • Burocracia nas parcerias entre universidades e governo. • Recursos insuficientes e lentos na chegada à academia. • Baixo investimento em ciência básica e HPC. • Falta de clareza sobre prioridades tecnológicas. • Dificuldade de integração entre pesquisa e mercado. • Dependência de tecnologias e dados estrangeiros.
Oportunidades	Ameaças
• Uso do SUS como modelo de soberania digital e IA aplicada. • Missões estratégicas em saúde, agro e sustentabilidade. • Criação de FIPs e novos instrumentos de fomento à IA. • Expansão de centros de P&D e hubs regionais. • Desenvolvimento de IA ética e treinada com dados brasileiros. • Fortalecimento das parcerias público-privadas e sandbox regulatórios. • Ampliação de investimentos em projetos de longo prazo.	• Descontinuidade de políticas públicas entre governos. • Concentração global de tecnologia em poucos países. • Evasão de talentos e fuga de cérebros. • Riscos geopolíticos e dependência de infraestrutura externa. • Subfinanciamento crônico da pesquisa científica. • Falta de coordenação entre ciência, indústria e Estado. • Competição global por dados e controle tecnológico.

Figura 42. Matriz SWOT – Painel 8

Painel 9: Reflexão sobre Desafios, Lacunas e Oportunidades para a Pilha de IA Nacional

Este painel tem como objetivo refletir sobre os principais desafios, lacunas e oportunidades identificadas ao longo do Seminário e propor caminhos concretos para o avanço do Brasil no desenvolvimento de tecnologias em todas as camadas da Pilha de IA. A discussão busca articular recomendações práticas, próximos passos, estratégias de articulação institucional e prioridades para políticas públicas, pesquisas e investimento.

Moderadora: Juliana Daguano (*Diretora do CTI Renato Archer*)

Carlos Rodrigo

Coordenador do Centro de Excelência de Ciência de Dados e Inteligência Artificial (SERPRO)

Já estão sendo delineadas respostas para enfrentar os problemas recorrentes identificados ao longo do evento. A fragmentação de iniciativas é uma questão crítica, resultando em gastos duplicados para resolver os mesmos desafios, o que exige mudança estrutural. Com esse propósito, foi iniciada uma articulação entre entidades de pesquisa, com o estabelecimento de uma linha de trabalho que deve começar ainda nesta semana. A proposta prevê uma parceria entre o MCTI, o LNCC, o Serpro e outras instituições para o desenvolvimento de modelos fundacionais voltados à pesquisa e à experimentação.

O objetivo central é fomentar o desenvolvimento coordenado da inteligência artificial, unindo esforços e compartilhando recursos já disponíveis. A atuação conjunta entre centros de pesquisa busca criar uma base concreta e colaborativa de inovação. Para garantir resultados efetivos, é imprescindível a definição de indicadores de desempenho (KPIs) e o monitoramento contínuo dos resultados. Sem acompanhamento sistemático, as iniciativas correm o risco de permanecer como intenções ou documentos sem continuidade. Ademais, é necessária uma coordenação central responsável por consolidar, acompanhar e divulgar metas e resultados, como o número de empresas e *startups* apoiadas, profissionais formados e projetos concluídos.

Outro ponto de destaque é o letramento em IA, que deve iniciar-se ainda no ensino médio. O letramento tardio limita a formação de competências técnicas e científicas nas áreas que compõem a pilha de IA. Foram identificadas, durante o evento, diversas ações de curto prazo, incluindo a proposta de acompanhamento sistemático e divulgação pública de resultados, com apoio do Serpro.

Quanto à soberania digital, os dados do Serpro permanecem restritos ao seu próprio ambiente, sem acesso externo, o que assegura soberania operacional e confidencialidade, uma vez que toda a infraestrutura e execução técnica são internas.

Paralelamente, busca-se consolidar a soberania tecnológica, fortalecendo a base da pilha de IA com tecnologias nacionais e múltiplas opções de fornecimento de *hardware*, mitigando riscos diante de possíveis cenários geopolíticos adversos. Assim, a nuvem soberana de governo já é uma realidade em expansão, ampliando sua capacidade de atender à população brasileira e consolidando a infraestrutura tecnológica necessária à autonomia digital do Estado.

Cristiano Maciel

Professor da Universidade Federal do Mato Grosso (UFMT) e Vice-presidente da Sociedade Brasileira de Computação (SBC)

Há um cenário complexo que temos diante de nós, que se estende por todas as camadas da pilha. Foi explicado mais sobre isso na palestra inaugural com o professor Edson Borin, e durante o evento refletimos sobre essa complexidade. O termo “tropicalização da pilha” é importante para pensar em como criar um ecossistema nacional, voltado à soberania tecnológica. Na parte de *hardware*, foi abordada a questão dos supercomputadores, sua necessidade de uso, manutenção, o papel do *open source*, dos consórcios e da união entre instituições, além da indústria de chips, que se coloca como um ponto crítico no qual precisamos continuar avançando.

Na parte de governança, chamou atenção para a necessidade de políticas públicas voltadas à regulação em prol da cidadania, à exploração responsável dos dados e à segurança como requisito emergente. Também foi discutida a importância de modelos de financiamento sustentáveis. Alguns dizem que há recursos, outros dizem que não, mas é preciso saber onde estão esses recursos e como construir projetos com eles. Temos novos modelos de fomento para *startups*, novos formatos de bolsas, e é preciso fortalecer a autonomia em IA para coordenar os multissetores e contratos plurais. Também é preciso explorar melhor o papel dos BRICS em termos de governança internacional e pensar em como integrar as camadas da pilha de maneira coordenada e eficiente.

A parte de educação é de grande interesse para a Sociedade Brasileira de Computação (SBC). Há a necessidade por eles de integrar melhor os currículos, e embora estejam surgindo novos cursos específicos de IA, isso não diminui o valor dos cursos já existentes na área de tecnologia, porque a IA é transversal. É preciso discutir a capacitação e retenção de talentos, investindo no letramento digital desde o ensino fundamental, não apenas no médio. As crianças já estão imersas no uso de tecnologia, e é preciso trabalhar aspectos de pensamento crítico desde cedo. Igualmente importante, é preciso discutir a questão de gênero e raça. A IA reflete os vieses existentes na sociedade, e, portanto, é preciso atuar pela diversidade para evitar que esses vieses se perpetuem nos sistemas tecnológicos.

Ainda sobre recursos humanos, destaque para a diversidade de profissionais, e é preciso potencializar ainda mais o trabalho conjunto em prol de todas as camadas da pilha de IA. Observa-se o aumento expressivo de trabalhos acadêmicos sobre IA, em diversas áreas do conhecimento, o que representa um desafio positivo. É preciso aumentar o uso de IA não apenas na academia, mas também na indústria, de forma coordenada e ética. Muitas empresas já utilizam IA, até mesmo com exemplos apresentados durante o evento, mas é preciso sistematizar e organizar essas iniciativas nos atentando a questões éticas.

Por fim, destaque para a importância de ferramentas desenvolvidas com segurança *by design*, projetadas desde o início com a segurança como requisito essencial. Mas é preciso garantir a segurança dos dados, em que esses sejam certificados e utilizados de forma confiável. O Serpro assegura que os dados dos brasileiros estão protegidos, mas, ao mesmo tempo, observa-se dados sendo colocados em sistemas de IA a todo momento. Desta forma, é preciso encontrar um equilíbrio entre abertura e proteção, para que o uso de dados promova inovação, mas sem comprometer a privacidade e a soberania.

Fábio Borges

Diretor e Professor do Laboratório Nacional de Computação Científica (LNCC)

Um grande desafio atualmente, muito comum na academia, é o de unir esforços. Nas universidades, essa união nem sempre é simples, há departamentos de matemática aplicada, matemática pura e outras áreas que, muitas vezes, falam linguagens diferentes. No entanto, essa integração é essencial, porque trabalhamos em uma área interdisciplinar, com muitos desafios e interdependências. Há mais de dois anos, a discussão foi pautada entre uma máquina de grande porte ou em várias menores, decidindo-se pela grande. Os argumentos mudam, mas as questões voltam: é possível comprimir modelos de IA e rodar com menos capacidade? A resposta é não. Sempre será preciso maior capacidade de processamento, e, por isso, é necessário ter uma máquina diferenciada.

Os CENAPADS espalhados pelo Brasil têm diferentes tamanhos e potenciais de processamento, e todos são máquinas poderosas. Não há falta de recurso, mas carência de demanda, com vários lugares com recursos disponíveis, com fundos existentes, mas que geralmente estão nas camadas mais altas do sistema. Além disso, as demandas dependem de estratégias, com definição de prioridades, elencando modelos e aplicações. Muitas dessas iniciativas não se viabilizam ou não atraem grande volume de investimentos, então é fundamental que exista uma política pública nacional que as sustente. No âmbito do PBIA, todos os ministérios deveriam estar envolvidos e aportar recursos ao plano, não o contrário. O Fundo Nacional de Desenvolvimento da Educação (FNDE), por exemplo, é muito maior do que o FNDCT e poderia contribuir com o PBIA. As Forças Armadas também têm diversos projetos relacionados à soberania nacional e

poderiam participar mais ativamente. Isso vale para o setor do agronegócio, que já realiza trabalhos relevantes com veículos autônomos.

Desta forma, recomenda-se descrever claramente essa pilha, das camadas mais altas às mais baixas, definir quanto de recurso vai para cada camada e construir esse planejamento de forma estruturada. O LNCC é um braço do MCTI e vem cumprindo sua missão institucional em parceria com universidades e outras instituições federais, por meio do diálogo e dispostos a colaborar em tudo o que for necessário para fortalecer essa agenda nacional.

Mateus Pierre

Diretor de Pesquisa e Desenvolvimento - Instituto ELDORADO

O seminário discutiu em vários momentos a importância de estratégias de longo prazo e de políticas de Estado para garantir a soberania tecnológica. O PBIA é um ótimo pontapé inicial, mas é necessário olhar para o futuro e planejar ações sustentadas no tempo, com investimentos contínuos e objetivos claros. Em relação à pilha de IA, é essencial dominar as tecnologias que a compõem, desde sistemas operacionais até *frameworks* de *machine learning*, e com isso focar no desenvolvimento de modelos que sejam relevantes para áreas estratégicas do país.

O SUS deve ser uma prioridade, dada a diversidade e o volume de dados disponíveis. Além disso, outros setores estratégicos, como energia e agronegócio, se destacam no Brasil. E é necessário acompanhar movimentos globais, como o da hiperpersonalização, a adaptação de dispositivos e serviços aos perfis individuais de uso. Para isso, instrumentos adequados de fomento são fundamentais, além de investimentos nas camadas inferiores da pilha e atração de projetos que possam escalar para as camadas superiores.

A carência de infraestrutura de HPC é um ponto crítico, em que a aquisição de um supercomputador representa um marco para a soberania tecnológica. Da mesma forma, o fortalecimento da rede de centros SINAPAD e a interconexão entre esses centros de processamento são aspectos fundamentais. Na questão da soberania dos dados, um ponto-chave é a implementação de *data spaces*, nos quais o dono do dado controla o acesso, delimitando restrições. Esse conceito é extremamente relevante e precisa entrar de forma estruturante na agenda.

Outro eixo essencial é a capacitação de profissionais, com a necessidade de uma política nacional de formação em IA, voltada a dominar tecnologias, especialmente modelos *open source* e bibliotecas fundamentais, que possam servir de base para o desenvolvimento de modelos próprios. É importante dominar tecnologias-chave e estratégicas para o país, como a conversão de modelos para o mundo embarcado, onde há limitações de memória e processamento. Também é preciso criar condições para

reter talentos, considerando no papel da soberania para reter recursos humanos altamente qualificados. Por fim, na questão dos projetos estruturantes da pilha de IA, é preciso definir conexões e arranjos colaborativos que permitam complementar conhecimentos. É fundamental estabelecer arranjos que integrem competências e viabilizem o desenvolvimento das tecnologias necessárias para a construção da pilha nacional de inteligência artificial.

Guilherme de Paula Corrêa

Coordenador Geral de Tecnologias Digitais do Ministério da Ciência, Tecnologia e Inovação (MCTI)

Há cerca de dois meses foi realizado um evento voltado aos modelos de linguagem (LLMs), que serviu de base para a evolução até o encontro atual. O seminário da pilha mostrou-se mais robusto, consistente e participativo, com maior engajamento do público. Houve ampla recepção de elogios, críticas e sugestões construtivas, elementos fundamentais para o aprimoramento contínuo das próximas edições. O formato foi amplamente debatido, optando-se por um maior número de palestrantes com tempo reduzido de exposição. Embora essa escolha tenha gerado diferentes percepções entre os participantes, a definição do formato depende dos objetivos específicos de cada encontro. Está prevista a realização de novos eventos, possivelmente com formatos variados, incluindo encontros técnicos e reuniões menores, em formato de *petit comité*.

A principal conclusão é a importância da continuidade desses encontros e da interação entre os participantes. O evento demonstrou o valor do diálogo e da colaboração entre os diferentes atores envolvidos no ecossistema de inteligência artificial. De forma simbólica, foi ressaltada a necessidade de consolidar políticas de apoio à “pilha nacional de IA”, com programas integrados de fomento, como bolsas específicas no CNPq, chamadas temáticas na Finep e linhas de crédito no BNDES. A mensagem final reforçou a relevância do engajamento e da cooperação entre as instituições e profissionais, destacando que o êxito das ações depende da participação ativa e do comprometimento coletivo. O MCTI reconhece e agradece a parceria e dedicação de todos os envolvidos.

André Rafael

Coordenador de Políticas de Ciência, Tecnologia e Inovação Digital do Ministério da Ciência, Tecnologia e Inovação (MCTI)

O Plano Brasileiro de Inteligência Artificial (PBIA) contempla diversos temas estratégicos cuja relevância já era reconhecida, embora ainda houvesse incertezas sobre a forma de atuação, entre eles a pilha nacional de IA. Recentemente, a Casa Civil atribuiu ao MCTI a responsabilidade de definir diretrizes para esse eixo, o que demandou ampla reflexão e coordenação. Diante da complexidade do tema e da importância do assunto para o país,

a direção do Ministério orientou a realização de um seminário nacional, envolvendo academia, empresas e parceiros institucionais, com o objetivo de discutir caminhos e propor ações concretas.

O evento contou com a colaboração de diversas instituições e resultará em um relatório detalhado, que será encaminhado à alta direção para deliberação sobre as próximas etapas. Por se tratar de um tema presente na agenda presidencial, há necessidade de respostas rápidas e estruturadas. Diferentemente de outros países, que realizaram debates restritos, o Brasil optou por um formato aberto e online, ampliando a participação e a pluralidade de perspectivas sobre o futuro da pilha de IA nacional.

Os resultados foram amplamente positivos, superando expectativas iniciais. Embora já houvesse anúncios previstos, o evento assumiu maior relevância e consolidou a decisão de realizar um novo encontro sobre a pilha de IA no próximo ano, possivelmente no primeiro semestre. Entre as lições aprendidas, destacou-se a necessidade de ampliar a representação da sociedade civil nas próximas edições. A principal diretriz que emerge desse processo é a definição de um mote unificador para orientar as ações e políticas futuras: a inteligência artificial deve ser desenvolvida e aplicada para o bem comum. Esse princípio sintetiza o compromisso coletivo e a responsabilidade compartilhada na construção de uma IA ética, inclusiva e socialmente orientada.

Considerações finais

172

O painel 9 fez um balanço franco dos desafios, lacunas e oportunidades para a pilha nacional de IA, apontando a urgência de coordenação central, metas e indicadores claros para evitar esforços isolados e desperdício de recursos. Dentre os principais apontamentos sobre o seminário, destacam-se:

- Fragmentação de iniciativas e duplicidade de esforços, gerando desperdício de recursos, com necessidade de coordenação central para consolidar metas, monitorar resultados e divulgar indicadores (KPIs); governança baseada em políticas públicas voltadas à cidadania, segurança de dados e financiamento sustentável; proposta de parceria interinstitucional entre MCTI, LNCC, Serpro e outras entidades para desenvolver modelos fundacionais nacionais.
- Importância do letramento em IA desde o ensino médio, para formação antecipada de competências técnicas; integração de educação e IA nos currículos desde o ensino fundamental, promovendo letramento digital precoce; valorização da diversidade de gênero e raça, para mitigar vieses nos sistemas de IA.
- Fortalecimento da soberania digital e tecnológica, com infraestrutura nacional, controle de dados e uso diversificado de *hardware*; ampliação da nuvem soberana de governo, consolidando a autonomia digital do Estado. Avanço em

hardware e supercomputadores, com ênfase no papel do *open source*, consórcios e indústria nacional de chips.

- Desenvolvimento de ferramentas seguras *by design*, com certificação e uso ético dos dados; implementação de *data spaces* com controle de acesso e soberania de dados.
- Identificação de carência de demanda, não de recursos; fundos e equipamentos estão subutilizados; necessidade de “tropicalização da pilha”, adaptando a estrutura tecnológica às realidades e potencialidades nacionais.
- Reforço da importância dos supercomputadores e da rede SINAPAD como pilares da infraestrutura nacional de IA; maior integração entre os CENAPADS; estímulo a arranjos colaborativos entre instituições.
- Necessidade de planejamento estruturado da pilha, com definição de recursos por camada (do *hardware* às aplicações); domínio das camadas da pilha de IA, do sistema operacional aos *frameworks* de *machine learning*; e *roadmap* nacional que define as prioridades por camadas e setores.
- Continuidade nas políticas e discussões, transformando o PBIA em política de Estado, com garantia de investimentos de longo prazo; realização de outros eventos e seminários para aprofundar no tema.

Dentre as contribuições, destaque para o compromisso de manter o diálogo aberto, com novos formatos de eventos, maior presença da sociedade civil e a diretriz de que a IA no Brasil deve ser, efetivamente, para o bem de todos.

173

Matriz SWOT – Painel 9

Forças	Fraquezas
• Articulação entre MCTI, LNCC, Serpro e outras instituições. • Avanço na criação da nuvem soberana de governo. • Base científica e técnica em HPC e IA. • Envolvimento ativo da SBC e da comunidade acadêmica. • Compromisso com segurança e soberania de dados. • Reconhecimento da IA como política de Estado. • Espírito colaborativo e disposição para diálogo entre setores.	• Falta de coordenação central e metas definidas. • Iniciativas isoladas e duplicação de esforços. • Déficit de letramento digital e formação inicial. • Carência de infraestrutura HPC em escala suficiente. • Dificuldade de integração entre academia, governo e empresas. • Ausência de planejamento de longo prazo e roadmap nacional. • Escassez de políticas consistentes para retenção de talentos.
Oportunidades	Ameaças
• Criação de um ecossistema nacional “tropicalizado” de IA. • Fortalecimento do PBIA com visão integrada e sustentável. • Ampliação da formação e do letramento digital desde o ensino básico. • Expansão da infraestrutura HPC e rede SINAPAD. • Cooperação internacional via BRICS e consórcios científicos. • Implementação de <i>data spaces</i> para governança de dados. • Avanço de políticas de diversidade e inclusão em IA.	• Dependência tecnológica externa e concentração global. • Descontinuidade de políticas públicas entre governos. • Evasão de talentos e fuga de cérebros. • Falta de financiamento estável e previsível. • Fragmentação institucional e sobreposição de esforços. • Vulnerabilidade de dados e riscos geopolíticos. • Dificuldade de cooperação efetiva entre governo, academia e setor produtivo.

Figura 43. Matriz SWOT – Painel 9

ANEXO 1. Propostas consolidadas a partir dos subsídios do Seminário Nacional da Pilha de IA

1. Novo Sinapad

Fortalecer a infraestrutura nacional de HPC e nuvem soberana, integrando os centros do Sinapad e sustentando novos investimentos em supercomputadores de ponta. A meta deve ser disponibilizar poder de cálculo como serviço (AlaaS), com foco em simplicidade de uso, interoperabilidade e democratização do acesso, permitindo que startups, pesquisadores e governos explorem IA avançada sem barreiras tecnológicas.

2. Aliança para a Pilha Nacional de IA e Centro de Excelência

Instituir Centro de Excelência, articulado a rede de ICTs, especializado(a) em pilha de IA Soberana, encarregado de estruturar, manter e certificar a Pilha Nacional de IA (*hardware, frameworks, bibliotecas, modelos fundacionais, protocolos de interoperabilidade*), além de manter a aliança estratégica para desenvolver a pilha de IA Nacional. Este centro deve operar como cérebro técnico nacional, garantindo independência tecnológica e alinhamento com padrões globais.

174

3.(SinIA) Dados e Modelos Estratégicos

Criar plataforma nacional de dados e modelos confiáveis, com repositórios abertos, sintéticos e federados, além de nós regionais que conectem universidades, empresas e governos locais. Essa rede de dados e modelos deve gerar valor econômico e social no território, reduzir assimetrias regionais e assegurar que o Brasil controle e valorize seus próprios ativos informacionais.

4. Missões Nacionais de IA para o Desenvolvimento

Criar forças tarefa para lançar grandes missões temáticas de IA aplicada em áreas críticas e agregadoras de demanda (governo, saúde, clima, biodiversidade, educação, indústria, finanças e agro), com recursos do FNDCT e de instrumentos inovadores (encomenda tecnológica, PPPs, coinvestimento privado). Cada missão deve ter metas claras de impacto e transformar IA em vetor de produtividade, inclusão e sustentabilidade.

5. Capacitação Massiva e Retenção de Talentos

Implementar um Programa Nacional de Formação em HPC e Pilha de IA, alinhado às missões e às demandas de mercado. Isso inclui especialização avançada (PhD, engenheiros de sistemas, arquitetos de nuvem), requalificação da força de trabalho e atração/ retenção de cérebros. O objetivo é criar massa crítica de talentos, com domínio e capacidade de aperfeiçoamento de soluções abertas.

6. Programa Nacional de IA em dispositivos e sistemas embarcados

Estruturar rede de ICTs e linhas de fomento específicas para IA na borda e embarcada. O programa deve focar no domínio de técnicas de otimização, compressão, quantização e poda (pruning) de modelos, além de desenvolver e apoiar plataformas de *hardware* e *software* que criem a base para embarcar aplicações de IA nacionais em produtos industrializados, dada a penetração cada vez maior da IA embarcada.

7. Alavancar Cooperações Estratégicas com o Sul Global

Priorizar cooperações internacionais que permitam a coprodução de tecnologia em termos vantajosos. Fortalecer a parceria Índia-Paraná, utilizando sua experiência em supercomputadores soberanos (SEDAC) para construir uma rede de HPCs com transferência de tecnologia. Liderar a integração na América Latina através de iniciativas como o Scalac, e usar o BRICS como plataforma para desenvolver uma nuvem federada e promover a visão brasileira de código aberto, mudando a mentalidade de "país colonizado".

175

8. Garantia de Financiamento Perene e Multissetorial

Ampliar e assegurar a continuidade do PBIA com modelos de financiamento de longo prazo (décadas), utilizando o FNDCT descontingenciado como um pilar. Articular este recurso com linhas de crédito do BNDES e da FINEP, que já oferecem bônus para projetos em colaboração com ICTs. Além disso, criar um mecanismo para que todos os ministérios que se beneficiarão da IA aportem recursos no PBIA de forma coordenada (ANP, Aneel, Funttel, Fust, Fundeb, Redata), retroalimentando os recursos do FNDCT.

9. Fomento à Cultura de Segurança por Design (*Security by Design*)

Integrar a segurança como pilar central na concepção da pilha de IA, e não como uma etapa posterior. Desenvolver padrões de certificação próprios para o Brasil, que garantam que os sistemas sejam auditáveis em todas as camadas, desde a coleta de dados até a inferência do modelo, para assegurar soberania e confiança.

10. Programa Nacional de Design de Chips (*Fabless*) em RISC-V

Lançar uma política pública para incentivar o design de chips utilizando a arquitetura aberta RISC-V, seguindo o modelo *fabless* (sem fábrica própria). O programa deve usar o poder de compra do Estado para criar demanda inicial por esses chips nacionais, mitigando a dependência de *hardware* estrangeiro. Envolver a Petrobrás e CEITEC, além de outros parceiros nacionais de peso.

11. Aceleração da Regulamentação para Parcerias Academia-Empresa

Criar um grupo de trabalho para simplificar e acelerar os processos burocráticos dentro das universidades para a formalização de parcerias com empresas. A burocracia excessiva, como em disputas por percentuais de royalties, foi citada como um motivo para a perda de projetos.

12. Disponibilizar LLMs soberanos para públicos prioritários (governo, estudantes e pesquisadores)

Desenvolver em parceria com LNCC, Serpro e RNP o instanciamento e aperfeiçoamento de modelos abertos de ponta, viabilizando infra soberana para oferta desses modelos para governo e ICTs. Incluir o desenvolvimento de rede nacional para aperfeiçoar e especializar LLMs de interesse nacional.

ANEXO 2. Programação do Seminário

DIA 1º de outubro (quarta-feira)

Link: [Gravação Dia 1](#)

8:00 – Boas vindas e abertura do evento

Henrique Miguel (Secretário de Ciência e Tecnologia para a Transformação Digital)

Alexandre Gonçalves de Amorim (Diretor-Presidente Serpro)

Cristiano Maciel (Vice-Presidente da Sociedade Brasileira de Computação - SBC)

Rogério da Veiga (CITDigital/Casa Civil)

8:50 - Apresentação 1: Pilha de IA: Oportunidades e Barreiras

Apresentador: Prof. Edson Borin (UNICAMP)

177

09:15 - Painel 1: Quais os caminhos para sustentar a IA com autonomia de infraestrutura tecnológica nacional?

Carlos Rodrigo (SERPRO)

Prof. Guido Araújo (UNICAMP)

Leandro Rosa dos Santos (Positivo)

Prof. Luca Belli (FGV-RJ)

Marcelo Pinheiro (Versatus)

Moderador: Renata Mielli (CGI.br)

10:15 - Break

10:30 - Painel 2: Como o Brasil pode atuar nas camadas superiores da Pilha de IA?

Cleilton Rocha (Atlântico)

Igor Régis (Banco do Brasil)

Profa. Karine Ferreira (INPE)

Norberto Alves Ferreira (CPQd)

Prof. Samuel Xavier de Souza (UFRN)

Moderador: Caetano Penna (CGEE)

11:30 - Painel 3: Como o Brasil está posicionado no desenvolvimento de IA para dispositivos móveis e embarcados?

Ariston Brandão (SIDI)

Fernando dos Santos (Venturus)

Prof. Leandro Villas (UNICAMP)

Mateus Pierre (Instituto ELDORADO)

Moderador: Eduardo Matos (Embrapii)

12:30 Almoço

14:15 - Painel 4: Ciência aberta, *open source* e cooperações internacionais: limites e potenciais

Painelistas:

Bianca Amaro (MCTI)

Danielle Ayres (GSI)

Guilherme Fitzgibbon (MRE)

Jorge Edison (Fundação Araucária)

Prof. Philippe Olivier Navaux (UFRGS)

Moderadora: Adriana Thomé (MCTI)

Este painel discutiu desafios e oportunidades relacionados à adoção de práticas de ciência aberta, uso de tecnologias *open source* e parcerias internacionais na construção da Pilha de IA nacional. O objetivo foi refletir como essas abordagens podem acelerar o desenvolvimento tecnológico, ampliar a colaboração e reduzir dependências, ao mesmo tempo em que se avaliam os riscos associados à soberania, segurança, governança e sustentabilidade dos projetos.

15:15 - Painel 5: Avanços PBIA: construindo o novo Sinapad e o Sistema Nacional de Tecnologia e Inovação em IA (SinIA)

Painelistas:

André Rafael (MCTI)

Fabio Borges (LNCC)

Hugo Valadares (MCTI)

José Ferreira (RNP)

Moderador: Washington Luís Ribeiro (Ibict)

16:15 – Break

16:30 - Painel 6: Capacitação e formação técnica para a construção de uma Pilha de IA soberana:

Profa. Alba Melo (UnB)

Prof. Anderson Soares (UFG)

Evânio Araújo (MEC)

Luciana Gomes (CNPq)

Otávio Ciribelli (Petrobras)

Prof. Wagner Meira (UFMG)

Moderador: Guilherme de Paula Corrêa (MCTI)

179

17:30 – Encerramento Dia 1

DIA 2 de outubro (quinta-feira)

Link: [Gravação Dia 2](#)

9:00 - Painel 7: Conexões que sustentam a IA: Onde estão as alianças/colaborações estratégicas nacionais para a construção de uma Pilha de IA soberana? :

Carla Osthoff (LNCC)

Deybson de S. Cipriano (Assespro)

Kiko Reis (Magalu Cloud)

Renan Gaya (MGI)

Profa. Teresa Ludemir (UFPE)

Moderadora: Beatriz Vasconcellos (Casa Civil)

10:00 - Painel 8: Como retroalimentar o Plano Brasileiro de Inteligência Artificial, conectando ciência, indústria e governo no financiamento da soberania digital brasileira em IA? :

Ana Estela Haddad (MS)

Prof. João Paulo Papa (UNESP)

Rafael Moreira (Bertha Capital)

Luiza Sidônio (BNDES)

Joana Souza de Meirelles (FINEP)

Moderador: Hugo Valadares (MCTI)

11:15 - Break

11:30 - Painel 9: Próximos passos para o desenvolvimento da Pilha de IA Nacional

André Rafael (MCTI)

Carlos Rodrigo (SERPRO)

Prof. Cristiano Maciel (SBC)

Fabio Borges (LNCC)

Guilherme Corrêa (MCTI)

Mateus Pierre (Instituto ELDORADO)

Moderador: Juliana Daguano (CTI Renato Archer)

12:30 – Encerramento Dia 2

ANEXO 3. Contribuições do Youtube

O seminário contou com ampla participação e engajamento do público ao longo dos dois dias de evento. Os participantes expressaram satisfação com a condução e a relevância das discussões, e ao final destacaram aspectos positivos como o acompanhamento em LIBRAS e a qualidade das apresentações

Discussões sobre políticas e estratégias em IA: As discussões no chat revelaram preocupação com o cenário brasileiro de ciência, tecnologia e inovação, especialmente em relação à formação de profissionais, à integração entre setores e à valorização científica e tecnológica. Várias mensagens mencionaram o Plano Brasileiro de Inteligência Artificial (PBIA), com sugestões, como a inclusão de uma “*Bolsa Pilha de IA*” e o incentivo a projetos de pesquisa vinculados à “*Pilha de IA Nacional*”. Entre os pontos mais debatidos, destacou-se a necessidade de cooperação entre governo, academia e mercado. Comentários como “*Há um gap entre governo, academia e mercado. É necessário estreitar relações.*” e “*As instituições representadas nesse seminário cooperam entre si para alavancarem uma pilha de IA funcional no Brasil?*” reforçam a urgência dessa integração.

Críticas e reflexões sobre o contexto nacional: Muitos participantes direcionaram preocupações sobre o desenvolvimento científico no país. Foram recorrentes as menções à desvalorização profissional e à fuga de talentos.

181

Entre as observações, destacam-se:

- “*O ensino básico do Brasil é bem ruim, por isso formar o profissional STEM aqui é muito complexo.*”
- “*Valorização é a palavra-chave.*”
- “*É triste ver tantos brasileiros de alto nível servindo outros países.*”

Tecnologia, mercado e soberania digital: As discussões também abordaram temas como o uso de tecnologias abertas, a soberania digital e o papel do setor privado. Alguns participantes manifestaram críticas à concentração tecnológica e à dependência de serviços estrangeiros de IA, defendendo a construção de uma infraestrutura nacional.

Comentários como “*Ao menos o fortalecimento da pilha de IA nacional é com tecnologia aberta.*” e “*Usar IA via API continua sendo usar externamente; os dados já foram até lá.*” revelam preocupações sobre soberania tecnológica.

Houve ainda debates sobre as barreiras impostas por modelos de negócios que priorizam lucro e isenções fiscais, com menções à necessidade de equilíbrio entre interesses públicos e privados no desenvolvimento de soluções de IA.

Setores estratégicos e aplicações: O agronegócio surgiu como um dos setores de destaque para aplicação de IA. Comentários como “*Podemos começar chamando a*

Embrapa” indicam o reconhecimento do potencial da IA para aprimorar a produtividade e a sustentabilidade do setor.

Outros participantes mencionaram aplicações voltadas à fiscalização e à gestão pública, reforçando o papel da IA como instrumento de modernização administrativa e eficiência institucional.

Propostas e iniciativas inovadoras: Alguns comentários trouxeram sugestões de projetos. Exemplo: *“Conheçam o projeto Knowledge3D, que visa mudar o paradigma, saindo de espaços de altas dimensões e baixa densidade especializada para espaços de baixa dimensão e alta densidade multimodal.*

Também foram mencionadas contribuições teóricas sobre arquitetura e memória de modelos de linguagem (LLMs), o que demonstra o nível avançado de parte do público participante.

Reflexões sociais e educacionais: As falas destacaram a importância da educação como base do desenvolvimento tecnológico sustentável. Comentários como *“Nada valem as discussões se o país continuar a desrespeitar os profissionais e não investir na educação básica e superior.”* reforçam a consciência coletiva sobre a relação entre investimento educacional e soberania tecnológica.

Conclusão Geral

O conjunto dos comentários dos dois dias do seminário evidencia:

- Engajamento expressivo e diversificado do público, com presença de representantes de instituições públicas, acadêmicas e de pesquisa;
- Interesse técnico e político nas discussões sobre inteligência artificial, políticas públicas e soberania digital;
- Críticas construtivas sobre educação, valorização profissional e integração entre setores;
- Reconhecimento da relevância do evento como espaço de diálogo e articulação para o fortalecimento da IA no Brasil;
- Valorização da acessibilidade, da inclusão e da comunicação transparente

ANEXO 4. Contribuição via Strateegia

O Strateegia foi utilizado como plataforma colaborativa do evento, permitindo que participantes registrassem comentários, perguntas, ampliando o engajamento, a transparência e o aproveitamento das contribuições para subsidiar a sistematização dos debates e a elaboração deste relatório.

1. Diagnóstico geral e senso de urgência

Os participantes expressaram consenso quanto à defasagem do país em infraestrutura e capacidade computacional, alertando que o Brasil corre o risco de ficar à margem da nova economia da IA.

Entre os principais pontos levantados:

- Dependência de infraestrutura estrangeira (GPU, nuvens privadas, *frameworks*).
- Necessidade de planejamento coordenado e fomento plurianual.
- Importância de modelos médios (7–13B parâmetros) em português, otimizados para RAG e agentes inteligentes como caminho mais viável para o curto prazo.

“O Brasil está atrasado. Precisamos parar de discutir se vale a pena e começar a construir.”

2. Hardware, HPC e soberania tecnológica

Os comentários destacam que não é realista produzir GPUs/TPUs de ponta no curto prazo, mas é possível:

- Investir em design fabless e *hardware* aberto (RISC-V).
- Criar laboratórios conjuntos universidade–empresa.
- Estabelecer fundos específicos para semicondutores aplicados à IA.

Há ênfase em infraestrutura federada, com centros regionais interligados (LNCC, RNP, universidades) e acesso transparente via marketplace nacional de computação.

3. SinIA – Sistema Nacional de Inteligência Artificial

O SinIA foi reconhecido como eixo estruturante da política de IA no Brasil, com papel de articulação, governança e financiamento integrado.

Principais elementos:

- Governança federada com comitês técnicos temáticos (infraestrutura, dados, *frameworks*, ética).
- Fundos plurianuais articulando FNDCT, Finep e BNDES.
- Compras públicas de inovação como tração de mercado.

- Infraestrutura federada (LNCC, RNP, hubs regionais).
- Bibliotecas abertas e modelos pré-treinados em português.

Síntese: o SinIA deve ser o orquestrador da Pilha Nacional de IA, conectando academia, governo e setor produtivo em ecossistema coeso.

4. Infraestrutura e *software* abertos

A atualização do Sinapad e a criação do SinIA são vistas como oportunidade de consolidar uma plataforma aberta e interoperável, com:

- Uso de containers e orquestração open source (Kubernetes, Slurm, Apptainer).
- Financiamento de bibliotecas livres e modelos fundacionais abertos.
- Licenciamento que preserve soberania e rastreabilidade.
- Contrapartidas obrigatórias de ciência aberta para projetos financiados.

“Sem *software* aberto e infraestrutura compartilhada, não há soberania possível.”

5. Pilha Nacional de IA: desafios e caminhos de superação

Desafios técnicos e institucionais incluem:

- Falta de interoperabilidade entre *datasets* e *frameworks*.
- Ausência de financiamento contínuo e governança unificada.
- Carência de equipes permanentes de engenharia de *software* científica.

184

Caminhos propostos:

- Infraestrutura federada e marketplace nacional de recursos.
- Repositórios FAIR com curadoria e versionamento.
- Financiamento plurianual e reconhecimento acadêmico para *software* e dados abertos.

6. Formação de talentos e capacitação nacional

A formação profissional é tratada como pilar estratégico da soberania em IA. As contribuições propõem:

- Trilhas técnicas modulares por camadas da pilha (infraestrutura, dados, *frameworks*, modelos, aplicações).
- Certificações nacionais em IA, reconhecidas por MCTI/MEC e setor produtivo.
- Residências tecnológicas e parcerias com empresas.
- Capacitação desde a infância, ampliando a cultura STEAM.

“Precisamos formar profissionais desde cedo. A motivação começa na escola, não só na pós-graduação.”

Também são sugeridos programas de formação em HPC nos moldes do CI-Expert, e rede federada de capacitação, com polos regionais no Norte e Nordeste.

7. Métricas de impacto e avaliação de políticas

Para decisões de investimento, os comentários sugerem métricas integradas de impacto tecnológico, científico e social:

- Tecnológico: capacidade computacional, eficiência energética, *software* e interoperabilidade.
- Científico: produção acadêmica, replicabilidade, formação de talentos.
- Social e econômico: acessibilidade, inclusão, geração de empregos, sustentabilidade.

8. Governança e arranjos institucionais

A governança é entendida como tripartite (governo, academia, indústria), com arranjos federados regionais.

Propostas centrais:

- Conselho Nacional de IA com comitês técnicos temáticos.
- Hubs regionais focados em vocações locais (agro, clima, saúde, indústria).
- Fundos plurianuais e editais encadeados (bolsa + GPU + dados + OSS).
- Painéis públicos de transparência sobre execução e impacto.

185

9. Fomento e financiamento encadeado

As contribuições defendem a criação de instrumentos integrados que conectem:

- Formação de talentos,
- Infraestrutura HPC,
- Bibliotecas e *frameworks*,
- Modelos e datasets,
- Aplicações setoriais.

Com mecanismos como:

- Fundos plurianuais com *matching* público-privado.
- Compras públicas de inovação como tração.
- Microbolsas permanentes para manutenção de bibliotecas abertas.
- *Stage-gates* (ideia → protótipo → piloto → adoção).

10. Critérios técnicos e estratégicos para seleção de projetos

Critérios técnicos:

- Abertura e reprodutibilidade;
- Governança de dados e conformidade LGPD;
- Eficiência energética e segurança;
- Interoperabilidade e escalabilidade.

Critérios estratégicos:

- Redução de dependência externa;
- Formação de talentos e inclusão regional;
- Sustentabilidade e impacto socioeconômico.

Recomenda-se *scorecard* padronizado, revisado por comitê tripartite, com transparência pública.

11. Formação e ciência aberta

A formação técnica deve incorporar:

- Ciência aberta e ética como eixo transversal.
- Currículo baseado em tecnologias open source.
- Publicação obrigatória de datasets e modelos abertos.
- Reconhecimento acadêmico para *software* e dados como produtos científicos.

186

12. Integração com defesa e políticas de Estado

Várias contribuições destacam a necessidade de alinhar o desenvolvimento de IA com projetos estratégicos de Defesa (PROSUB, SISFRON, CIEMB), evitando duplicidade de esforços e assegurando visão de Estado, não de governo.

Também é enfatizada a integração com Institutos Federais, aproveitando sua capilaridade regional.

ANEXO 5. Respostas às perguntas levantadas por painel

Painel 1: Quais os caminhos para sustentar a IA com autonomia de infraestrutura tecnológica nacional?

1. Quais são os requisitos mínimos para uma infraestrutura capaz de suportar modelos avançados de IA, e como esses requisitos se comparam aos padrões internacionais?

Os requisitos mínimos incluem:

- *Hardware* de alto desempenho: GPUs, TPUs, *clusters* de servidores, sistemas de armazenamento paralelos.
- *Software* especializado: *frameworks* de IA, sistemas operacionais otimizados (como Linux), ferramentas de supervisão e observabilidade.
- Energia e conectividade: infraestrutura energética robusta e redes de alta velocidade.
- Segurança e governança de dados: proteção de dados sensíveis e soberania sobre o processamento.

Comparação com padrões internacionais:

- O Brasil ainda depende fortemente de importações de componentes críticos (GPUs, TPUs), majoritariamente dos EUA e da China.
- Países como China e Índia avançam com arquiteturas abertas (ex: RISC-V) e modelos fabless (design local, fabricação externa).
- A Europa, embora atrasada em relação à China, está investindo em European Chips Act e parcerias estratégicas.

187

Marcelo Pinheiro (Versatus HPC):

“Hoje, um equipamento de alto desempenho é, na verdade, uma rede com vários PCs. Claro que não são PCs comuns, são servidores, mas o cerne da tecnologia é oriundo de PCs, consoles de videogame e celulares. O processador mais poderoso da Nvidia, por exemplo, tem um processador ARM, o mesmo usado em celulares, junto com GPUs. Tudo isso foi bancado por indústrias de alta escala. Um supercomputador é definido por *software*, e a tecnologia que usamos é de PC.”

Leandro Rosa (Positivo Tecnologia):

“A Positivo fornece os maiores supercomputadores da América Latina em parceria com empresas líderes mundiais como Supermicro, Nutanix e Even. Temos competência para entregar soluções de HPC e IA generativa de última geração no Brasil, desde pequenas empresas até *hyperscalers*. A soberania nacional exige domínio sobre os dados, sobre a operação do *hardware* e sobre as aplicações que rodam nele.”

Carlos Rodrigo (SERPRO):

“A capacidade de processamento é latente e inerente à capacidade de entender o problema e oferecer a melhor solução da forma mais barata. A gente treinava modelos por 40 dias com CPU, e com GPU conseguimos reduzir para 4 horas. Desde então, investimos anualmente em aquisição de GPUs.”

2. Como você classifica o grau de soberania tecnológica do Brasil nas áreas de HPC/IA hoje? Quais deficiências críticas em *hardware* e *software* precisam de atenção prioritária, e em que janela de tempo elas poderiam ser superadas?

Grau de soberania atual: baixo a médio, com avanços pontuais em *software* e infraestrutura, mas alta dependência externa em *hardware*.

Deficiências críticas:

- *Hardware*: ausência de produção nacional de GPUs/TPUs e semicondutores avançados.
- *Software*: pouca contribuição brasileira em projetos open source críticos (ex: sistemas de arquivos paralelos).
- Capacitação: falta de profissionais altamente qualificados para operar e desenvolver soluções de IA.

188

Janela de superação:

- Curto prazo (1–3 anos): fortalecimento de *software* open source, capacitação técnica, uso estratégico de parcerias.
- Médio prazo (4–8 anos): desenvolvimento de chips básicos, infraestrutura nacional de data centers.

Longo prazo (10+ anos): produção nacional de componentes críticos e consolidação de soberania digital.

Luca Belli (FGV):

“O Brasil depende de CPUs, GPUs e TPUs importadas, produzidas por pouquíssimas empresas, basicamente todas nos Estados Unidos e, cada vez mais, na China. É irreal pensar que o Brasil criará GPUs que competem com a H100 da Nvidia no curto prazo. Mas podemos focar em semicondutores mais básicos, alavancando o open source e arquiteturas abertas como RISC-V.”

Marcelo Pinheiro:

“Em *hardware* é muito mais difícil. Uma das peças fundamentais é o sistema de armazenamento, os sistemas de arquivos paralelos. São supercomplexos e fundamentais para o funcionamento de um supercomputador. Os melhores projetos são

open source, mas o Brasil não tem grupos contribuindo neles. Seria importante dominar essa tecnologia.”

Guido Araújo (Unicamp):

“O Brasil pode construir uma indústria sem ter todos os componentes. A Embraer não fabrica motores nem aviônicos, mas tem uma indústria aeronáutica competitiva. Isso vale para HPC. Podemos desenvolver componentes e agregar valor com parceiros sólidos. Desenvolver uma GPU nos próximos 3 a 5 anos não é possível, mas desenvolver IPs e servidores nacionais é totalmente factível.”

3. Considerando a rede de universidades, empresas e ICTs do país, qual o papel de cada uma destas instituições e que mecanismos você proporia para que esta rede possa concretamente equacionar as deficiências levantadas na pergunta anterior?

Papéis:

- Universidades e ICTs: pesquisa básica e aplicada, formação de talentos, desenvolvimento de *software* e design de *hardware*.
- Empresas: produção, integração de soluções, gestão de data centers, comercialização.
- Governo: coordenação estratégica, financiamento, regulação e articulação de políticas públicas.

189

Mecanismos propostos:

- Plano nacional estruturado com metas de curto, médio e longo prazo.
- Fomento à inovação via editais e compras públicas.
- Criação de centros de competência e laboratórios nacionais.
- Estímulo à formação de startups e spin-offs universitárias.

Parcerias internacionais estratégicas com países do Sul Global e consórcios abertos.

Guido Araújo:

“A universidade é a raiz, a fonte e o grande nutriente da soberania. Sem um ITA, uma UFRJ, uma Unicamp, não existiria nada disso. Mas ela não mantém a soberania do Estado. São empresas e laboratórios nacionais que mantêm isso vivo.”

Carlos Rodrigo:

“É fundamental que haja ligação entre institutos de pesquisa, universidades, empresas, startups e governo. O papel do MCTI é essencial em colocar todos à mesa para entender os problemas e propor soluções conjuntas.”

Leandro Rosa:

“O Brasil tem boas competências hoje. Empresas públicas e privadas são capazes de complementar e se estruturar num programa nacional coordenado. O open source é uma forma de cooperar quando não se é líder.”

4. É possível desenvolver ou fabricar nacionalmente componentes críticos como GPUs e TPUs? Quais seriam os caminhos, desafios e limitações?

Possível? Sim, mas não no curto prazo.

Caminhos:

- Investir em design de chips (modelo fabless).
- Desenvolver IPs (propriedade intelectual) para componentes específicos.
- Focar em aceleradores especializados para inferência, onde há mais viabilidade.

Desafios:

- Alto custo de P&D.
- Escassez de mão de obra especializada.
- Dependência de fundições externas (ex: TSMC em Taiwan).

Limitações:

- Risco geopolítico (ex: Taiwan).
- Escala de produção.

Complexidade técnica e tempo de maturação (10+ anos).

Luca Belli:

“Podemos apostar no desenvolvimento de uma indústria de design de chips mais básicos e no chamado packaging (encapsulamento) no Brasil. No médio e longo prazo, podemos subir na cadeia de valor. Mas isso exige estabilidade e política industrial.”

Guido Araújo:

“Desenvolver uma GPU nos próximos anos não é possível, mas desenvolver componentes que possam ser vendidos no mercado internacional, como propriedade intelectual, é viável. A ARM faz isso há décadas sem fabricar chips.”

Marcelo Pinheiro:

“Em *hardware* é mais difícil, mas não impossível. O Estado precisa entrar diretamente, como fez com a Petrobras e a Embraer. É preciso um projeto de longo prazo, com empresas grandes e médias, criando um ecossistema.”

5. Como o Brasil poderia fomentar a fabricação de chips de HPC segundo arquitetura aberta, como por exemplo a RISC-V? Como estabelecer metas e cooperações nesse sentido?

Fomento à RISC-V:

- Apoio governamental a projetos de design de chips baseados em RISC-V.
- Parcerias com universidades e centros de pesquisa para formar especialistas.
- Criação de consórcios nacionais com empresas e ICTs.
- Investimento em infraestrutura fabless e packaging local.

Estabelecimento de metas:

- Definir *roadmaps* tecnológicos com entregas por etapas.
- Incentivos fiscais e financiamento via BNDES.
- Participação em fóruns internacionais de arquitetura aberta.

Estímulo à adoção em aplicações públicas (educação, saúde, segurança).

Guido Araújo:

“A China, em 2015, adotou como estratégia de Estado se juntar à RISC-V International. A Europa chegou tarde, mas está se movimentando com o European Chips Act. O Brasil pode desenvolver IPs nacionais baseados em RISC-V e estabelecer parcerias com universidades e empresas.”

Luca Belli:

“A aposta mais acertada é utilizar padrões e arquiteturas abertas como RISC-V, talvez não na produção, mas no design, como o modelo fabless. A Nvidia, por exemplo, não fabrica seus chips, apenas os projeta.”

6. Como as iniciativas públicas e privadas podem se articular para viabilizar investimentos em infraestrutura tecnológica nacional de IA?

Articulação possível via:

- Programas como o PBIA, que já apontam direções estratégicas.
- Contratações públicas orientadas à soberania digital.
- Modelos de PPPs (Parcerias Público-Privadas).
- Fomento à indústria nacional via compras governamentais.

Criação de ecossistemas regionais de inovação com ICTs, empresas e governos locais.

Leandro Rosa:

“É preciso um plano estruturado que unifique os melhores atores internos e externos e coopere para manter isso pujante no curto, médio e longo prazo. Sem coordenação e orquestração, falharemos.”

Carlos Rodrigo:

“O governo tem que estar presente para atender à população e regular. É fundamental a participação dos institutos de pesquisa, universidades e, mais importante ainda, o fomento a startups e novas empresas.”

Guido Araújo:

“O Estado brasileiro pode viabilizar infraestrutura de IA para empresas e startups, reduzindo custos e permitindo que elas gerem valor. Isso é essencial para reter talentos e fomentar inovação.”

7. Temos um foco grande na aquisição de supercomputadores de ponta. No entanto, a experiência internacional e nacional mostra que a mera existência do *hardware* não garante o avanço científico. Como transformar a aquisição de supercomputadores de ponta em um ecossistema sustentável, acessível, inclusivo e eficiente para a comunidade científica nacional?

Transformações necessárias:

- Capacitação massiva de profissionais (não apenas técnicos, mas também especialistas em aplicações).
- Acesso federado e democrático à infraestrutura para universidades e startups.
- Foco em aplicações de impacto social (saúde, educação, segurança).
- Desenvolvimento de *software* nacional para reduzir dependência de licenças proprietárias.
- Criação de centros de competência e redes de colaboração.
- Governança participativa e transparente da infraestrutura.

192

Painel 2: Como o Brasil pode atuar nas camadas superiores da Pilha de IA?

1. Em quais nichos ou camadas específicas da pilha de *software* (bibliotecas, *frameworks*, modelos pré-treinados, MLOps) o Brasil teria maior potencial competitivo para focar seus esforços de pesquisa e inovação?

O Brasil tem maior potencial competitivo nas seguintes camadas:

- Modelos pré-treinados especializados: Adaptados à realidade brasileira, com foco em setores estratégicos como saúde, agricultura, biodiversidade, energia e clima.

- Ferramentas de MLOps e governança: Observabilidade, auditabilidade e gestão de modelos e dados, com destaque para *frameworks* como Kubeflow.
- Interfaces e plataformas de IA como serviço: Ferramentas no-code/low-code que democratizam o uso da IA por pesquisadores não especializados.
- *Frameworks* e bibliotecas intermediárias: Incorporando diversidade linguística e cultural, com foco em nichos específicos e aplicações setoriais.

Essas áreas permitem avanços rápidos e aproveitam a capacidade brasileira de desenvolver *software* de alta qualidade, como destacado por Cleiton Rocha e Norberto Ferreira.

Norberto Alves Ferreira (CPQD):

“Nos casos de uma pilha de *software* de IA, dividindo aqui em seis camadas, que é uma visão interna nossa, a gente tem pouquíssimo trabalhado em *frameworks* básicos. A gente entende que não tem muita coisa para ser feita aqui. As oportunidades são de acompanhar e desenvolver algumas coisas pontuais. A coisa começa a criar maiores perspectivas de desenvolvimento e trabalho nas camadas mais superiores mesmo, em bibliotecas intermediárias, onde já podemos trazer algumas dessas bibliotecas junto com a nossa cultura, o nosso português, a nossa diversidade cultural, trazendo algumas diferenciações em determinadas áreas, em nichos específicos. Isso tem gerado resultados interessantes.”

Cleilton Rocha (Instituto Atlântico):

“Nós temos um potencial muito grande no Brasil de resolvermos essa problemática, porque exportamos mão de obra altamente capacitada para países desenvolvidos, onde o Brasil já é reconhecido mundialmente por fazer uma excelente entrega de produtos de *software*. Estados Unidos, Canadá, União Europeia brigam por talentos brasileiros, em detrimento de outros talentos de outras nações emergentes, porque nós temos uma maestria muito grande de desenvolver solução e entregar em alta qualidade esses *softwares* em produção.”

2. Quais são os principais desafios técnicos para desenvolver bibliotecas, *frameworks* e modelos pré-treinados no Brasil, e como superá-los?

Desafios técnicos:

- Falta de cultura de open source e interoperabilidade.
- Baixa densidade de startups de alta tecnologia.
- Escassez de repositórios públicos e padrões comuns.
- Dificuldade de reprodutibilidade entre instituições.
- Complexidade da integração entre dados, modelos e código.

Como superar:

- Criar plataformas compartilhadas com templates e repositórios públicos.
- Fomentar startups deep tech com apoio governamental.
- Estabelecer padrões de integração e colaboração.
- Investir em capacitação técnica e comunidades de prática.

Promover cooperação internacional e nacional coordenada.

Cleilton Rocha:

“Atualmente nós temos uma baixa cultura de open source system e poucas startups de alta densidade tecnológica. Isso, infelizmente, impacta na interoperabilidade e numa alta granularidade quando fazemos o desenvolvimento de soluções de IA para as nossas áreas temáticas, utilizando tecnologia nacional. Porém, nós temos um potencial muito grande no Brasil de resolvermos essa problemática. Precisamos garantir que esses modelos possam ser auditáveis, observáveis, assim como os dados também possam ter aplicada uma devida governança para que então consigamos auditar esses resultados e resolver alguns problemas que temos, como data drift e model drift.”

Norberto Alves Ferreira:

“As novas ferramentas de Kubernetes, Kubeflow, de MLOps têm criado oportunidades interessantes para aumentar e melhorar a observabilidade dos modelos, gerando registros para verificar LGPD, para perceber viés e tomar ações ao longo do tempo. Uma área muito interessante para a gente se envolver também e cuidar dentro da pilha.”

194

3. Que estratégias podem ser adotadas para garantir que modelos pré-treinados desenvolvidos no Brasil incorporem diversidade linguística, cultural e regional, promovendo a inclusão e soberania digital?

- Uso de dados nacionais abertos e representativos, como os do INPE.
- Treinamento de modelos médios (SLMs) com foco em português e contextos locais.
- Desenvolvimento de agentes setoriais especializados.
- Incorporação de multimodalidade (texto, imagem, som) com dados brasileiros.
- Fomento à pesquisa em NLP e multimodalidade com foco regional.

Parcerias com universidades e centros de pesquisa em diferentes regiões.

Norberto Alves Ferreira:

“O uso de modelos existentes e adaptados a modelos médios, configurados para o português e nossa cultura, prevendo em breve a integração da multimodalidade também nessas integrações, pode trazer resultados muito interessantes para soluções nacionais, incorporando as questões do português, da nossa cultura. Muitos desses modelos

poderiam ser muito úteis e produtivos para aumentar a eficiência de determinados setores do país.”

4. Quais mecanismos regulatórios e de governança são necessários e precisam ser desenvolvidos para garantir transparência, auditabilidade e confiabilidade nas ferramentas de aplicação de IA desenvolvidas no país?

- Ferramentas de MLOps com observabilidade e rastreabilidade.
- Governança de dados e modelos, incluindo LGPD e mitigação de vieses.
- Auditoria contínua de modelos em produção (ex: monitoramento de data/model drift).
- Regulação sobre uso de dados públicos e sensíveis.
- Modelos de governança colaborativa entre instituições públicas e privadas.

Infraestrutura de suporte à reprodutibilidade e interoperabilidade.

Cleilton Rocha:

“Precisamos garantir que esses modelos possam ser auditáveis, observáveis, assim como os dados também possam ter aplicada uma devida governança para que então consigamos auditar esses resultados e resolver alguns problemas que temos, como data drift e model drift. O governo precisa agir como catalisador e disseminador de uma cultura de sistemas open source e alavancar e ser o indutor de fomentar a criação de novas startups de densidade tecnológica para que então consigamos fazer a tropicalização de serviços.”

Igor Regis (Banco do Brasil):

“A lição principal, que eu acho que é o principal desafio que nós temos hoje, é quem administra essa camada que vai oferecer esta pilha num nível de excelência de serviço, como Samuel falou, que vai além de simplesmente entregar, seja processamento, seja biblioteca, mas entregar também capacitação, orientação do uso, a governança sobre quem é que pode fazer uso. O desafio é quem vai liderar isso aqui ter a visão de cliente, ter a visão de que essa camada precisa ser administrada como se fosse um serviço, um produto ofertado por uma população consumidora especializada, seja SUS, seja centro de pesquisa. Mas eles são clientes dessa camada que vai ofertar esse serviço.”

5. Considerando a inferência de modelos, quais cenários são mais relevantes para o desenvolvimento econômico e social do nosso país?

- Saúde pública (SUS): Diagnóstico, triagem, gestão hospitalar.
- Agricultura e biodiversidade: Monitoramento ambiental, uso de imagens de satélite.
- Educação: Ferramentas de apoio ao ensino e personalização.

- Energia e clima: Previsões, otimização de recursos.
- Justiça e políticas públicas: Análise de dados jurídicos e sociais.

Norberto Alves Ferreira:

“Os trabalhos têm sido muito intensivos e interessantes e têm se voltado para a criação de agentes setoriais. É onde a gente acha que o país vai ganhar mais: credenciar esses agentes para assuntos específicos, domínios específicos, facilitando o desenvolvimento de aplicações. E finalmente, é claro, a camada de cima, onde o Brasil inclusive já tem empresas e soluções sendo exportadas. E aqui estamos falando de chatbots, de portais, de muitos mecanismos de predição e diagnóstico, usando tudo que foi feito nas camadas de baixo.”

6. Que tipo de investimento estratégico o Brasil precisa fazer no desenvolvimento de plataformas, ferramentas de *software* livre e interfaces de usuário que efetivamente 'democratizem' o acesso às nossas infraestruturas de ponta? Como viabilizamos que um cientista de qualquer área possa, de forma intuitiva, executar seu experimento de IA em um supercomputador nacional?

- Plataformas de IA como serviço (IAaaS) com interfaces intuitivas.
- Ferramentas no-code/low-code baseadas em linguagem natural.
- Automação de toda a esteira de IA, incluindo AutoML.
- Infraestrutura compartilhada com suporte técnico especializado.
- Investimento em capacitação e suporte humano para pesquisadores não técnicos.
- Modelos de financiamento sustentáveis e colaborativos entre instituições.

196

Samuel Xavier de Souza:

“Uma das premissas do NPAD é que a gente torne a supercomputação uma ferramenta cotidiana do pesquisador. O resultado disso é que hoje temos centenas de publicações em revistas internacionais que foram apoiadas pelo NPAD e publicadas por pesquisadores de áreas que não são da computação. Isso só foi possível porque na criação do NPAD houve uma visão da reitoria de que precisaríamos ter uma equipe para dar suporte a esses pesquisadores. Essa visão foi uma visão que eu tentei passar para a reitoria, eles acataram e com isso conseguimos realmente alavancar o uso de poucas máquinas que conseguimos comprar com recursos FINEP, com recursos de projetos de infraestrutura, mas fazíamos um uso eficiente dessas máquinas.”

Cleilton Rocha:

“Com uma plataforma como serviço, entregando serviços de IA, conseguimos oferecer recurso para criação de sistemas inteligentes utilizando hoje interfaces que sejam no-

code ou que sejam low-code, facilitando assim a democratização e utilização da IA para os demais setores.”

7. Qual deve ser o papel dos centros de computação nacionais? Deveriam evoluir para se tornarem 'plataformas de IA como serviço'?

Sim. Os centros devem evoluir para:

- Oferecer ambientes pré-configurados e interfaces amigáveis.
- Atuar como provedores de IA como serviço, não apenas de horas de processamento.
- Fornecer suporte técnico e capacitação.
- Facilitar o uso por startups, pesquisadores e instituições públicas.
- Reduzir atritos e acelerar o ciclo de inovação.

Exemplos como o NPAD da UFRN e o INPE mostram que esse modelo já está em prática e é eficaz.

Samuel Xavier de Souza:

“Não basta termos apenas *software* e *hardware* funcionando. Precisamos de pessoas que consigam ajudar pessoas. Você não vai querer que essas startups virem especialistas em MLOps, que saibam usar essas máquinas, como configurar e executar um job específico de treinamento ou configurar um serviço de inferência. É preciso uma equipe que consiga ajudar, que compreenda a interface entre computação e a necessidade do pesquisador, para que ele possa usufruir da infraestrutura.”

Igor Régis:

“A plataforma do Banco do Brasil entrega tudo pronto, montado, com as escolhas difíceis e técnicas já definidas, garantindo a governança, a segurança da informação. O objetivo é entregar o acesso com todas as escolhas técnicas definidas, garantindo a governança, a segurança da informação. A plataforma também fomenta uma comunidade interna, entrega cursos, parcerias com universidades, treinamentos sobre as bibliotecas escolhidas. O objetivo é facilitar, mas não engessar.”

8. Que tipo de investimento em padronização e em plataformas compartilhadas precisamos fazer para criar um ecossistema de pesquisa em IA mais colaborativo e eficiente? Como garantimos que um experimento desenvolvido na UNICAMP possa ser replicado sem dor de cabeça na UFRN ou no SENAI Cimatec?

- Criação de repositórios públicos e templates de esteiras de IA.
- Padronização de ambientes e ferramentas (ex: Jupyter, Python, R).
- Infraestrutura interoperável entre instituições.
- Modelos de governança compartilhada e financiamento colaborativo.

- Ferramentas que garantam reprodutibilidade e auditabilidade.
- Apoio à ciência aberta e *software* livre.

Cleilton Rocha:

“Podemos desenvolver uma plataforma que seja compartilhada com repositórios públicos, com templates de esteira que possam alcançar múltiplos usuários. Como foi falado aqui pelo professor Edson Borin, temos diversos tipos de problemáticas onde a IA e seus diversos tipos de aprendizagem podem alcançar e entregar valor fazendo a diferença já lá na ponta. Então, termos repositórios públicos, a criação de padrões comuns de integração e de colaboração e de templates de esteira, aproveitando essa ambiência que estamos agora no ‘verão da IA’, para que possamos dizer não somente o que fazer, mas também como fazer.”

9. Como podemos usar nossa infraestrutura computacional nacional para fortalecer a indústria? Que tipo de 'pontes' são necessárias para que uma empresa brasileira veja mais valor em usar os recursos nacionais para treinar seus modelos do que simplesmente contratar uma plataforma de nuvem estrangeira?

Como fortalecer a indústria:

- Oferecer IA como serviço com suporte técnico e interfaces amigáveis.
- Reduzir custos e complexidade de uso.
- Garantir segurança e soberania dos dados.
- Criar incentivos fiscais e regulatórios para uso de infraestrutura nacional.
- Estabelecer parcerias público-privadas com foco em inovação.
- Fomentar startups deep tech com acesso facilitado à infraestrutura.

198

Pontes necessárias:

- Modelos de governança colaborativa entre centros públicos e empresas.
- Contratações públicas orientadas à inovação.
- Programas de aceleração e incubação com acesso à infraestrutura.
- Campanhas de valorização da infraestrutura nacional como diferencial competitivo.

Karine Ferreira:

“Soberania nacional: afirmo aqui que hoje o INPE tem soberania nacional nessa parte de observação da Terra. Organizamos esse big data na nossa instituição, temos todo esse centro de dados para o processamento e toda essa pilha, um ambiente de ciência de dados com métodos de machine learning e deep learning para processar. Hoje não somos dependentes da AWS, do Google Cloud, da Microsoft, do Planetary Computer e

de iniciativas europeias como a ESA. Isso considero uma soberania que o governo federal tem que realmente perseguir.”

Cleilton Rocha:

“O governo precisa agir como catalisador e disseminador de uma cultura de sistemas open source e também alavancar e ser o indutor de fomentar a criação de novas startups de densidade tecnológica para que então consigamos fazer a tropicalização de serviços. Recentemente tivemos o investimento de uma gigante global, a Databricks, em uma startup local aqui, que foi a Indício. Temos visto que quando o Brasil começa a se desenvolver, quando começa a ter serviços de qualidade, começamos a ter a atração de investidores globais e de outras nações. Isso prova nossa competência e também que podemos agir em colaboração, para que essas startups, para que essas consultorias possam se desenvolver localmente e fazer a entrega para os nossos consumidores locais.”

Painel 3: Como o Brasil está posicionado no desenvolvimento de IA para dispositivos móveis e embarcados?

1. Quais os *frameworks* e bibliotecas estão sendo desenvolvidos ou adaptados no Brasil para suportar inferência de IA na borda, e como garantir sua compatibilidade com padrões internacionais e segurança?

199

No Brasil, estão sendo utilizados e adaptados:

- *Frameworks* tradicionais otimizados para borda, como TensorFlow Lite, PyTorch Mobile e ONNX Runtime.
- Bibliotecas ultraleves para microcontroladores, dispensando sistema operacional e otimizadas para consumo mínimo de memória e energia.
- Projetos como o RIAC (Unicamp) têm desenvolvido ambientes de aprendizado federado e testbeds para inferência distribuída, com foco em compatibilidade e segurança.

A compatibilidade com padrões internacionais é garantida por:

- Adoção de tecnologias open source amplamente utilizadas.
- Participação em redes internacionais de pesquisa.

Desenvolvimento de modelos auditáveis e explicáveis, especialmente em áreas críticas como saúde.

Leandro Villas (Unicamp):

“Em relação a *frameworks* e bibliotecas para inteligência artificial, nós temos *frameworks* tradicionais da inteligência artificial adaptados para borda, que visam otimizar e ter portabilidade. Esses *frameworks* já suportam técnicas de otimização, como quantização,

poda, aceleração via *hardware*. Eles são muito utilizados para aplicações móveis e IoT. Também temos bibliotecas ultra leves para microcontroladores, que são bibliotecas otimizadas, dispensando o uso de sistema operacional. Dessa forma, reduzimos o consumo de memória e energia, sendo bastante utilizadas em projetos de TinyML.”

2. Quais são os principais desafios técnicos para executar modelos de IA em dispositivos com restrições de memória, processamento e energia, e como a pesquisa nacional tem avançado em técnicas como quantização, pruning e destilação de modelos?

Desafios técnicos:

- Memória e CPU limitadas (ex: microcontroladores com <1MB de RAM).
- Consumo energético crítico, especialmente em dispositivos alimentados por bateria.
- Latência e conectividade intermitente, mesmo em regiões próximas a centros urbanos.

Avanços da pesquisa nacional:

- Compressão de modelos: Redução de tamanho sem perda significativa de acurácia.
- Quantização, pruning e destilação: Aplicadas com sucesso em projetos como os da Venturos e RIAC.
- Otimização de modelos: Foco em simplicidade e eficiência, com resultados superiores em alguns casos.
- Explicabilidade: Aplicações em saúde com redes neurais que mostram áreas de atenção em imagens médicas.

Fernando dos Santos (Venturos):

“Quando estamos falando a respeito de dispositivos móveis embarcados, o que queremos de fato é que sejam modelos que rodem com uma capacidade computacional menor. Não teremos super servidores disponíveis na borda. Essas aplicações precisam ter um tempo de resposta menor para que tenhamos eficiência na operação das empresas e do usuário. Podemos responder isso por meio de técnicas como compressão de modelos, onde pegamos modelos altamente complexos e, através de vários mecanismos, conseguimos reduzir essa complexidade. A performance dos modelos comprimidos é basicamente a mesma dos modelos complexos. Outro aspecto importante é a otimização de modelos. A abordagem mais simplificada deu resultados melhores do que a abordagem mais complexa. Isso possibilita que modelos complexos, como redes neurais, possam ser explicáveis, o que é altamente atraente para a área da saúde.”

Leandro Villas:

“Quais são os desafios técnicos para inteligência artificial na borda? Memória e processamento limitados, especialmente em microcontroladores e sensores embarcados que operam com menos de 1 MB de RAM e capacidades de CPU restritas. O consumo de energia é crítico, pois a maioria desses dispositivos depende de bateria. Dado esses desafios, temos técnicas de otimização como quantização de modelos, poda, destilação de conhecimento. O principal é reduzir o modelo para diminuir o uso de memória e o custo computacional, com pequenas perdas de acurácia. Já realizamos vários trabalhos que mostram essa pequena perda de acurácia.”

3. Quais arquiteturas e protocolos estão sendo explorados no Brasil para viabilizar inferência distribuída em ambientes de edge computing, considerando latência, consumo energético e confiabilidade?

- TinyML: Para execução de modelos em dispositivos com recursos extremamente limitados.
- Federated Learning: Treinamento distribuído com dados locais, preservando privacidade e reduzindo dependência da nuvem.
- Testbeds de aprendizado federado: Desenvolvidos no RIAC para simular e validar soluções distribuídas.
- Protocolos de comunicação eficientes: Adaptados para ambientes com conectividade limitada, como áreas rurais e industriais.

201

Leandro Villas:

“Temos o TinyML e o Federated Learning, que são tecnologias complementares que concretizam a visão de inteligência artificial na borda. O Federated Learning permite que o modelo seja treinado nos dispositivos locais, onde os dados são mantidos localmente, com interação entre os clientes e o servidor no envio do modelo atualizado a cada rodada, no qual ele é agregado e retornado ao cliente até ter um modelo treinado. Desenvolvemos e estamos aprimorando no RIAC um *testbed* para aprendizado federado, que será disponibilizado nos próximos meses para a comunidade poder desenvolver, analisar e testar soluções federadas.”

4. Quais setores — como saúde, agricultura, segurança pública e educação — podem se beneficiar da IA embarcada, e como políticas públicas podem acelerar a adoção dessas soluções em escala nacional?

Setores com alto potencial:

- Saúde: Assistência médica em áreas remotas com LLMs embarcados em tablets.
- Agricultura: Monitoramento de máquinas e precisão no campo, mesmo sem conectividade.
- Segurança pública: Câmeras com IA embarcada para análise em tempo real.

- Educação: Plataformas digitais com IA embarcada para personalização do ensino.

Políticas públicas recomendadas:

- Definição de setores prioritários com indicadores claros.
- Uso de instrumentos como Lei de TICs, FINEP, Embrapii, PPIA.
- Fomento à capacitação técnica e à criação de demanda nacional.
- Governança estratégica com comitês dedicados à IA embarcada.

Mateus Pierre (Instituto ELDORADO):

“Podemos pensar em duas vertentes. Primeiro, com o olhar de serviços públicos que o Brasil oferta hoje para a população. A riqueza de dados que o SUS tem hoje é única. Como desenvolver modelos para levar mais inteligência para a saúde, especialmente para regiões mais distantes, sem depender de conectividade ou nuvem, mas que um tablet possa auxiliar um médico em áreas remotas. Fora outros segmentos como o agronegócio. Pensando na parte de empresas, com as megatendências de hiperpersonalização, o dispositivo ser capaz de recomendar serviços com base no perfil específico de uso da pessoa. Temos instrumentos como a Lei de TICs, Embrapii e outras ferramentas para incentivar as empresas a fazer pesquisa no país, desenvolver modelos no país, trazer para o Brasil desenvolvimentos que hoje são feitos fora.”

Fernando dos Santos:

“Na Venturos, mapeamos oportunidades nos últimos anos. Quatro itens que focamos: *hardware* nacional com desenvolvimento de robôs; análise de vídeos, aplicável ao setor de óleo e gás, por exemplo, para respostas imediatas em caso de acidentes; voz e linguagem, especialmente para o português brasileiro, com modelos comprimidos em dispositivos embarcados; e agentes com RAG, modelos que funcionam com processamento de texto em dispositivos como celulares.”

Leandro Villas:

“Na parte de cidades inteligentes, posso dar como exemplo a segurança pública, que pode fazer uso de câmeras com inteligência artificial embarcada. Certamente teremos uma restrição maior em termos de *hardware* e poder de processamento para embarcar essas soluções, comparado com outros países. No estado de São Paulo, temos a cana-de-açúcar com várias fazendas próximas à região metropolitana, inclusive de Campinas, que não têm conectividade. A agricultura de precisão já possui plantadeiras, tratores com telemetria e sensores, mas deixa de prover serviços por falta de conectividade. Seria necessária uma parceria com operadoras para levar conectividade ao campo.”

5. Como os programas nacionais como a Lei de TICs, PADIS e a "Nova Indústria Brasil" (NIB) poderiam ser aperfeiçoados/fortalecidos para priorizar a criação de

hardware especializado para IA, uma vez levantadas as prioridades nacionais nesse tema?

- Estímulo à demanda nacional por chips especializados, criando mercado interno.
- Inclusão de design de circuitos integrados como prioridade, além do encapsulamento.
- Incentivo à adoção de arquiteturas abertas como RISC-V, com projetos já em andamento.
- Inserção de *hardware* de IA em políticas de PPB (Processo Produtivo Básico), como foi feito com memória.
- Fomento à propriedade intelectual nacional em microeletrônica.

Mateus Pierre:

“A indústria da microeletrônica é muito forte na parte de encapsulamento e menos na parte de design de circuito integrado. O Eldorado já desenvolveu chips para fora, com dois projetos nesse sentido. Para o Brasil, falta estímulo para criar um mercado que consuma esses projetos de chip de IA. Já tem muita coisa pronta lá fora, então por que uma indústria procuraria o Brasil para desenvolver? Um exemplo é a memória, que foi estimulada no Brasil com a inclusão no PPB. Com chip de IA, o caminho é parecido. Não vamos concorrer com GPUs da Nvidia, mas podemos desenvolver tecnologia nacional para casos específicos. O RISC-V é uma grande opção para isso. Temos instrumentos como PADIS, Lei de TICs, Embrapii, FINEP. O passo que falta é criar esse mercado para fomentar o desenvolvimento no país.”

203

6. Como equilibrar privacidade de dados com a capacidade de fazer inferência local no dispositivo? Devemos enxergar a borda como um aliado da proteção de dados ou seria apenas como um alvo de ataques?

- Uso de Federated Learning: Treinamento local sem envio de dados brutos.
- Inferência embarcada: Dados permanecem no dispositivo, reduzindo riscos.
- Modelos explicáveis e auditáveis: Especialmente em áreas sensíveis como saúde.
- Governança de dados e compliance com LGPD: Integrada à arquitetura de IA embarcada.

Leandro Villas:

“A inteligência artificial na borda permite que os dados sejam mantidos localmente nos dispositivos e não sejam transmitidos para a nuvem. Com isso, conseguimos uma latência reduzida, uma privacidade aprimorada, funcionalidades offline, o que é muito importante para o contexto do Brasil em termos de conectividade.”

Fernando dos Santos:

“A simplicidade do modelo possibilita que redes neurais possam ser explicáveis, o que é altamente atraente para a parte da saúde, por exemplo, em que temos imagens de câncer de mama e podemos ver o que a rede neural está prestando atenção.”

Mateus Pierre:

“Ter uma governança dessa estratégia é essencial. Um comitê de IA especialmente para definir a direção, setores prioritários com indicadores claros, instrumentos para viabilizar o desenvolvimento e focar na capacitação para termos pessoas que saibam desenvolver e aplicar essas tecnologias em demandas do setor público e privado.”

7. Devemos enxergar a borda como um aliado da proteção de dados ou seria apenas como um alvo de ataques?

A borda deve ser vista como aliada da proteção de dados, desde que:

- Modelos e dispositivos sejam auditáveis e seguros.
- Infraestruturas embarcadas sejam projetadas com foco em privacidade.
- Capacitação técnica inclua segurança cibernética embarcada.
- Políticas públicas incentivem boas práticas de segurança desde o design.

Embora a borda possa ser alvo de ataques, seu potencial para reduzir exposição de dados sensíveis é estratégico para soberania digital.

204

Painel 4: Ciência aberta, open source e cooperações internacionais: limites e potenciais

1. Quais são os principais requisitos técnicos e institucionais para garantir que projetos de ciência aberta em IA sejam responsáveis, sustentáveis, auditáveis e interoperáveis?

Requisitos técnicos:

- Infraestruturas interoperáveis: Sistemas que dialoguem entre si, com padrões abertos e metadados estruturados.
- Segurança by design e by default: A segurança deve ser parte integrante da arquitetura desde o início.
- Auditorias técnicas: Ferramentas e processos que permitam rastreabilidade e verificação dos modelos e dados.
- Padrões técnicos nacionais e internacionais: Certificações e protocolos que garantam confiabilidade e compatibilidade.

Requisitos institucionais:

- Governança de dados estruturada: Com equilíbrio entre abertura e proteção, como destacou Bianca Amaro.

- Capacitação de pesquisadores: Formação em práticas de ciência aberta, gestão de dados e uso de repositórios.
- Valorização na avaliação científica: Incentivo à adoção de práticas abertas como critério de mérito.
- Infraestrutura pública distribuída: Como a rede de HPC do Paraná e o SINAPAD, com governança federada.

Bianca Amaro (MCTI):

“A ciência aberta está baseada inicialmente, classicamente, nesses pilares. Quando se fala em ciência aberta, hoje em dia, se relaciona muito diretamente que ciência aberta é dados abertos de pesquisa. Bom, sim, é também, mas não é só. O primeiro desafio é disseminar a cultura e a prática da ciência aberta. Existe ainda uma grande resistência em praticar a ciência aberta. A ideia sempre calcada da ciência aberta é que aquilo que é desenvolvido com recursos públicos deve tornar-se público, salvaguardadas exceções que devem ser apenas exceções. Temos que criar as infraestruturas necessárias para suportar a ciência aberta. Capacitar os pesquisadores nas práticas de ciência aberta. Considerar as práticas de ciência aberta como um fator positivo na avaliação desses mesmos pesquisadores. A arquitetura de dados deve contemplar infraestruturas organizadas e que dialoguem entre si. Tornar esses dados estruturados e interoperáveis. Criar uma governança de dados de forma aberta, porém também soberana.”

205

Danielle Ayres (UFSC / GSI):

“Dois pontos que hoje são centrais quando a gente pensa não só numa ciência aberta, num open source, numa cooperação, mas quando a gente pensa em todo um ecossistema do que será o desenvolvimento da inteligência artificial para o Estado. Primeiro deles: *security by design, security by default*. O segundo é criar elementos capazes de auditar sistemas e sistema capazes de serem auditados. Incorporar isso, ganhamos em segurança e em conscientização do que importa para nós nessa auditoria. A governança deve proteger nossos dados, mas também criar padrões para certificação dos nossos sistemas, com base nos padrões internacionais e nas nossas necessidades.”

2. Como as iniciativas em redes de pesquisa, consórcios internacionais e programas multilaterais podem ser estruturadas para fortalecer a presença brasileira na construção da pilha global de IA?

- Participação ativa em fóruns multilaterais: Como ONU, G20, BRICS, OCDE, promovendo a visão brasileira de soberania digital e ciência aberta.
- Criação de consórcios regionais: Exemplo do SCALAC (Sistema de Computação de Alto Rendimento da América Latina e Caribe), que conecta centros de HPC.

- Programas como o PBIA: Com eixos voltados à cooperação internacional e compartilhamento de infraestrutura.
- Parcerias com países complementares: Como a cooperação Brasil–Índia em HPC, que envolve transferência de tecnologia e montagem local.
- Promoção da diplomacia da inovação: Articulada pelo Itamaraty, conectando ecossistemas brasileiros a oportunidades globais.

Phillippe Olivier Navaux (UFRGS / SCALAC):

“O PBIA tem um eixo de infraestrutura e desenvolvimento na área de compartilhamento de infraestrutura brasileira com outros países em desenvolvimento e tem disponibilizado 50 milhões para o desenvolvimento de cooperação na América Latina e Caribe com o Brasil. O SCALAC é o sistema de computação de alto rendimento da América Latina e Caribe, que seria o Sinapad latino-americano. A ideia é expandir esse trabalho do SCALAC, com o LNCC junto, para integrar essa cooperação entre os Sinapads e estender isso para toda a América Latina.”

Guilherme Fitzgibbon (MRE):

“A resolução da ONU criou tanto o Diálogo Global sobre Governança da IA, que será um fórum de debate multissetorial, como também um painel de especialistas em IA, composto por 40 especialistas de todo o mundo, para dar uma função consultiva aos países para que desenvolvam suas próprias capacidades em inteligência artificial. Isso está de acordo com nossa visão brasileira de desenvolvimento soberano das capacidades de IA do jeito que nós podemos fazer.”

206

3. Como o Brasil pode contribuir para o desenvolvimento de padrões abertos e protocolos técnicos que garantam interoperabilidade entre sistemas de IA desenvolvidos em diferentes países?

- Desenvolvendo certificações nacionais: Como sugerido por Daniele Ayres, com base em padrões internacionais, mas adaptados às necessidades brasileiras.
- Participando da definição de padrões multilaterais: Em fóruns como a ONU, OCDE e BRICS.
- Apoiando redes como a LA Referencia e Rede CLARA: Que já trabalham com interoperabilidade de metadados e repositórios na América Latina.
- Criando e compartilhando repositórios abertos e auditáveis: Com dados confiáveis e estruturados para treinar modelos de IA.
- Estimulando o uso de *software* livre e open source: Como parte da pilha nacional de IA.

Bianca Amaro:

“A arquitetura de dados deve contemplar infraestruturas organizadas e que dialoguem entre si. Houve um boom de criação de sistemas de informação de forma desordenada, que não são interoperáveis. Está presente esse termo hoje em dia: interoperabilidade. Temos que tornar esses dados estruturados e interoperáveis. Criar uma governança de dados para uma ciência cidadã.”

Danielle Ayres:

“A governança deve proteger nossos dados, mas também criar padrões dos quais queremos que nossos dados e sistemas sejam certificados. Por que não ter uma certificação que seja nossa, de Estado, que nós vamos certificar o nosso sistema, a partir dos padrões internacionais e também daquilo que entendemos ser necessário para nós?”

4. Quais são os critérios técnicos e éticos que devem orientar a participação brasileira em projetos internacionais de IA, especialmente no compartilhamento de dados, modelos e infraestrutura computacional?

Critérios técnicos:

- Interoperabilidade e auditabilidade.
- Segurança cibernética e proteção de dados sensíveis.
- Infraestrutura soberana ou controlada nacionalmente.
- Capacidade de adaptação (tropicalização) das soluções.

207

Critérios éticos:

- Soberania sobre os dados gerados no território nacional.
- Equidade no acesso e uso dos dados compartilhados.
- Transparência nos acordos de cooperação.
- Respeito à diversidade cultural e regional.

A participação deve ser guiada por interesses estratégicos e pela capacidade de gerar valor para o país, como enfatizado por vários painelistas.

Bianca Amaro:

“A ciência deve ser tão aberta quanto possível e tão fechada quanto necessário. Existem dados sensíveis, dados de soberania nacional, dados que devem ser classificados e que não serão exportados ou visibilizados para todo o mundo. Esses dados, em relação a todos os dados da ciência, são uma pequena parcela. A soberania de dados refere-se ao fato de que um país deve ter autoridade sobre seus dados que são gerados, coletados, armazenados e processados dentro do seu território.”

Danielle Ayres:

“A resposta é por essência cooperativa. A resposta tem que ser multidimensional. Os arranjos cooperativos não podem ser feitos sem que tenhamos consciência efetiva das nossas demandas e dos nossos interesses. Devem ser arranjos coletivos que atendam aquilo que temos estrategicamente como demanda enquanto Estado, do ponto de vista protetivo, produtivo e tecnológico.”

5. Como o Brasil deve posicionar sua cooperação em IA com parceiros da América Latina, do Sul Global e com países desenvolvidos? Qual o potencial de colaboração com blocos como os BRICS?

- América Latina e Sul Global: Priorizar parcerias com países que compartilham desafios e demandas semelhantes. Exemplo: proposta de banco de dados sul-americano.
- Países desenvolvidos: Buscar cooperação com foco em complementaridade e transferência de conhecimento, evitando dependência tecnológica.
- BRICS: Fortalecer parcerias com países como Índia, que oferecem cooperação tecnológica real e não apenas comercial. Exemplo: cooperação em HPC e microeletrônica.
- Multilateralismo ativo: Defender a soberania digital e padrões abertos em fóruns como ONU, G20, OCDE.
- Diplomacia da inovação: Promover a imagem do Brasil como parceiro confiável e líder regional em ciência aberta e IA.

208

Guilherme Fitzgibbon:

“Temos que começar a procurar parcerias que nos ajudem a desenvolver capacidades locais, não necessariamente produzir essas capacidades locais. Temos uma rede chamada SECTEC, setores de ciência, tecnologia e inovação nas nossas embaixadas e consulados no exterior. Um exemplo é a parceria com a Malásia para desenvolver o setor de semicondutores brasileiros. A declaração do BRICS sobre inteligência artificial fala bastante sobre ciência aberta, código aberto, inovação aberta. O Brasil tem interesse em firmar parcerias, mas temos regras e uma maneira própria de fazer isso.”

Jorge Edison:

“Se a questão for soberania, dependência, temos que escolher melhor nossos parceiros. Trabalhar com a Índia é uma coisa bem diferente. É uma parceria real, uma vontade de trabalhar junto. É uma via de mão dupla com os indianos. É uma oportunidade para nós entrarmos na área de RISC-V de maneira mais séria. Eles já têm processadores *multicore* desenvolvidos na Índia, como o Shakt e o Vega.”

Danielle Ayres:

“Privilegiar relações com a América Latina, criar um grande banco de dados sul-americano seria algo muito interessante do ponto de vista da cooperação. Não na ideia

de líder impositor, mas na ideia de que estamos desenvolvendo tecnologia e seríamos muito importantes nesse setor. Os BRICS podem ser uma das alternativas, como trouxe o nosso colega do Paraná, o HPC deles foi feito com a Índia.”

Philippe Olivier Navaux:

“Nós aqui todos — professores, instituições, governo — temos que fazer um esforço muito grande para colocar na nova geração um sentimento de capacidade. Os nossos alunos são bons, a formação deles é boa. Nós temos que mudar essa mentalidade de que o que é bom é lá fora. Nós como formadores temos que mudar essa mentalidade.”

6. O modelo de governança federado de infraestrutura em HPC para atender demandas da Ciência, como o usado na Europa, seria viável no Brasil e/ou na América Latina?

Sim, e já existem iniciativas nesse sentido:

- SCALAC: Rede latino-americana de centros de HPC, semelhante ao SINAPAD brasileiro.
- Projeto do Paraná com a Índia: Rede de HPC com governança distribuída entre universidades estaduais.
- PBIA e SINAPAD: Estrutura nacional que pode ser expandida para uma governança federada regional.
- Viabilidade técnica e política: Requer coordenação entre instituições, financiamento sustentável e alinhamento estratégico.

A governança federada permite compartilhamento de recursos, otimização de investimentos e fortalecimento da soberania científica regional.

Philippe Olivier Navaux:

“O SCALAC é o sistema de computação de alto rendimento da América Latina e Caribe, que seria o Sinapad latino-americano. A ideia é expandir esse trabalho do SCALAC, com o LNCC junto, para integrar essa cooperação entre os Sinapads e estender isso para toda a América Latina. Estamos trabalhando para juntar essa cooperação entre os Sinapads aqui e com autenticação do SINAPES expandir isso para América Latina toda, tentando integrar essa cooperação.”

Jorge Edison:

“Nosso projeto planeja construir uma rede de HPCs com tecnologia indiana, transferida a uma empresa nacional. Todos os computadores serão montados inicialmente no Brasil. O projeto prevê uma rede de nove supercomputadores interligados por uma rede de alta velocidade a 400 Gbps, instalados nas sete universidades estaduais do Paraná e em um instituto de pesquisa. Isso é uma prova de conceito, uma interação Sul-Sul realmente acontecendo nessa área de tecnologia da informação.”

Painel 5: Avanços PBIA: construindo o novo Sinapad e o Sistema Nacional de Tecnologia e Inovação em IA (SinIA)

1. Quais métricas de impacto tecnológico, científico e social devem ser priorizadas para orientar decisões estratégicas sobre investimento em IA?

As métricas prioritárias mencionadas ou inferidas no painel incluem:

- Eficiência computacional: Medida por tempo de processamento, consumo energético (ex: *flops* por joule), e otimização de recursos computacionais.
- Utilização da infraestrutura: Taxa de ocupação de GPUs e supercomputadores, tempo de fila, e ociosidade.
- Produção científica e inovação: Número de publicações, patentes, e modelos desenvolvidos com base na infraestrutura nacional.
- Capacitação e formação: Número de pesquisadores treinados, pós-graduandos, e usuários atendidos.
- Impacto social e setorial: Aplicações em saúde, educação, agricultura, segurança pública, e indústria.
- Sustentabilidade financeira e energética: Custo de operação, manutenção, e viabilidade de longo prazo.

210

Fábio Borges (LNCC):

“A camada, quando a gente fala das camadas superiores, normalmente é onde tem recurso do comércio, é onde tem transações financeiras, e as camadas inferiores normalmente são subsidiadas pelo governo. As camadas de baixo normalmente precisam ser subsidiadas porque elas não têm valor econômico agregado, mas são fundamentais. Quando a gente está nas camadas mais baixas, não dá para ter um algoritmo guloso. Você tem que ter algoritmos bem pensados, bem formulados, bem trabalhados com muitos detalhes matemáticos, com teoremas. Outra frase interessante é: ‘Quem não computa não compete’. Essa frase está sendo muito usada pela Comunidade Europeia hoje. E realmente a gente precisa computar.”

André Rafael (MCTI):

“O Sinapad não é só a provisão de serviços de processamento, é também desenvolver a rede de pesquisa e de produção de produtos de alto desempenho. O plano de investimentos precisa ser adotado com inteligência. O Brasil ficou atrasado e agora a gente tem que correr atrás do prejuízo. O PBIA é o começo. Precisamos sinalizar quais são as outras agendas de desenvolvimento que precisam buscar mais recurso.”

2. Como o SinIA pode ser estruturado para articular políticas públicas, financiamento à inovação e a Pilha de IA Nacional, promovendo sinergia entre academia, governo e setor produtivo?

O SinIA pode ser estruturado com:

- Governança federada e colaborativa, envolvendo universidades, centros de pesquisa, empresas e governo.
- Articulação com o Sinapad para integrar camadas inferiores (infraestrutura) e superiores (aplicações e serviços).
- Mapeamento de demandas setoriais para orientar investimentos e desenvolvimento de modelos especializados.
- Fomento à inovação aberta e ciência aberta, com repositórios públicos e interoperáveis.
- Capacitação técnica e suporte contínuo para uso eficiente da infraestrutura.
- Modelos de financiamento sustentáveis, incluindo créditos computacionais, parcerias público-privadas e uso de nuvens privadas.

André Rafael:

“O Sinapad vai pensar muito nas camadas baixas, com foco nas camadas mais baixas e preparando para as camadas mais altas. O SinIA, como as necessidades de serviços de infraestrutura dos setores são muito variadas, vai pensar um pouco mais na organização dessas camadas mais altas. O sistema do SinIA é mais amplo, mas ainda está em projeto e discussão. Estamos observando experiências internacionais, como da Índia, dos Estados Unidos, da Europa. Acreditamos que em um médio prazo conseguimos oferecer uma capacidade de computação de alto desempenho à altura do que o Brasil precisa.”

211

3. Quais têm sido as principais influências e demandas da indústria nacional para o PBIA? Como elas têm sido atendidas?

Demandas da indústria:

- Capacidade de processamento de alto desempenho, especialmente para setores como petróleo, química, e manufatura.
- Acesso a infraestrutura nacional com menor custo.
- Modelos especializados e APIs nacionais para evitar dependência de soluções estrangeiras.
- Eficiência energética e sustentabilidade.

Como têm sido atendidas:

- Expansão do Sinapad com novos centros e governança mais próxima do MCTI.

- Projetos como o do LNCC e RNP que oferecem infraestrutura compartilhada e conectividade.
- Planejamento de integração com o setor privado via créditos computacionais e uso de nuvens privadas.
- Desenvolvimento de modelos nacionais com dados locais e soberanos.

André Rafael:

“Estamos sendo cobrados da presidência para oferecer esse serviço para a comunidade, inclusive integrando com outros atores, não só federais, mas também privados. O ministério está sendo cobrado para prover capacidades para o setor privado, trabalhar junto do setor privado. A ideia é que o Sinapad seja expandido nesse sentido. Temos que pensar em todas as camadas, tanto a estratégia de investimentos para aquisição dos produtos que já estejam no mercado, quanto a melhor forma de gestão.”

4. Como garantir que o novo Sinapad e o SinIA promovam tecnologias open source e a soberania tecnológica da Pilha de IA nacional?

- Fomento à ciência aberta e repositórios públicos.
- Desenvolvimento de *frameworks* e bibliotecas nacionais com base em *software* livre.
- Capacitação em tecnologias abertas e incentivo à contribuição em projetos open source.
- Infraestrutura pública interoperável e auditável, com governança transparente.
- Evitar dependência de APIs e modelos proprietários estrangeiros, oferecendo alternativas nacionais.
- Políticas públicas que priorizem o uso de tecnologias abertas em compras e projetos estratégicos.

212

Fábio Borges:

“O LNCC tem trabalhado para fortalecer esse sistema. Ganhamos um projeto de 50 milhões que não ficou 1 centavo para o LNCC, só para as unidades Sinapad. Isso mostra o quanto a equipe trabalha para o Brasil. A maioria das aplicações que usamos hoje vieram de fora. Uma empresa vai fazer um app, usa uma API para uma IA estrangeira porque é mais barato. Se você tem uma infraestrutura nacional que oferece essa API de graça para as empresas, elas vão usar aqui. Essa frase da Comunidade Europeia é importante: ‘Quem não computa não compete’.”

5. Como balancear a eficiência dos modelos estrangeiros fechados com o incentivo à construção de modelos nacionais?

- Oferecer APIs nacionais gratuitas ou subsidiadas para empresas e pesquisadores.

- Investir em infraestrutura nacional para treinamento e inferência.
- Desenvolver modelos médios (SLMs) especializados em português e contextos locais.
- Fomentar reuso e adaptação de modelos open source internacionais com tropicalização.
- Criar incentivos para uso de modelos nacionais em serviços públicos e aplicações estratégicas.
- Promover benchmarking e avaliação comparativa entre modelos nacionais e estrangeiros.

Fábio Borges:

“Uma empresa vai fazer um app, usa uma API para uma IA estrangeira porque é muito mais barato. O custo de fazer uma LLM, uma API, é caríssimo. Se você tem uma infraestrutura nacional que oferece essa API de graça para as empresas, elas vão usar aqui. Para que ela vai pagar um centavinho lá? Eles estão usando os nossos dados, treinando mais, usando isso em robôs, em várias outras coisas. Se você fizer bem a conta, vai ver que o custo é muito maior do que parece.”

6. Quais são os desafios técnicos e institucionais para consolidar uma plataforma colaborativa de dados, modelos, bibliotecas e *frameworks* de IA para a construção da Pilha entre instituições brasileiras?

213

Desafios técnicos:

- Interoperabilidade entre sistemas e repositórios.
- Conectividade de alta velocidade entre centros (solução proposta pela RNP).
- Escalonamento eficiente de tarefas e compartilhamento de recursos.
- Gestão de dados sensíveis e compliance com LGPD.
- Capacitação para uso de ferramentas como MLFlow, Jupyter, etc.

Desafios institucionais:

- Governança federada e transparente.
- Sustentabilidade financeira e energética.
- Coordenação entre múltiplos atores (academia, governo, empresas).
- Fomento à cultura de colaboração e ciência aberta.
- Evitar fragmentação e duplicação de esforços.

A proposta de uma plataforma federada com governança compartilhada e suporte técnico contínuo foi destacada como caminho viável.

José Ferreira (RNP):

“A RNP interconecta universidades federais, estaduais, centros de pesquisa, parques tecnológicos, num número superior a 600 instituições. A RNP já vem oferecendo computação em nuvem, contêineres, e agora infraestrutura de GPUs para pesquisadores. Além da oferta, a parte mais importante talvez seja a questão da conectividade. A proposta envolve não só conectividade, mas também recursos computacionais e armazenamento, que é extremamente importante para datasets e repositórios de dados. A governança envolve revisar de tempos em tempos, monitorar os usos dos recursos, otimizar ao máximo o uso dos recursos. A capacitação será focada no uso da infraestrutura: como extrair o máximo, quais bibliotecas, qual forma de trabalho mais bem adaptada ao seu caso. A sustentabilidade é algo que discutimos o tempo todo. Quando acabar o projeto, como é que a gente faz? Quem continua pagando a conta?”

Fábio Borges:

“O Santos Dumont é top 1 *green* na América Latina. A fila mais rápida demora 24 horas para começar a rodar. Tem muita gente que não quer rodar lá porque tem um problema pequeno e vai demorar. Quando distribuímos GPUs por aí em projetos, estamos perdendo não só as GPUs, mas energia. Se conseguimos agregar isso, temos otimização na refrigeração. Falta escalonar nossas tarefas. Estamos perdendo tempo de máquina e tempo de gente.”

214

Painel 6: Capacitação e formação técnica para a construção de uma Pilha de IA soberana

1. Quais competências técnicas são prioritárias para formar profissionais capazes de atuar nas diferentes camadas da pilha de IA, desde infraestrutura até aplicações avançadas?

As competências prioritárias incluem:

- Fundamentos de computação: Arquitetura de computadores, sistemas operacionais, redes, hierarquia de memória.
- Programação paralela e distribuída: OpenMP, MPI, CUDA, OpenCL, TVM.
- Infraestrutura de HPC e GPUs: Conhecimento prático em supercomputadores e arquiteturas como RISC-V.
- Matemática e estatística: Essencial para modelagem e compreensão de algoritmos de IA.
- Fundamentos de IA e aprendizado de máquina: Redes neurais, SVM, KNN, deep learning.
- Big Data e engenharia de dados: Manipulação, armazenamento e processamento de grandes volumes de dados.

- IA responsável e ética: Compreensão dos impactos sociais, culturais e éticos da IA.
- Capacidade de adaptação: Para lidar com a rápida evolução das ferramentas e técnicas.

Alba Melo (UnB):

“O profissional de IA deve ter conhecimento dos fundamentos: computação, microeletrônica, matemática e estatística. Tem que saber o que é um sistema computacional, os componentes CPU, hierarquia de memória, entrada e saída, GPU, interconexão de computadores (*clusters*), programação paralela (OpenMP, MPI, Socket), e programação de placas gráficas (CUDA, OpenCL, SYCL). Tem que conhecer inteligência artificial e aprendizado de máquina: KNN, SVM, redes neurais, redes neurais profundas. Tem que saber dimensionar a infraestrutura computacional e se adaptar ao surgimento de novas técnicas.”

Wagner Meira (UFMG):

“Qualidade é inegociável. Temos que continuar tendo bons cursos. Todos os bacharelados plenos reconhecidos pela SBC — ciência da computação, sistemas de informação, engenharia de computação, engenharia de *software*, ciência de dados e inteligência artificial — são completamente pertinentes porque já são *AI-enabled*. O mercado diz que nossos alunos ainda se formam muito ingênuos. É necessário ter estratégias mais agressivas de integração com problemas reais e dados reais.”

215

2. As universidades e centros de pesquisa estão preparados para atender às demandas emergentes da soberania tecnológica em IA? Como eles podem adaptar seus currículos e programas de pós-graduação?

Situação atual:

- Algumas universidades, como UnB, UFG e UFMG, já criaram ou estão criando bacharelados em IA.
- Há iniciativas de ênfases em IA dentro de cursos tradicionais.
- Pós-graduações estão sendo adaptadas com linhas específicas em IA e HPC.

Adaptações recomendadas:

- Integração entre cursos: Mobilidade entre áreas como ciência da computação, engenharia, estatística e matemática.
- Projetos reais e interdisciplinares: Para conectar teoria com prática e fomentar inovação.
- Internacionalização: Estágios e cotutelas com centros de excelência internacionais.

- Atualização curricular constante: Para acompanhar a volatilidade da pilha de IA.
- Formação em IA responsável e ciência aberta.

Evânio Araújo (MEC):

“Temos três frentes de trabalho: promoção dos cursos de graduação em IA, fomento à pós-graduação com bolsas de mestrado e doutorado, e construção de uma infraestrutura nacional de dados da educação. Estamos discutindo com a AGU a criação de um *sandbox* regulatório para desenvolvimento de IA na educação. A criação da Secretaria de Gestão da Informação, Inovação e Avaliação de Políticas Educacionais no MEC foi justamente para pensar dados, inovação e uso de tecnologia nos processos educacionais.”

Anderson Soares (UFG / SEIA):

“Criamos o bacharelado em IA em 2019. Reduzimos conteúdo e aumentamos projetos, acreditando que no mundo atual as pessoas podem aprender até sozinhas. Temos 3% de evasão num curso difícil de exatas. Os alunos têm acesso à variedade de projetos e infraestrutura de ponta, como o chip da Blackwell. Isso dá condições de trabalho e opções para que o aluno pense em ficar, mesmo tendo ofertas internacionais.”

3. Como fomentar a capacitação a partir da criação de trilhas técnicas e certificações nacionais em IA, alinhadas com as demandas para a construção da Pilha Nacional de IA e necessidades do setor produtivo?

216

- Trilhas técnicas por nível de complexidade:
 - *Letramento digital*: Para crianças e jovens.
 - *Cursos de curta duração (40–80h)*: Técnicas de prompt, uso de ferramentas.
 - *Residência TIC e pós-graduação*: Formação profunda em IA e HPC.
- Certificações nacionais:
 - Alinhadas com competências da pilha (infraestrutura, *frameworks*, modelos, aplicações).
 - Reconhecidas por setor produtivo e academia.
- Programas como CI Expert (microeletrônica) adaptados para IA e HPC.
- Parcerias com SENAI, universidades e empresas para formação prática e certificação.

Luciana Gomes (CNPq):

“O CNPq atua com bolsas de vocação tecnológica, iniciação científica, mestrado e doutorado. Destaco o programa de mestrado e doutorado acadêmico para inovação, onde o projeto do aluno é de interesse do setor empresarial. Também temos programas

de fixação de mestres e doutores, como o RHAE, o DCR (em parceria com FAPs), e o Inova Talentos (em parceria com o IEL/CNI), que já tem bolsas de até R\$ 12.000. Estamos atentos à necessidade de flexibilizar a dedicação exclusiva e ampliar parcerias com o setor privado.”

Otávio Ciribelli (Petrobras):

“Hoje temos 15 supercomputadores em operação. Em 2026, planejamos aumentar de 3% para 13% a capacidade dedicada à IA. Temos ações concretas como o programa Autonomia e Renda Petrobras Tecnologia, em parceria com o SENAI, para capacitar 10.000 pessoas online e 420 presencialmente, com foco em minorias. Internamente, já temos 2.000 profissionais capacitados na nossa pilha de IA e 9.000 usando o hub de modelos da companhia.”

4. Como garantir que a formação técnica em IA esteja alinhada com práticas de ciência aberta, uso de tecnologias open source e desenvolvimento ético e responsável?

- Inserção de temas como IA responsável, ética e soberania digital nos currículos.
- Uso de ferramentas open source em disciplinas e projetos.
- Capacitação em repositórios públicos e ciência aberta.
- Fomento à pesquisa em IA explicável e auditável.
- Parcerias com centros como LNCC, SCALAC e Rede Clara para acesso a dados e infraestrutura aberta.
- Programas de bolsas e projetos com foco em ciência aberta e soberania tecnológica.

Alba Melo:

“O profissional de IA deve conhecer ferramentas abertas como OpenCL e SYCL, além de CUDA. A formação deve incluir fundamentos e não apenas uso de ferramentas. A capacitação deve ser alinhada com ciência aberta e uso de tecnologias open source. Sugiro estágios em centros de supercomputação nacionais e internacionais, como o LNCC e o Barcelona Supercomputing Center.”

Otávio Ciribelli:

“Estamos tentando atacar a dependência de *software* proprietário. A pluralidade de arquiteturas é uma prioridade, incluindo RISC-V. Nosso modelo Petronemo é treinado com dados internos, sem chamadas de API externas. Isso é soberania: estar no controle, poder escolher quando instalar, atualizar ou flexibilizar.”

5. Como criar uma rede nacional de capacitação contínua e estruturada para cientistas e especialistas de outras áreas estratégicas, especialmente no Norte e Nordeste?

- Programas de capacitação interdisciplinares: Voltados para profissionais de saúde, química, direito, agronomia, etc.
- Infraestrutura compartilhada: Acesso remoto a supercomputadores e GPUs (ex: Santos Dumont, Pegaso).
- Cursos modulares e escaláveis: Online, híbridos, com certificação.
- Parcerias com universidades regionais e ICTs.
- Fomento a projetos de capacitação regional: Via CNPq, FAPs, MEC e MCTI.
- Criação de polos de excelência e programas de verão em centros como CEIA, LNCC, UFMG.

Evânio Araújo:

“Precisamos pensar numa grande política nacional de requalificação. Alguns profissionais terão ganhos enormes com a complementariedade da IA, mas outros precisarão chegar a um novo patamar. Estamos discutindo uma nova política nacional de educação superior, mais aderente às inovações tecnológicas.”

6. No contexto de capacitação em HPC, não seria interessante a criação de um programa nos moldes do CI Expert (microeletrônica), onde alunos de graduação aprendessem sobre a infraestrutura de supercomputadores, incluindo a pilha de *software*, e a programação em GPUs?

218

Sim, e foi sugerido diretamente no painel:

- Programa nacional de formação em HPC para IA, inspirado no CI Expert.
- Módulos práticos em centros de supercomputação (LNCC, Santos Dumont, Petrobras).
- Estágios em empresas e centros internacionais.
- Foco em programação paralela, otimização, uso de GPUs e *frameworks* como TVM.
- Integração com programas de pós-graduação e certificações técnicas.

Alba Melo:

“Sugiro que estudantes de graduação em ciência da computação e engenharia sejam treinados em infraestrutura de HPC para IA, nos moldes do programa CI Expert. Sugiro estágios em centros de supercomputação nacionais como o LNCC e internacionais. Isso permitiria que os alunos usassem ferramentas em problemas reais e não apenas em problemas toy.”

Otávio Ciribelli:

“A prática do HPC acadêmico já entregou quatro supercomputadores nacionalmente. Qualquer cidadão pode ter acesso a essas máquinas. A nova máquina

que será entregue em 2025 terá arquitetura mais plural, incluindo RISC-V. Estamos investindo em capacitação com problemas reais, comunicados publicamente via GitHub.”

7. Como equilibrar a formação em fundamentos de IA (aprendizado de máquina, redes neurais, big data) com a capacitação em ferramentas específicas, que mudam rapidamente?

- Priorizar fundamentos sólidos: Matemática, estatística, algoritmos, redes neurais.
- Ensinar capacidade de adaptação: Para lidar com novas ferramentas e paradigmas.
- Projetos práticos com problemas reais: Para aplicar fundamentos em contextos dinâmicos.
- Atualização curricular constante: Com envolvimento de academia, indústria e governo.
- Formação em ciência aberta e uso crítico de ferramentas: Evitar dependência de soluções proprietárias.
- Combinar cursos teóricos com capacitações técnicas de curta duração.

219

Alba Melo:

“Não adianta conhecer o que tem agora. Os profissionais têm que ser capazes de se adaptar ao que virá. A formação deve priorizar fundamentos e complementar com capacitação em ferramentas específicas.”

Wagner Meira:

“A pilha não é uma pilha, é algo muito mais volátil. A LLM já virou commodity. A parte mais complexa da IA talvez não sejam os modelos, mas a IA responsável. A pilha dobra de tamanho quando se fala em IA responsável. A formação sólida é o que vai nos dar capacidade de resistir à volatilidade da pilha.”

Painel 7: Conexões que sustentam a IA: Onde estão as alianças/colaborações estratégicas nacionais para a construção de uma Pilha de IA soberana?

1. Quais são os principais arranjos de colaboração nacional atualmente em curso que contribuem para o desenvolvimento de diferentes camadas da Pilha de IA, e como podem ser ampliados ou replicados?

Arranjos em curso:

- Consórcio liderado pelo LNCC: Foco na substituição de camadas da pilha de HPC por versões open source, envolvendo cerca de 200 pesquisadores de diversas instituições nacionais.
- SCALAC (Sistema de Computação de Alto Rendimento da América Latina e Caribe): Rede federada de supercomputadores latino-americanos, com memorandos de entendimento já assinados.
- BRICS HPC & AI Working Group: Cooperação internacional com países do BRICS para desenvolvimento de infraestrutura e pesquisa em IA.
- Parcerias com universidades e centros de excelência: Unicamp, UFRN, UFPE, UFSC, UnB, entre outras, atuando em áreas como compiladores, sistemas operacionais, segurança, tolerância a falhas, cloud federada, etc.
- Magalu Cloud: Iniciativa privada nacional que oferece infraestrutura de nuvem e colabora com institutos de pesquisa em open source e virtualização.
- CNPq e FNDCT: Apoio a INCTs e bolsas de pesquisa, embora com limitações orçamentárias.

Ampliação e replicação:

- Criação de mini supercomputadores para pesquisa: Permitir experimentação em camadas da pilha sem afetar ambientes de produção.
- Mapeamento nacional de grupos de pesquisa em HPC e IA: Identificação de 102 grupos em 75 instituições para orientar futuros editais.
- Fomento à ciência aberta e infraestrutura compartilhada: Para democratizar o acesso e estimular colaborações.
- Integração com polos regionais e hubs de inovação: Para capilarizar o desenvolvimento e evitar concentração geográfica

Carla Osthoff (LNCC):

“Hoje temos cerca de 170 projetos de pesquisa em 12 estados e cerca de 2.000 usuários no supercomputador Santos Dumont. A pilha de *software* utilizada é baseada em *software* aberto, mas com distribuições customizadas por fabricantes. Propomos um consórcio para substituir cada uma dessas camadas por distribuições abertas. Já temos colaborações com grupos de excelência como o grupo do professor Navaux (UFRGS), o CEFET-RJ, Unicamp, UFRN, UnB, UFPE, UFSC, RNP, entre outros. Estamos fazendo um levantamento com a SBC de 102 grupos de pesquisa em HPC em 75 instituições para mapear projetos relacionados à pilha de IA.”

Renan Gaya (MGI):

“Nos últimos três anos, unimos MCTI, FINEP, MGI e ENAP para configurar problemas da área de governo e lançar editais para identificar startups nacionais que possam endereçar

esses problemas com IA. Já temos 25 projetos em andamento em órgãos como Saúde, Anvisa, MAPA, ANS. Estão previstos novos editais, inclusive na área de meio ambiente.”

Deybson Cipriano (Confederação ASSESPRO):

“A ASSESPRO tem mais de 3.000 empresas associadas, 95% delas micro, pequenas e médias. Temos capilaridade nacional e presença em todos os estados. Contribuímos com disseminação de conhecimento, promoção de debates, representação do setor junto ao governo e conexão entre empresas e startups. Apoiamos centros acadêmicos como o SEIA (UFG), Unicamp, UFPE, entre outros.”

2. Como o Brasil pode estruturar mecanismos de governança colaborativa entre universidades, institutos/centros de pesquisa e órgãos governamentais para garantir continuidade e impacto das parcerias em IA?

Propostas e caminhos discutidos:

- Criação de um conselho nacional de IA (como o proposto “Conapia” ou o futuro “SinIA”): Inspirado no modelo do CGI.br, com gestão tripartite (governo, academia, sociedade civil).
- Governança federada e multissetorial: Com representação formal de universidades, centros de pesquisa, empresas e governo.
- Planejamento de longo prazo: Com metas claras, roadmap tecnológico e financiamento contínuo (10–15 anos).
- Articulação interministerial: Via Comitê de Transformação Digital da Casa Civil, envolvendo MCTI, MEC, MGI, FINEP, CNPq, etc.
- Uso estratégico de compras públicas e encomendas tecnológicas: Para estimular a demanda e garantir perenidade dos projetos.

Teresa Ludermir (UFPE):

“Para autonomia em IA, precisamos de coordenação multissetorial: governo, academia e sociedade civil. Sem ciência de primeira linha, não há soberania tecnológica. A universidade tem papel fundamental: fazemos pesquisa de excelência, formamos talentos e entregamos produtos. Proponho uma governança tripartite, inspirada no modelo do CGI.br, com participação de governo, sociedade civil e academia. Precisamos de financiamento e planejamento de longo prazo — não se forma um pesquisador em cinco anos.”

Renan Gaya:

“A soberania não será uma inércia, ela precisa ser política pública. O governo já criou instrumentos como o Marco Legal das Startups, a nova Lei de Licitações, a Lei de Governo Digital. Mas esses instrumentos ainda são pouco utilizados. Precisamos investir mais em governança para fazer o ‘match’ entre demanda e oferta. O governo planejou gastar

R\$ 15 bilhões em tecnologia em 2024, a maior parte em infraestrutura, serviços e *software*.”

3. Como integrar iniciativas regionais e locais — como polos de inovação e redes estaduais — ao esforço nacional de construção da pilha de IA, garantindo capilaridade e inclusão?

Propostas e ações discutidas:

- Criação de hubs regionais de IA: Especialmente no Norte e Nordeste, com infraestrutura, formação e projetos locais.
- Residências tecnológicas e programas de capacitação descentralizados: Para formar talentos em regiões menos atendidas.
- Infraestrutura compartilhada e federada: Como o Santos Dumont e futuras máquinas do Sinapad, acessíveis nacionalmente.
- Parcerias com redes estaduais e universidades locais: Para fomentar pesquisa aplicada e desenvolvimento regional.

Inclusão digital e científica: Com foco em diversidade, equidade e acesso a dados e ferramentas.

Teresa Ludermir:

“O plano tem que pensar em hubs regionais, especialmente no Norte e Nordeste. É necessário criar nós de dados, residências tecnológicas e infraestrutura computacional para formar pessoas e garantir inclusão. Os recursos financeiros do PBIA parecem muito poucos para a academia. Não estamos vendo os recursos chegarem aonde deveriam, especialmente nas regiões Norte e Nordeste.”

Deybson Cipriano:

“A infraestrutura computacional compartilhada de IA pode oportunizar instituições e empresas que não possuem capacidade de investir altos valores. A ASSESPRO está presente em centros acadêmicos de referência e leva demandas de mercado para aproximar a academia da realidade da população.”

4. Para construir uma Pilha de IA soberana e duradoura, que tipo de arranjos institucionais e modelos de governança de longo prazo precisamos fomentar? Como podemos criar um *framework* onde a academia, o LNCC e a indústria não apenas colaborem, mas estejam formalmente integrados na gestão e na definição estratégica do uso de nossa infraestrutura tecnológica nacional?

Modelos sugeridos:

- Criação de uma “Petrobras da IA” ou “Embraer da IA”: Uma instituição central com missão estratégica de desenvolver e coordenar a infraestrutura nacional de IA.

- *Framework* de integração formal:
 - Academia: Produção científica, formação de talentos, pesquisa de base.
 - LNCC e centros de HPC: Infraestrutura, coordenação técnica, desenvolvimento da pilha.
 - Indústria: Aplicações, demanda, inovação, financiamento.
- Governança colaborativa e perene:
 - Conselhos com representação de todos os setores.
 - Planejamento estratégico com metas e indicadores.
 - Financiamento contínuo via FNDCT, parcerias público-privadas e compras públicas.
- Foco em soberania tecnológica e ciência aberta: Com desenvolvimento de *software* nacional, uso de dados locais e infraestrutura auditável.

Teresa Ludermir:

“Precisamos de uma estrutura de governança colaborativa, com gestão tripartite. Sugiro a criação de um conselho nacional para IA (ConaPIA), inspirado no CGI.br. Esse conselho deve definir estratégias de longo prazo, garantir financiamento contínuo e promover a inclusão regional. A academia está pronta para contribuir em todas as camadas da pilha, da infraestrutura ao desenvolvimento de ciência.”

223

Carla Osthoff:

“O projeto que propomos envolve cerca de 200 profissionais acadêmicos. A ideia é substituir camadas da pilha por equivalentes abertos, com desempenho similar. Precisamos de uma estratégia nacional clara: quais objetivos queremos atingir, em quanto tempo, e quais são as prioridades — HPC, sistemas embarcados, etc.”

Renan Gaya:

“As empresas públicas como SERPRO e DATAPREV têm papel fundamental. Elas garantem perenidade em soluções tecnológicas personalizadas para o governo. Precisamos utilizar mais os instrumentos já criados e fomentar a articulação entre governo, academia e setor privado.”

Kiko Reis (Magalu Cloud):

“O maior desafio é a falta de demanda brasileira. Os grandes sistemas de IA são detidos por poucas empresas estrangeiras. O Brasil precisa fomentar grandes agregadores de demanda. Se houver demanda, o recurso aparece. A China resolveu isso quebrando a internet e criando um ecossistema local. O Brasil precisa fazer apostas estruturantes e fomentar o aparecimento de grandes agregadores de demanda para alinhar a pilha tecnológica.”

Painel 8: Como retroalimentar o Plano Brasileiro de Inteligência Artificial, conectando ciência, indústria e governo no financiamento da soberania digital brasileira em IA?

1. Quais são os principais gargalos nos atuais instrumentos de financiamento à IA, e como novos modelos de fomento público-privado podem superá-los?

Principais gargalos identificados:

- Falta de continuidade e perenidade dos investimentos: Muitos editais são pontuais e não sustentam projetos de longo prazo.
- Burocracia excessiva nas universidades e fundações: Dificulta parcerias com empresas e desestimula pesquisadores.
- Baixa integração entre academia, governo e setor privado: Ainda existem abismos institucionais que dificultam colaboração.
- Dificuldade de acesso ao crédito por pequenas empresas e startups: Especialmente em projetos de alto risco tecnológico.
- Falta de critérios claros para seleção de projetos estratégicos.

Modelos que podem superar esses gargalos:

- Fomento em cascata: Como sugerido pela professora Ana Estela, com missões estratégicas (ex: IA para saúde pública) guiando o financiamento.
- Editais em fluxo contínuo: Já adotados pela FINEP, permitem maior flexibilidade e reenvio de propostas.
- Bonificação por integração com ICTs: FINEP oferece taxas menores para empresas que contratam ICTs ou mestres/doutores.
- Fundos de investimento em IA: Como o fundo de R\$ 100 milhões anunciado pela Berta Capital com apoio do MCTI.
- Parcerias público-privadas com foco em infraestrutura e dados: Ex: centros de IA em saúde com dados do SUS.

Joana Souza de Meirelles (FINEP):

“Quando falamos de financiamento, temos os gargalos clássicos: crédito, garantia, acesso das empresas menores aos recursos, e a própria disponibilidade de recursos não reembolsáveis. A integração é o melhor caminho. Temos ampliado os FIPs e os fundos para investimento em empresas de base tecnológica. A FINEP tem atuado com editais em fluxo contínuo, permitindo maior tempo para arranjos entre empresas e ICTs, maior flexibilidade na documentação e possibilidade de ressubmissão de projetos.”

João Paulo Papa (UNESP / Instituto Eldorado):

“Muitos colegas têm perdido projetos com empresas por conta da burocracia das universidades. A fundação quer 30%, depois vem o imposto de renda, e o professor acaba ganhando R\$ 0,00 para desenvolver um projeto de cinco anos. Precisamos mudar essa mentalidade e desburocratizar. Professores não são filantropos. Há gargalos que desestimulam a colaboração com o setor privado.”

Ana Estela Haddad (Ministério da Saúde):

“Os desafios e gargalos são muitos e de várias naturezas. No SUS, por exemplo, temos que trocar as asas do avião com ele voando. A migração para uma nuvem de governo exige segurança e continuidade. Às vezes, temos que dar passos mais lentos, mas seguros. A arquitetura de interoperabilidade e a Rede Nacional de Dados em Saúde (RNDS) são avanços importantes, mas exigem investimentos contínuos e planejamento cuidadoso.”

2. Como o Plano Brasileiro de IA pode incorporar mecanismos de financiamento contínuo e multissetorial para sustentar projetos nas camadas da Pilha de IA, da infraestrutura ao nível de aplicação?

Propostas discutidas:

- Integração com a Nova Indústria Brasil (NIB): Missões industriais orientadas por IA, como transformação digital da indústria.
- Uso estratégico do FNDCT: Com recursos não reembolsáveis e crédito via FINEP e BNDES.
- Criação de instrumentos híbridos: Subvenção + crédito + investimento de risco.
- Articulação entre ministérios e agências: MCTI, MGI, FINEP, BNDES, Casa Civil, etc.
- Planejamento de longo prazo com metas claras: Evitar “soluções” de política pública e garantir continuidade.
- Fomento à infraestrutura nacional de dados e nuvem pública: Como o projeto Inspire (R\$ 390 milhões).

Joana Souza de Meirelles:

“A decisão de investir em IA precisa ser perene, passando por todos os governos. O PBIA, encomendado diretamente pelo presidente Lula, traz eixos bem estruturados e indicação de fontes e objetivos. O FNDCT foi descontingenciado, permitindo recursos não reembolsáveis contínuos. A aprovação do PL 847 possibilitou acesso ao estoque do FNDCT, mais R\$ 22 bilhões aplicáveis ao crédito.”

Luíza Sidônio (BNDES):

“O BNDES atua na Nova Indústria Brasil (NIB), que tem missões como a transformação digital da indústria. O plano Mais Produção busca ser o ‘plano safra da indústria’, com

continuidade e perenidade. São R\$ 611 bilhões em quatro anos, com participação de BNDES, FINEP, Caixa, Banco do Brasil, BNB, BASA. O BNDES apoia principalmente infraestrutura e integradores, onde há maior adensamento empresarial e barreiras de entrada mais baixas.”

3. Quais são os critérios técnicos e estratégicos para selecionar projetos de IA que contribuam efetivamente para a soberania digital brasileira, e como garantir que esses critérios estejam alinhados com as prioridades nacionais?

Critérios sugeridos:

- Alinhamento com missões estratégicas nacionais: Ex: saúde pública, educação, segurança, meio ambiente.
- Uso de dados brasileiros e infraestrutura nacional.
- Integração com ICTs e contratação de mestres/doutores.
- Potencial de impacto social e econômico.
- Capacidade de gerar conhecimento e inovação de base.
- Contribuição para a autonomia tecnológica e redução de dependência externa.

Garantia de alinhamento:

- Governança multissetorial: Com participação de academia, governo e setor produtivo.
- Avaliação técnica e estratégica dos projetos.
- Mapeamento de competências e demandas regionais.
- Monitoramento e avaliação contínua dos resultados.

Joana Souza de Meirelles:

“Nos editais de subvenção, exigimos que a empresa proponente traga uma ICT como parceira. No crédito, oferecemos bônus de taxa (TR + 2,5%) para empresas que contratem ICTs ou mestres/doutores. Projetos devem ter relevância, externalidade e risco tecnológico. A integração entre ciência, academia e empresas é fundamental para que projetos estruturantes saiam do papel.”

Ana Estela Haddad:

“O SUS pode ser um estudo de caso. Missões como IA para redução da mortalidade materno-infantil na Amazônia Legal, vigilância epidemiológica em tempo real, personalização do tratamento oncológico com algoritmos treinados em dados brasileiros. O SUS pode adotar soluções, gerar novos dados e alimentar novas pesquisas, criando um ciclo virtuoso.”

4. Como integrar os instrumentos de financiamento à formação de talentos, à infraestrutura computacional, ao desenvolvimento de bibliotecas/*frameworks*, e à

produção científica, criando um ecossistema sustentável e competitivo de IA no Brasil?

Propostas discutidas:

- Fomento à ciência básica e aplicada: Como destacado pelo professor João Paulo Papa.
- Centros de excelência com financiamento contínuo: Ex: Centro de IA em cardiologia da UFMG com dados do SUS.
- Editais que exigem parceria com ICTs: FINEP e BNDES já adotam esse modelo.
- Bonificação por contratação de mestres/doutores: Estímulo à formação e retenção de talentos.
- Infraestrutura compartilhada e nacional: Nuvem pública, data centers verdes, supercomputadores.
- Programas de residência tecnológica e pós-graduação em IA.

Joana Souza de Meirelles:

“A FINEP estimula a contratação de mestres e doutores nos projetos financiados. A integração entre instrumentos — subvenção, crédito, investimento — é essencial. Editais em fluxo contínuo permitem maior flexibilidade e tempo para formação de arranjos entre empresas e ICTs.”

227

João Paulo Papa:

“Precisamos identificar grupos que precisam de IA, mas não têm conhecimento técnico, e conectá-los com quem desenvolve IA básica. A ciência básica precisa de fomento perene. Os melhores ecossistemas envolvem empresa, universidade e governo. Projetos com o Ministério da Saúde e a Fundação Gates têm funcionado bem, mas exigem aporte financeiro e agilidade.”

Rafael Moreira (Bertha Capital):

“Abert Capital investe em startups com professores universitários como sócios e conselheiros. Criamos o primeiro fundo de investimento em participações para startups de IA com R\$ 100 milhões, usando recursos da Lei de Informática. O Brasil precisa conectar estratégias empresariais, academia e políticas públicas. A IA é uma mudança de paradigma tecnológico e exige coordenação multissetorial.”

5. Existem centros de pesquisa no Brasil que já demonstraram um sucesso notável em atrair milhões de reais em investimento privado para o uso de sua infraestrutura computacional. Como podemos escalar esses casos de sucesso e transformá-los em uma política de fomento? Que instrumentos podemos desenvolver para alavancar o investimento público e estimular parcerias público-privadas que

garantam a sustentabilidade financeira e a modernização contínua desses centros tecnológicos, que são cruciais para a nossa soberania em IA?

Casos de sucesso mencionados:

- Centro de IA em cardiologia da UFMG: Com 11 milhões de laudos e financiamento contínuo.
- Instituto Eldorado: Parcerias com empresas e governo, atuação em ciência aplicada.
- Berta Capital: Fundo de R\$ 100 milhões para startups de IA com apoio do MCTI.

Instrumentos sugeridos para escalar:

- Fomento em rede e arranjos colaborativos: Empresas + ICTs + governo.
- FIPs temáticos em IA: Com recursos da Lei de Informática e FNDCT.
- Editais de centros de P&D com contrapartida privada: Como os lançados pelo BNDES.
- *Sandbox* regulatórios e compras públicas para inovação.
- Modelos de financiamento híbrido e flexível: Subvenção + crédito + investimento.
- Mapeamento e replicação de boas práticas: Identificar centros com alto impacto e criar políticas de incentivo.

228

Ana Estela Haddad:

“O núcleo de Telessaúde da UFMG, com apoio do Ministério da Saúde, já realizou 11 milhões de laudos de eletrocardiograma. Estamos investindo em um centro de IA e predição em cardiologia com base na nossa própria população. Esse é um exemplo de financiamento contínuo e parceria público-acadêmica que pode ser replicado.”

Joana Souza de Meirelles:

“A FINEP está estruturando editais voltados à pilha de *software* e tecnologias digitais, com foco em subvenção econômica e integração com ICTs. A atuação conjunta com o BNDES e outros parceiros é essencial para escalar casos de sucesso e garantir sustentabilidade.”

Luíza Sidônio:

“O BNDES está estudando a criação de um FIP específico para IA. Já apoiamos centros de P&D com chamadas públicas, como a recente para centros tecnológicos, que teve demanda enorme e já aumentamos o valor de R\$ 3 bilhões para mais de R\$ 9 bilhões.”

Painel 9: Próximos passos para o desenvolvimento da Pilha de IA Nacional

1. Quais são os principais consensos e divergências que emergiram ao longo dos painéis sobre os caminhos para a construção de uma pilha de IA soberana no Brasil?

Consenso:

- Necessidade de soberania tecnológica: Há ampla concordância sobre a importância de o Brasil dominar tecnologias estratégicas em todas as camadas da pilha de IA — da infraestrutura ao nível de aplicação.
- Importância da coordenação multissetorial: Governo, academia, indústria e sociedade civil devem atuar de forma integrada.
- Fomento contínuo e estruturado: É essencial garantir financiamento perene, com instrumentos adequados para ciência básica, aplicada e inovação.
- Capacitação e retenção de talentos: Formação técnica e científica em IA deve começar desde o ensino fundamental, com foco em pensamento crítico e diversidade.
- Infraestrutura nacional robusta: A ampliação e modernização da rede Sinapad e a aquisição de supercomputadores são vistas como estratégicas.
- Governança colaborativa: A criação de mecanismos de acompanhamento e governança é vista como essencial para garantir a efetividade das ações.

Divergências:

- Percepção sobre recursos disponíveis: Enquanto alguns participantes afirmam que há recursos disponíveis, outros apontam escassez e dificuldades de acesso.
- Foco da atuação nacional: Divergências sobre se o Brasil deve desenvolver tecnologias próprias em todas as camadas ou focar em dominar e adaptar tecnologias existentes.
- Formato dos eventos e processos participativos: Houve críticas e sugestões sobre o formato dos seminários, tempo de fala e representatividade de diferentes setores, como sociedade civil e segurança pública.

Carlos Rodrigo (SERPRO):

“Problema de iniciativas isoladas é uma coisa que me incomoda bastante. Tem dinheiro sendo gasto para resolver o mesmo problema várias vezes. Já começamos a traçar estratégias para reunir entidades de pesquisa, laboratórios, academia. A ausência de letramento é uma preocupação. O conhecimento está restrito a poucas pessoas. Esse letramento tem que começar lá no ensino médio.”

Cristiano Maciel (UFMT / SBC):

“Uns dizem que temos dinheiro, outros dizem que não temos. Eu só quero saber onde ele está e como fazer projeto com ele. Temos novos modelos para startups, novos tipos de

bolsas. Temos que integrar melhor os currículos. Cursos de IA são importantes, mas os cursos existentes também devem continuar tendo seu valor. A capacitação e retenção de talentos é essencial. A indústria já tem utilizado IA, mas precisamos coordenar melhor e sistematizar esse uso.”

Mateus Pierre (Instituto ELDORADO):

“O PBIA é um ótimo começo. Agora precisamos de um olhar de longo prazo. Não queremos reescrever *frameworks* como TensorFlow, mas dominar essas tecnologias. O SUS tem que ser prioridade número um. A aquisição de um supercomputador é um marco para a soberania tecnológica. Data spaces são fundamentais para garantir a soberania dos dados. Precisamos criar condições para reter os talentos e evitar a fuga de cérebros.”

Fábio Borges (LNCC):

“Falta demanda, não falta recursos. Os recursos estão nas camadas mais altas. A demanda depende de estratégia. Todos os ministérios deveriam estar envolvidos no PBIA e aportar recursos. Precisamos descrever a pilha com detalhes das camadas mais altas às mais baixas, com metas e alocação de recursos.”

André Rafael (MCTI):

“O presidente da República está cobrando respostas rápidas. O evento foi pensado para abrir o debate e entender os caminhos. Ficamos com receio de fazer algo que não fosse do tamanho e da importância que o Brasil requer. O próximo seminário será sobre dados e nuvem: como construir valor e resiliência em território nacional.”

Guilherme de Paula Correa (MCTI):

“Todo mundo me deu o feedback de que precisamos nos reunir mais vezes. O evento foi mais robusto, mas também recebeu críticas. Precisamos pensar em diferentes formatos para os próximos encontros. A ideia de ‘minha pilha, minha vida’ ou ‘pilha para todos’ mostra o desejo coletivo de participação e apropriação da tecnologia.”

2. Como refinar e transformar as recomendações do Seminário em diretrizes concretas para políticas públicas e programas de fomento?

Propostas para refinamento e transformação:

- Síntese estruturada das contribuições: A equipe do MCTI iniciará a sistematização das recomendações em um documento orientador para o PEBIA e outras políticas.
- Definição de metas e KPIs: Estabelecer indicadores claros para cada ação proposta, com acompanhamento e divulgação pública dos resultados.
- Identificação de projetos estruturantes: Priorizar iniciativas com alto impacto e potencial de articulação entre instituições.

- Articulação com outros ministérios e fundos: Exemplo: FNDE, Forças Armadas, setor agropecuário, para ampliar o financiamento e a abrangência das ações.
- Criação de instrumentos específicos: Como “Bolsa Pilha”, “Chamada Pilha”, “Crédito Pilha”, e editais temáticos em fluxo contínuo.
- Planejamento de longo prazo: Transformar o PEBIA em uma política de Estado, com continuidade e resiliência institucional.

Carlos Rodrigo:

“Qualquer iniciativa precisa definir KPIs e ser acompanhada. O MCTI tem papel fundamental na coordenação central. As metas devem ser acompanhadas e divulgadas. Quantas pessoas estão usando os modelos fundacionais? Quantas empresas estão sendo apoiadas? Quantos profissionais foram formados?”

André Rafael:

“Nosso diretor já cobrou um relatório para apresentar os resultados do evento. Precisamos de respostas rápidas. O seminário foi aberto e online para ampliar o debate. A ideia é gerar um documento norteador para orientar políticas públicas e ações do PBIA.”

Guilherme de Paula Correa:

“A gente já está pensando no primeiro seminário do Sinapad no começo do ano que vem. Também queremos realizar o segundo seminário da pilha no primeiro semestre. Precisamos continuar nos reunindo, discutindo e trabalhando juntos.”

231

3. Quais mecanismos de governança e acompanhamento podem ser criados para garantir o efetivo desenvolvimento da Pilha Nacional de IA?

Mecanismos sugeridos:

- Conselho Nacional da Pilha de IA: Inspirado em modelos como o CGI.br, com representação de governo, academia, indústria e sociedade civil.
- Comitê técnico permanente: Para avaliação contínua das ações, revisão de metas e integração entre projetos.
- Sistema de monitoramento com KPIs: Indicadores de uso, impacto, formação, inovação e soberania.
- Eventos regulares e fóruns técnicos: Seminários, workshops e reuniões técnicas para manter o diálogo e ajustar estratégias.
- Integração com o Sinapad e o SinIA: Como estruturas de coordenação e execução das ações em infraestrutura e inovação.
- Plataforma pública de transparência: Para acompanhamento dos investimentos, resultados e dados abertos.

Carlos Rodrigo:

“Sugiro acompanhamento e divulgação dos resultados. O SERPRO está disponível para ajudar nisso. Temos soberania operacional e estamos buscando soberania tecnológica. A nuvem soberana de governo é realidade e está em uso.”

Cristiano Maciel:

“Precisamos de governança institucionalizada. A SBC tem 27 comissões especiais e pode contribuir com coordenação das ações conjuntas. Segurança deve ser by design. A regulação da IA nas universidades é urgente. Precisamos de pensamento crítico desde o ensino fundamental.”

Mateus Pierre:

“Projetos estruturantes devem ser definidos com arranjos colaborativos. Ninguém vai fazer nada sozinho. A conectividade entre centros de processamento é essencial. A soberania dos dados exige controle de acesso e tempo de uso.”

Fábio Borges:

“Precisamos descrever a pilha com metas e alocação de recursos por camada. O LNCC está à disposição para contribuir com o MCTI e as instituições federais.”

André Rafael:

“O próximo seminário será sobre dados e nuvem. A governança deve garantir que a IA seja para o bem de todos, como deseja o presidente da República.”