

E-CIÊNCIA:

PROJETANDO SOLUÇÕES
COLABORATIVAS PARA OS
DESAFIOS DO BRASIL

RELATÓRIO DE RESULTADOS

{ozovits
conteúdo essencial

 **cg ee**
Centro de Gestão e Estudos Estratégicos
Ciência, Tecnologia e Inovação



E-CIÊNCIA

E-CIÊNCIA: PROJETANDO SOLUÇÕES COLABORATIVAS PARA OS DESAFIOS DO BRASIL

RELATÓRIO DE RESULTADOS



Março de 2025
Rio de Janeiro, Brasil





Organização: Centro de Gestão e Estudos Estratégicos (CGEE)
Diretor responsável: Anderson Stevens Leonidas Gomes

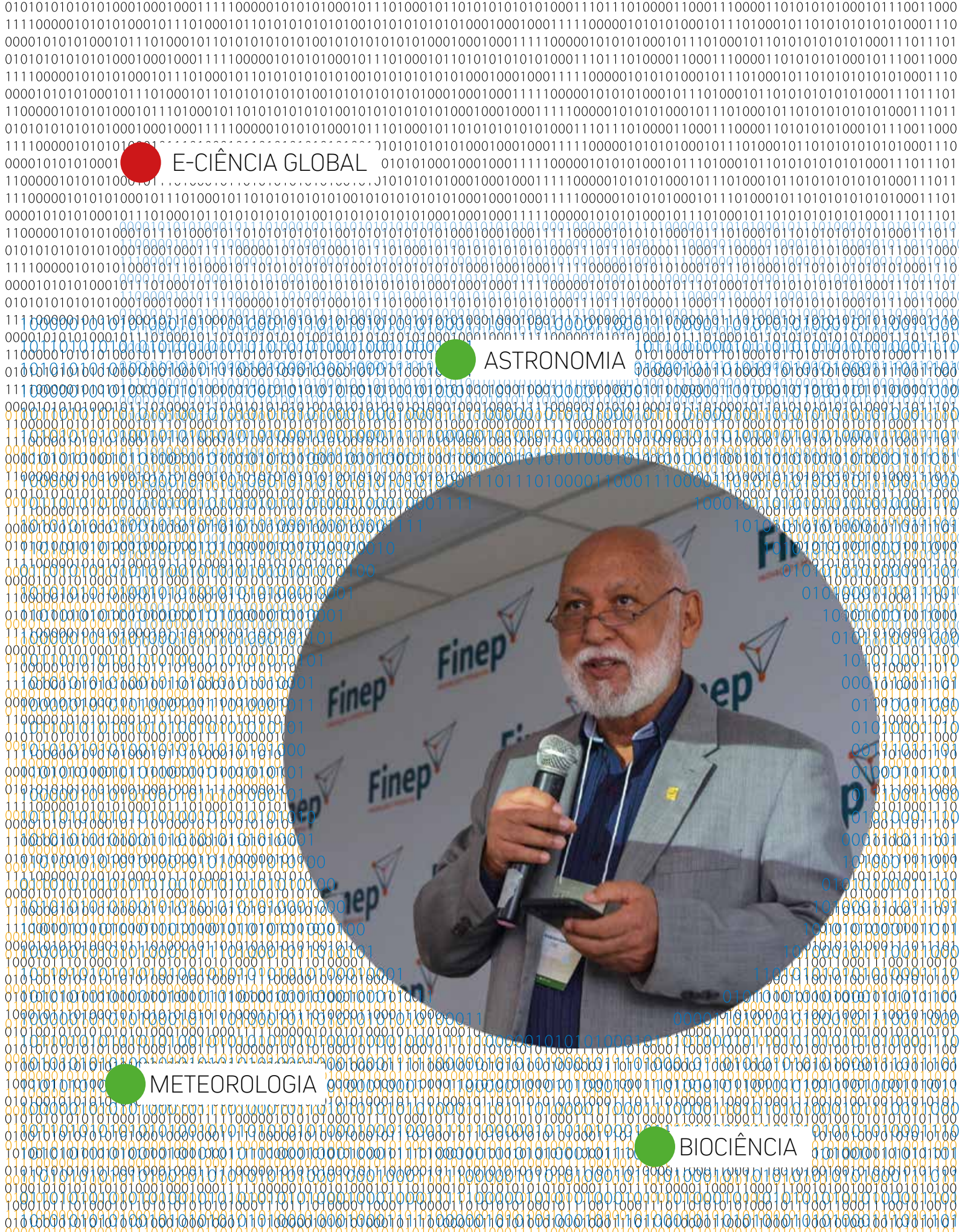
Design da experiência, facilitação do design sprint, moderação
do workshop e elaboração do relatório:
Editora e Produtora Kozovits

Facilitadores: Cris Kozovits e Eduardo Murad
Co-facilitador: Heitor Murad

Fotografias do evento: Hek Guerra

Sumário

1.	Apresentação	5
2.	Introdução	7
	2.1. Objetivos	7
	2.2. Programação	8
	2.3. Metodologia	9
3.	Apresentação dos cenários da e-ciência no mundo e no brasil - sinopses	11
	3.1. Cenário mundial	11
	3.2. Cenário brasileiro	11
4.	Apresentações das áreas - Sinopses	13
	4.1. Ana Tereza Vasconcelos, do LNCC/MCTI	13
	4.2. Gilvan Sampaio - INPE	14
	4.3. Helio Rocha Pinto - SAB	15
	4.4. Ignácio Bediaga - CBPF	15
5.	Ciclo 1 design sprint - Mapeamento dos desafios das áreas e do brasil	17
	5.1. Desafios específicos das áreas participantes	17
	5.2. Os 10 principais desafios comuns às áreas participantes	20
	5.3. Os 5 avanços que o Brasil precisa alcançar em e-Ciência	21
6.	Exemplos	23
	6.1 O exemplo da Astronomia	23
	6.2. O exemplo da meteorologia	24
7.	Estudo do caso LIneA	27
8.	Ciclo 2 design sprint - Estudo do caso LIneA	31
9.	Sinopse da palestra Modelos de Institucionalização - Wilson Coury	35
10.	Ciclo 3 design sprint - Qual o melhor modelo de institucionalização para um centro de e-Ciência no Brasil	37
11.	Ciclo 4 do design sprint - Formulação das recomendações ao MCTI	41
	Curto prazo	42
	Médio prazo	43
	Longo prazo	43
12.	Conclusão	45
13.	Galeria de fotos	46



1	APRESENTAÇÃO
---	--------------

A e-Ciência global tem passado por uma transformação exponencial, impulsionada pelo crescente volume de dados gerados por instrumentos cada vez mais sofisticados. Na Astronomia, por exemplo, observa-se um salto significativo: se na década de 1980 os dados produzidos por telescópios eram de ordem de megabytes, hoje falamos em terabytes produzidos em uma única noite de observação. Essa realidade não se restringe apenas à Astronomia. Áreas como a Biociência, Meteorologia e Agricultura, dentre muitas outras, também enfrentam desafios semelhantes, lidando com conjuntos de dados heterogêneos que chegam a petabytes e exabytes, demandando inovações nas abordagens tecnológicas e organizacionais.

Essa nova realidade dos grandes volumes de dados trouxe mudanças profundas não apenas na escala, mas também na metodologia científica. A e-Ciência, que anteriormente era conduzida por pesquisadores individuais acessando quantidades limitadas de dados, evoluiu para modelos colaborativos em que diversos pesquisadores acessam grandes bancos de dados compartilhados, impulsionando assim uma nova forma de realizar pesquisas. Esse cenário exige tecnologias avançadas de armazenamento e processamento de dados, como contêineres gerenciados por sistemas como Kubernetes e

plataformas híbridas em nuvem, que garantem escalabilidade, segurança e eficiência.

Diante dessa evolução acelerada e globalizada, o Brasil encontra-se em um momento crítico de definição estratégica para o futuro da sua infraestrutura de e-Ciência. Embora as ciências brasileiras já produzam resultados compatíveis com padrões internacionais, o país precisa urgentemente consolidar um modelo integrado que considere as necessidades comuns e específicas de todas as áreas. Essa integração não só otimizará recursos financeiros, técnicos e humanos, como também fortalecerá parcerias estratégicas, incluindo cooperações internacionais que já vêm sendo estabelecidas por instituições como a FINEP e o Laboratório LineA, por exemplo.

É nesse contexto que se insere a iniciativa do Centro de Gestão e Estudos Estratégicos (CGEE) em realizar o workshop “e-Ciência: Projetando Soluções Colaborativas para os Desafios do Brasil”. O evento objetiva justamente construir recomendações concretas, tanto imediatas quanto de médio e longo prazo, para orientar o Ministério da Ciência, Tecnologia e Inovação (MCTI) na construção de um sistema nacional robusto e integrado de e-Ciência, capaz de garantir a soberania científica e tecnológica do país frente às demandas crescentes impostas pela nova ordem mundial em ciência e tecnologia.



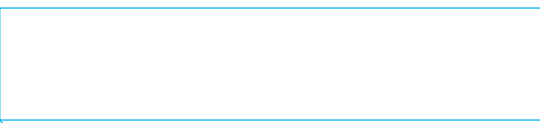
2	INTRODUÇÃO
---	------------

O workshop **“e-Ciência: Projetando Soluções Colaborativas para os Desafios do Brasil”** foi realizado nos dias 18 e 19 de março de 2025 no Rio de Janeiro, no Auditório José Pelúcio - FINEP, com uma carga horária total de 16 horas.

O evento, organizado pelo CGEE, contou com a participação de 40 pessoas, entre palestrantes, visitantes e representantes de 19 organizações, abrangendo 8 áreas de atuação. Destacaram-se representantes de importantes unidades de pesquisa vinculadas ao Ministério da Ciência, Tecnologia e Inovação (MCTI), como CBPF, INPE, LNCC, ON E RNP. Registra-se que o Laboratório Nacional de Astrofísica (LNA) foi convidado e confirmou presença, mas não enviou representantes ao evento.

2.1. OBJETIVOS

- Mapear os desafios comuns e específicos das áreas que geram, processam e armazenam grandes volumes de dados;
- Analisar iniciativas e oportunidades que se apresentam nos cenários de curto, médio e longo prazo para o desenvolvimento de soluções integradas;
- Debater e elaborar ao MCTI uma recomendação de qual seria o melhor modelo de institucionalização e operação de um centro de e-ciência integrado no Brasil.



2.2. PROGRAMAÇÃO

18 DE MARÇO	
08:30 às 09:10	Credenciamento e acolhimento
09:10 às 09:40	Abertura Oficial – Anderson Gomes (CGEE), Alexandre Barragat (FINEP), Andrea Latge (MCTI)
09:40 às 10:40	Palestras "Cenário global e brasileiro da e-Ciência" -
	William O’Mullane, Rubin Observatory
	Anderson Gomes, CGEE
10:40 às 12:30	Apresentações individuais das áreas
	Bioinformática (Ana Tereza Vasconcelos, LNCC, Brasil)
	Biofísica (Pedro Pascutti, UFRJ)
	Meteorologia (Gilvan Sampaio, INPE, Brasil)
	Astronomia (Helio Rocha Pinto, SAB, Brasil)
	Física de Altas Energias (Ignácio Bediaga, CBPF, Brasil)
14:00 às 15:30	Design Sprint – Ciclo 1: Mapeamento dos desafios gerais e específicos das áreas em e-Ciência
15:30 às 16h30	O exemplo da Astronomia - participações internacionais
	Andrew Connolly, eScience Institute da Universidade de Washington.
	George Beckett responsável pela construção do Rubin Independent Data Access Center (IDAC) do Reino Unido.
16h30 às 17h30	O exemplo da Meteorologia - Haroldo Campos Velho - INPE
19 DE MARÇO	
09:00 às 10:00	Estudo de caso LineA - Luiz Nicolaci da Costa
10:00 às 12:30	Design Sprint – Ciclo 2: O que temos a aprender com o LineA
14:00 às 14:40	Palestra Modelos de Institucionalização - Wilson Coury - Presidente do Conselho de Administração do LineA
14:40 às 15h20	Design Sprint Ciclo 3: Qual o melhor modelo para um centro de e-Ciência integrado no Brasil?
15:20 às 15:40	Apresentação dos resultados dos ciclos anteriores - Cris Kozovits (Facilitadora)
15:40 às 16:40	Design Sprint Ciclo 4: Elaboração das recomendações ao MCTI
16:40 às 17:00	Alinhamento final sobre as recomendações - Andrea Latge - MCTI
	Encerramento e próximos passos - Thyrso Villela - CGEE

2.3. METODOLOGIA

O desenho da experiência do workshop contou com expertise dos profissionais da Editora e Produtora Kozovits, empresa contratada para moderar o evento e facilitar as dinâmicas em grupo. A metodologia utilizada foi o Design Sprint, aplicada em ciclos na lógica de storytelling do evento.

- Ciclo 1 - Mapeamento dos desafios em e-Ciência comuns e específicos das áreas participantes
- Ciclo 2 - Análise do caso LineA

- Ciclo 3 - Modelo de institucionalização de um centro de e-Ciência
- Ciclo 4 - Elaboração das recomendações ao MCTI

Durante os dois dias, os participantes foram organizados em grupos multidisciplinares, com membros de diferentes áreas e instituições, garantindo múltiplas perspectivas e maior profundidade nas discussões e recomendações produzidas. Os resultados dos trabalhos em grupos nos ciclos do Design Sprint serão apresentados a seguir.



CENÁRIOS

3

APRESENTAÇÃO DOS CENÁRIOS DA E-CIÊNCIA NO MUNDO E NO BRASIL

SINOPSES

3.1. CENÁRIO MUNDIAL

Na palestra de abertura, o convidado internacional, William O’Mullane, responsável pelo gerenciamento de dados do Observatório Rubin, destacou em sua palestra o crescimento exponencial na geração e processamento de dados científicos, apontando para uma profunda mudança de paradigma. Ele ilustrou esta evolução através da linha do tempo dos dados astronômicos, que passaram de centenas de megabytes nos anos 1980 (como no caso da missão Hipparchus) para 20 terabytes por noite que serão produzidos pelo Observatório Rubin. Segundo O’Mullane, desde o ano 2000, o volume de dados astronômicos dobra anualmente, uma tendência semelhante observada também em outras áreas científicas, como a Biologia e a Medicina, que têm registrado crescimento comparável a cada 40 meses. Este avanço implica novos métodos colaborativos e tecnológicos, com a adoção de soluções híbridas e escaláveis, incluindo computação em nuvem e o uso de contêineres gerenciados por sistemas como Kubernetes, redefinindo profundamente o ambiente de trabalho científico.

3.2. CENÁRIO BRASILEIRO

Analisando o cenário da e-Ciência no Brasil, o palestrante Anderson Gomes, Diretor do Centro de Gestão de Estudos Estratégicos (CGEE) destacou em sua fala o papel estratégico do CGEE como um “centro de estudos e pesquisas estratégicas” responsável por apoiar diversos ministérios e instituições brasileiras na tomada de decisões estratégicas. Enfatizou a importância da e-Ciência, descrevendo-a como uma área de pesquisa científica que depende fortemente de infraestrutura computacional de larga escala para lidar com grandes volumes de dados, destacando tecnologias essenciais como computação em grid e redes distribuídas que possibilitam colaborações interdisciplinares. Ele também mencionou a situação atual da infraestrutura brasileira, ressaltando a presença do supercomputador Pegasus da Petrobras e do supercomputador Santos Dumont do LNCC, que recentemente alcançou a capacidade de 20 petaflops (unidade que mede a capacidade de processamento em quatrilhões de operações por segundo). Anderson enfatizou ainda a necessidade urgente de formar profissionais qualificados, com uma preocupação especial pela inclusão de mulheres nas áreas STEM. Finalizou ressaltando a multidisciplinaridade do tema e a importância da integração das plataformas e bancos de dados existentes no país, especialmente na área climática e biodiversidade, visando otimizar o uso dos recursos científicos disponíveis no Brasil.



4

APRESENTAÇÕES DAS ÁREAS

SINOPSES

Segundo a programação, representantes das áreas presentes no workshop apresentaram os cenários de e-Ciência, volumes de dados, instrumentos, exemplos e evoluções em suas esferas de atuação. A rica troca de experiências e os esclarecimentos mútuos gerados nesta etapa do workshop propiciaram uma visão mais transversal do tema, enriquecendo também o repertório de conhecimentos para os ciclos de Design Sprint realizados na sequência.

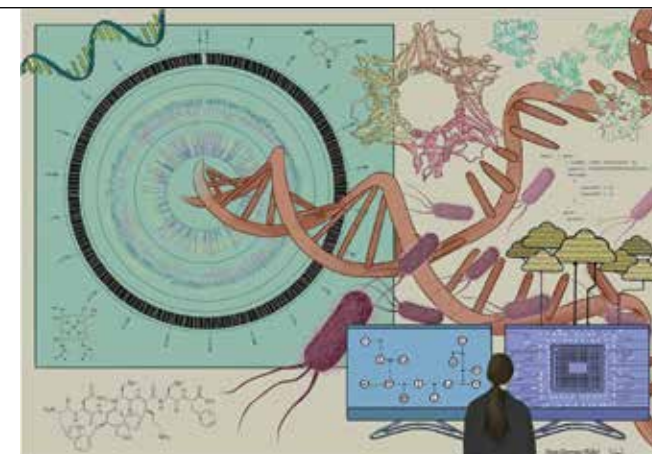
4.1. ANA TEREZA VASCONCELOS, DO LNCC/MCTI

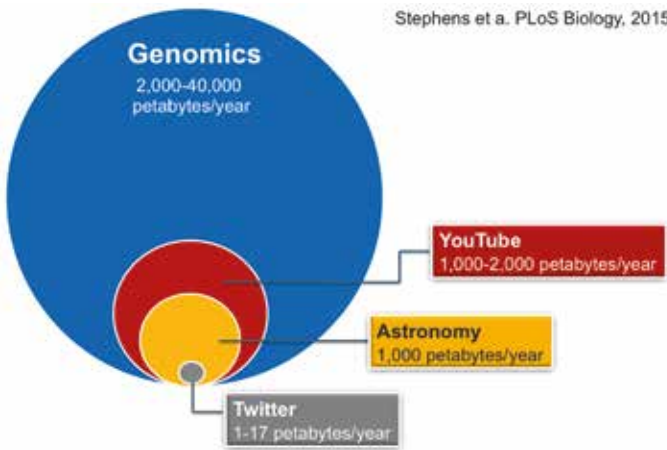
Apresentou um panorama sobre a aplicação da e-Ciência na Biologia, destacando o uso intensivo de tecnologias digitais avançadas e computação de alto desempenho para gerenciamento e análise de grandes volumes de dados biológicos (Big Data). Exemplificou a explosão no crescimento de dados do GenBank, que evoluiu de aproximadamente 680.000 pares de bases

em 1982 para cerca de 2,5 trilhões de pares de bases em 2023. A palestra destacou também a relevância da Bioinformática, das ferramentas de simulação como AlphaFold e das plataformas colaborativas internacionais, que facilitam a integração global e aceleram as descobertas científicas em áreas como medicina de precisão, farmacogenômica e biodiversidade.

E-science na Biologia:

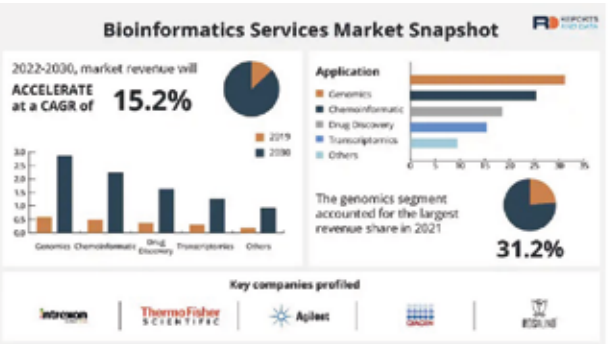
Ômicas e bioinformática: navegando por um dilúvio de dados.





4.2. GILVAN SAMPAIO, DO INPE

Apresentou as contribuições do Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) para a e-Ciência no Brasil, destacando sua liderança em previsão numérica de tempo e clima na América Latina. Ressaltou o alto volume de dados gerados diariamente por satélites e radares (235 GB/dia e 43 TB/ano, respectivamente), além da recente aquisição



The global bioinformatics services market size in 2021 was USD 2.50 Billion.

<https://www.reportsanddata.com/report-detail/bioinformatics-services-market>

de um supercomputador de 1,7 petaflops com capacidade inicial de armazenamento de 20 petabytes. Destacou ainda a plataforma BIG (Base de Informações Georreferenciadas), que disponibiliza cerca de 2 petabytes de dados, ressaltando a importância de sistemas mais eficientes para gestão e visualização desses dados.



4.3. HELIO ROCHA PINTO, SAB

Destacou o caráter histórico e colaborativo da Astronomia, mencionando a primeira grande colaboração internacional ocorrida durante o trânsito de Vênus em 1761, que possibilitou a medição precisa do tamanho do sistema solar através da paralaxe do planeta. Salientou que, atualmente, a Astronomia enfrenta desafios semelhantes em termos de colaboração internacional, diante do crescimento exponencial do volume de dados gerados, o que torna inviável sua análise por pesquisadores ou instituições isoladas. Ele também enfatizou a importância estratégica de o Brasil possuir centros capazes de processar e analisar grandes

volumes de dados astronômicos de forma autônoma, garantindo soberania científica. Ressaltou que a colaboração internacional é essencial, porém alertou sobre a necessidade do país ter independência na manipulação e análise de seus próprios dados. Citou como exemplo histórico bem-sucedido de integração nacional o surgimento do Observatório Astrofísico Brasileiro (atual Laboratório Nacional de Astrofísica), como modelo a ser seguido em outras áreas da e-Ciência. Reforçou ainda a relevância da integração multidisciplinar das áreas da e-Ciência para o avanço conjunto da ciência no Brasil.

4.4. IGNÁCIO BEDIAGA, REPRESENTANTE DO CENTRO BRASILEIRO DE PESQUISAS FÍSICAS (CBPF)

Destacou a relevância da física de altas energias no contexto da e-Ciência, apresentando três projetos-chave: o sistema GRID consolidado desde 2008, que opera com cerca de 5 mil núcleos computacionais atendendo grandes experimentos internacionais como LHC-B e ALICE; um novo sistema de computação em nuvem para análise de dados científicos, associado ao INCT CERN Brasil, que visa ampliar

a capacidade e eficiência para pesquisadores brasileiros; e um projeto pioneiro de sistema de trigger remoto para o Large Hadron Collider (LHC), que poderá processar e selecionar dados diretamente no Brasil, otimizando recursos energéticos e tecnológicos nacionais e contribuindo significativamente para a colaboração internacional em física de altas energias.



MAPEAMENTO

DESAFIOS

5

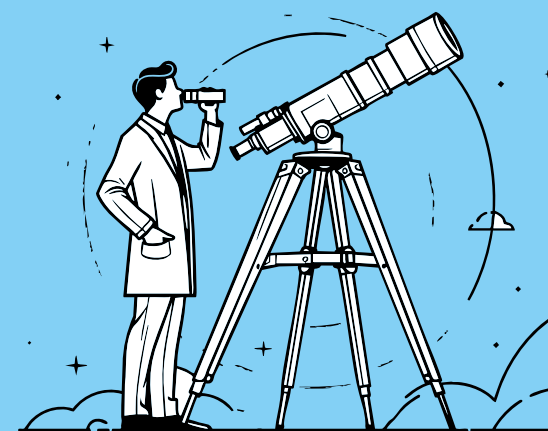
CICLO 1 DESIGN SPRINT

MAPEAMENTO DOS DESAFIOS DAS ÁREAS E DO BRASIL

No primeiro ciclo do Design Sprint, os grupos discutiram e integraram percepções sobre os principais desafios comuns e específicos que as áreas enfrentam no campo da e-Ciência, à luz das mudanças sinalizadas nas palestras de abertura e nas apresentações das áreas.

A seguir, apresentamos os resultados consolidados das contribuições dos grupos para este tema, sobre três perspectivas: Desafios comuns, desafios específicos das áreas e os avanços necessários para o Brasil no campo da e-ciência.

5.1. DESAFIOS ESPECÍFICOS DAS ÁREAS PARTICIPANTES



5.1.1. Astronomia

- Infraestrutura de rede e transferência de dados: necessidade de alta largura de banda, baixa latência e melhor conectividade (especialmente Rio-LNCC), garantindo alta vazão e eficiência na transmissão de grandes volumes de dados científicos.

- Recursos humanos qualificados e capacitação: escassez e dificuldade de retenção de profissionais especializados em TI científica e e-Ciência, perda de talentos para o mercado privado e fuga de cérebros do país, além da falta de treinamento especializado para pesquisadores e técnicos.
- Computação de Alto Desempenho (HPC): necessidade urgente de investimentos previsíveis, aumento da capacidade de processamento disponível, e maior experiência prática dos usuários em HPC.
- Armazenamento, gestão e disponibilização dos dados: desafios em garantir armazenamento local adequado, bem como melhorias na gestão, armazenamento eficiente, disponibilização dos dados e plataformas científicas integradas.

- Acesso e colaboração internacional: dificuldades no acesso a grandes volumes de dados provenientes de colaborações internacionais e necessidade de ampliar o apoio a iniciativas que desenvolvem plataformas e pipelines científicos integrados.

Planejamento estratégico e coordenação nacional: necessidade de coordenação eficiente das iniciativas nacionais, planejamento estratégico integrado e melhorias na gestão institucional da infraestrutura e serviços de TI científica.



5.1.2. Bio

- Recursos humanos e capacitação: necessidade de ampliar recursos humanos especializados em TI científica, garantindo bolsas para formação e retenção de pessoal qualificado em todos os níveis.
- Infraestrutura local: necessidade de ampliar infraestrutura local para armazenamento, processamento eficiente dos dados e melhorias significativas na banda larga disponível.
- Gestão institucional e apoio técnico: necessidade de reduzir burocracias

enfrentadas pelas unidades federais, ampliar o apoio efetivo das equipes institucionais de TI e melhorar a gestão de projetos científicos de longo prazo (ex.: INCT).

- Coordenação e planejamento estratégico nacional: necessidade urgente de coordenação das iniciativas nacionais em e-Ciência e planejamento estratégico integrado.

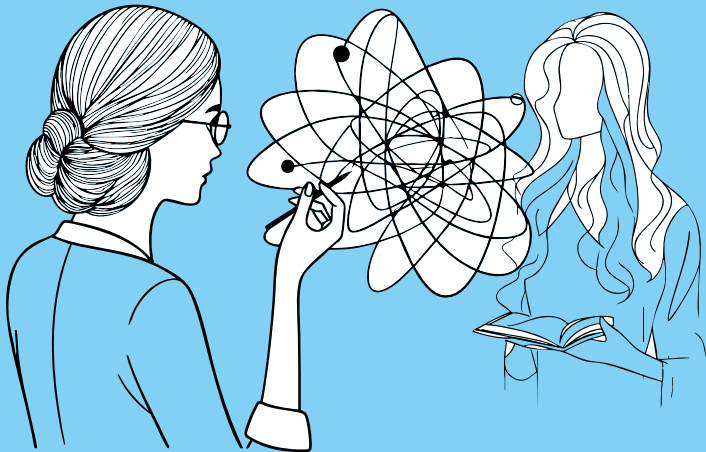


5.1.3. Meteorologia

- Plataforma e gestão de dados: necessidade de desenvolvimento e implementação de uma plataforma integrada para acesso, tratamento e visualização científica dos dados meteorológicos, considerando suas diferentes sazonalidades.
- Infraestrutura e soberania tecnológica: necessidade de melhorar a geração própria de dados (ex.: satélite próprio estacionário), armazenamento eficiente, e transferência adequada desses dados para garantir soberania científica e tecnológica.
- Recursos humanos e capacitação: necessidade urgente de recursos

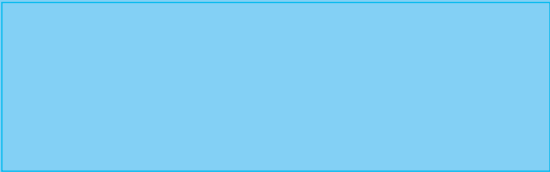
humanos qualificados, treinamento específico em e-Ciência e capacitação para uso adequado da infraestrutura disponível.

- Gestão integrada dos dados: necessidade de middleware, curadoria especializada e criação de um ecossistema efetivo para gestão e processamento dos dados científicos.
- Planejamento estratégico: necessidade de desenvolvimento de um planejamento estratégico específico para e-Ciência na Meteorologia.

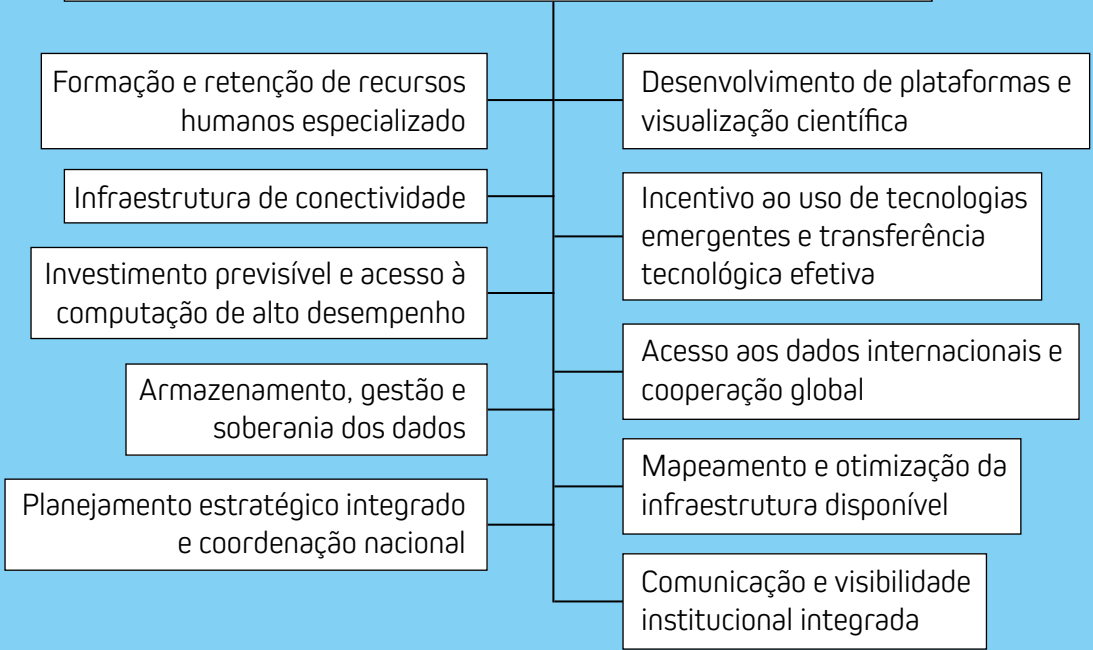


5.1.4. Física de Altas Energias

- Retenção de conhecimento
- Retenção de pessoal especializado
- Dificuldade de formação de pessoal
- Falta de interlocução com os órgãos financiadores
- Falta de treinamento dos usuários pesquisadores



5.2. OS 10 PRINCIPAIS DESAFIOS COMUNS ÀS ÁREAS PARTICIPANTES

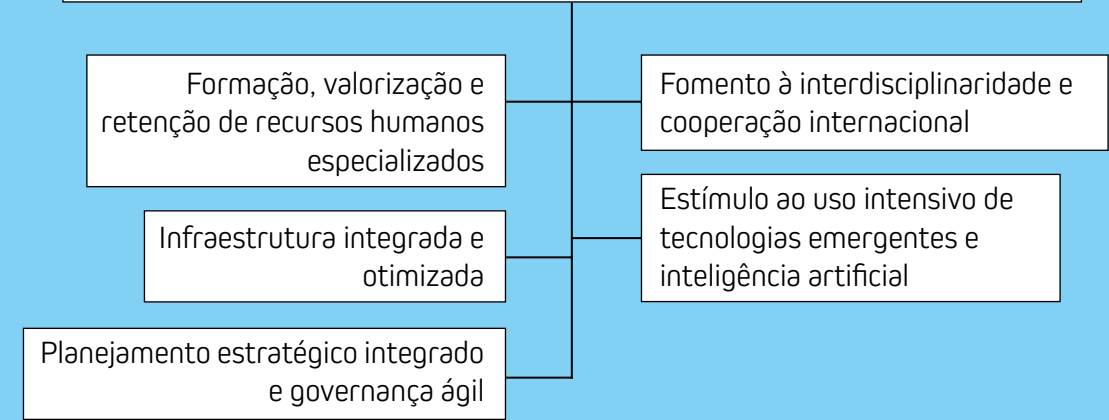


- 1. Formação e retenção de recursos humanos especializados:** escassez de profissionais qualificados em TI científica, pouca capacitação especializada e dificuldade de retenção de talentos.
- 2. Infraestrutura de conectividade:** largura de banda limitada e insuficiente, com baixa latência e capacidade inadequada para transmissão eficiente de grandes volumes de dados científicos em todo o país.
- 3. Investimento previsível e acesso à computação de alto desempenho (HPC):** necessidade de investimento contínuo e previsível, ampliando o acesso das áreas científicas ao processamento de alto desempenho.
- 4. Armazenamento, gestão e soberania dos dados:** insuficiência na infraestrutura local de armazenamento, gestão inadequada dos dados gerados nacionalmente, dificultando a soberania tecnológica.
- 5. Planejamento estratégico integrado e coordenação nacional:** ausência de um planejamento estratégico claro, coordenado e integrado entre instituições e projetos nacionais de e-Ciência.
- 6. Desenvolvimento de plataformas e visualização científica:** falta de plataformas integradas e softwares especializados que facilitem o tratamento, análise e visualização científica dos dados.
- 7. Incentivo ao uso de tecnologias emergentes e transferência tecnológica efetiva:** baixo incentivo ao uso de tecnologias avançadas como IA e pouca efetividade na transferência tecnológica entre academia e mercado.
- 8. Acesso aos dados internacionais e cooperação global:** dificuldade de acesso efetivo aos dados de projetos internacionais e limitações na integração desses dados às pesquisas nacionais.

9. **Mapeamento e otimização da infraestrutura disponível:** ausência de mapeamento detalhado e otimização estratégica das infraestruturas nacionais existentes para melhor aproveitamento.

10. **Comunicação e visibilidade institucional integrada:** baixa visibilidade dos resultados das instituições científicas e falta de comunicação estruturada entre instituições e sociedade.

5.3. OS 5 PRINCIPAIS AVANÇOS QUE O BRASIL PRECISA ALCANÇAR EM E-CIÊNCIA



- 1. Formação, valorização e retenção de recursos humanos especializados:** Formação contínua e capacitação específica para o novo perfil de cientista de dados, valorização profissional por meio de salários competitivos, flexibilização nas contratações, mobilidade interna de pessoal, treinamento amplo da comunidade científica e criação de políticas claras de contratação.
- 2. Infraestrutura integrada e otimizada:** Criação e expansão de infraestrutura computacional de alto desempenho, armazenamento robusto e rede de alta velocidade com abrangência nacional; otimização e melhor aproveitamento das infraestruturas existentes (centros multiusuários), com plataformas unificadas para processamento e acesso aos dados científicos.
- 3. Planejamento estratégico integrado e governança ágil:** Implementação de modelos de governança mais ágeis e inovadores específicos para a e-Ciência,

com planejamento estratégico contínuo e de longo prazo, previsão orçamentária clara e coordenação eficiente dos centros nacionais dedicados à e-Ciência.

- 4. Fomento à interdisciplinaridade e cooperação internacional:** Reconhecimento explícito da interdisciplinaridade, envolvendo profissionais de TI, cientistas de dados, engenheiros e pesquisadores das áreas finalísticas; maior participação brasileira em colaborações internacionais estratégicas, especialmente com países do Sul Global.
- 5. Estímulo ao uso intensivo de tecnologias emergentes e inteligência artificial:** Investimentos significativos em programas específicos que promovam a implementação e o desenvolvimento contínuo da Inteligência Artificial em todas as áreas científicas, além da criação de um Observatório Nacional de Tecnologias em e-Ciência para acompanhar e disseminar inovações tecnológicas.

6

EXEMPLOS

Para embasar as próximas produções em grupo no processo do design sprint, duas das áreas presentes, que possuem um caminho percorrido, lições aprendidas e experiências a compartilhar, no que tange a institucionalização e formação de parcerias nacionais e internacionais, apresentaram seus cases.

6.1. O EXEMPLO DA ASTRONOMIA

Com as participações internacionais de George Beckett, responsável pela construção do Rubin Independent Data Access Center (IDAC) britânico; e Andrew Connolly, diretor do eScience Institute da Universidade de Washington.

- George Beckett falou sobre a participação no projeto LSST envolvendo todas as 36 universidades do Reino Unido atuando em astronomia, apresentando as contribuições in-kind sendo oferecidas e o plano de operação até 2036 com um orçamento previsto de 72.6 milhões de libras e os produtos científicos a serem gerados pelo UK-IDAC.
- Andrew Connolly apresentou a evolução da ciência de dados e de inteligência artificial na Universidade e o impacto que teve ao longo dos últimos 16 anos em termos de temas, do número de

professores envolvidos no programa o número de pós-graduandos o número de programas de pos-graduacao e o número de alunos de graduação sendo formados na área. Chamou a atenção que em 2008 a formação de cientistas de dados ocorria apenas na indústria e não na área acadêmica e a contribuição em 2007 de Jim Gray que cunhou o termo eScience como uma nova forma de se fazer ciência quando se tem grande volume de dados. Andrew comentou ainda as várias soluções encontradas para estimular o crescimento da área como definir um espaço apropriado para estimular a troca de conhecimento, a organização de uma equipe interessada, diferentes formas de engajar a comunidade e o envolvimento de engenheiros de software no programa para resolver problemas de escalabilidade o grande desafio na era de Big Data. Completou sua apresentação comentando sobre projetos realizados nas mais variadas áreas, a importância de se desenvolver uma expertise no uso de computação nas nuvens, a evolução da educação formal na área de eScience, finalizando com o exemplo do Observatório Rubin na geração de grandes volumes de dados.

ASTRONOMIA

METEOROLOGIA

6.2. O EXEMPLO DA METEOROLOGIA
- HAROLDO CAMPOS VELHO,
REPRESENTANDO O INPE

Haroldo Campos Velho destacou a evolução do papel dos dados ao longo da história da ciência, ressaltando que, no século XXI, a ciência dos dados ocupa uma posição central na Meteorologia, com desafios associados à assimilação de grandes volumes de dados (Big Data). Ele apontou que as técnicas modernas de mineração de dados, aprendizado de máquina e assimilação são essenciais para enfrentar esses desafios, citando exemplos como a previsão numérica do tempo e as previsões climáticas sazonais, particularmente para a América do Sul.

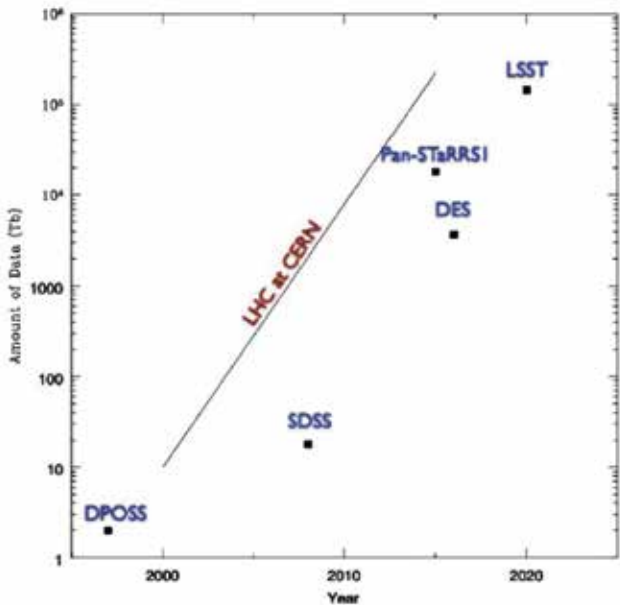
No contexto da Meteorologia atual, Velho enfatizou a importância dos modelos híbridos, que integram modelos físicos baseados em equações diferenciais e abordagens baseadas

em dados (Machine Learning). Destacou, especialmente, o uso bem-sucedido de redes neurais artificiais, comparando sua eficiência computacional e precisão às técnicas tradicionais de assimilação de dados meteorológicos. Exemplificou que as redes neurais podem ser até centenas de vezes mais rápidas que métodos tradicionais como o LETKF, mantendo qualidade comparável nas previsões.

Finalmente, salientou a relevância da cooperação nacional e internacional, citando parcerias com instituições brasileiras e estrangeiras, como a colaboração do INPE com o Imperial College London e com o CMCC (Itália), como fundamentais para avançar na aplicação dessas novas tecnologias e métodos na previsão do tempo e do clima, destacando como crucial a integração entre diferentes áreas científicas e tecnológicas para resolver os desafios atuais da e-Ciência.



Increase of astronomical data



7

ESTUDO DO CASO LINEA

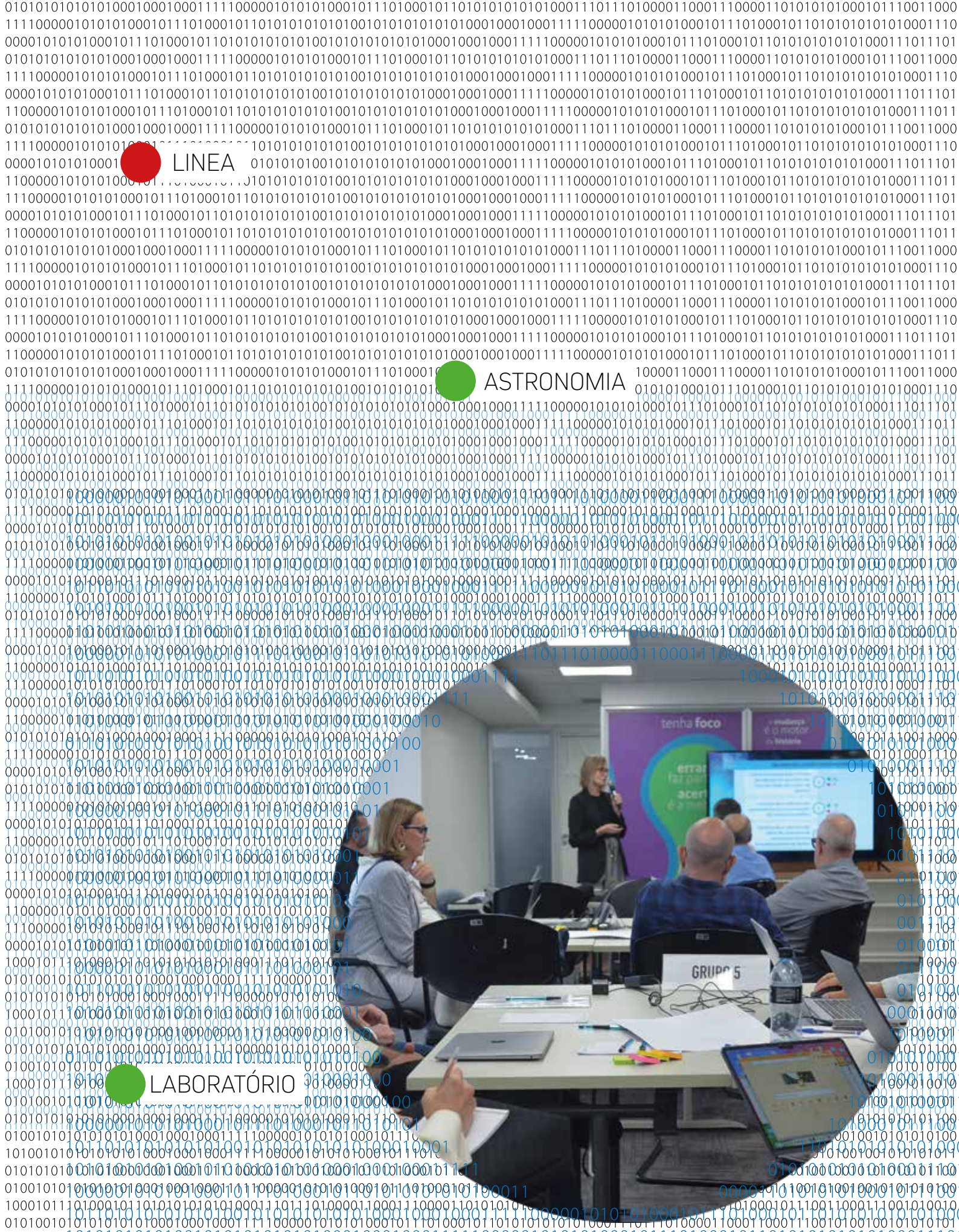
O diretor do Laboratório Interinstitucional de e-Astronomia (LineA), Luiz Nicolaci da Costa apresentou o cenário atual de desafios e oportunidades do Laboratório da Astronomia, esclarecendo a urgência de uma tomada de decisão a respeito da manutenção do projeto. Levantou a possibilidade de, inclusive, o LineA receber aporte imediato para prosseguir com suas atividades e servir como incubadora de uma experiência piloto de outra área (ex: CENABIO). Por sua experiência, infraestrutura, equipe e relacionamento, o LineA pode diminuir a curva de aprendizado no processo de formatação de um modelo de Centro de e-Ciência integrado no Brasil, vindo depois a compor um dos elos do Hub nacional.

Em sua fala, Luiz Nicolaci ressaltou que o LineA é um centro de e-ciência que opera um laboratório multiusuário internacional, promovendo, por meio de contribuições in-kind, a participação de pesquisadores brasileiros em grandes projetos internacionais. São exemplos dessas participações o Dark Energy Survey, o Sloan Digital Sky Survey, Dark Energy Spectroscopic Instrument e o Legacy Survey of Space Time, que colabora na formação de jovens pesquisadores e tecnólogos em e-ciência organizando workshops e cursos na área de e-ciência, além de manter inúmeras parcerias nacionais e internacionais.

Também comentou sobre o que o laboratório realiza, ressaltando que, atualmente,

sua principal missão é a implantação e o comissionamento de um Independent Access Center (IDAC) para o projeto LSST no Brasil, como parte de um acordo com SLAC National Accelerator Laboratory, operado pela Universidade de Stanford em nome do Office for Science do Departamento de Energia do governo dos Estados Unidos. Nicolaci apresentou um breve histórico sobre o laboratório, lembrando do seu início informal em 2006 e alguns marcos importantes em sua longa história, incluindo a criação do ICT privado Associação LineA, seguindo sugestão de um grupo de trabalho formado pelo MCTI. Comentou ainda que as principais fontes de financiamento nos últimos 10 anos foram a FINEP e o programa INCT (CNPq, FAPERJ), tendo sido investido um total de R\$64 milhões.

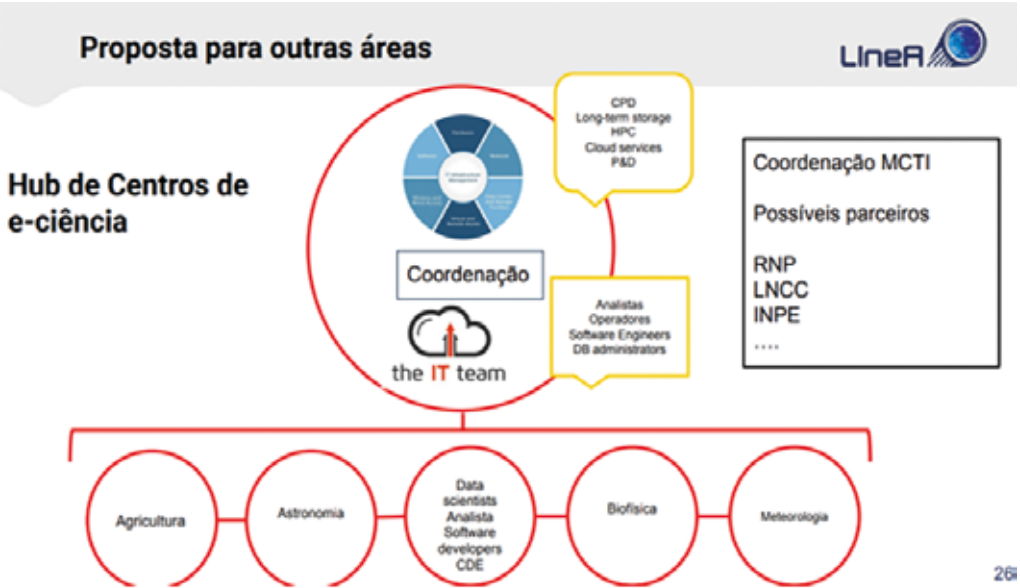
Em sua apresentação lembrou a importância do acordo com o SLAC e a difícil situação em que o laboratório se encontra, possuindo recursos para a compra dos equipamentos necessários à implantação do IDAC, mas sem recursos para pagar a equipe necessária para levar adiante o trabalho. Alertou que os recursos disponíveis para pessoal terminam em maio de 2025, e caso uma solução emergencial não seja encontrada, o LineA terá que notificar a saída do acordo internacional, ameaçando a participação brasileira no projeto e tendo que encerrar imediatamente as suas atividades. A



consequente perda de todo o conhecimento e experiência acumulada ao longo dos últimos 20 anos e a perda de toda a infraestrutura computacional montada é um desfecho lamentável, especialmente num momento em que o Brasil precisa evoluir em ciência de dados.

Ao final de sua apresentação, Nicolaci deixou, como sugestão, um modelo

organizacional para a área de e-ciência, a qual é composta por um centro de suporte a centros de e-ciência (e.g. RNP) que apoiaria centros temáticos, seguindo o modelo do LIneA. A principal ideia é ter meios de racionalizar gastos e compartilhar ao máximo infraestrutura, talentos e experiências, ao mesmo tempo reconhecendo as especificidades de cada área.



26



8

CICLO 2 DO DESIGN SPRINT

ESTUDO DO CASO LINEA

Neste ciclo do Design Sprint, os grupos responderam a algumas questões apresentadas na matriz online, a saber:

- A. O trabalho do Linea pode ser replicado para outras áreas? Por quê?
- B. Pontos fortes do Linea: o que manter e o que ampliar.

- C. Pontos Frágeis ou de Melhoria: o que corrigir, ajustar ou eliminar
- D. Inovações: o que ainda não existe e poderia ser acrescentado.

Abaixo, apresentamos os resultados consolidados das contribuições dos grupos nesses quesitos:

A. O trabalho do Linea pode ser replicado para outras áreas? Por quê?

Três respostas “sim” e uma resposta “parcialmente” sinalizaram que o trabalho do Linea pode ser replicado em outras áreas, desde que consideradas algumas particularidades. Abaixo, o compilado das respostas a esta questão:

Infraestrutura tecnológica: Grande parte da infraestrutura (estimada entre 80-90%) é agnóstica e replicável em outras áreas. Isso inclui serviços genéricos como clusters Kubernetes, autenticação federada, serviços de transferência de dados, infraestrutura TIC, arquitetura e operação de dados, criação de portais e utilização de ferramentas como Jupyter Notebooks.

Software e serviços: Camadas do tipo Software as a Service (SAAS), sistemas de autorização, bancos de dados, e aplicações comuns são amplamente replicáveis.

Recursos humanos e competências técnicas: Profissionais como analistas, engenheiros DevOps e de software são transferíveis entre diferentes áreas científicas. No entanto, especialistas como cientistas de dados, profissionais ligados à interface com o usuário final e desenvolvedores de modelos específicos de IA precisam de conhecimento especializado e particular das áreas em questão.

Experiência e cooperação: O conhecimento acumulado e as experiências em cooperação internacional também são replicáveis e benéficos às outras áreas.

Considerações adicionais:

A replicação é facilitada se houver um claro catálogo de serviços, definindo o que é específico e o que pode ser aproveitado genericamente.

A aplicação específica e final, incluindo o uso de Inteligência Artificial e arquitetura de metadados, pode requerer ajustes e conhecimentos particulares das áreas científicas específicas.

B. Pontos fortes do LIneA: o que manter e o que ampliar:
Recursos Humanos e Capacitação
Equipe altamente qualificada e capacitada.
Formação contínua e especializada de recursos humanos.
Experiência acumulada ao longo dos anos em e-ciência.
Parcerias e Cooperações
Capacidade de estabelecer rapidamente acordos de cooperação nacionais e internacionais.
Experiência reconhecida em cooperação internacional.
Desenvolvimento de Software e Infraestrutura Tecnológica
Desenvolvimento de soluções genéricas e customizáveis em software científico.
Experiência no desenvolvimento de softwares específicos para atender demandas particulares (astronomia).
Capacidade comprovada na criação de bancos de dados, portais, e aplicações específicas.
Modelo robusto e eficiente de infraestrutura e operação de dados.
Agilidade e Inovação
Agilidade e flexibilidade devido ao modelo institucional (ICT privada).
Capacidade de desenvolvimento de soluções tecnológicas inovadoras para acesso e processamento de dados.
C. Pontos Frágeis ou de Melhoria: o que corrigir, ajustar ou eliminar
Orçamento e Recursos Financeiros
Imprevisibilidade e falta de sustentabilidade orçamentária a longo prazo.
Dificuldades recorrentes na obtenção de recursos financeiros estáveis, especialmente para despesas de custeio e administrativas.
Descontinuidade financeira provocada por mudanças políticas e governamentais.
Recursos Humanos e Equipe
Equipe reduzida e sobrecarregada em relação à amplitude dos projetos.
Vulnerabilidade pela falta de redundância ou pares para cada perfil de atuação técnico.
Dificuldade na manutenção e contratação da equipe especializada para operação da infraestrutura já instalada.
Articulação Institucional
Necessidade de melhorar a articulação e cooperação com outras áreas de conhecimento e instituições.

D. Inovações: o que ainda não existe e poderia ser acrescentado
Expansão e Replicação do Modelo
Expandir e replicar o modelo do LineA para outras áreas de conhecimento, favorecendo a multidisciplinaridade.
Gestão de Identidade e Segurança
Desenvolver soluções avançadas para gestão de identidade, autenticação e autorização.
Fluxos de Trabalho Científicos
Desenvolvimento e compartilhamento de workflows científicos, ampliando o aproveitamento das experiências adquiridas.
Compartilhamento de Conhecimento
Intensificar o compartilhamento e a disseminação do conhecimento técnico e científico já acumulado.
Inteligência Artificial Interpretável
Desenvolver uma camada de Inteligência Artificial Interpretável (XAI), alinhada com regulamentações previstas para 2026, especialmente aplicável a áreas críticas.



ESTRUTURA HÍBRIDA

9

MODELOS DE INSTITUCIONALIZAÇÃO

SINOPSE

A partir de tudo que foi visto e analisado ao longo do workshop, os grupos receberam, nesta etapa, informações sobre os modelos de institucionalização de um centro de e-Ciência e sobre o que diz a legislação a respeito do tema. O palestrante, Wilson Coury, Presidente do Conselho de Administração do LineA, apresentou vantagens e desvantagens de cada modelo, visando subsidiar o debate dos grupos sobre qual o modelo que mais se adequa às necessidades das áreas, numa visão de um hub integrado.

Wilson Coury abordou os principais desafios e soluções para viabilizar o funcionamento do IDAC (Infrastructure Data Access Center) do LSST (Legacy Survey of Space and Time), especialmente no contexto brasileiro. Destacou inicialmente a necessidade emergencial de concluir o IDAC-LSST até junho de 2025, por meio da contratação de pessoal qualificado e da estruturação da infraestrutura tecnológica adequada para processamento de Big Data.

Coury enfatizou que, no curto prazo, o instrumento jurídico mais adequado para viabilizar o repasse de recursos públicos ao LineA é o Convênio para Pesquisa, Desenvolvimento e Inovação (PD&I). Este convênio, amparado legalmente pela Lei nº 10.973/2004, alterada pelo Marco Legal

da Ciência, Tecnologia e Inovação (Lei nº 13.243/2016) e regulamentado pelo Decreto nº 9.283/2018, permite que órgãos da administração pública transfiram recursos financeiros públicos para Instituições Científicas, Tecnológicas e de Inovação (ICTs), públicas ou privadas, sem a necessidade de licitação prévia. Esse modelo foi reforçado pelo Parecer nº 00004/2024 da Advocacia-Geral da União, que legitima expressamente essa forma de convênio para ICTs privadas.

Como modelo de longo prazo, Coury sugeriu uma estrutura híbrida para gestão de grandes volumes de dados governamentais, composta por uma infraestrutura centralizada de processamento e distribuição de dados, com bases descentralizadas mantidas pelos órgãos responsáveis pela coleta. Este modelo híbrido apresenta vantagens como menor custo operacional, maior segurança jurídica (alinhada à Lei Geral de Proteção de Dados – LGPD), melhor interoperabilidade entre sistemas, além de resiliência administrativa por evitar concentração excessiva de poder ou risco de influências políticas indevidas. Coury ressaltou que o sucesso desse modelo depende essencialmente de governança sólida, padrões claros de qualidade de dados e uma infraestrutura robusta de interoperabilidade.

10

CICLO 3 DO DESIGN SPRINT QUAL O MELHOR MODELO DE INSTITUCIONALIZAÇÃO DE UM CENTRO DE E-CIÊNCIA NO BRASIL?

À luz da palestra proferida pelo especialista Wilson Coury, os grupos discutiram e se pronunciaram em relação ao modelo de institucionalização a ser adotado na elaboração de um centro de e-ciência integrado, que venha atender às necessidades de todas as áreas. Não houve consenso entre os grupos neste tópico, portanto, apresentamos as diferentes posições nos quesitos de discordância.

Os grupos estão em consenso sobre:

- Médio e longo prazo: Um modelo descentralizado, composto por uma coordenação geral vinculada ao MCTI e múltiplos centros de e-Ciência distribuídos no país.
- Curto prazo: No caso específico do LInEA, recomendam que permaneça como uma Instituição Científica e Tecnológica (ICT) privada operando por meio de convênio.

MODELOS SUGERIDOS PARA MÉDIO E LONGO PRAZO:

1. ICT Privada (Organização Social - OS):

- Garantia de maior autonomia administrativa, financeira e técnica.
- Maior flexibilidade para contratação e manutenção de recursos humanos qualificados.
- Menos burocracia para a aquisição de equipamentos e captação de recursos financeiros.

2. Integração como unidade especializada da RNP (Rede Nacional de Ensino e Pesquisa):

- Aproveitamento da infraestrutura consolidada da RNP.
- Compartilhamento de experiência em gestão de infraestrutura científica de dados.

3. Modelo híbrido inspirado na experiência do IMPA/RNP, com as seguintes características específicas destacadas:

- Facilidade e flexibilidade para contratação (incluindo modalidades CLT e Pessoa Jurídica – PJ).
- Remuneração competitiva com o mercado.
- Flexibilidade na atribuição e progressão de níveis de carreira.
- Menor burocracia em processos administrativos e financeiros.
- Facilidade na expansão ou redução ágil da equipe técnica conforme demandas.
- Infraestrutura física adequada e expansível.
- Simplificação na aquisição de equipamentos.
- Adoção de rebranding institucional com uma identidade mais ampla e inclusiva

INSTITUCIONALIZAÇÃO

PRAZOS

MODELO

CONDIÇÕES COMPLEMENTARES E CONSIDERAÇÕES ADICIONAIS APRESENTADAS:

A adoção do modelo ideal deve considerar o contexto específico de cada área, permitindo adaptações conforme necessidades locais.

Destaca-se a necessidade urgente de garantir sustentabilidade financeira a médio e longo prazo.

Reforçada a importância da transparência administrativa e financeira independentemente do modelo escolhido.

Ressaltado que o modelo escolhido precisa assegurar agilidade e competitividade para retenção de talentos especializados no setor de e-Ciência.





RECOMENDAÇÕES

LINEA/MCTI

11

CICLO 4 DO DESIGN SPRINT

FORMULAÇÃO DAS
RECOMENDAÇÕES AO MCTI

Como objetivo final da iniciativa, para uma visão de médio e longo prazo, destaca-se que toda a programação foi desenhada para gerar subsídios e clarear as perspectivas das diferentes áreas sobre o futuro da ciência de dados no Brasil. Assim, o CGEE acredita que terá o embasamento necessário para a elaboração de um projeto robusto e consistente de âmbito nacional que venha a beneficiar todas as áreas que processam e armazenam grandes volumes de dados. Além disso, a discussão e os desdobramentos necessários sobre este tema são questões que impactam a competitividade e a autonomia e a soberania do país em relação a toda sua produção de dados científicos, abrangendo também a formação e especialização de pessoas na área para dar vazão à complexidade e escala global da geração, processamento, armazenamento e compartilhamento de dados, cada vez mais velozes e precisos.

Numa visão de curto prazo, o workshop trouxe à luz os desafios atuais do LInEA, que exigem celeridade ímpar para que o laboratório não seja desmobilizado, impactando fortemente as pessoas que nele atuam, as pesquisas e parcerias nacionais e internacionais que são efetivadas e executadas pelo LInEA.

O LInEA vive hoje o dilema de ter ao mesmo tempo que levar adiante a instalação de novos equipamentos recém chegados para o IDAC, planejar novas compras com recursos recém obtidos da FINEP, organizar na Academia Brasileira Ciências (ABC) o evento celebrando o início das operações no dia 17 de junho (Rubin First Look) ao mesmo tempo que tem que discutir o plano de desmobilização e encerramento das atividades do LInea por falta de recursos e compromissos trabalhistas.

Tendo em vista estas duas dimensões de análise, os grupos formularam as seguintes recomendações ao Ministério de Ciência, tecnologia e INovação (MCTI):

A MAIORIA ABSOLUTA DOS GRUPOS RECOMENDOU:

Que o MCTI demande ao CGEE o desenvolvimento de um Plano Nacional de e-Ciência, a partir dos resultados desta oficina e de novas rodadas de design sprint, visando ouvir outras áreas, instituições e atores para embasar um plano estratégico consistente e representativo.

Utilizar a infraestrutura desenvolvida para o LInEA como um projeto piloto para impulsionar o desenvolvimento de iniciativas semelhantes em outras áreas científicas, destinando de forma emergencial recursos financeiros do FNDCT ao LInEA para que este desempenhe o papel de centro-piloto nacional de e-Ciência.

Sensibilizar o MCTI para pautar uma proposta de projeto de encomenda junto ao Conselho Diretor do FNDCT, visando assegurar o apoio financeiro à continuidade da participação da comunidade brasileira no LSST.
Mobilizar a CONFAP e as Fundações Estaduais de Apoio à Pesquisa (FAPs) para que estas também apresentem propostas de encomendas específicas, garantindo recursos para manter a participação brasileira no LSST.
Realizar uma ação direta do MCTI para estruturação do projeto Independent Data Access Center (IDAC), por meio da assinatura de convênio com a Associação LIneA, assegurando suporte financeiro necessário.
Destinar recursos financeiros via encomenda FINEP especificamente para a implementação do IDAC.
Promover ações claras de comunicação junto às instituições públicas de Ciência & Tecnologia, destacando os ganhos potenciais e estratégicos do modelo proposto.
Garantir ampla publicização e comunicação transparente das decisões e dos resultados alcançados para toda a comunidade científica brasileira, envolvendo entidades representativas como SAB, ABC e SBF.
Manter rigorosa transparência financeira nas ações e projetos implementados.

CONSIDERAÇÕES ADICIONAIS E COMPLEMENTARES DOS GRUPOS:

Foi reforçado que as ações emergenciais devem ser acompanhadas por medidas estratégicas de longo prazo para garantir a sustentabilidade.
Alguns grupos ressaltaram a importância de criar mecanismos de governança ágeis e efetivos para otimizar a aplicação dos recursos públicos.
A necessidade de envolver diretamente as comunidades científicas em processos decisórios foi destacada, visando o fortalecimento da cooperação e da interdisciplinaridade no âmbito nacional.

LINHA DO TEMPO DAS RECOMENDAÇÕES

AÇÕES DE CURTO PRAZO	Destinar de forma emergencial recursos financeiros do FNDCT ao LIneA, garantindo seu papel como centro-piloto nacional de e-Ciência.
	Sensibilizar o MCTI para pautar imediatamente uma proposta de projeto de encomenda junto ao Conselho Diretor do FNDCT, visando assegurar apoio financeiro à participação brasileira no LSST.
	Mobilizar a CONFAP e as FAPs para apresentação urgente de propostas específicas de encomendas que garantam recursos imediatos à participação brasileira no LSST.
	Realizar ação direta do MCTI para estruturação do projeto Independent Data Access Center (IDAC), por meio da assinatura urgente de convênio com a Associação .

AÇÕES DE MÉDIO PRAZO	Destinar de forma emergencial recursos financeiros do FNDCT ao LIneA, garantindo seu papel como centro-piloto nacional de e-Ciência.
	Sensibilizar o MCTI para pautar imediatamente uma proposta de projeto de encomenda junto ao Conselho Diretor do FNDCT, visando assegurar apoio financeiro à participação brasileira no LSST.
	Mobilizar a CONFAP e as FAPs para apresentação urgente de propostas específicas de encomendas que garantam recursos imediatos à participação brasileira no LSST.
	Realizar ação direta do MCTI para estruturação do projeto Independent Data Access Center (IDAC), por meio da assinatura urgente de convênio com a Associação .
	Promover ações claras e contínuas de comunicação junto às instituições públicas de Ciência & Tecnologia, destacando os ganhos potenciais e estratégicos do modelo proposto.
	Garantir ampla publicização e comunicação transparente sobre as decisões tomadas e os resultados alcançados para toda a comunidade científica brasileira, incluindo SAB, ABC e SBF.
	Manter rigorosa transparência financeira nas ações e projetos implementados.
	Criar mecanismos ágeis e efetivos de governança, para otimizar e dar celeridade à aplicação dos recursos públicos.
	Envolver diretamente as comunidades científicas em processos decisórios, fortalecendo a cooperação e a interdisciplinaridade no âmbito nacional.
	Utilizar a infraestrutura desenvolvida para o LIneA como modelo para projetos semelhantes em outras áreas científicas, visando uma expansão consistente do conceito de e-Ciência no país.

AÇÕES DE LONGO PRAZO	Implementar um projeto robusto e consistente de âmbito nacional, baseado nos subsídios obtidos durante o processo, beneficiando todas as áreas que processam e armazenam grandes volumes de dados, visando autonomia e soberania tecnológica e científica.
	Garantir sustentabilidade financeira e institucional, assegurando que ações emergenciais sejam acompanhadas de medidas estratégicas duradouras.





12	CONCLUSÃO
----	-----------

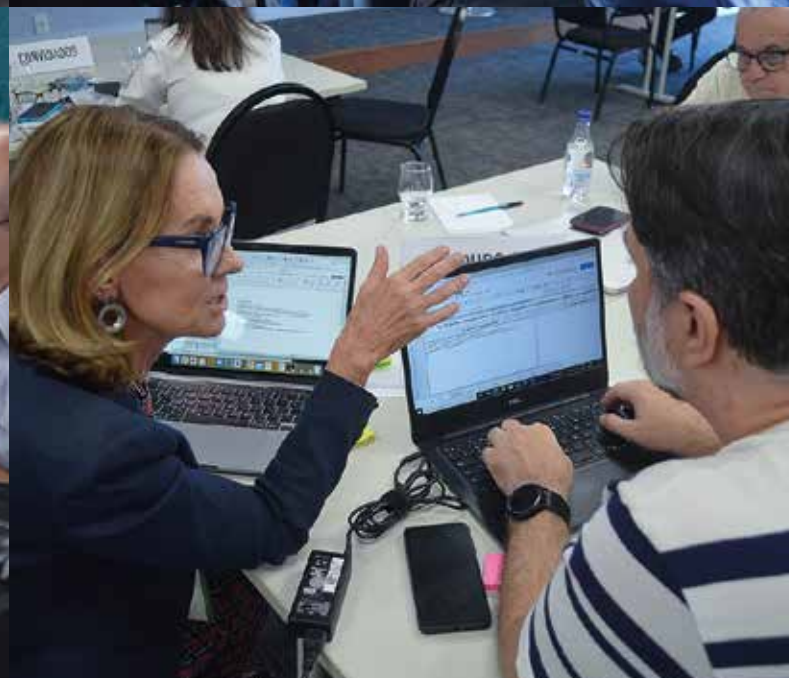
Ao realizar o processamento final do workshop, a **secretária de Políticas e Programas Estratégicos do MCTI, Andrea Latgé** disse que a realização do workshop representou um passo fundamental para o avanço da e-Ciência no Brasil, destacando a importância da colaboração interinstitucional e interdisciplinar. Ela evidenciou que, apesar dos desafios de curto prazo, que precisam de uma resposta célere, iniciativas como esta são essenciais para definir estratégias sustentáveis e efetivas a longo prazo. Em sua análise, as contribuições dos participantes foram decisivas para traçar caminhos possíveis em direção a um Planejamento Estratégico

robusto, e promissor para o futuro da e-Ciência brasileira. Segundo ela, “o comprometimento demonstrado pelos participantes reforça a importância da união e da mobilização em prol da ciência nacional, fortalecendo as capacidades internas e a soberania tecnológica do país”.

Não esquecendo os próximos passos, o **representante do CGEE, Thyrso Villela**, agradeceu a presença e disponibilidade de todos, esclarecendo que o CGEE vai estudar a possibilidade de expandir o diálogo para outras áreas, de forma a ampliar a representatividade e o repertório de informações para a formulação de uma proposta efetiva ao MCTI.

GALERIA DE FOTOGRAFIAS









01010100010001000111110000
101000101101010101010001
10101010101000100010001111
01010100010001000111110000
10101010100010001000111110
101000101101010101010001
10101010101000100010001111
01010100010001000111110000
10101010100010001000111110
01010101010001000100011111
01010100010001000111110000



Centro de Gestão e Estudos Estratégicos
Ciência, Tecnologia e Inovação

kozovits
conteúdo essencial

E-CIÊNCIA

kozovits
conteúdo essencial

conteúdo essencial