

LIneA: construindo uma infraestrutura para astronomia brasileira na era de Big Data – CCAD-Astro

C. Adean, R. Bouffleur, L. da Costa, M. Fernandes, N. Garcia, J. Gschwend, L. Silva, C. Singulani, E. Troyack, e G. Vila-Verde

¹ Laboratório Interinstitucional de e-Astronomia (LIneA). e-mail: contato@linea.org.br

Recebido em 29/09/2026; aceito em 05/10/2026

Abstract. Astronomy is undergoing a profound transformation driven by the production of unprecedented volumes of scientific data. Large observational surveys have come to require not only increasingly larger and more sophisticated telescopes and detectors, but also infrastructures capable of storing, curating, distributing, combining, and analyzing enormous datasets. In this context, the Interinstitutional Laboratory of e-Astronomy (LIneA) has promoted the establishment of a strategic scientific infrastructure for Brazilian participation in the era of data-driven astronomy. Over nearly two decades, the laboratory has developed competencies in high-performance computing, data curation, scientific software development, and the operation of analysis platforms, enabling, since 2006, Brazil's integration into international projects such as the Dark Energy Survey (DES), the Sloan Digital Sky Survey (SDSS), the Dark Energy Spectroscopic Instrument (DESI), and the Legacy Survey of Space and Time (LSST). This trajectory culminated in the deployment of IDAC-BR (Independent Data Access Center – Brazil), one of the authorized international centers for accessing data from the Vera C. Rubin Observatory, and paved the way for the creation of the Center for Astronomy Data Curation and Analysis (CCAD-Astro). CCAD-Astro was conceived not only to make this infrastructure available to the entire Brazilian astronomical community but also to foster the use of Artificial Intelligence and Machine Learning in the analysis of large volumes of data. The proposal directly responds to the needs identified by the community in the National Astronomy Plan (PNA), particularly regarding the strengthening of the national infrastructure for curation, processing, and high-level scientific analysis. This article presents this evolution and discusses its significance for the future of data-driven astronomy in the country.

Resumo. A astronomia vive uma profunda transformação impulsionada pela produção de volumes sem precedentes de dados científicos. Grandes levantamentos observacionais passaram a exigir não apenas telescópios e detectores cada vez maiores e mais sofisticados, mas também infraestruturas capazes de armazenar, curar, distribuir, combinar e analisar enormes acervos de dados. Nesse contexto, o Laboratório Interinstitucional de e-Astronomia (LIneA) vem promovendo a criação de uma infraestrutura científica estratégica para a participação brasileira em grandes projetos na era da astronomia orientada a dados. Ao longo de quase duas décadas, o laboratório desenvolveu competências em computação de alto desempenho, curadoria de dados, desenvolvimento de software científico e operação de plataformas de análise, permitindo, a partir de 2006, a inserção do Brasil em projetos internacionais como o Dark Energy Survey (DES), o Sloan Digital Sky Survey (SDSS), o Dark Energy Spectroscopic Instrument (DESI) e o Legacy Survey of Space and Time (LSST). Essa trajetória culminou na implantação do IDAC-BR (Independent Data Access Center – Brasil), um dos centros internacionais autorizados para acesso aos dados do Observatório Vera C. Rubin, e abriu caminho para a criação do Centro de Curadoria e Análise de Dados em Astronomia (CCAD-Astro). O CCAD-Astro foi concebido não apenas para disponibilizar essa infraestrutura a toda a comunidade astronômica brasileira, mas também para impulsionar o uso de Inteligência Artificial e Machine Learning na análise de grandes volumes de dados. A proposta responde diretamente às necessidades identificadas pela comunidade no Plano Nacional de Astronomia (PNA), particularmente no fortalecimento da infraestrutura nacional para curadoria, processamento e análise científica de alto nível. Este artigo apresenta essa evolução e discute seu significado para o futuro da astronomia orientada a dados no país.

Palavras-chave. Astronomia orientada por dados; e-Ciência; curadoria de dados; computação de alto desempenho; LSST; infraestrutura científica.

1. Introdução

A astronomia atravessa uma transformação profunda, impulsionada pela produção de volumes de dados sem precedentes. Nas últimas décadas, levantamentos sistemáticos do céu como o Sloan Digital Sky Survey (SDSS, York et al. (2000)), o Dark Energy Survey (DES, Sánchez (2016)), o Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration (2016)) e, mais recentemente, o Legacy Survey of Space and Time (LSST, Ivezić et al. (2019)), do Observatório Vera C. Rubin, passaram a produzir catálogos contendo bilhões de objetos e acervos que atingirão dezenas de petabytes. Nesse novo cenário, a capacidade de curar, armazenar, combinar, compartilhar e analisar dados tornou-se tão estratégica para a pesquisa quanto a própria construção dos telescópios e instrumentos científicos.

Essa transformação alterou profundamente a própria infraestrutura da pesquisa astronômica. A exploração científica desses grandes acervos passou a depender de ambientes capazes de aproximar o processamento dos dados (next-to-data), reduzindo a movimentação de enormes volumes de informação e permitindo que milhares de pesquisadores utilizem simultaneamente recursos computacionais avançados. Como consequência, infraestrutura computacional, desenvolvimento de software, curadoria de dados e plataformas científicas deixaram de ser atividades de apoio para se tornarem componentes essenciais da pesquisa astronômica contemporânea. A capacidade de transformar grandes volumes de dados em conhecimento científico tornou-se tão estratégica quanto a capacidade de produzi-los.

Foi nesse contexto que diversas comunidades astronômicas passaram a construir infraestruturas para gestão e análise

de dados para apoiar grandes levantamentos internacionais. No exterior, destacam-se o Astro Data Lab, operado pelo U.S. National Science Foundation National Optical-Infrared Astronomy Research Laboratory (NOIRLab), o Canadian Astronomy Data Centre (CADC), o Barbara A. Mikulski Archive for Space Telescopes (MAST) hospedado no Space Telescope Science Institute (STScI), e a mais recente plataforma European Open Science Cloud (EOSC). No Brasil, essa estratégia materializou-se com a criação em 2010 do Laboratório Interinstitucional de e-Astronomia (LIneA). Desde sua origem, o laboratório adotou uma estratégia baseada em contribuições tecnológicas (in-kind), desenvolvendo infraestrutura computacional, plataformas científicas e serviços especializados em troca da participação brasileira nas colaborações e do acesso antecipado aos seus dados.

A maturidade alcançada por essa infraestrutura foi reconhecida internacionalmente com a seleção do LIneA para sediar o IDAC-BR (Independent Data Access Center – Brasil), um dos centros internacionais autorizados (Figura 1) para prover acesso aos dados do Observatório Vera C. Rubin. Mais do que ampliar a participação brasileira em uma das maiores colaborações científicas da atualidade, essa conquista consolidou um conjunto de competências que une infraestrutura avançada de TI, computação científica, curadoria de dados e operação de plataformas, que vão além das necessidades de um único projeto. Essa experiência abriu caminho para a criação do Centro de Curadoria e Análise de Dados em Astronomia (CCAD-Astro), concebido para disponibilizar essa infraestrutura a toda a comunidade astronômica brasileira.

Este artigo apresenta essa trajetória, mostrando como a participação brasileira em grandes levantamentos internacionais levou à construção de uma infraestrutura científica para a Astronomia orientada por dados. Mais do que relatar a evolução institucional do LIneA, o presente artigo procura demonstrar como investimentos realizados para atender projetos específicos resultaram na formação de competências, plataformas e serviços que culminaram na criação do CCAD-Astro.

2. Da participação em grandes levantamentos à construção do LIneA

A origem do LIneA está diretamente associada à transformação da Astronomia em uma ciência orientada por grandes levantamentos e grandes volumes de dados. Em meados da década de 2000, tornou-se evidente que a participação em projetos internacionais de fronteira dependeria não apenas da excelência científica dos grupos de pesquisa, mas também da capacidade de contribuir com infraestrutura computacional, desenvolvimento de software e serviços especializados. Ao mesmo tempo, diversas colaborações passaram a adotar um novo modelo de participação, no qual parte dos custos de adesão poderia ser compensada por contribuições tecnológicas e científicas (in-kind), criando uma oportunidade para países com competências técnicas mas recursos financeiros limitados. Essa mudança inaugurou uma nova forma de inserção internacional, na qual infraestrutura computacional, desenvolvimento de software e operação de serviços científicos passaram a constituir ativos estratégicos para a participação em projetos de fronteira.

Foi nesse contexto que, em 2006, iniciaram-se os esforços para viabilizar a participação brasileira no Dark Energy Survey. Rapidamente ficou claro que o sucesso dessa estratégia exigia a formação de equipes multidisciplinares capazes de desenvolver

e operar bancos de dados científicos, pipelines de processamento, plataformas de acesso aos dados e ferramentas de análise. Essa necessidade motivou a criação em 2010 do Laboratório Interinstitucional de e-Astronomia (LIneA), reunindo inicialmente o Observatório Nacional, o Laboratório Nacional de Computação Científica e o Centro Brasileiro de Pesquisas Físicas. Em 2012, a Rede Nacional de Ensino e Pesquisa (RNP) passou a integrar a colaboração, ampliando sua capacidade de atuação nas áreas de infraestrutura de redes avançadas, segurança e serviços distribuídos, mediante um acordo institucional que consagrou a natureza transversal da iniciativa. A criação do LIneA representou, assim, a institucionalização de uma estratégia de longo prazo voltada ao desenvolvimento de competências fundamentais para a Astronomia orientada por dados.

Desde sua criação, o LIneA adotou uma estratégia baseada na conversão de competências tecnológicas em oportunidades científicas para a comunidade brasileira. Em vez de limitar sua atuação ao desenvolvimento de software para projetos específicos, o laboratório passou a construir uma infraestrutura operacional capaz de atender sucessivas gerações de levantamentos astronômicos. Essa estratégia viabilizou a participação brasileira nos projetos DES, SDSS, DESI e, posteriormente, no LSST, reduzindo substancialmente os custos financeiros de adesão e permitindo que pesquisadores brasileiros participassem ativamente dessas colaborações internacionais. Cada novo projeto passou a contribuir não apenas para seus próprios objetivos científicos, mas também para ampliar a infraestrutura e as competências disponíveis para as colaborações seguintes.

No âmbito do DES, o LIneA assumiu responsabilidades que abrangeram praticamente todo o ciclo de vida dos dados científicos. Foram desenvolvidos sistemas para controle de qualidade das observações, ferramentas de inspeção automática dos dados produzidos pela câmera DECam, ambientes para validação dos produtos gerados pelos pipelines de processamento e serviços destinados à visualização e exploração de imagens e catálogos científicos. Essas atividades proporcionaram à equipe experiência prática na operação contínua de sistemas distribuídos e no desenvolvimento de plataformas científicas utilizadas por centenas de pesquisadores. Elas marcaram a consolidação das primeiras competências do LIneA em operação contínua de plataformas científicas, processamento de grandes volumes de dados e curadoria de produtos científicos.

Entre as contribuições realizadas durante esse período, destaca-se o desenvolvimento do DES Science Portal (Fausti et al., 2018; Gschwend et al., 2018), projetado como um ambiente integrado para preparação e análise científica dos dados do levantamento. Mais do que um portal de acesso, o sistema reunia bancos de dados, recursos computacionais, algoritmos e ferramentas especializadas em uma única plataforma, permitindo que diferentes etapas da análise científica fossem realizadas diretamente sobre os dados armazenados (Aguena et al., 2021; Benoist et al., 2025). Essa arquitetura reduz significativamente a necessidade de transferir grandes volumes de informação entre instituições e introduz um ambiente colaborativo que favorece a reprodutibilidade dos resultados e o compartilhamento de metodologias entre diferentes grupos de pesquisa.

Embora o conceito de aproximar o processamento dos dados já fosse discutido internacionalmente, o DES Science Portal demonstrou, na prática, as vantagens dessa abordagem para levantamentos de grande escala. Muitos dos princípios desenvolvidos nesse ambiente anteciparam o paradigma hoje conhecido como next-to-data, que posteriormente se tornou um dos pilares da infraestrutura computacional do Observatório Vera C. Rubin. O DES Science Portal também introduziu mecanismos para a gestão de produtos científicos e de sua proveniência, assegurando

¹ <https://knutago.github.io/idac-dashboard/>

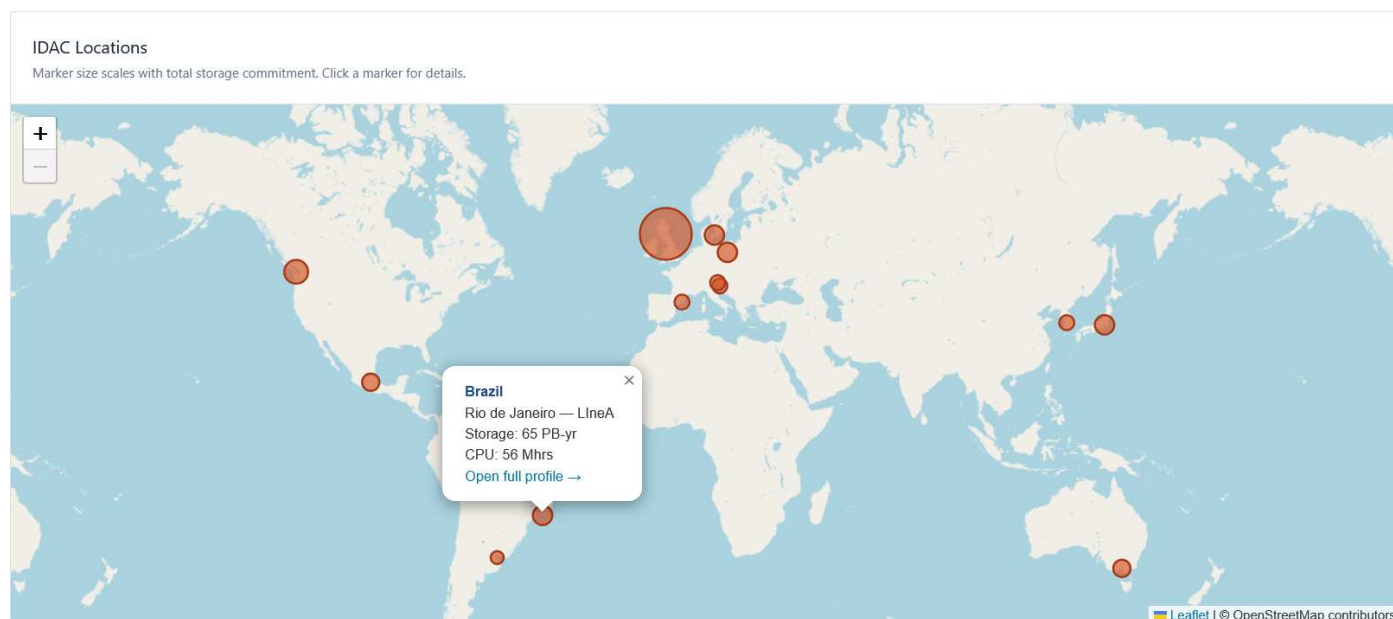


FIGURA 1. Distribuição geográfica da rede de IDACs onde os dados do projeto LSST poderão ser acessados. O raio do círculo reflete o tamanho da infraestrutura sendo oferecida como contribuição in-kind por cada país. Fonte: Rubin Observatory - IDAC Dashboard¹

a rastreabilidade das etapas de processamento, a reprodutibilidade das análises e o compartilhamento eficiente dos resultados. Diversos desses conceitos nortearam abordagens que posteriormente passaram a integrar documentos técnicos e de arquitetura do LSST (e.g., DMTN-049²), onde a infraestrutura utilizada para a produção de redshifts fotométricos do portal foi citada nominalmente como um exemplo de recurso tecnológico e de gestão desejável para o projeto. Mais importante, o DES Science Portal demonstrou que plataformas científicas poderiam tornar-se parte integrante da infraestrutura de pesquisa, aproximando dados, software e recursos computacionais em um ambiente único para exploração científica. Essa experiência estabeleceu as bases conceituais e tecnológicas sobre as quais seriam posteriormente construídas as atuais plataformas científicas e a infraestrutura operacional do IDAC-BR.

A experiência acumulada no DES foi rapidamente ampliada para outras colaborações. No SDSS, o LIneA implantou e operou um espelho nacional do banco de dados do levantamento e participou da instalação e monitoramento da infraestrutura computacional para os sistemas de aquisição de dados do projeto APOGEE-Sul. No DESI, desenvolveu ferramentas para inspeção em tempo real da qualidade dos milhares de espectros produzidos a cada exposição. Embora distintos em seus objetivos científicos, esses projetos contribuíram para consolidar competências em bancos de dados científicos, processamento distribuído, visualização de grandes acervos e operação contínua de serviços de apoio à pesquisa. Com isso, o laboratório ampliou suas competências incorporando experiência em bancos de dados científicos, processamento distribuído e serviços operacionais de missão crítica.

Paralelamente ao desenvolvimento tecnológico, o LIneA desempenhou um papel importante na preparação da comunidade astronômica brasileira para a nova era da ciência orientada por dados. Ao longo desse período, foram organizados cursos, workshops, bootcamps, escolas e seminários dedicados aos grandes levantamentos astronômicos, contribuindo para a capacitação de estudantes e jovens pesquisadores em áreas como computação científica, gestão de dados, desenvolvimento de software e análise

de grandes bases de dados. O apoio dado pelo INCT do e-Universe e pela FINEP foi decisivo para consolidar essa capacidade, viabilizando a expansão da infraestrutura computacional e a qualificação de equipes altamente especializadas. Essas iniciativas contribuíram para fortalecer uma comunidade capaz de explorar as novas infraestruturas científicas que começavam a ser construídas no país.

À medida que a infraestrutura crescia e as responsabilidades do laboratório se ampliavam, tornou-se evidente que sua continuidade exigia um modelo institucional mais adequado à operação contínua de serviços científicos. Tendo em vista a natureza interdisciplinar da iniciativa, para viabilizar a sua sustentabilidade, o Ministério da Ciência, Tecnologia e Inovação (MCTI) criou um grupo de trabalho para sugerir possíveis cenários administrativos para o laboratório. A recomendação desse grupo à Secretaria Executiva do MCTI levou, em 2020, à fundação da Associação LIneA, uma instituição privada sem fins lucrativos. Essa transformação representou uma significativa mudança da principal missão do laboratório, que inicialmente era voltada ao desenvolvimento de soluções para projetos específicos e se tornou uma instituição responsável pela operação de uma infraestrutura computacional e de um vasto acervo de dados com a oferta de diversos serviços para a comunidade astronômica.

O novo modelo proporcionou maior flexibilidade administrativa e agilidade para a aquisição de equipamentos e a contratação de especialistas, ampliou a capacidade de estabelecer acordos nacionais e internacionais e criou as condições necessárias para que o laboratório assumisse compromissos institucionais de longo prazo. Permaneceu, contudo, pendente a definição de um modelo de financiamento de longo prazo para a sustentabilidade do laboratório, uma vez que, naquele período, seus custos operacionais eram absorvidos pelo INCT do e-Universe e ainda não estava definido como seria financiada uma estrutura dessa natureza após o término desse apoio do e-Universe e por projetos FINEP enquanto foram procuradas várias alternativas junto ao MCTI.

Essa trajetória preparou naturalmente o LIneA para assumir uma nova responsabilidade: integrar a infraestrutura internacional de acesso aos dados do Observatório Vera C. Rubin por

² <https://dmtn-049.lsst.io>

meio da implantação do IDAC-BR. A Figura 2 ilustra esse processo, mostrando como a participação brasileira em grandes levantamentos internacionais impulsionou a construção de competências, infraestrutura e serviços científicos que culminaram na implantação do IDAC-BR e, posteriormente, na criação do CCAD-Astro como uma infraestrutura científica compartilhada para a comunidade astronômica brasileira.

A participação brasileira em grandes levantamentos astronômicos internacionais impulsionou a construção de competências em computação científica, engenharia de software, curadoria de dados e operação de plataformas. Essas competências foram consolidadas com a implantação do IDAC-BR, integrando o LIneA à infraestrutura operacional do Observatório Vera C. Rubin, e evoluíram para o CCAD-Astro, concebido como uma infraestrutura científica compartilhada para apoiar a pesquisa, a formação de recursos humanos e a colaboração da comunidade astronômica brasileira na era da ciência intensiva em dados.

3. O IDAC-BR: da participação científica à operação de uma infraestrutura internacional

A maturidade técnica e institucional construída ao longo da trajetória descrita na seção anterior foi colocada à prova quando o Observatório Vera C. Rubin lançou, em 2019, uma chamada internacional para a implantação de uma rede distribuída de Independent Data Access Centers (IDACs). Esses centros foram concebidos para complementar a infraestrutura principal de dados do projeto, ampliando a capacidade mundial de acesso, processamento e exploração científica dos produtos do Legacy Survey of Space and Time (LSST) durante toda a sua vida operacional.

Para o LIneA, essa chamada representava uma oportunidade distinta de todas as experiências anteriores. Nas colaborações em que havia participado até então, o laboratório contribuiu principalmente com software, plataformas científicas e serviços especializados. A implantação de um IDAC, entretanto, exigia assumir a operação contínua de uma infraestrutura integrante do ecossistema oficial do Observatório Vera C. Rubin. O desafio deixava de ser apenas desenvolver soluções tecnológicas para uma colaboração científica e passava a envolver a responsabilidade de oferecer aos pesquisadores um ambiente operacional completo, integrado aos sistemas internacionais de distribuição de dados e suporte científico.

A proposta submetida pelo LIneA foi aprovada em 2022, reconhecendo a maturidade alcançada pela instituição na operação de infraestruturas para ciência intensiva em dados. Em 2024, a assinatura do acordo com o SLAC National Accelerator Laboratory, organização afiliada ao Departamento de Energia dos Estados Unidos, formalizou a criação do IDAC-BR e tornou o LIneA a primeira instituição internacional a estabelecer esse tipo de acordo diretamente com a organização responsável pela operação científica do Observatório Vera C. Rubin. Pela primeira vez, uma instituição brasileira passava a integrar oficialmente a infraestrutura operacional de um dos maiores projetos científicos da atualidade, assumindo responsabilidades permanentes pela disponibilização dos dados do LSST à comunidade brasileira durante toda a vida do levantamento.

A implantação do IDAC-BR exigiu uma profunda modernização da infraestrutura do LIneA que ainda está em andamento e ao longo dos 10 anos de operação serão necessárias pelo menos duas atualizações dos equipamentos envolvidos segundo acordado com o SLAC³. Com o apoio do programa INCT do e-Universe e da FINEP, foram incorporados novos recursos de

armazenamento e computação de alto desempenho, a arquitetura de rede foi redesenhada para atender aos requisitos de desempenho e segurança definidos pela colaboração e foram implantados mecanismos de autenticação federada compatíveis com os padrões internacionais adotados pelo Rubin Observatory. Paralelamente, diversos sistemas desenvolvidos ao longo da trajetória do laboratório foram adaptados e integrados às novas plataformas utilizadas pelo projeto.

Mais importante do que a expansão da infraestrutura física foi a consolidação, em escala operacional, do paradigma next-to-data. Os volumes de dados produzidos pelo LSST tornam inviável o modelo tradicional, no qual cada pesquisador transfere para sua instituição os dados necessários às suas análises. Em seu lugar, os algoritmos passam a ser executados diretamente sobre os dados armazenados no IDAC, utilizando recursos computacionais localizados junto aos próprios acervos científicos. Essa abordagem reduz significativamente a movimentação de grandes volumes de informação pelas redes, melhora o desempenho das análises, favorece a reprodutibilidade dos resultados e assegura o cumprimento das políticas de acesso durante o período de embargo dos dados.

Embora muitos desses conceitos já estivessem presentes em iniciativas anteriores do LIneA, particularmente no DES Science Portal, a implantação do IDAC-BR representou sua consolidação em uma infraestrutura operacional de larga escala. O laboratório passou a integrar armazenamento de grande capacidade, computação de alto desempenho, bancos de dados científicos, plataformas interativas baseadas em Jupyter Notebooks e um conjunto crescente de serviços especializados para exploração de imagens, catálogos e produtos científicos. Dessa forma, o IDAC-BR deixou de representar apenas um centro de acesso aos dados do Rubin Observatory para constituir um ambiente completo de pesquisa científica.

A implantação do IDAC também motivou o desenvolvimento do LIneA Science Platform (LSP) (Gschwend et al., 2025), concebido para oferecer uma interface integrada de acesso aos diversos serviços disponibilizados pelo laboratório para o LSST e a comunidade como um todo. Reunindo ferramentas de consulta, visualização, exploração de imagens e catálogos, ambientes interativos de análise e acesso aos recursos computacionais do CCAD, a plataforma simplificou significativamente a experiência dos usuários e estabeleceu uma arquitetura suficientemente flexível para incorporar continuamente novos serviços científicos.

O desenvolvimento dessa infraestrutura ocorreu em estreita colaboração com parceiros internacionais. O LIneA estabeleceu acordos de cooperação com outros IDACs e centros de dados do Rubin Observatory, ampliou sua participação na LSST Discovery Alliance e fortaleceu colaborações com organizações dedicadas ao desenvolvimento de software científico e tecnologias para ciência intensiva em dados, como a LINCC Frameworks⁴. Paralelamente, passou também a integrar a rede nacional de e-Ciência coordenada pela RNP, ampliando sua inserção em iniciativas voltadas ao desenvolvimento de infraestruturas digitais compartilhadas.

Além da operação do IDAC-BR, o laboratório passou a contribuir diretamente para diferentes componentes da infraestrutura científica do Rubin Observatory. Entre essas atividades destacam-se: 1) o desenvolvimento, implementação e operação do serviço Transformed Engineering and Facility Database (EFD, SQR-058)⁵ no âmbito do Metadata Consolidated Database

³ <https://tinyurl.com/acordoslac>

⁴ Clique para acessar o link: LINCC Frameworks.

⁵ <https://sqr-058.lsst.io/>

Evolução da Infraestrutura Científica construída pelo LIneA

De contribuições tecnológicas em grandes levantamentos internacionais à infraestrutura compartilhada para toda a Astronomia brasileira



FIGURA 2. Evolução da infraestrutura científica construída pelo LIneA.

(ConsDB, DMTN-227)⁶; 2) a participação no processamento de produtos avançados do LSST e no desenvolvimento de produtos derivados, como estimativas de redshifts fotométricos para os data releases oficiais, gerados a partir da própria infraestrutura do Rubin Observatory hospedada no SLAC National Accelerator Laboratory (SITCOMTN-154⁷, RTN-124⁸), e; 3) estimativas probabilísticas de redshifts fotométricos geradas na infraestrutura do IDAC-BR (BRA-LIN-S4)⁹; a disponibilização de catálogos científicos complementares, como catálogos de aglomerados de galáxias e predições de ocultações estelares por pequenos corpos do Sistema Solar; e contribuições para projetos de software open source, incluindo o pipeline RAIL (RAIL team, 2026) e a biblioteca LSDB (Caplar et al., 2025).

Essa infraestrutura também ampliou o alcance das ações de extensão e ciência cidadã desenvolvidas pelo laboratório. Serviços como as predições de ocultações estelares permitem a participação coordenada de redes de observadores distribuídos geograficamente, incluindo astrônomos amadores, fortalecendo a interação entre a pesquisa profissional e a sociedade.

Os primeiros testes em condições reais ocorreram com a disponibilização do Data Preview 1 (DP1), em 2025. Diferentemente dos futuros Data Releases, destinados à exploração científica do levantamento, os Data Previews utilizam dados de comissionamento para validar a infraestrutura operacional. O DP1 permitiu verificar toda a cadeia de recepção, catalogação, armazenamento, autenticação e disponibilização dos dados, além de identificar gargalos e aperfeiçoar processos. Posteriormente, o recebimento do Data Preview 2 (DP2), cujo volume de dados em uma única entrega é comparável ao produzido pelo Dark Energy Survey durante seus seis anos de

operação, confirmou a capacidade operacional da infraestrutura e preparou o IDAC-BR para o início dos Data Releases científicos do LSST.

Mais do que um marco para a participação brasileira no Observatório Vera C. Rubin, a implantação do IDAC-BR consolidou uma infraestrutura operacional para ciência orientada por dados. A combinação de recursos computacionais, plataformas científicas, serviços especializados e equipes altamente qualificadas construída para atender às exigências do LSST ultrapassou as necessidades de uma única colaboração internacional e criou as bases sobre as quais seria concebido o Centro de Curadoria e Análise de Dados em Astronomia (CCAD-Astro), tema da próxima seção.

4. Do IDAC-BR ao CCAD-Astro: da operação de uma infraestrutura internacional a uma facility nacional

A implantação do IDAC-BR representou um ponto de inflexão na trajetória do LIneA. Mais do que assegurar o acesso da comunidade brasileira aos dados do Observatório Vera C. Rubin, a operação do centro exigiu a construção de uma infraestrutura completa para armazenamento, processamento, distribuição e análise de grandes volumes de dados científicos, além da implantação de serviços de autenticação federada, gestão de identidades, monitoramento operacional e suporte contínuo aos usuários. Esse processo também impulsionou a adoção de boas práticas de governança e segurança da informação, alinhando a instituição aos requisitos de ambientes internacionais de computação distribuída. A constatação de que essa capacidade poderia atender simultaneamente diferentes projetos científicos motivou a ideia do Centro de Curadoria e Análise de Dados em Astronomia (CCAD-Astro).

Ao longo dessa implantação tornou-se evidente que as necessidades específicas do IDAC correspondiam apenas a uma fração das capacidades desenvolvidas. Para garantir robustez op-

⁶ <https://dmtn-227.lsst.io/>

⁷ <https://sitcomtn-154.lsst.io/>

⁸ <https://rtn-124.lsst.io/>

⁹ https://linea-it.github.io/pz-lsst-inkind-doc/s4_4/

eracional, escalabilidade e sustentabilidade ao longo da década de operação do LSST, o LIneA precisou construir uma infraestrutura significativamente mais ampla e flexível do que aquela estritamente requerida pelo projeto. O resultado foi uma plataforma científica de propósito geral, composta por recursos computacionais, armazenamento, plataformas de análise, serviços de dados e ferramentas colaborativas capazes de atender simultaneamente diferentes levantamentos astronômicos e comunidades científicas. Dessa forma, o investimento realizado para viabilizar o IDAC-BR produziu uma infraestrutura que poderia ser compartilhada por outros projetos praticamente sem necessidade de novos desenvolvimentos estruturais.

Essa constatação motivou a criação do Centro de Curadoria e Análise de Dados em Astronomia (CCAD-Astro), uma evolução natural da infraestrutura implantada para o IDAC-BR. Seu objetivo é transformar essa capacidade operacional em uma facility nacional aberta à comunidade astronômica brasileira, oferecendo serviços especializados de curadoria, preservação, processamento e análise de dados científicos.

A missão do CCAD-Astro está também alinhada às necessidades identificadas pela própria comunidade astronômica brasileira no Plano Nacional de Astronomia (PNA), particularmente àquelas relacionadas ao fortalecimento da infraestrutura nacional para uma Astronomia cada vez mais intensiva em dados e à ampliação do acesso da comunidade aos grandes conjuntos de dados e aos recursos necessários para seu processamento e análise. O CCAD-Astro busca traduzir essas necessidades em uma infraestrutura concreta e compartilhada, integrando dados, computação, software e serviços científicos em um mesmo ambiente. Assim, a proposta não parte apenas da disponibilidade de uma infraestrutura já existente, mas da convergência entre uma capacidade tecnológica construída ao longo de quase duas décadas e demandas estratégicas identificadas pela comunidade.

Ao disponibilizar esses serviços de forma compartilhada, o CCAD-Astro amplia significativamente o acesso da comunidade nacional à infraestrutura necessária para a Astronomia orientada por dados. Essa infraestrutura beneficia tanto grandes colaborações quanto projetos conduzidos por pesquisadores individuais e pequenos grupos, democratizando o acesso a recursos computacionais avançados e a plataformas científicas antes restritos a poucas instituições. Além disso, favorece a reprodutibilidade científica, o reúso de dados e cria oportunidades para iniciativas de ciência cidadã, aproximando estudantes, professores e astrônomos amadores das modernas plataformas de exploração de dados astronômicos.

No contexto do CCAD-Astro, o conceito de curadoria de dados vai muito além do armazenamento de grandes acervos. A curadoria compreende todo o ciclo de vida dos dados científicos, desde a identificação de conjuntos de interesse, provenientes de grandes levantamentos observacionais, campanhas de pequeno e médio porte ou produtos científicos gerados pelos próprios pesquisadores, até sua transferência, organização, documentação, preservação e disponibilização à comunidade. Esse processo é orientado pelos princípios FAIR¹⁰, garantindo que os dados sejam encontráveis, acessíveis, interoperáveis e reutilizáveis (Wilkinson et al. 2016).

Os dados são disponibilizados por meio de plataformas científicas que integram ferramentas de visualização, consultas interativas, cruzamento de catálogos, exploração científica e processamento distribuído, combinando bancos de dados relacionais, formatos otimizados para análise em larga escala, ambientes baseados em Jupyter Notebooks e recursos de computação de alto desempenho. Ao concentrar dados, software e capacidade

computacional em um único ambiente, a infraestrutura reduz significativamente o esforço necessário para preparar experimentos científicos, eliminando boa parte das tarefas associadas à instalação de softwares, gerenciamento de dependências, atualização de bibliotecas e manutenção de ambientes computacionais. Dessa forma, os pesquisadores podem dedicar maior parte de seu tempo às questões científicas propriamente ditas, aumentando a produtividade e favorecendo a reusabilidade dos dados e a reprodutibilidade dos resultados.

A infraestrutura reúne atualmente dezenas de levantamentos internacionais, produtos científicos derivados, simulações numéricas e acervos produzidos por grupos brasileiros. A coexistência desses diferentes conjuntos de dados em um ambiente único favorece o compartilhamento de metodologias, reduz a duplicação de esforços, elimina gargalos associados à transferência de grandes volumes de dados e facilita o cruzamento de catálogos em larga escala, possibilitando pesquisas em astronomia multi-mensageira e de múltiplos comprimentos de onda. Além disso, a organização e sanitização desses grandes acervos cria condições para a aplicação de técnicas avançadas de mineração de dados e inteligência

O CCAD-Astro também estabelece uma oportunidade estratégica de integração com o Laboratório Nacional de Astrofísica (LNA). A infraestrutura desenvolvida pelo LIneA poderá ampliar o modelo de curadoria atualmente implementado em colaboração com o grupo de Sistema Solar para outros programas científicos do LNA, incluindo a preservação e valorização do acervo de observações acumulado ao longo de mais de quatro décadas de operação do Observatório do Pico dos Dias (OPD). Em uma etapa posterior, essa integração poderá incorporar também dados produzidos em instalações internacionais utilizadas pela comunidade brasileira, como SOAR e Gemini, reunindo em um ambiente único diferentes acervos observacionais para preservação, curadoria e análise científica segundo os princípios FAIR, dando maior visibilidade à produção científica brasileira e possibilitando o reúso dos dados na era do LSST.

A demanda por uma infraestrutura nacional compartilhada tornou-se evidente durante a reunião anual da Sociedade Astronômica Brasileira, em 2025, quando diversos grupos manifestaram interesse em utilizar os recursos computacionais do LIneA para seus próprios projetos científicos. Esse interesse revelou a existência de uma demanda reprimida por infraestrutura especializada e motivou o lançamento, no final daquele ano, de uma chamada pública (SINCADA)¹¹, projeto piloto do CCAD-Astro. O objetivo era validar a infraestrutura em condições reais de operação, identificar as necessidades da comunidade e antecipar os desafios decorrentes da futura chegada dos grandes volumes de dados do LSST.

A resposta da comunidade confirmou plenamente essa expectativa. Aproximadamente 52 pesquisadores e estudantes, pertencentes a sete instituições, submeteram cerca de vinte projetos cobrindo uma ampla variedade de temas científicos. As propostas demandavam desde armazenamento de grandes acervos até ambientes de computação de alto desempenho para análise de dados e desenvolvimento de software científico. O projeto piloto confirmou que a demanda por infraestrutura compartilhada para ciência intensiva em dados vai muito além das grandes colaborações internacionais, abrangendo grupos de diferentes portes e áreas de pesquisa da Astronomia brasileira.

Recentemente, foi realizada uma enquete com os pesquisadores que participaram da chamada pública, com o objetivo de avaliar a experiência de uso, identificar necessidades computacionais e orientar a evolução do ambiente. Os projetos

¹⁰ <https://www.go-fair.org/fair-principles/>

¹¹ Comunicações da SAB, Nº 1121 – 14 de novembro de 2025.

abrangem diferentes áreas da Astronomia, incluindo dinâmica orbital e corpos do Sistema Solar, astrofísica estelar e de altas energias, jatos relativísticos, raios cósmicos, formação planetária e análise de meteoros. As demandas concentram-se em computação de alto desempenho para simulações e processamento científico, armazenamento de dados e execução de códigos e softwares especializados, incluindo aplicações MPI e métodos de Monte Carlo. Ao todo 22 pesquisadores de 10 projetos responderam a enquete e os resultados estão disponíveis¹².

Os resultados deste experimento indicam que o CCAD-Astro já funciona como um ambiente compartilhado de pesquisa, no qual computação, armazenamento, software e suporte são combinados para viabilizar o trabalho científico. A experiência já resultou em publicações e apresentações científicas, incluindo eventos internacionais, e permitiu o acesso de pesquisadores sem experiência prévia em HPC. O principal resultado da avaliação não é a satisfação com a infraestrutura, mas o reconhecimento por parte dos usuários de que ela produz valor científico.

O SINCADA também tem contribuído para a evolução do CCAD-Astro em três eixos principais. Primeiro, permite dimensionar a infraestrutura a partir do uso real, fornecendo evidências sobre demanda por processamento, armazenamento, softwares e diferentes tipos de recursos computacionais. Segundo, possibilita testar e ajustar a operação em escala, incluindo mecanismos de acesso, alocação de recursos, suporte aos usuários, documentação e monitoramento, antes da expansão do serviço para uma comunidade mais ampla. Terceiro, oferece elementos para estabelecer políticas de uso, permitindo definir critérios de acesso, prioridades, limites de recursos, formas de acompanhamento e mecanismos de avaliação dos projetos. Dessa forma, o SINCADA funciona como um ambiente de experimentação operacional que reduz a incerteza associada à implantação de uma infraestrutura nacional e fornece uma base concreta para o planejamento e a governança do CCAD-Astro.

Desde então, o CCAD-Astro vem ampliando continuamente sua base de usuários. Atualmente, a infraestrutura atende mais de 420 usuários registrados, pertencentes a mais de 105 organizações, apoiando pesquisadores de diversos estados brasileiros. Os projetos apoiados abrangem áreas que vão da instrumentação astronômica, astrometria e física solar até cosmologia, estrutura em grande escala, simulações numéricas e aplicações que se estendem para além da própria Astronomia apoiando outras áreas científicas (e.g. biofísica) e start-ups, procurando explorar diferentes modelos de sustentabilidade.

A diversidade das instituições participantes e a distribuição geográfica dos usuários demonstram que a infraestrutura desenvolvida pelo LIneA já atende uma comunidade nacional ampla e distribuída. Esse resultado evidencia que a demanda por serviços compartilhados de curadoria, computação e análise de dados não se restringe às grandes colaborações internacionais, mas constitui uma necessidade transversal da Astronomia brasileira (Figura 3).

O mapa mostra o alcance nacional da infraestrutura, reunindo pesquisadores de diferentes instituições e regiões do país. Essa distribuição demonstra que os serviços desenvolvidos pelo LIneA já atendem uma comunidade ampla e diversificada, reforçando o papel do CCAD-Astro como uma infraestrutura científica compartilhada para a Astronomia brasileira.

Cada projeto incorporado à infraestrutura passa por um processo de integração que inclui treinamento, definição dos ambientes computacionais, organização dos dados e acompanhamento técnico por parte da equipe do LIneA. Essa experiência demonstrou que o sucesso de uma infraestrutura dessa natureza depende

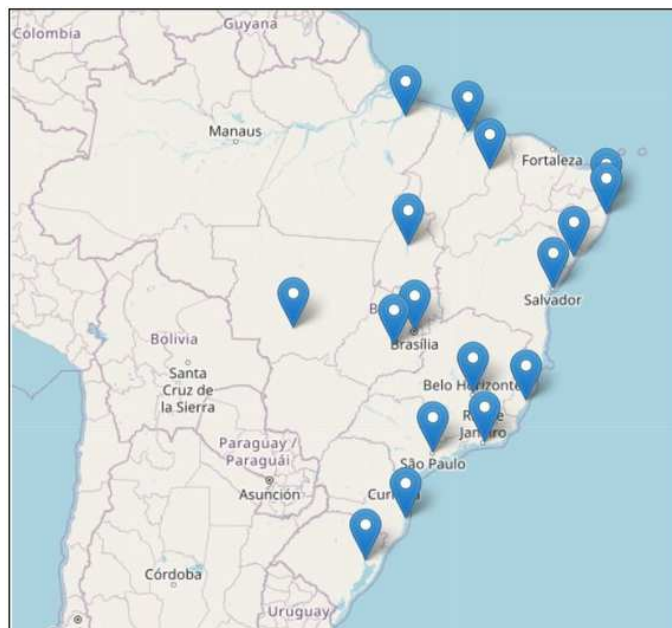


FIGURA 3. Distribuição geográfica das instituições e usuários da infraestrutura científica do LIneA/CCAD-Astro.

não apenas da disponibilidade de recursos computacionais, mas também da existência de documentação adequada, equipes multidisciplinares e suporte especializado capazes de orientar os usuários na utilização eficiente da plataforma e na adoção de boas práticas de desenvolvimento de software científico. Em outras palavras, a operação de um centro de curadoria envolve tanto a gestão da infraestrutura quanto a capacitação da comunidade para sua utilização.

Os primeiros meses de operação confirmaram a elevada taxa de utilização da infraestrutura e demonstraram que diferentes projetos podem compartilhar simultaneamente recursos computacionais, serviços de dados e plataformas científicas sem perda de flexibilidade. Esse modelo permite otimizar investimentos públicos, evitar duplicações entre instituições e ampliar significativamente o acesso da comunidade nacional a recursos avançados de computação científica.

Idealizado inicialmente para atender às necessidades da Astronomia, o CCAD-Astro foi desenvolvido para incorporar continuamente novas tecnologias, incluindo aplicações baseadas em inteligência artificial, aprendizado de máquina e análise distribuída de grandes volumes de dados. Embora tenha surgido a partir das demandas da Astronomia brasileira, seu modelo de operação responde a desafios comuns a diversas áreas da ciência intensiva em dados, demonstrando como investimentos realizados para um projeto científico específico podem resultar em uma infraestrutura compartilhada e de interesse estratégico para todo o sistema nacional de pesquisa gerando economia escala e otimização de recursos públicos.

O CCAD-Astro representa, assim, a culminação de uma estratégia iniciada quase duas décadas antes com a participação do LIneA em grandes colaborações internacionais. O que começou como o desenvolvimento de ferramentas para projetos específicos evoluiu para uma infraestrutura nacional dedicada à curadoria, preservação e análise de dados científicos, preparada para sustentar a ciência orientada por dados nas próximas décadas e servir de referência para outras áreas do conhecimento.

¹² <https://tinyurl.com/feedbacksincada>

5. Perspectivas

O início das operações científicas do Observatório Vera C. Rubin inaugura uma nova etapa para a Astronomia orientada por dados. Durante a próxima década, o LSST produzirá um fluxo contínuo de imagens, catálogos e alertas que transformarão profundamente a maneira como a pesquisa astronômica é conduzida. A exploração eficiente desse volume de informações dependerá cada vez menos da capacidade técnica individual de cada grupo de pesquisa e cada vez mais da existência de infraestruturas compartilhadas capazes de integrar armazenamento, processamento, curadoria de dados e ferramentas avançadas de análise.

Nesse contexto, o principal desafio do LIneA deixa de ser a implantação do IDAC-BR e passa a ser sua operação sustentável e sua evolução contínua. A infraestrutura deverá acompanhar o crescimento progressivo dos volumes de dados do Rubin Observatory, incorporar novos serviços científicos, atualizar suas plataformas computacionais e responder às demandas de uma comunidade em constante expansão, já que os dados liberados anualmente ficarão disponibilizados para toda a comunidade com dois anos de diferença após a liberação para membros. Ao mesmo tempo, será necessário integrar novos levantamentos, produtos derivados e bases de dados produzidas pela própria comunidade astronômica, consolidando o CCAD-Astro como uma infraestrutura científica compartilhada para a Astronomia brasileira. Mais do que ampliar a capacidade computacional disponível, o desafio passa a ser transformar grandes acervos de dados em um ambiente científico integrado, no qual dados, software, algoritmos e serviços evoluem conjuntamente ao longo da próxima década.

A rápida incorporação de Inteligência Artificial (IA) e Aprendizado de Máquina (ML) está ampliando significativamente a capacidade de exploração científica dos grandes levantamentos astronômicos. Esses métodos desempenham um papel cada vez mais importante na classificação automática de bilhões de objetos, na identificação de eventos transientes, na estimativa de parâmetros físicos e na descoberta de fenômenos raros. Entretanto, seu desempenho depende fundamentalmente da disponibilidade de dados bem curados, sanitizados, documentados e interoperáveis. A preparação desses grandes acervos — envolvendo curadoria, cruzamento de catálogos e controle de qualidade — tornou-se um dos principais desafios técnicos da Astronomia orientada por dados.

Nesse contexto, o CCAD-Astro desempenha um papel estratégico ao disponibilizar coleções científicas previamente curadas, harmonizadas e enriquecidas por meio de processos especializados de gestão de dados. Integrado ao paradigma next-to-data, o centro reúne, em um mesmo ambiente, dados, recursos computacionais e ferramentas de análise, permitindo que algoritmos de IA sejam executados diretamente sobre os acervos, reduzindo a movimentação de dados e aumentando a eficiência, a reprodutibilidade e a escalabilidade das análises. Nesse contexto, a curadoria de dados deixa de ser uma atividade de apoio e passa a constituir um componente essencial da infraestrutura necessária à aplicação eficiente de técnicas modernas de IA na Astronomia.

A rápida evolução dos modelos de linguagem e dos sistemas baseados em agentes cria novas oportunidades para simplificar o acesso aos grandes acervos científicos e automatizar parte das atividades de exploração e curadoria de dados. Nesse cenário, agentes inteligentes poderão atuar como assistentes científicos, auxiliando pesquisadores na formulação de consultas complexas, na descoberta e integração de conjuntos de dados, na orquestração de pipelines de processamento, na seleção de algoritmos de análise e na exploração interativa dos resultados. Ao reduzir a complexidade técnica associada ao acesso e ao processamento

de grandes volumes de dados, esse ambiente tem o potencial de democratizar o uso de metodologias avançadas e ampliar significativamente a capacidade de inovação científica da comunidade astronômica brasileira.

Além do sucesso no enfrentamento dos desafios tecnológicos, a consolidação do CCAD-Astro dependerá igualmente do fortalecimento da comunidade usuária e da construção de mecanismos institucionais capazes de garantir sua sustentabilidade ao longo do tempo.

Outra perspectiva importante é a consolidação do CCAD-Astro como um ambiente multiusuário e colaborativo para toda a comunidade astronômica brasileira. Além de hospedar acervos científicos e disponibilizar recursos computacionais, o centro já atua como espaço para desenvolvimento colaborativo de software, compartilhamento de algoritmos, hospedagem de plataformas desenvolvidas pela própria comunidade (e.g., GravDyn¹³, MackSun¹⁴, RADIORUM¹⁵), treinamento de estudantes e formação continuada de pesquisadores em ciência de dados. Ao compartilhar conhecimento, ferramentas e infraestrutura, o CCAD-Astro contribui para reduzir assimetrias entre instituições e ampliar a capacidade científica da comunidade astronômica nacional como um todo.

Embora o foco deste artigo seja a Astronomia, a experiência acumulada pelo LIneA passou também a ser reconhecida como referência para além dessa área. Essa conclusão foi corroborada pelo Design Sprint promovido pelo Centro de Gestão e Estudos Estratégicos (CGEE), que em seu relatório (Kozovits e Murad, 2025) identificou o LIneA como um caso bem-sucedido de infraestrutura científica orientada por dados e recomendou que sua experiência servisse de referência para a criação de centros semelhantes em outras áreas estratégicas. Esse resultado também confirma a visão já apresentada no Design Sprint organizado pela RNP em 2018, que antecipava a necessidade de uma rede nacional de centros de e-Ciência capazes de oferecer serviços compartilhados de computação, curadoria e análise de grandes volumes de dados para diferentes comunidades científicas. A trajetória do LIneA e a proposta do CCAD-Astro representam, portanto, uma evolução natural dessa visão. Nesse sentido, o CCAD-Astro demonstra como competências construídas ao longo da participação brasileira em grandes colaborações internacionais podem ser transformadas em uma infraestrutura científica, disponível para toda a comunidade astronômica nacional.

Assim como telescópios, observatórios e laboratórios nacionais constituem infraestruturas permanentes para a produção de conhecimento científico, centros de curadoria e análise de dados tornam-se componentes igualmente essenciais da Astronomia na era dos grandes levantamentos. Seu valor não está apenas nos equipamentos instalados, mas na capacidade de integrar equipes multidisciplinares, desenvolver tecnologias, preservar acervos científicos e disponibilizar serviços especializados para uma comunidade ampla e diversificada. A consolidação dessa visão dependerá, entretanto, da adoção de mecanismos de governança e financiamento capazes de assegurar sua operação, evolução tecnológica e retenção de equipes altamente qualificadas ao longo de todo o ciclo de vida dos grandes levantamentos astronômicos. Somente dessa forma será possível consolidar o CCAD-Astro como uma infraestrutura científica essencial para a Astronomia brasileira.

¹³ <https://gravdyn.linea.org.br/>

¹⁴ <https://www.macksun.org/index.html>

¹⁵ <http://plataforma.radiorum.com.br/>

6. Conclusões

A trajetória do LIneA reflete a transformação da própria astronomia em uma ciência intensiva em dados. Criado para viabilizar a participação brasileira em grandes levantamentos internacionais por meio de contribuições tecnológicas, o laboratório evoluiu gradualmente de um centro de desenvolvimento de software e serviços científicos para uma infraestrutura operacional de computação, curadoria e análise de dados. Essa evolução foi impulsionada por uma estratégia consistente de investimento em competências técnicas, formação de recursos humanos especializados, treinamento da comunidade com cursos e workshops oferecidos de forma gratuita, suporte técnico com atendimento por vídeo-chamadas, e construção de plataformas compartilhadas, permitindo ao Brasil participar de forma competitiva de algumas das principais colaborações astronômicas internacionais. Vale a pena destacar a estratégica contratação de mestres e doutores em astronomia com profundo conhecimento sobre dados astronômicos e competências técnicas reconhecidas para atuar na interface com a comunidade, traduzindo seus requisitos científicos em design e planejamento para o desenvolvimento de soluções tecnológicas customizadas para a astronomia.

A implantação do IDAC-BR representou um importante passo neste processo. Ao integrar a infraestrutura operacional do Observatório Vera C. Rubin, o LIneA passou a desempenhar um papel importante em um dos maiores projetos científicos da atualidade, demonstrando que as competências desenvolvidas no país atingiram um nível de maturidade compatível com os requisitos das grandes colaborações internacionais.

Ao mesmo tempo, a infraestrutura construída para atender ao LSST revelou um potencial muito mais amplo do que aquele previsto inicialmente. A criação do CCAD-Astro transforma esse investimento em uma infraestrutura científica compartilhada, disponibilizando à comunidade astronômica brasileira serviços de curadoria, armazenamento, processamento e análise de grandes acervos científicos. Mais do que apoiar projetos específicos, o centro cria as condições necessárias para fortalecer a colaboração entre instituições, ampliar o acesso a recursos computacionais avançados e preparar a comunidade para os desafios científicos da próxima década.

A criação do CCAD-Astro está também alinhada às prioridades identificadas pela comunidade astronômica brasileira no Plano Nacional de Astronomia (PNA), particularmente à necessidade de fortalecer a infraestrutura nacional para uma Astronomia cada vez mais intensiva em dados e ampliar o acesso da comunidade aos recursos necessários para a curadoria, processamento e análise de grandes acervos científicos.

O principal legado da trajetória do LIneA, entretanto, vai muito além da infraestrutura física construída. Ao longo de quase duas décadas, o laboratório formou equipes multidisciplinares altamente especializadas e consolidou competências em computação científica, engenharia de software, gestão de dados e operação de infraestruturas complexas.

As plataformas, os serviços e, sobretudo, o conjunto de conhecimentos desenvolvidos ao longo desse período constituem hoje um patrimônio estratégico para a Astronomia brasileira e representa um dos ativos mais duradouros construídos a partir da participação do país em grandes colaborações internacionais.

A experiência do LIneA demonstra que investimentos realizados para apoiar grandes levantamentos internacionais podem gerar benefícios que ultrapassam seus objetivos científicos imediatos. Além de viabilizar o acesso a dados e a participação em projetos de fronteira, essas colaborações promovem a transferência de conhecimento, o desenvolvimento tecnológico e a formação de recursos humanos altamente qualificados, criando

capacidades que permanecem no país muito além da duração de cada projeto.

Ao transformar esse conhecimento acumulado em uma infraestrutura científica como o CCAD-Astro, o LIneA ampliou o impacto desses investimentos para toda a comunidade astronômica brasileira. A consolidação desse patrimônio dependerá agora de sua ampla utilização pela comunidade científica e da adoção de mecanismos de governança e financiamento capazes de assegurar sua evolução, atualização tecnológica e sustentabilidade ao longo das próximas décadas.

Em uma era em que a capacidade de transformar dados em conhecimento é tão estratégica quanto a própria observação do Universo, consolidar infraestruturas nacionais de curadoria e análise de dados deixa de ser apenas um requisito técnico e passa a ser um elemento essencial para a soberania sobre os dados científicos, autonomia científica nacional e para a competitividade internacional da pesquisa brasileira.

O CCAD-Astro representa, assim, mais do que a evolução institucional do LIneA. Ele demonstra que é possível transformar competências adquiridas em grandes colaborações internacionais em uma infraestrutura científica consolidada, compartilhada e voltada ao fortalecimento de toda a comunidade astronômica brasileira. Em um cenário em que os dados passam a ocupar um papel central na produção do conhecimento, iniciativas dessa natureza tornam-se tão estratégicas para a pesquisa quanto os próprios observatórios e instrumentos científicos.

Agradecimentos. Este trabalho foi apoiado pela FINEP por meio dos auxílios nº 0311/16 e nº 0883/24, e pelo programa INCT do e-Universo, com recursos do CNPq (auxílio nº 465376/2014-2) e da FAPERJ (auxílios nº E-26/200.965/2018, E-26/201.681/2019, E-26/210.010/2018 e E-26/211.013/2019).

Declaração sobre o uso de Inteligência Artificial. A revisão do texto contou com o apoio de ferramentas de Inteligência Artificial.

References

- Aguena, M., C. Benoist, L. N. da Costa, R. L. C. Ogando, J. Gschwend, H. B. Sampaio-Santos, M. Lima, et al. 2021. "The WZP Galaxy Cluster Sample of the Dark Energy Survey Year 1." *Monthly Notices of the Royal Astronomical Society* 502 (3): 4435–4456. <https://doi.org/10.1093/mnras/stab264>.
- Benoist, C., M. Aguena, L. N. da Costa, J. Gschwend, S. Allam, O. Alves, F. Andrade-Oliveira, et al. 2025. "The Final WZP Galaxy Cluster Catalog of the Dark Energy Survey and Comparison with SZE Data." *arXiv preprint arXiv:2507.05360*. <https://doi.org/10.48550/arXiv.2507.05360>.
- Caplar, N., W. Beebe, D. Branton, S. Campos, A. Connolly, M. DeLucchi, D. Jones, et al. 2025. "Using LSDB to Enable Large-Scale Catalog Distribution, Cross-Matching, and Analytics." *arXiv preprint arXiv:2501.02103*. <https://doi.org/10.48550/arXiv.2501.02103>.
- DESI Collaboration. 2016. "The DESI Experiment Part I: Science, Targeting, and Survey Design." *arXiv e-prints*. <https://doi.org/10.48550/arXiv.1611.00036>.
- Fausti Neto, A., L. N. da Costa, A. Carnero, J. Gschwend, R. L. C. Ogando, F. Sobreira, M. A. G. Maia, et al. 2018. "DES Science Portal: Creating Science-Ready Catalogs." *Astronomy and Computing* 24: 52–69. <https://doi.org/10.1016/j.ascom.2018.01.002>.
- Gschwend, J., A. C. Rossel, R. L. C. Ogando, A. F. Neto, M. A. G. Maia, L. N. da Costa, M. Lima, et al. 2018. "DES Science Portal: Computing Photometric Redshifts." *Astronomy and Computing* 25: 58–80. <https://doi.org/10.1016/j.ascom.2018.08.008>.
- Gschwend, J. e LIneA team. 2026. "LIneA Science Platform: Tools for Scientific Exploration of LSST Data and Other Astronomical Archives." *Boletim da Sociedade Astronômica Brasileira* 37 (1): 198–200. <https://sab-astro.org.br/wp-content/uploads/2026/04/JuliaGschwend.pdf>.
- Ivezić, Ž., S. M. Kahn, J. A. Tyson, B. Abel, E. Acosta, R. Allsman, D. Alonso, et al. 2019. "LSST: From Science Drivers to Reference Design and Anticipated Data Products." *The Astrophysical Journal* 873 (2): 111. <https://doi.org/10.3847/1538-4357/ab042c>.
- Kozovits, C. e E. Murad. 2025. *Relatório: Workshop de Design Sprint e Ciência*. Rio de Janeiro: LIneA / CGEE. Acessado em 5 de agosto de 2026. <https://www.linea.org.br/admin/wp-content/uploads/2025/04/Relatorio-Workshop-de-Design-Sprint-e-Ciencia.pdf>.

- Sánchez, E. 2016. "The Dark Energy Survey: Status and First results." *Nuclear and Particle Physics Proceedings* 273–275: 302–308. <https://doi.org/10.1016/j.nuclphysbps.2015.09.042>.
- The RAIL Team, J. L. van den Busch, E. Charles, J. Cohen-Tanugi, A. Crafford, J. F. Crenshaw, S. Dagoret, et al. 2026. "Redshift Assessment Infrastructure Layers (RAIL): Rubin-Era Photometric Redshift Stress-Testing and at-Scale Production." *The Open Journal of Astrophysics* 9 (fevereiro). <https://doi.org/10.33232/001c.158200>.
- Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3: 160018. <https://doi.org/10.1038/sdata.2016.18>.
- York, D. G., J. Adelman, J. E. Anderson Jr., et al. 2000. "The Sloan Digital Sky Survey: Technical Summary." *The Astronomical Journal* 120: 1579–1587. <https://doi.org/10.1086/301513>.

Saiba mais sobre o trabalho desenvolvido pelo LIneA em: www.linea.org.br