



Relatório Anual 2020

INCT do e-Universo / LineA

	2
Sobre este documento	3
I. Introdução	3
II. Atividades em 2020	5
III. Principais Resultados	16
Ciência	16
Cientistas de Dados do LIneA	16
Colaboradores Externos	18
SDSS	18
TON	19
DESI	21
DES & LSST	21
Produção Científica	24
Outras Atividades	26
Educação	28
Divulgação	32
Serviços	36
Apoio aos Usuários	36
Operação	37
Desenvolvimento de software	37
Infraestrutura	38
IV. Discussão	41
V. Conclusão	44

Sobre este documento

Este é o relatório do conjunto de atividades do LIneA e do INCT do e-Universo, tendo em vista que a missão de ambas as organizações é conjunta com o INCT financiando a rede de pesquisa e parte da infraestrutura e com o LIneA fornecendo o apoio para que a pesquisa baseada no grande volume de dados dos levantamentos sendo apoiados seja executada.

O ano de 2020 apesar de ter apresentado grandes dificuldades, foi um ano importante, com a fundação da associação LIneA o que representa um importante passo no estabelecimento de um novo modelo institucional mais flexível e adaptado a nova era de *Big Science* e *Big Data* que requer equipes inter e multidisciplinares nas áreas de ciência da computação, física e astronomia, trabalhando colaborativamente em rede. O LIneA que celebrou em 2020 o seu aniversário de 10 anos tem demonstrado o seu potencial transformativo sendo um agente facilitador para acelerar no Brasil a pesquisa baseada em acervos legados, principalmente na era do projeto *Legacy Survey of Space and Time* (LSST), que irá revolucionar a área da astronomia.

Este relatório descreve as principais realizações alcançadas em 2020 nas várias áreas de atividade do LIneA e indica as metas prioritárias para 2021. Entre estas está a preparação da comunidade brasileira para o aproveitamento do acervo de dados do LSST, sem precedentes em termos de qualidade, profundidade, cobertura espacial e cadência. Este acervo abre novas fronteiras de pesquisa além da possibilidade de responder questões fundamentais da física moderna como a natureza da matéria e energia escura e possíveis limitações da teoria da relatividade geral.

I. Introdução

Este documento descreve as atividades de 2020 da recém criada associação Laboratório Interinstitucional de e-Astronomia (LIneA). O LIneA é um instituto privado de ciência e tecnologia fundado para atender novas demandas provenientes do envolvimento de pesquisadores brasileiros em projetos internacionais de *Big Science* que geram grandes volumes de dados (*Big Data*). A associação foi criada seguindo a sugestão do Grupo de Trabalho indicado pelo MCTIC para propor uma forma de institucionalização do esforço iniciado em 2006 com a entrada do Brasil no projeto *Dark Energy Survey* (DES), o primeiro de uma série de projetos internacionais negociados pelo LIneA. Originalmente um projeto estruturante do Observatório Nacional de 2006, o LIneA evoluiu a partir de 2010 para um acordo entre várias unidades de pesquisa do Ministério (CBPF, LNCC, ON, e RNP), refletindo o caráter multidisciplinar de suas atividades. Mais recentemente foi incubado na RNP, reconhecida sua atuação na área de e-ciência e portanto útil na modelagem de um Centro de Suporte de e-Ciência sendo discutido pelo LNCC e a RNP, até Setembro de 2020 quando terminaram os recursos liberados pelo MCTI.

Várias tentativas foram feitas a partir de Outubro de 2019 junto ao MCTI para complementar os recursos originalmente disponibilizados em Dezembro de 2018 pelo então Secretário-Executivo. Estas tentativas incluíram reuniões presenciais em Brasília do coordenador do LIneA em Outubro e Novembro de 2019, em carta do Conselho de Administração da recém fundada organização endereçada ao Ministro em Maio de 2020, e em reuniões remotas do Comitê de Usuários do LIneA e do Comitê Gestor do INCT do e-Universo, que depende do LIneA para suas atividades, com o Subsecretário das Unidades Vinculadas.

Diante das dificuldades encontradas, o LIneA a partir de Dezembro de 2019 implantou um programa de redução de pessoal que ao todo, levou a dispensa de sete indivíduos e da firma responsável há 10 anos pela manutenção dos mais de 120 equipamentos pertencentes ao LIneA localizados no PoP-RJ na Urca e, em sua grande maioria no centro de processamento do LNCC, em Petrópolis. A falta de recursos, a significativa redução do time, a busca de soluções alternativas e a pandemia fizeram do ano de 2020 o mais difícil de sua história ao completar oficialmente 10 anos de sua criação. Felizmente, graças ao apoio do INCT do e-Universo, com recursos liberados pelo CNPq e FAPERJ, da FINEP e da compreensão e dedicação da equipe do LIneA, foi possível chegar ao fim do ano com importantes realizações descritas na próxima seção. Apesar de alguns atrasos, causados pela redução do time, foi possível cumprir em grande parte o cronograma e alcançar as metas do planejamento realizado no início do ano, entre elas:

1. A consolidação da associação criando os conselhos responsáveis pela governança da organização, a preparação de um plano diretor apresentando a visão de futuro, a

definição da equipe necessária para atingir as metas propostas e os meios que garantam o financiamento adequado e previsível.

2. O cumprimento de compromissos assumidos pelos acordos celebrados com as colaborações científicas internacionais em particular com o SDSS e o DES.
3. A manutenção da infraestrutura de hardware e software adequada para que os pesquisadores associados às colaborações científicas apoiadas pudessem executar sua pesquisa de forma eficiente e competitiva.
4. Progresso na implantação de novos serviços que serão disponibilizados à toda a comunidade nacional para que esta se beneficie do repositório de dados e das ferramentas desenvolvidas e mantidas pelo LIneA.
5. A implantação e realização de atividades de educação e divulgação procurando: i) formar uma nova geração de jovens pesquisadores preparados para atuar internacionalmente; ii) capacitar pesquisadores e tecnologistas para enfrentar os problemas de *big data*; iii) treinar professores e alunos do ensino médio; e iv) estimular o público em geral a se interessar pelas atividades científicas compartilhando informações e resultados obtidos pelo trabalho realizado pela associação.

Finalmente, é importante enfatizar que 2020 foi um ano de transição com projetos como o DES e o SDSS-IV completando suas observações, enquanto os projetos DESI e LSST iniciaram seus períodos de verificação científica. Em particular, o projeto LSST representa um grande desafio pelo grande volume de dados que irá gerar e as demandas de armazenamento e processamento que isto significa. Sem dúvida, isto irá dominar as atividades do LIneA pelos próximos anos.

II. Atividades em 2020

Seguindo a recomendação do GT, mencionado acima, possíveis alternativas de institucionalização foram examinadas com assessoria jurídica e discutidas com o GT. Este concluiu que a melhor forma seria a criação de uma associação sem fins lucrativos credenciada como um Instituto de Ciência e Tecnologia. Isto levou a criação, em 17 de Janeiro de 2020, da Associação LIneA com a indicação de Luiz Nicolaci da Costa como seu Diretor-Geral e de Julio Camargo como Diretor Adjunto. A associação teve seu estatuto registrado em cartório no dia 12 de Março de 2020. Para sua governança o LIneA-ICT conta com um Conselho de Administração (CA)¹ formado por representantes de entidades governamentais, da academia e da sociedade civil. Também foram criadas uma Comissão Técnico-Científica (CTC)² com

¹ Membros do CA: Augusto Gadelha (presidente) Carlos Alberto Aragão, Eliete Bouskela, Jailson de Andrade, Luiz Fernando Palmer da Fonseca, Roberto Martins, Thyrso Vilella, Wilson Coury, Wanderley de Souza

² Membros do CTC: Beatriz Barbuy, Iara Machado, Márcia Barbosa, Othon Winter, Pedro Dias.

membros de reconhecida relevância nas áreas afins, um Comitê Fiscal e um Comitê de Usuários (CdU)³.

A Assembleia do LIneA aprovou a constituição do CA e do Conselho Fiscal remotamente em 29/04/2020. O CA teve sua primeira reunião em 31 de Março, quando aprovou a proposta de trabalho apresentada pelo Diretor-Geral, a composição da equipe e a tabela salarial proposta. Uma importante decisão foi o apoio dado à associação junto ao MCTI em carta enviada ao Sr. Ministro em 09/04/2020. Também em sessão remota, o CA aprovou a composição do CTC e do CdU, que realizaram suas primeiras reuniões (remotamente) em 20/05/2020 e 02/06/2020, respectivamente. Estes também se manifestaram a favor dos planos apresentados pela direção do LIneA, sendo que o CdU teve um encontro remoto no dia 06/07/2020 com o Subsecretário das Unidades Vinculadas solicitando apoio financeiro do MCTI para esta iniciativa.

A proposta apresentada pelo Diretor-Geral foi desenvolvida com o apoio de uma especialista contratada, logo após a fundação da associação, para coordenar a preparação de um plano diretor para o laboratório. Como parte desse processo, durante o mês de Fevereiro foram realizadas reuniões, envolvendo a equipe do LIneA e membros do INCT, para construir um cronograma alinhado com os compromissos internacionais assumidos e os interesses dos pesquisadores afiliados, e assim, ter elementos para estimar o número FTEs (*full-time equivalent* ou funcionários trabalhando em tempo integral) necessários para atender a demanda imposta pelas diferentes atividades do laboratório (administração, ciência, formação de pessoal, desenvolvimento de software, operação e serviços, divulgação) e os diferentes perfis profissionais (cientista de dados, engenheiros de software, programadores científicos, analistas, arquitetos, programadores de *front-end* e especialistas em educação e divulgação) exigidos para realização do trabalho.

A figura 1 ilustra a visão estratégica da missão do laboratório e sua singularidade, sendo o apoio à pesquisa e educação e divulgação seus eixos principais sustentados por inúmeros serviços que incluem apoio ao usuário, o desenvolvimento e operação de plataformas científicas, a curadoria de acervos de dados e acesso ao processamento de alto desempenho. O apoio à pesquisa, que até agora se limitou aos pesquisadores envolvidos diretamente nos projetos apoiados pelo LIneA, está sendo repensado. Durante o ano de 2020, um grande esforço foi feito para a implantação de: i) *kubernetes* cluster para o gerenciamento dos recursos de máquinas; ii) sistema de monitoramento dos recursos de máquinas; e iii) sistema de autenticação e autorização. A implantação destes sistemas permitirá o acesso de toda a comunidade aos dados e aos serviços mantidos pelo LIneA de forma sustentável. Complementando este esforço, o LIneA contratou uma especialista em educação e divulgação para preparar cursos de treinamento na familiarização dos dados disponíveis nos repositórios locais e no uso das várias ferramentas do LIneA para acessar, visualizar e analisar esses

³ Membros do CdU: Carlos Wuenschen, Carlos ADean, Fabio Porto, Gustavo Dias, Joao Canalle, Julia Gschwend, Marcelo Emilio, Mariana Penna Lima, Rogerio Riffel, Sandro Rembold.

dados. Esses cursos serão abertos a todos os interessados e estão previstos para o segundo trimestre de 2021. Experiências nesse sentido foram realizadas pelo LIneA no passado na forma de um [bootcamp](#) presencial em Setembro de 2019 e, mais recentemente, de um [tutorial](#) remoto para a equipe da RNP em Junho de 2020. Essas experiências e o material compilado nessas atividades formam a base dos cursos e tutoriais sendo preparados para 2021.

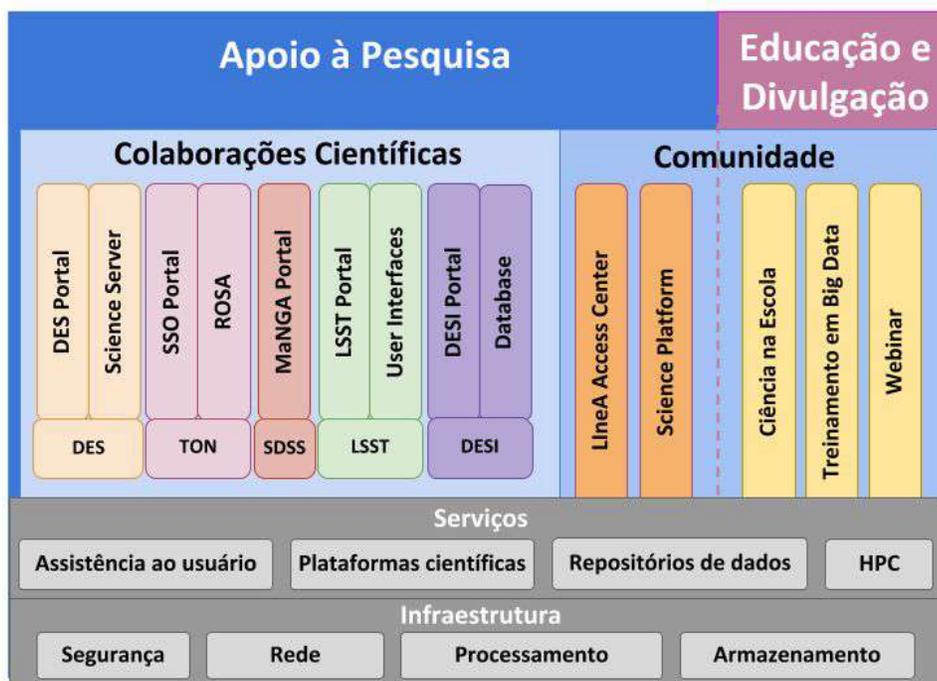


Figura 1 - Visão das atividades do LIneA. A figura também mostra alguns dos principais projetos de desenvolvimento de software em andamento para as cinco colaborações científicas sendo apoiadas e os projetos de interesse da comunidade.

A proposta para o LIneA-ICT é ter um pequeno time (~25 pessoas), formado por cientistas de dados, tecnologistas de TI e especialistas em educação, focado em atender as necessidades de: i) pesquisadores atuando nos cinco projetos sendo apoiados pelo LIneA (DES, SDSS, TON, DESI e LSST); ii) grupos emergentes; iii) pesquisadores individuais; e iv) alunos espalhados pelo Brasil. O objetivo é que todos tenham acesso aos acervos de dados mantido pelo LIneA, acesso a modernas ferramentas e algoritmos de análise e a sistemas de processamento, previamente reservados aos grandes centros de pesquisa, para explorar cientificamente estes dados. O LIneA vai ainda oferecer o treinamento e o apoio de uma equipe de especialistas. Com esta missão clara e específica o LIneA pretende complementar o trabalho mais focado no ensino e pesquisa realizado pelos atuais institutos de pesquisa e departamentos universitários.

Para o [DES](#), durante o ano de 2020, a equipe do LIneA continuou o seu esforço no desenvolvimento do portal científico que hospeda os algoritmos usados na análise dos dados (figura 2). O principal foco foi revisar os pipelines usados na preparação de catálogos prontos

para a análise científica, adaptando-os para lidar com os dados finais do projeto e as novas quantidades introduzidas pela equipe do NCSA, e reexaminando suas estruturas visando melhorar seus desempenhos.

Esses códigos foram executados no final do ano para criar produtos avançados como amostras espectroscópicas para treinamento de algoritmos usados para estimar redshifts fotométricos, para calcular redshift fotométricos para toda amostra de objetos (da ordem de 400 milhões) identificadas nas imagens e os mais de 100 mapas descrevendo a variação das condições observacionais, incluindo o mapa de profundidade, na região observada pelo levantamento para cada um dos cinco filtros.

Esses produtos foram usados por membros do consórcio DES-Brazil em suas análises científicas e compartilhados com a colaboração. Entre os tópicos abordados em 2020 pela equipe brasileira estão:

1. A identificação de 370 mil objetos do sistemas solar nos 4,2 milhões CCDs observados ao longo dos seis anos de observação. As posições desses objetos ao longo do tempo serão usadas para calcular órbitas e quando comparado com a posição de estrelas do catálogo GAIA de alta precisão astrométrica predizer eventos de ocultação estelar, quando estes objetos passam em frente a uma estrela e a sombra produzida na terra pode ser usada para determinar o tamanho e forma.
2. O estudo da estrutura da Via Láctea determinando os parâmetros que descrevem alguns de seus componentes como disco, bojo e halo a partir da comparação direta entre modelos e observações.
3. A identificação de aglomerados de galáxias nos dados dos seis anos de observação utilizando o algoritmo desenvolvido pela equipe do DES-Brazil e integrado ao portal científico baseado na distribuição espacial e de redshifts fotométricos.
4. A evolução da função de luminosidade usando os dados do levantamento como um todo e os dados de regiões menores mas onde as observações são significativamente mais profundas.

É importante salientar que a maioria destes trabalhos estão sendo desenvolvidos usando os algoritmos disponíveis no portal científico que servirá de legado para o projeto LSST sendo, portanto, um importante exercício na preparação da equipe para o início deste projeto previsto para 2023.

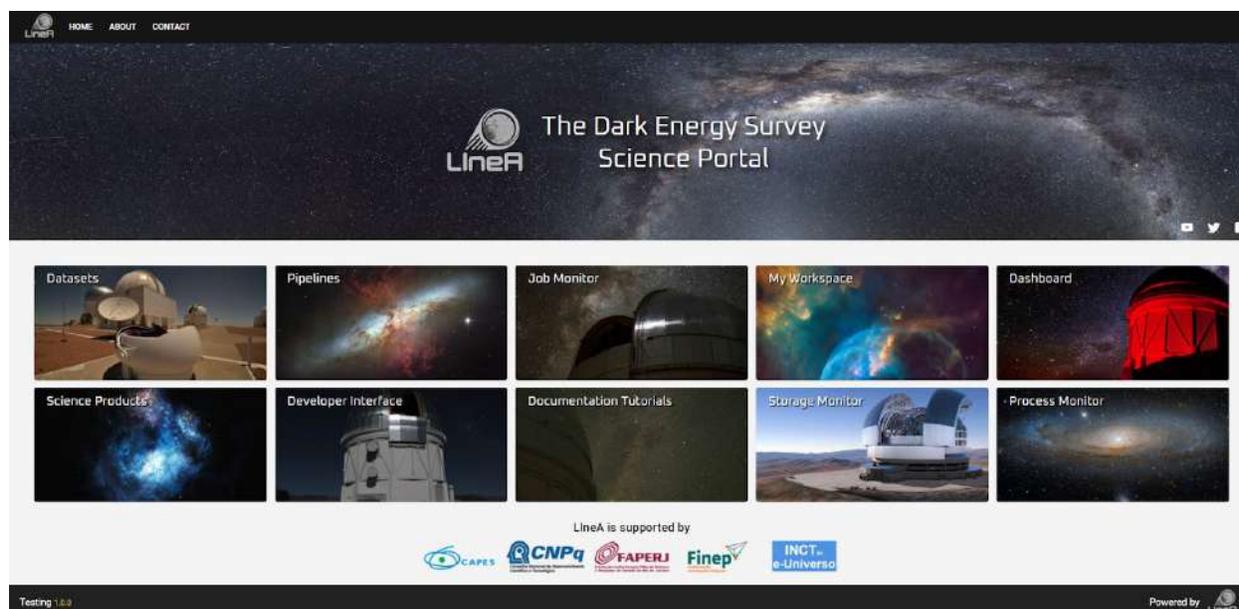


Figura 2 - Página principal (em desenvolvimento) do portal científico do DES, mostrando os vários serviços disponíveis.

Em paralelo, como parte dos compromissos assumidos pelo LIneA para a participação do DES-Brazil, a plataforma científica conhecida como *LIneA Science Server* (figura 3) também foi revista em preparação do lançamento público dos dados acumulados pelo DES ao longo dos seis anos (DR2), previsto para Janeiro de 2021. O Science Server desenvolvido pelo LIneA, e em operação desde 2018 no NCSA, é uma das interfaces disponíveis para membros da colaboração e para o público de acessar os dados do DES para a visualização de imagens, catálogos e de listas de objetos gerados por buscas no banco de dados. No momento, o LIneA também está empenhado, a exemplo do que acontece com o SDSS, de manter a curadoria dos dados deste levantamento e para isto está transferindo da ordem de 70 TB de dados para o LNCC. Esse serviço estará disponível para a comunidade brasileira a partir de Fevereiro de 2021.

Este esforço é parte do trabalho de preparação para a implantação de um centro de acesso aos dados do LSST sendo proposto pelo LIneA, uma das metas prioritárias para 2021. Este trabalho foi tema de uma das reuniões organizadas pelo LIneA, com a participação da comunidade, para avaliar diferentes aspectos do trabalho realizado por seu time de TI. A reunião de [avaliação](#) do Science Server foi realizada no dia 15/12 e serviu para validar várias novas funcionalidades, verificar medidas corretivas em preparação para a liberação da nova versão da plataforma em Janeiro para o DR2 do DES. O [relatório final](#) já está disponível e será publicado no site do LIneA.

O material apresentado também será usado para o curso de treinamento desta ferramenta prevista para 2021. O LIneA pretende usar estas reuniões para estimular o engajamento da comunidade no trabalho sendo realizado pela equipe de TI do LIneA convocando membros externos para participar da comissão avaliadora.

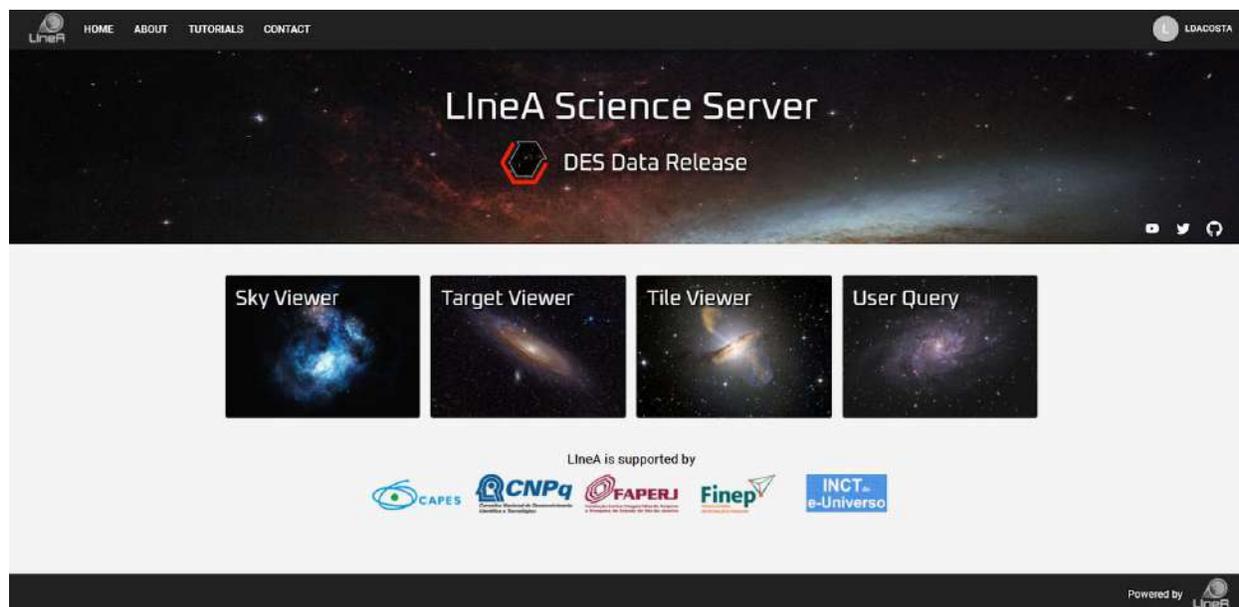


Figura 3 - Página de entrada da plataforma *LIneA Science Server* que dá acesso aos dados do DES. A plataforma desenvolvida pela equipe de TI do LIneA engloba serviços de visualização de imagens, catálogos e uma interface para buscas no banco de dados dos catálogos. A plataforma está em funcionamento no NCSA e em breve terá uma versão local para facilitar o acesso de pesquisadores brasileiros. Este é um primeiro passo para a transformação do LIneA para atuar como um centro de dados na era do LSST.

As plataformas científicas desenvolvidas pelo time de TI para o DES fazem parte de publicações em revistas especializadas ([Fausti Neto et al 2018](#), [Gschwend et al 2018](#)) e de um capítulo do livro “*The Dark Energy Survey: The Story of a Cosmological Experiments*” publicado este ano pela World Scientific com lançamentos em reunião remota a partir de Londres no dias 13 de Outubro e 25 de Novembro com a participação do Diretor do LIneA num dos painéis de autores. Um outro bom exemplo da utilidade do portal científico foi o trabalho desenvolvido por [Aguena e colaboradores](#) (2020) para criar um catálogo de aglomerados de galáxias alternativo ao que estava sendo utilizado pela colaboração baseado na hipótese da existência de uma sequência de galáxias vermelhas no espaço de cor-magnitude caracterizando a população de galáxias membros. Este trabalho que exigiu inúmeros testes ao longo de vários anos se beneficiou enormemente das facilidades oferecidas pelo portal.

A equipe científica brasileira no DES inclui pesquisadores do LIneA, ON, USP, UNICAMP, UNESP, e UFRGS cobrindo o sistema solar, a estrutura de nossa galáxia, astronomia extragaláctica e cosmologia. Essas atividades serão descritas na próxima seção.

O apoio dado ao **SDSS** por parte do LIneA consistiu da disponibilização dos dados do DR16, penúltima do SDSS-IV, realizada em Fevereiro de 2020, e o monitoramento remoto do sistema de computação responsável pela operação do espectrógrafo APOGEE-2 no Observatório de Las Campanas (LCO) no Chile. Lembrando que tanto a manutenção do repositório dos dados no Brasil e o apoio técnico no LCO fazem parte do acordo estabelecido entre o LIneA e o SDSS para a participação de pesquisadores brasileiros. Deste projeto participam vários pesquisadores, pós docs e alunos da USP, UFRGS e UFSM extremamente atuantes nos projetos MaNGA e APOGEE produzindo várias publicações sobre núcleos ativos e sobre o bojo Galáctico.

Para facilitar o trabalho sendo realizado do grupo de AGNs o LIneA vem desenvolvendo uma plataforma que permite a visualização dos 10.000 cubos de dados obtidos com o espectrógrafo de campo integral (IFU) utilizado nas observações. O [portal MaNGA](#) (figura 4) permite ao usuário visualizar os mapas mostrando o fluxo e espectro em cada ponto da galáxia, assim como mapas mostrando a variação espacial de dezenas de quantidades que podem ser derivadas a partir dos espectros obtido em cada fibra (e.g., metalicidade, dispersão de velocidades) colocada em cada ponto da galáxia. Uma reunião para a entrega desta plataforma para uso do time trabalhando no projeto MaNGA está prevista para segunda quinzena de Janeiro.

LineA MaNGA Portal

IFU Viewer

LineA is supported by

CAPES CNPq FAPERJ Finep INCT-e-Universo

Teetling 1.0.0 Powered by LineA

Figura 4 - Página de entrada do [portal MaNGA](#) desenvolvido pela equipe de TI do LIneA para os pesquisadores do grupo brasileiro de participação (BPG-SDSS) trabalhando neste levantamento.

O LIneA também contribuiu para a participação de um pequeno número de pesquisadores brasileiros da USP e da UNICAMP no projeto *Dark Energy Spectroscopic Instrument (DESI)*, graças ao desenvolvimento de um protótipo para uma plataforma para avaliar a qualidade 15.000 espectros de 5.000 projetos entregue a colaboração em 2019. O projeto, liderado pelo *Lawrence Berkeley National Laboratory (LBNL)*, talvez seja um dos mais promissores para determinar se a aceleração da expansão do Universo se deve à energia escura ou a limitações em grande escala da Teoria da Relatividade.

O progresso deste projeto tem sido extraordinário sendo que o período de verificação científica teve início em Março de 2020 quando foi interrompido pela pandemia. As observações foram retomadas com sucesso em Dezembro quando foram observadas 12 regiões com alvos de diferentes tipos, representando um importante passo para o início do levantamento. Este ano o Diretor do LIneA transferiu a representatividade no [Institutional Board](#) (54 instituições e consórcios) por uma pesquisadora da UNICAMP.

Em 2020 também continuou o seu apoio ao projeto *Transneptunian Occultation Network (TON)* contribuindo na preparação da proposta para a construção de uma rede de pequenos telescópios robóticos distribuídos de norte a sul do Brasil como ilustrado na figura 5.

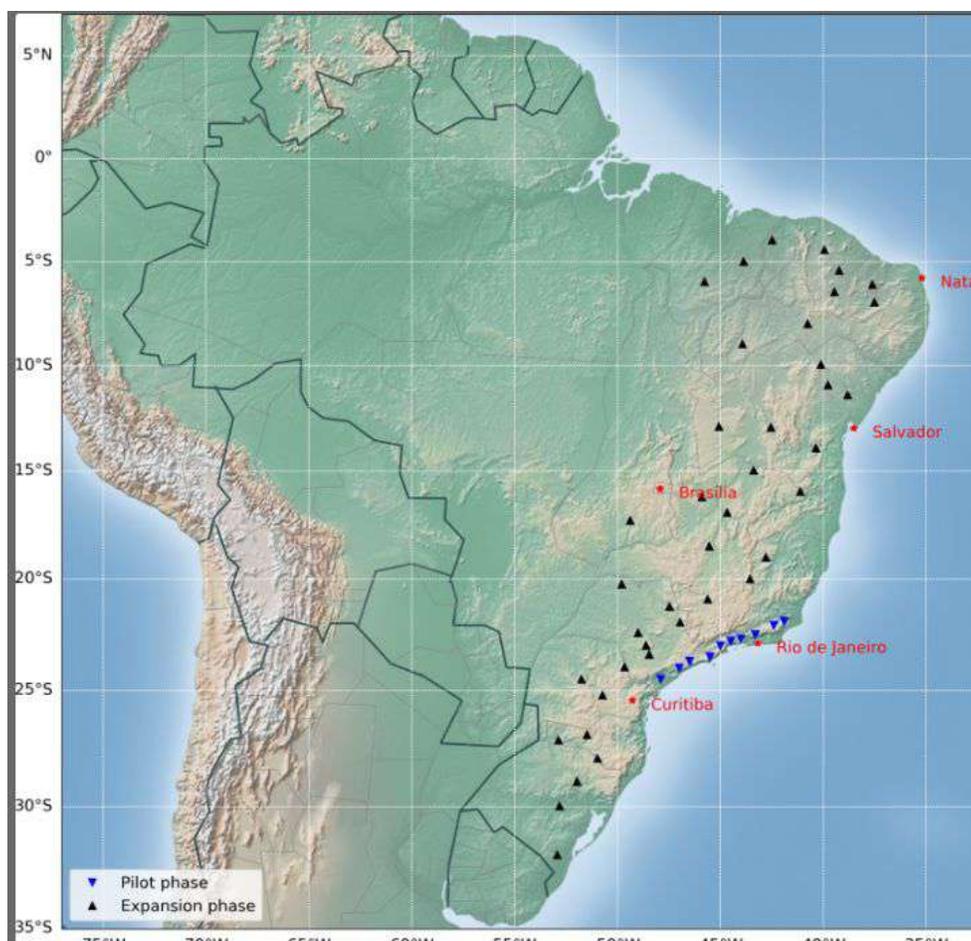


Figura 5 - Ilustração da Rede de Ocultação Sul-Americana (ROSA) que prevê a instalação de 50 telescópios em pequenas localidades do Brasil. A proposta é realizar o projeto em duas fases sendo a primeira a instalação de 10 telescópios com um espaçamento menor entre o Rio de Janeiro e Curitiba.

O projeto procura capitalizar no fato de que o LSST irá descobrir um grande número de novos objetos do sistema solar com órbitas bem determinadas, graças à cadência das observações onde cada ponto do céu será observado a cada três dias. Este fato somado à grande precisão da posição de estrelas obtidas pelo projeto GAIA implica que o número de eventos de ocultação irá crescer exponencialmente, com cerca de mais de 10 eventos por noite. Para aproveitar ao máximo esta incrível oportunidade foi proposto a construção de uma rede de pequenos telescópios robóticos distribuídos de norte a sul do Brasil como ilustrado na figura 3. O projeto completo é estimado em USD 2 milhões dos quais USD 500 mil foram solicitados, com o apoio do *LSST Corporation* (LSSTC), à Fundação Templeton para a construção de 10 telescópios entre o Rio de Janeiro e Curitiba. Neste sentido, uma carta de intenção foi enviada à Fundação em Julho de 2019 para consideração, que resultou num pedido da Fundação para que fosse submetida uma proposta final. O uso da Rede de Ocultação Sul Americana (ROSA)

pela colaboração LSST foi também incluído na carta de intenção enviada ao LSST em Novembro de 2019 como uma possível contribuição in-kind do LIneA. A proposta foi submetida em Janeiro de 2020, com o resultado final de que o projeto não seria financiado em Junho de 2020. Mesmo assim, tendo em vista o grande impacto que o projeto poderia ter para o grupo, este continua pretendendo levar adiante o projeto construindo, pelo menos, um protótipo com recursos de outras fontes.

Também em 2020 a equipe completou a primeira parte do desenvolvimento de uma plataforma científica solicitada pelo projeto TON. Os objetivos desta plataforma são:

1. Fazer predições de eventos de ocultação a partir de dados de grandes levantamentos fotométricos.
2. Compilar, processar e obter curvas de luz a partir de observações feitas em diferentes locais inclusive por astrónomos amadores.
3. Analisar as curvas de luz obtidas e determinar as características físicas dos objetos observados.

O primeiro caso de uso considerado foi o de identificar objetos do sistema solar nas imagens acumuladas pelo projeto DES, obter órbitas para esses objetos, combinar com dados disponíveis no *Minor Planet Center* e prever eventos de ocultação.

O projeto teve início em Março de 2018, e envolveu mais de 12 profissionais entre pesquisadores e tecnólogos, correspondendo a aproximadamente quatro pessoas trabalhando em tempo integral. A primeira versão do sistema foi entregue em Dezembro de 2019 e fez parte de uma tese de doutorado apresentada no Observatório Nacional. Como esta versão apresentou dificuldades de ser sustentada operacionalmente, foi feita uma reestruturação do sistema ao longo de 2020 que foi entregue para avaliação do grupo em Dezembro de 2020. A principal característica da nova versão é separar a fase de preparação de dados da predição propriamente dita que deve depender do interesse do usuário. As diferentes etapas da preparação dos dados envolve a identificação dos CCDs contendo objetos do sistema solar, o download de imagens ou catálogos associadas a essas imagens, a determinação da posição dos objetos de interesse nestas imagens e finalmente a construção da órbita. Essas diferentes etapas podem ser executadas separadamente pela página de acesso da plataforma apresentada na figura 6. A [reunião de avaliação](#) foi realizada em 10/12 quando foram apresentados os resultados da análise de mais de 180 mil exposições acumuladas ao longo dos seis anos de observação. Na reunião, importantes recomendações foram feitas que vão orientar as próximas implementações e uma visão de como o sistema poderá ser usado pelo grupo do TON e pela comunidade em geral. A comissão contou com a participação de pesquisadores do ON, Observatoire de Paris-Meudon e UTFPR e produziu o [relatório](#) que será disponibilizado no site do LIneA.



Figura 6 - Página de entrada para o portal *Solar System Objects* (SSO) desenvolvido pelo time de TI do LIneA em estreita colaboração com os pesquisadores envolvidos no projeto TON.

O LIneA se prepara agora para um novo e grande desafio representado pelo projeto LSST. Usando a experiência acumulada, ele está apoiando a participação de 50 pesquisadores neste projeto que formam o Grupo de Participação Brasileiro (BPG, em inglês). O projeto agora conhecido como o [Large Survey of Space and Time](#) irá observar mais da metade do céu por 10 anos e se estenderá até 2034. Cada posição do céu terá 800 visitas, produzindo ao mesmo tempo as imagens mais profundas já obtidas de metade do céu em seis filtros e um filme da evolução do céu no período. Todo este acervo ficará à disposição de todos em forma digital e poderá ser revisitado a qualquer momento. Isto servirá como referência para o futuro da Astronomia de onde amostras de alvos para diferentes estudos serão selecionadas por várias décadas. Esses dados serão usados para estudos que vão desde a formação do sistema solar à origem e evolução do Universo, em particular procurando responder questões básicas da física moderna como a natureza da matéria escura e da energia escura, esta última uma componente que corresponde a 70% da composição do Universo atual.

O LIneA tenta agora expandir o número de pesquisadores no BPG propondo possíveis contribuições *in-kind* permitindo assim a entrada de novos PIs. Neste sentido em Novembro de 2019 uma carta de intenção solicitada pelo projeto LSST foi enviada oferecendo os serviços do LIneA para:

1. O desenvolvimento de sistemas de software para a avaliação da qualidade das imagens brutas e reduzidas.

2. A geração de produtos científicos avançados para uso da colaboração como photo-z, catálogos de aglomerados de galáxias, de sistemas estelares e de galáxias anãs, previsões de eventos de ocultação estelar.
3. A geração de simulações para serem usadas no estudo de nossa galáxia e na análise cosmológica.
4. A implantação de um Centro de Acesso aos Dados do LSST, conhecido como *independent Data Access Center (iDAC)*.
5. A implantação de uma rede com 50 pequenos telescópios robóticos espalhados pelo Brasil.

A carta de intenção foi analisada pelo projeto e revista pelo *National Science Foundation (NSF)* e Departamento de Energia (DOE) do governo americano, as agências que financiam o projeto, que avaliaram os projetos de maior interesse para o LSST e solicitaram que uma proposta formal fosse submetida. Segue abaixo parte da resposta:

Notification of Approval to Further Develop

Proposed In-Kind Contributions in Return for LSST Data Rights

We are happy to inform you that your Letter of Interest (LOI) outlining possible in-kind contributions to Rubin Observatory has been **reviewed by NSF and DOE and approved for further development**. Following our email from February 17, 2020, we can confirm that the deadline for developing a full proposal has been moved to the end of June this year.

A proposta final foi preparada pela equipe do LIneA e submetida pelo seu Diretor em 25/09/2020 e inclui os seguintes itens:

- A implantação de um IDAC no Brasil.
- A participação de pesquisadores brasileiros no desenvolvimento de pipelines para o Grupo de Trabalho de Energia Escura (DESC).
- A geração local de produtos científicos avançados para serem compartilhados com toda a colaboração através da rede federada de IDACs.

Em paralelo, o LIneA submeteu um pedido para sua entrada na *LSST Corporation*, uma organização de apoio ao projeto LSST cuja principal função é a de angariar fundos privados e financiar atividades para a preparação da comunidade para o aproveitamento dos dados do LSST que serão sem precedentes. Entre suas iniciativas estão a organização de workshops internacionais, reuniões técnicas e bolsas para estudantes. A corporação conta com mais de 30 instituições e a lista completa pode ser vista [aqui](#). Mesmo sem ainda ser membro, o LIneA teve o apoio da corporação na submissão da proposta à Fundação Templeton e foi convidado para participar de uma chamada para concorrer a uma fellowship financiada pela corporação.

A entrada do LIneA na corporação foi recentemente aprovada em [carta](#) à direção do LIneA. Este é um importante passo para a integração do BPG e do seu time de TI a outros grupos que participam do projeto. Além disso, o Brasil fazer parte desta organização será mais uma forma de integrar a comunidade brasileira nas atividades patrocinadas pelo LSST.

III. Principais Resultados

III.1 Ciência

Cientistas de Dados do LIneA

Para atrair pesquisadores de qualidade para trabalhar no LIneA como cientistas de dados, a proposta é que estes possam dedicar 25% do seu tempo para sua pesquisa e os demais 75% para tarefas de infraestrutura que podem ser ou não alinhadas ou necessárias para a sua ciência. Pelo menos neste primeiro ano esta estratégia tem sido bem sucedida, com os três cientistas de dados contribuindo com serviços relevantes ao mesmo tempo produzindo artigos científicos de qualidade dos quais são os autores principais. Um destes trabalhos ([Pieres et al. 2020](#)) descreve os resultados obtidos da comparação entre dados de estrelas do terceiro ano do *Dark Energy Survey* (DES) com modelos do *Trilegal*, na busca por melhores parâmetros estruturais para o disco espesso e para o halo da Via-Láctea. Outro trabalho submetido para publicação em 2020 ([Aguena et al 2021](#)), descreve o algoritmo *Wavelet Z-Photometric* (WaZP) usado para identificar aglomerados de galáxias, disponibiliza o catálogo obtido com os dados do primeiro ano do DES, a propriedade dos sistemas identificados, e os resultados da comparação deste catálogo com os de outros autores.

Na área dos estudos sobre a Via Láctea estão sendo feitas adaptações no portal científico para poder incluir na análise dados de outros levantamentos e desta forma expandir a cobertura da Galáxia e o desenvolvimento de um novo algoritmo para a busca de galáxias anãs de baixo brilho superficial e aglomerados estelares esparsos para ser implementado no portal científico do LIneA.

No primeiro semestre, foi iniciada a análise, ainda em andamento, da evolução das populações de galáxias do DES que são membros de aglomerados detectados pelo software WaZP ([Aguena et al. 2021](#)), em comparação com as galáxias membros de aglomerados detectados pelo software redMaPPer ([Rykoff et al., 2014](#)). Esse estudo utiliza como ferramenta o *Jupyter Notebook* e tem como principal objeto de análise o diagrama cor-magnitude, avaliado em diferentes faixas de *redshift* e para diferentes faixas de riqueza dos aglomerados. O objetivo é o de investigar diferenças na população de galáxias membros em função do *redshift* que poderia justificar um melhor desempenho do WAZP em relação ao redmapper já que este algoritmo não depende da presença de uma população de galáxias elípticas a alto *redshifts*.

No segundo semestre, foram feitas melhorias na produção de *redshifts* fotométricos que são utilizados como um dos principais ingredientes para detecção de aglomerados pelo WaZP. Houve atualização do catálogo de *redshifts* espectroscópicos utilizado como conjunto de treinamento em métodos de *machine learning* que calculam o *photo-z*. Houve também a refatoração do esquema de paralelização do pipeline que computa o *photo-z* no portal, reduzindo significativamente o tempo das rodadas e possibilitando que novos catálogos de galáxias sejam gerados sob demanda em tempo real.

Em preparação para o LSST, foi iniciado o refatoramento do *pipeline* que calcula a Função de Luminosidade de galáxias, com os objetivos de: atualização da linguagem de programação e bibliotecas necessárias (Python 2 → Python 3); inclusão do cálculo da Função de Massa derivada da distribuição de probabilidade conjunta de massa e *redshift* fornecida por algoritmo de *machine learning* (Mucesh et al., 2020); adequação ao modelo de classes desenhado para o novo portal que o LIneA está desenvolvendo para processar os dados do LSST. Como contribuição *in-kind* para a colaboração DESC, foi iniciado o desenvolvimento do módulo “*evaluation*” do software [RAIL](#), com base na refatoração dos códigos utilizados na análise do Data Challenge 1 (DC1, [Schmidt et al., 2020](#)).

Outro trabalho em progresso está relacionado às regiões de detecções mais profundas (*deep fields*) do terceiro ano de observações do DES, que foi processado pelo portal científico do LIneA e está sendo preparado para detecção de aglomerados. Além disso, os dados finais do sexto ano de observação do DES foram preparados para a detecção de cluster e, a partir da comparação de diversas detecções baseados nas variações de *redshift* fotométricos e classificação de objetos, um catálogo de aglomerados do WaZP otimizado foi produzido. Este catálogo se estende até *redshifts* mais altos ($z > 1$) e apresentou grande completeza em relação a catálogos de aglomerados massivos detectados pelo efeito Sunyaev Zel'dovich como o ACT e SPT, e cujos resultados estão num artigo em preparação. No contexto do LSST, a simulação DC2 foi integrada ao portal científico do LIneA, possibilitando a primeira detecção de aglomerados nessa simulação sem a suposição da sequência vermelha de galáxias. Este trabalho possibilitará uma análise aprofundada da performance do WaZP assim como da própria simulação, além de promover maior integração do grupo brasileiro na colaboração internacional.

O cientista de dados Adriano Pieres também contribuiu com uma [palestra](#) para a série Astronomia ao Entardecer organizada pelo INCT do e-Universo/LIneA em parceria com o Planetário do Rio de Janeiro. Essa foi uma re-edição remota de uma série presencial realizada em 2019. Pelo LIneA/INCT participaram os pesquisadores Márcio Maia, Rogemar Riffel e Júlio Camargo na comissão organizadora, e suporte técnico de Rodrigo Bouffleur.

Colaboradores Externos

[SDSS](#)

No SDSS o grupo formado por pesquisadores da UFRGS e UFSM atuam no projeto MaNGA, onde tem participado de uma série de trabalhos (12 artigos) analisando diversas propriedades físicas e químicas de galáxias do MaNGA. Em especial dois destes artigos foram liderados por pesquisadores do grupo. Em um dos artigos (de Conto Machado et al., submetido) as propriedades cinemáticas da amostra de 170 galáxias que hospedam um núcleo ativo foram analisadas e as comparadas com uma amostra de controle de galáxias inativas, pareada em propriedades da galáxia hospedeira. Pela primeira vez a potência e extensão dos outflows de gás ionizado foram estimados. Em Riffel e colaboradores (aceito para publicação) a mesma amostra de 170 AGNs foi utilizada para determinar as taxas de formação estelar em AGNs, algo fundamental em modelos de evolução de galáxias, visto que nas modelagens as massas estelares (tamanhos) são controlados pelo feedback produzido pelos AGNs. Uma vez que o gás é ionizado por ambos os mecanismos (AGN e/ou formação estelar) é impossível medir essa taxa em AGNs usando as linhas de recombinação do H. Além disso, uma ferramenta de processamento dos dados do MaNGA chamada Megacube e em uma interface no portal científico do LIneA foram desenvolvidas para a análise destes dados.

Outro levantamento que afiliados do LIneA participam no projeto SDSS-IV é o [APOGEE-2](#) que observa o registro arqueológico em centenas de milhares de estrelas para explorar a história da formação e evolução da Via Láctea. Este é uma extensão do projeto APOGEE do SDSS-III utilizando dois espectrógrafos de alta resolução no infravermelho um em cada hemisfério para estudar as propriedades químicas e cinemáticas de centenas de milhares de estrelas na nossa galáxia. Neste projeto atuam pesquisadores, pós-doutorando e alunos da UFRGS e IAG-USP graças ao acordo feito pelo LIneA em troca de contribuições financeiras e de serviço, neste caso mantendo um [site alternativo](#) para distribuição dos dados da última liberação (DR16 de 2020) e gerenciando o uso e a operação das máquinas do projeto localizadas no Observatório de Las Campanas. As observações acabaram de terminar e a contribuição brasileira foi reconhecida em [carta](#) enviada pelo Diretor do projeto.

Para este projeto foram desenvolvidos métodos e técnicas, com base em teoria bayesiana para determinação de distâncias de estrelas do bojo Galáctico, tendo sido concluído o código STARHORSE (Queiroz et al. 2018, MNRAS, 475, 2556). Com base nos resultados de distância assim obtidos, combinados com resultados de metalicidades do projeto APOGEE, e movimentos próprios a partir de dados do satélite Gaia, os seguintes projetos estão sendo desenvolvidos:

- Análise das populações estelares do bojo Galáctico, que mostram uma composição complexa; estrelas do bojo confinadas na barra, parte delas com órbitas suportando a

barra, e outras não suportando a barra; componente com contra-rotação, possivelmente resultante de material (gás e estrelas) acretado pobre em metais; estrelas de disco fino e de disco espesso (Queiroz A. et al. 2020, A&A, 638, A76; Queiroz, A.. et al. 2020b, [A&A, submetido](#)).

- Seleção de 50 estrelas com todas as características de bojo genuíno, para análise espectroscópica, refinando os resultados disponíveis no APOGEE. O objetivo é obter padrão de abundâncias do bojo original, que deve ser a componente mais velha da Galáxia (Razera et al., em andamento).

TON

Este projeto conta com a participação de pesquisadores do ON. UFRJ, UNESP e UTFPR em estreita colaboração com pesquisadores do Observatoire de Paris e do Instituto de Astrofísica de Andalucía.

. Entre as várias atividades realizadas em 2020 se destacam:

- A submissão do *Full Proposal* para implantar a primeira fase do projeto Rede de Ocultações Sul-Americana (ROSA). O pedido foi submetido a pedido da fundação Sir John Templeton por meio da LSST Corporation, após uma primeira fase de seleção. O financiamento solicitado estava orçado em 500 mil dólares e contemplava a aquisição de 10 telescópios, o desenvolvimento de hardware e software para robotização das operações, o custo de pessoal e de operação para os três primeiros anos. No mês de junho de 2020 recebemos a notícia de que o financiamento não foi aprovado. O projeto ROSA continua em desenvolvimento e tem por objetivo a instalação de 50 telescópios robóticos distribuídos longitudinalmente no Brasil para observar principalmente ocultações estelares em parceria com escolas secundárias.
- Trabalho de manutenção do pacote PRAIA usado como ferramenta astrométrica e fotométrica no tratamento de imagens do DES.
- Desenvolvimento do Solar System Object Portal ([SSO Portal](#)) criado para analisar dados de grandes levantamentos como os do DES. A plataforma tem como objetivo identificar os CCDs contendo objetos do sistema solar conhecidos, e usar a posição medida e a época das observações, em conjunto com informações de outras fontes (MPC, JPL), para calcular órbitas para esses objetos e comparando com posições precisas de estrelas medidas pelo satélite GAIA prever eventos de ocultação. Esta plataforma foi desenvolvida ao longo de mais de dois anos por membros do TON em colaboração com o time de TI do LIneA. Os resultados obtidos com uma primeira versão da plataforma foram apresentados no início de 2020 numa [tese de doutorado](#) de Martin Banda. Baseada nesta experiência uma nova versão foi desenvolvida e apresentada a uma comissão de avaliação no final de 2020. O parecer desta comissão está disponível neste [link](#).

- A fotometria multi filtro do DES obtida para os objetos do sistema solar identificados também são usados para determinar para esses objetos suas propriedades superficiais, magnitudes absolutas, períodos de rotação e classificação taxonômica (esta última, até o momento, para os objetos transnetunianos).
- O desenvolvimento do módulo python SORA - *Stellar Occultation Reduction and Analysis*. O módulo fornece funções otimizadas para analisar a fotometria obtida de ocultações estelares, seja utilizando scripts python seja utilizando a interface Jupyter.
- Treinamentos sobre o uso do SORA, em âmbito nacional e internacional, foram realizados após o primeiro release no primeiro semestre usando a infraestrutura disponibilizada pelo LIneA.
- Avanços importantes no estudo de séries temporais de pequenos corpos do Sistema Solar observados pelo Dark Energy Survey, de maneira mais significativa a partir da disponibilização das medidas single-epoch do Y6. Este estudo teve o apoio de um estudante de iniciação científica (Daniel Ferreira).
- Um dos pesquisadores deste grupo, J. Camargo, também coordena o grupo de trabalho “*Transient and Moving Objects* (TMO-WG) do projeto DES.
- Um dos membros do grupo (Felipe Ribas) é responsável por coordenar campanhas internacionais para observação de ocultações estelares, organizando e distribuindo as informações e dados coletados, além de orientar os alunos e contribuir com os colegas nas respectivas reduções e análises. Um [banco de dados](#) é mantido com os resultados de ocultações estelares por objetos do sistema solar exterior.
- Em 2020 um total de 112 noites foram alocadas ao grupo TON nos vários telescópios do Observatório do Pico dos Dias e 10 noites no SOAR. Do total de 122 noites, 72 noites foram aproveitadas. No ano 49 Ocultações por TNOs, Centauros, Troianos e satélites de planetas gigantes foram reportadas. Destas, 46 ocultações foram coordenadas pelo projeto [ERC LuckyStar](#), em colaboração com o TON. Estas observações obtiveram 174 cordas positivas o que permitiu deduzir: i) um tamanho mínimo para 25 objetos; ii) um tamanho médio e uma posição para 10 objetos; e iii) o tamanho, forma (elipse) e posição para 14 objetos.
- Destaque especial para o evento de ocultação do TNO 2002MS4, ocorrido no dia 08/08/2020 quando foram obtidas 61 cordas positivas e cuja análise está sendo liderada pela aluna de doutorado do INCT Flávia Rommel.

É importante enfatizar que a experiência sendo obtida com os dados do DES e o SSO portal podem ser levadas imediatamente para o LSST. O crítico será poder mobilizar equipes interessadas em participar do grande número de eventos de ocultação que ocorrerão na era do LSST. O projeto ROSA certamente seria a melhor solução, bem como o engajamento e treinamento de astrônomos amadores brasileiros.

Pesquisadores desse grupo mantêm ainda outras colaborações, entre elas:

- O pesquisador Atair Gomes também contribuiu para a Semana Nacional de Ciência e Tecnologia participando da série Astronomia ao Entardecer com sua [palestra](#) sobre o sistema solar.
- Rodrigo Bouffleur implementou a técnica *Singular Spectrum Analysis* para o estudo de variabilidade observada no limbo solar, determinou o período de rotação do satélite Elara, e deu continuidade ao estudo de binárias eclipsantes com medidas do CoRoT (Convection, Rotation and Planetary Transits) e Kepler.

DESI

A participação brasileira neste importante projeto é pequena com 2 pesquisadores seniores e 4 posições entre alunos e pós-doutorandos. E foi negociada pelo LIneA em troca pelo desenvolvimento de um protótipo para uma plataforma para avaliar a qualidade dos espectros sendo acumulados. No momento participam deste projeto pesquisadores do LIneA, UNESP e UNICAMP sendo que estes últimos são os mais ativos. É bom lembrar que o instrumento ainda está em fase de comissionamento e verificação científica.

Do grupo da UNICAMP participam um aluno de doutorado (Anderson Souza) e sua orientadora (Flávia Sobreira) que vem atuando ativamente na análise de dados do primeiro ano de observação (DESI Year-1). Neste projeto, medidas das oscilações acústicas de bárions (BAO) e das distorções no espaço *redshift* (RSD) de todos os alvos astronômicos (diversos tipos galáxias e quasares) serão combinadas a fim de criar um dos vínculos mais poderosos ao modelo cosmológico padrão, além de explorar algumas questões importantes que ainda estão em aberto, como o valor absoluto das massas dos neutrinos, a natureza da energia escura e teorias de gravidade modificada.

A tese de doutorado de A. Souza investiga tópicos de Física de Neutrinos através da Cosmologia Observacional. Responder a perguntas abertas sobre neutrinos (o valor absoluto das massas, o problema da hierarquia, espécies estéreis, interações não-padrão) requer o uso de levantamentos amplos e profundos de agrupamentos de galáxias, como os do DESI, além de conjuntos de dados externos complementares: com lentes gravitacionais, supernovas e a radiação cósmica de fundo. O estudante está construindo uma pipeline robusta para o fluxo de trabalho do projeto, cujo objetivo é restringir os parâmetros e fazer análises cosmológicas com todos os dados DESI Year-1.

Souza participou da reunião global da colaboração - *DESI Collaboration Virtual Meeting* - que ocorreu na primeira semana de dezembro de 2020, onde o estudante apresentou o seu trabalho intitulado *DESI-Y1 Forecasts and Analysis*. O estudante também participa do grupo de trabalho EPO (*Education & Public Outreach*), cujo objetivo é ampliar o alcance da divulgação científica ao público geral. Neste caso, A. Souza ajudou na tradução para a língua portuguesa de algumas páginas do site principal da colaboração. As páginas inéditas em português devem ser publicadas ainda no primeiro semestre de 2021.

[DES & LSST](#)

Nesta seção são reportados os trabalhos dos vários grupos trabalhando no projeto DES e/ou se preparando para participar do projeto LSST na área de cosmologia. Estes grupos incluem pesquisadores na USP, UNESP, UnB e UFES. Os dois primeiros engajados nos dois projetos e os dois últimos se preparando para o LSST. O grupo da USP e UNESP vêm atuando ativamente e entre seus principais resultados se destacam:

- Co-liderança no [trabalho](#) sobre modelagem e validação da matriz de covariância para a amostra de dados acumuladas até o terceiro ano de observações (DES-Y3). O pós-doutorando Felipe Oliveira é o responsável pelo desenvolvimento do código teórico de covariâncias a ser usado nas pipelines de análise cosmológicas. Em 2020 ele também trabalhou na refatoração do código numérico original de covariâncias gaussianas e na produção de testes unitários, visando as próximas etapas no desenvolvimento desse código.
- Trabalho de análise de oscilações acústicas bariônicas (BAO) em espaço harmônico no Y3 sendo liderado pelo pós-doutorando Hugo Camacho com a participação do estudante de mestrado Daniel Gomes do IF-USP.
- Geração de simulações lognormais de distribuição de galáxias e de seus lenteamentos fracos que serão usados por várias análises.
- Participação de Hugo Camacho, Marcos Lima e Rafael Gomes na análise de não-gaussianidade na distribuição de galáxias (fNL) e co-liderando a análise de distribuição de galáxias no espaço harmônico, com um trabalho baseado nas simulações e dados do Y1 já no processo de revisão interna da colaboração.
- Outro projeto desenvolvido é a aglomeração de galáxias no espaço harmônico em dados do primeiro ano do DES. Esta análise está sendo conduzida por Hugo Camacho, Felipe Oliveira, Antonino Troja e o estudante de mestrado Rafael Gomes fez contribuições ao trabalho, que está sob revisão interna do DES.
- Estudo de galaxy-galaxy lensing no espaço harmônico em dados do primeiro ano do DES, numa colaboração entre Marcos Lima, Felipe Oliveira e o estudante de mestrado Lucas Faga. O objetivo deste trabalho é eventualmente permitir uma análise completa de 3x2pt no espaço harmônico para dados do DES e também do LSST.
- O pós-doutorando Antonino Troja também desenvolveu código de Integração de Espectro de Potência (PSICo) que permite calcular o espectro de potência angular no espaço harmônico sem o uso da aproximação de Limber. Este código terá um desempenho melhor do que a aproximação de Limber nas maiores escalas do espectro de potência e está participando de um desafio promovido pelo grupo de trabalho LSST, comparando o desempenho e os resultados de diferentes códigos.

A experiência do grupo com o DES está agora servindo na preparação para o LSST, onde participam como membros do BPG como o apoio do INCT do e-Universo e do LIneA. No DESC-LSST o grupo está co-liderando o esforço de calcular a matriz de covariância desenvolvendo o framework conhecido como TJPCov e participam rotineiramente das telecons de LSS e TJP.

O grupo formado pelos pesquisadores Mariana Penna-Lima da UnB e Sandro Vitenti da Universidade Federal de Londrina, durante 2020 desenvolveram as seguintes atividades:

- Validação da biblioteca *Core Cosmology Library* (CCL) que é uma biblioteca desenvolvida por membros do DESC/LSST para realizar cálculos de diferentes quantidades cosmológicas (distâncias, espectro de potências, função de massa, etc). Essa biblioteca será utilizada para realizar as análises com os dados do LSST. Uma importante atividade é validar esta biblioteca comparando-a com outras existentes na literatura, definindo também a precisão com a qual que cada quantidade é calculada numericamente. Neste sentido uma série de notebooks (jupyter) estão sendo desenvolvidos para realizar estas comparações, como por exemplo, entre CCL e outros códigos como a *Numerical Cosmology Library*.
- Implementação dos algoritmos para calcular as quantidades de lenteamento gravitacional fraco que são escritas em função de um perfil de densidade de matéria.
- Desenvolvimento do código CLMM para estimar a massa de aglomerados de galáxias a partir de medidas de lenteamento gravitacional fraco. Este projeto está em fase final de desenvolvimento e o artigo correspondente será submetido até março de 2021. Espera-se usar este código com o novo catálogo de aglomerados identificados pelo algoritmo WAZP mencionado acima.
- Reestruturação do pipeline *FIRECROWN* sendo desenvolvido por Sandro Vitenti um dos pesquisadores do BPG e associado ao LIneA que foi selecionado pelo grupo de trabalho *Dark Energy Science Collaboration* (DESC) do LSST como um dos pipeline scientists responsáveis pelo desenvolvimento dos pipelines de análise.
- Apresentação do seminário "[Questões Cosmológicas Atuais](#)" na série de palestras Astronomia ao Entardecer patrocinada pelo LIneA/INCT do e-Universo e planetário do RJ como parte da Semana Nacional de Ciência e Tecnologia.

Pesquisadores e estudantes da UFES formam outro grupo que são membros do BPG e se preparam para o LSST. Tendo em vista que o LSST produzirá um catálogo de milhões de supernovas identificadas fotometricamente, o que torna essencial identificar corretamente o subconjunto de supernovas de tipo Ia que serão usadas como velas padrão para medir a expansão do universo. Com este objetivo o grupo liderado por Valerio Marra está desenvolvendo métodos de *Machine Learning* para esta importante tarefa. Tipicamente, as supernovas Ia são usadas para restringir o espaço de parâmetros dos modelos cosmológicos através da distância de luminosidade. O grupo está explorando a aglomeração de supernovas como um novo observável para estudar a formação de estruturas no universo (*Clustering of Standard Candles*). De fato, o LSST proporcionará uma oportunidade única para aplicar este método. Em particular, simulações numéricas calibradas ao futuro catálogo de supernovas do LSST estão sendo desenvolvidas. O objetivo é obter uma estimativa acurada do poder estatístico deste método. As simulações estão sendo executadas no supercomputador Santos Dumont pelo projeto 'EGGELINEA' do LIneA.

Uma das alunas de doutorado do grupo, Tássia Ferreira, está engajada no grupo de trabalho DESC do LSST participando de dois workshops, do comitê organizador do encontro virtual do DESC em Fevereiro de 2021, representando os jovens pesquisadores, e publicando um [artigo](#) submetido ao *Physical Review D*. Como parte deste trabalho está também desenvolvendo um método independente para a comparação de matrizes de covariância.

Além desses trabalhos, aqueles sendo desenvolvidos pelos cientistas de dados, pós-doutorandos e alunos do LIneA e do INCT nas áreas do sistema solar, via Láctea, astronomia extragaláctica e cosmologia mencionados devem também ser destacados tanto no contexto do projeto DES como em preparação para o LSST. Em particular, o grupo se prepara para participar do primeiro teste de distribuição de dados a ser feito pelo LSST em Junho de 2021. Nesta ocasião está prevista a liberação de dados de uma simulação numérica que ficará disponível para a colaboração.

Produção Científica

Em 2020 os pesquisadores do LIneA/INCT participaram de 73 artigos já publicados e com 31 artigos submetidos. A lista completa completa desses artigos pode ser vista [aqui](#). O grande número de artigos se deve em grande parte ao número de pesquisadores do DES-Brazil, considerados como *builders* (veja a lista de nomes do DES-Brazil) da colaboração. Estes são pesquisadores que a colaboração reconheceu sua contribuição com o equivalente a 24 meses de trabalho para a infraestrutura do projeto, em grande parte no desenvolvimento das plataformas mencionadas acima.

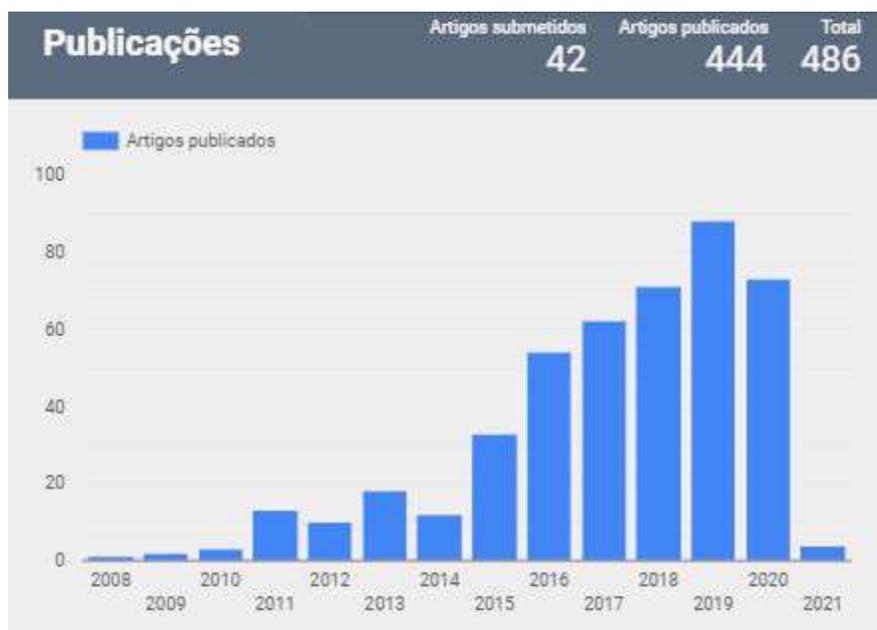


Figura 7 - Gráfico mostrando a evolução das publicações baseadas nos projetos apoiados pelo LIneA. A partir de 2017 também estão incluídas as publicações dos membros do INCT do e-Universo.

Em 2020 também foram completadas duas teses de doutorado no ON, uma sobre o sistema solar e outra sobre a luz difusa em aglomerados de galáxias, ambas baseadas em dados do DES e uma tese de mestrado teórica no IFT-UNESP de interesse para o DES e LSST. A figura 8 mostra a evolução do número de teses ao longo dos anos. Uma lista completa das teses de alunos associados ao LIneA e usando dados dos projetos apoiados pode ser vista [aqui](#). Ao todo são vinte e uma teses de doutorado e dezenove de mestrado.

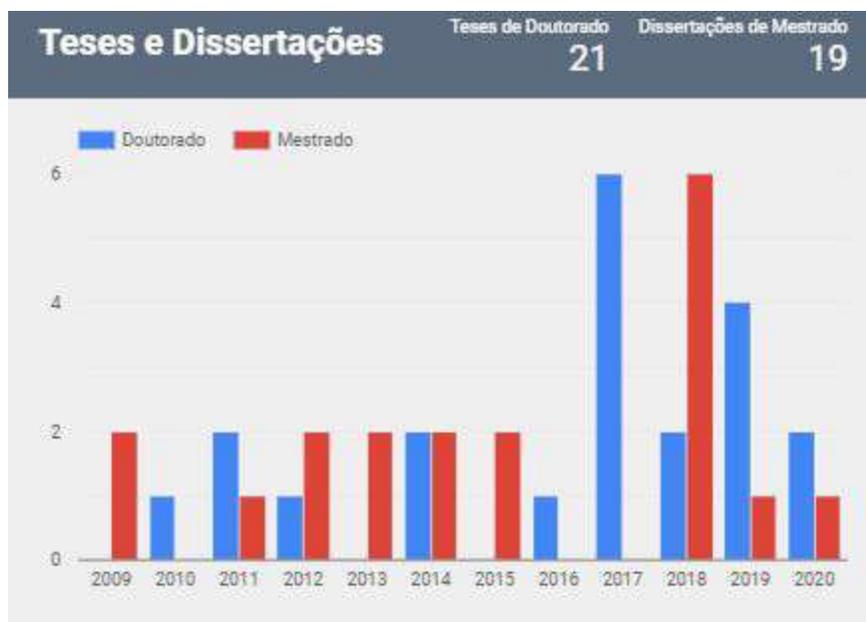


Figura 8 - Gráfico mostrando a evolução do número de teses de doutorado (azul) e mestrado (vermelho) baseadas nos projetos apoiados pelo LIneA.

Outras Atividades

Ao longo de 2020 foram realizados 41 webinars, um recorde histórico, cobrindo uma variedade de tópicos com a participação de membros de diferentes colaborações, indo desde jovens pesquisadores a importantes lideranças científicas. Para facilitar o acesso de todos além da participação pelo aplicativo *zoom*, liberada para toda a comunidade, foi implantado o *streaming* pelo youtube atingindo assim de 30 a 80 participantes por webinar. A série com mais de 300 webinars (figura 9) pode ser encontrada no [site](#) ou no [canal do youtube](#) do LIneA. Um relatório específico sobre o programa está sendo preparado com base numa pesquisa sendo feita junto à comunidade sobre a avaliação do programa.

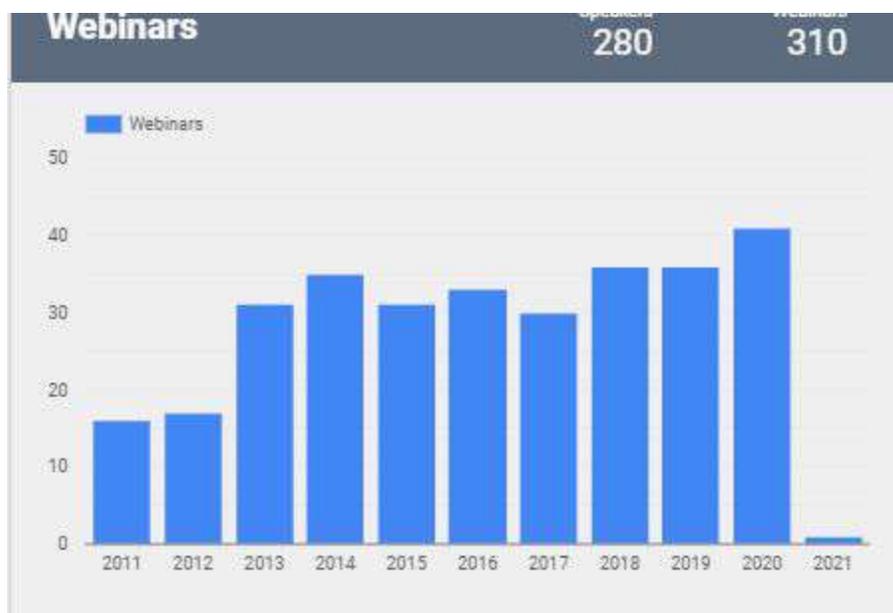


Figura 9 - Gráfico mostrando o número de webinars organizados pelo LInEA por ano. A grande maioria das palestras são de pesquisadores e tecnologistas internacionais procurando manter a comunidade brasileira atualizada das pesquisas e projetos em andamento no mundo.

Membros do LInEA participaram de um total de 69 eventos, incluindo presencialmente de reuniões internacionais no início do ano. Entre os principais eventos se destacam:

- Escola de Supercomputador Santos Dumont - Janeiro 2020 (Petrópolis, RJ, Brasil)
- DESC Winter - Janeiro 2020 (Tucson, USA)
- X Taller de Ciências Planetárias (Maldonado, Uruguai)
- Organização do Simpósio no. 359 da International Astronomical Union; Galaxy Evolution and Feedback across Different Environments - Março 2020 (Bento Gonçalves, RS, Brasil)
- DES Online Collaboration Meeting ([site](#)) - Maio 2020 (remoto)
- DESC Summer - Julho 2020 (remoto)
- Workshop: “From the Solar System to the Cosmos” (material educacional sendo preparado pelo LSST) - Julho 2020 (remoto)
- LSST Project & Community Workshop - Agosto 2020 (remoto)
- LSST - First Operations Bootcamp - Outubro 2020 (remoto)
- [ADASS](#) - Novembro 2020 (remoto)
- DESC Sprint Week - Dezembro 2020 (remoto)
- Community Engagement - Dezembro 2020 (remoto)
- [21a Workshop RNP](#) - Dezembro 2020 (remoto)

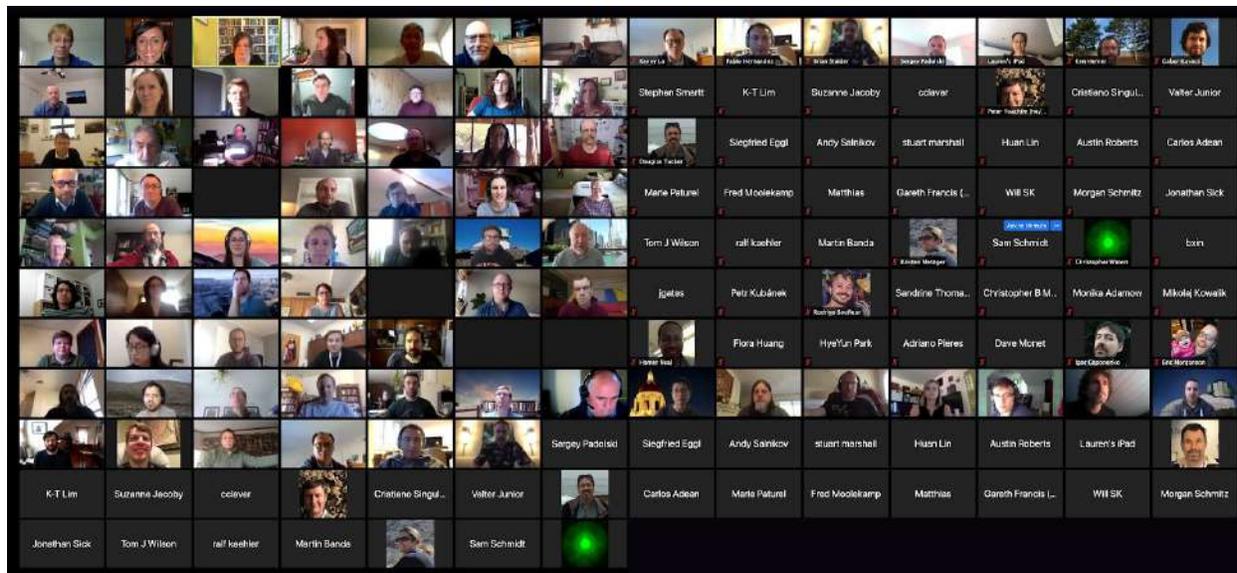


Figura 10 - Imagem de captura de tela dos participantes da reunião *First Operations Bootcamp*, que teve a participação de vários membros do INCT/LineA.

III.2 Educação

Para auxiliar na formação de jovens pesquisadores, o INCT/LineA promove semanalmente a reunião “Hotspot” cujo propósito é estimular a interação entre alunos e pós-docs de diferentes áreas, instituições e colaborações. O objetivo é de compartilhar informações obtidas nas inúmeras telecons mantidas pelas colaborações, apresentar trabalhos em andamento remotamente e desta forma se familiarizar com reuniões remotas, a principal forma de comunicação das colaborações internacionais.

O LineA também encoraja a participação de todos nas reuniões mensais *All-hands* realizadas na última sexta-feira do mês onde os principais acontecimentos e atividades são revistas, bem como as prioridades para o próximo mês são abordadas. O objetivo principal da reunião além de informar sobre o que acontece com o laboratório é visto como uma forma de engajar todos na vida do laboratório e divulgar a contribuição de seus membros.

A organização de um [curso de capacitação](#) para professores de física do ensino médio organizado em colaboração com ICTP-SAIFR para ser realizado no Colégio Andrews como parte do acordo de colaboração entre o INCT e o ICTP - suspenso pela pandemia.

Como mencionado anteriormente o INCT/LineA ofereceu ainda um tutorial de aproximadamente 3 horas sobre o uso da tecnologia Jupyter Lab para um grupo de tecnólogos da RNP e que está servindo de base para o curso sendo preparado para 2021.

https://docs.google.com/presentation/d/1-nzoV1Sr9T76fwEX405aV5ZL_U_iBdH7ZdX8K1n53dRj/presente?token=AC4wSVierzERDxRhC2CJV5w0f1vD0xFlng%3A15929500

Servidor JupyterHub

LineA

JupyterHub + Kubernetes

- » Cluster JupyterHub
- » K8S local ou na nuvem
- » Diferentes "flavors" de configuração

Diagram illustrating the JupyterHub + Kubernetes architecture. It shows a central LineA box connected to CManage and LineA vdp boxes. The LineA vdp box connects to a Kubernetes cluster of three servers, each running JupyterHub. The URL https://jupyter.linea.gov.br is shown below the servers, with arrows indicating access to the system.

22

Chrome File Edit View History Bookmarks People Tab Window Help

jupyter-tutorial/ notebook_3 notebook_2 jupyterhub64567987/ notebook_1 README.md

Not Secure | jupyter.linea.gov.br/user/julia/notebooks/jupyter-tutorial/notebook_3.ipynb?export

jupyter notebook_3 [unsaved changes] Logout Control Panel

File Edit View Insert Cell Kernel Widgets Help

Run Code

Net Trusted | science

4.2 Utilizando a biblioteca Seaborn

```
In [24]: 1 # Histograma simples não normalizado das redshifts (coluna 'zphot')
2 sns.distplot(df['zphot'], kde=False)
3 plt.show()
```

Out [24]:

In []: 1 # Histograma normalizado com suavização (KDE) das redshifts (coluna 'zphot')
2 sns.distplot(df['zphot'], kde=True)
3 plt.show()

Julia Gschwend

Figura 9 - Cenas do tutorial apresentado na RNP pelo analista Carlos Adean (na parte de cima) que discutiu sobre a instalação do sistema no LineA e a cientista de dados Julia Gschwend (na parte de baixo) sobre o uso de Jupyter notebooks.



CURSO Introdução à Astronomia Moderna

O Laboratório Interinstitucional de e-Astronomia e o Instituto Nacional de Ciência e Tecnologia do e-Universo convidam estudantes de ensino médio e graduandos na área de ciências exatas para participarem do Curso de Introdução à Astronomia. O curso será ministrado por cientistas de dados do LIneA e pesquisadores do INCT do e-Universo e serão formadas duas turmas para alunos de ensino médio e graduandos.

Cada turma terá no máximo 50 participantes.

Aprenda sobre a origem do sistema solar e os objetos que cercam a Terra, como estrelas nascem e morrem, entenda como a matéria está distribuída no Universo, o que é matéria e energia escura e sobre a origem e evolução do Universo. Se familiarize com os grandes projetos internacionais que procuram entender estas várias questões e as modernas técnicas e ferramentas sendo usadas para analisar grandes volumes de dados.

Inscrições estão abertas para estudantes de todo o Brasil até o dia 15 de Julho de 2021.

Para mais informações visite o site do curso e confira mais detalhes sobre o programa e como registrar.



No site será exibido o número de estudantes registrados.
linea.gov.br/cursos_online





Descrição

O Laboratório Interinstitucional de e-Astronomia (LIneA) oferece um curso remoto de introdução a Astronomia Moderna. O curso cobrirá diversas áreas da Astronomia sendo desenvolvida pelos pesquisadores atuando em grande projetos internacionais e envolvendo apresentações e atividades práticas.

O curso cobrirá os seguintes temas:

1. Introdução a linguagem Python
2. Introdução a plataforma Jupyter notebook
3. Sistema Solar (Julio Camargo, Rodrigo Boefleur)
3. Via Láctea (Basilio Santiago, Adriano Pieres)
5. Evolucao de Galaxias (Marcio Maia, Julia Gschwend)
6. Aglomerados de Galáxias (Marcos Lima, Michel Aguená)
7. Cosmologia (Rogerio Rosenfeld, Sandro Vitenti)
8. Projetos Internacionais

Isso será feito em sessões remotas de 3 horas de 15 em 15 dias que consistirão de uma apresentação de 30 minutos por parte de um pesquisador sênior seguido por um trabalho em grupo de 2:30 min liderados por utilizando como ferramenta a tecnologia de Jupyter notebook. Para que isso seja possível as duas primeiras sessões serão de introdução a linguagem Python e a plataforma Jupyter notebook.

O programa será feito em dois níveis – um para estudantes de ensino médio para despertar a curiosidade científica de alunos do ensino médio e outro para graduandos na áreas de ciência exatas em particular Física e Astronomia procurando introduzi-los a ferramentas modernas de pesquisa.

Esta é uma excelente oportunidade para despertar interesse na área e recrutar talentos para participarem dos grandes projetos internacionais em andamento.



Acesse
o Formulário



INCT
e-Universo



LIneA

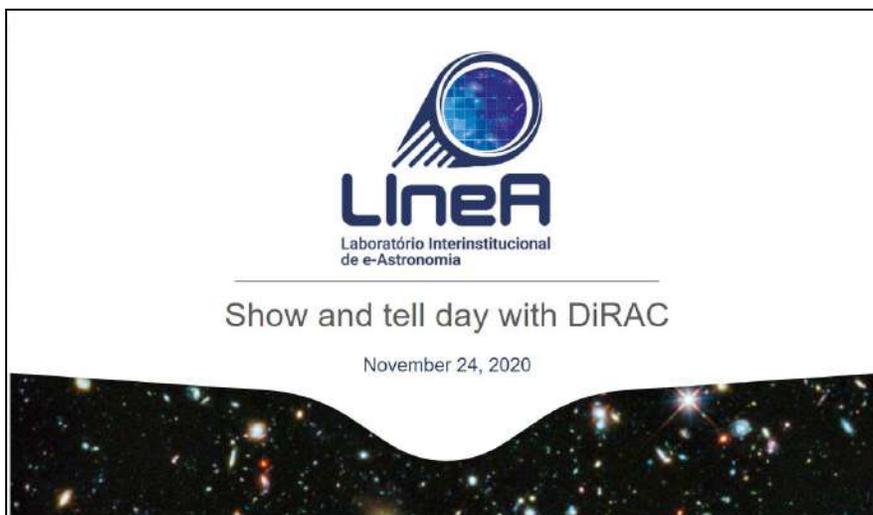
Figura 11 - Material para um dos cursos que será oferecido ao público para incentivar o engajamento da comunidade nos serviços do LIneA.

O LIneA também investe na formação de tecnólogos usando diferentes mecanismos como:

- Os webinars com a participação de convidados como Fábio Porto do LNCC, Kyle Chard da Universidade de Chicago, Luis Macedo da VMware Tanzu, Petar Zacevic da Universidade de Zagreb e Carolina Odman-Govender da Universidade de Western Cape e Institute for Data Intensive Astronomy.
- Apoio do INCT do e-Universo para a participação de tecnólogos em escolas e reuniões internacionais como ADASS e LSST Operations Bootcamp.
- A série informal “Conversando sobre TI” de debates com especialistas da ciência da computação como Antonio Tadeu, Carla Ostrowski e Roberto Souto do LNCC, Eduardo Ogasawara e Diego de Carvalho do CEFET, Daniel Oliveira da UFF, e Álvaro Coutinho e Marta Mattoso da COPEE/UFRJ.



- A reunião *Show and Tell* do time do LIneA com membros da organização *Data Intensive Research in Astronomy and Cosmology* (DIRAC) da Universidade de Washington procurando identificar áreas de cooperação.



- Organização da *Workshop ON the Future of Data Centers and eScience Institutes: Celebrating LineA's 10th Anniversary* contando com a participação de Andreas Wicener (University of Western Australia) Fabio Hernandez (IN2P3), George Beckett (University of Edinburgh), Mario Yuric (University of Washington) e William O'Mullane (LSST) no comitê organizador internacional. A reunião está prevista para Abril e o objetivo é discutir as novas tecnologias sendo consideradas para atender a crescente demanda de armazenamento e processamento.

III.3 Divulgação

Para marcar a fundação da associação LineA, uma firma foi contratada para desenvolver uma nova identidade visual incluindo um novo logo, um encarte impresso para divulgar as atividades do LineA (figura 12), material de escritório, assinaturas eletrônicas e slides. Encontra-se também em desenvolvimento um novo site para o LineA que incorpore duas importantes características - seja atraente e informativo para o grande público e seja um ponto de acesso à informação científica de interesse de profissionais da área e acesso aos serviços do LineA.



LineA: Um novo modelo institucional
na era de Big Science e Big Data



LineA (Laboratório Interinstitucional de e-Astronomia) é um laboratório científico que pesquisa o universo, apoiando e disseminando conhecimentos astronômicos para superar os desafios na era da Big Science e Big Data, promovendo a qualidade e o acesso a dados espaciais na comunidade científica internacional e gerando impactos econômico, social e intelectual que ultrapassam fronteiras.

O laboratório apoia o trabalho técnico/científico do INCT do e-Universo (Institutos Nacionais de Ciência e Tecnologia), cuja missão é contribuir na participação de pesquisadores brasileiros em grandes projetos mundiais, e vai além. Em 10 anos de atuação, tornou-se referência global na Astronomia, e a qualidade de seus sistemas foram reconhecidas por avaliações periódicas conduzidas por comitês internacionais.

Conheça a situação do laboratório e sua abrangente capacidade de promover benefícios à comunidade e à sociedade.

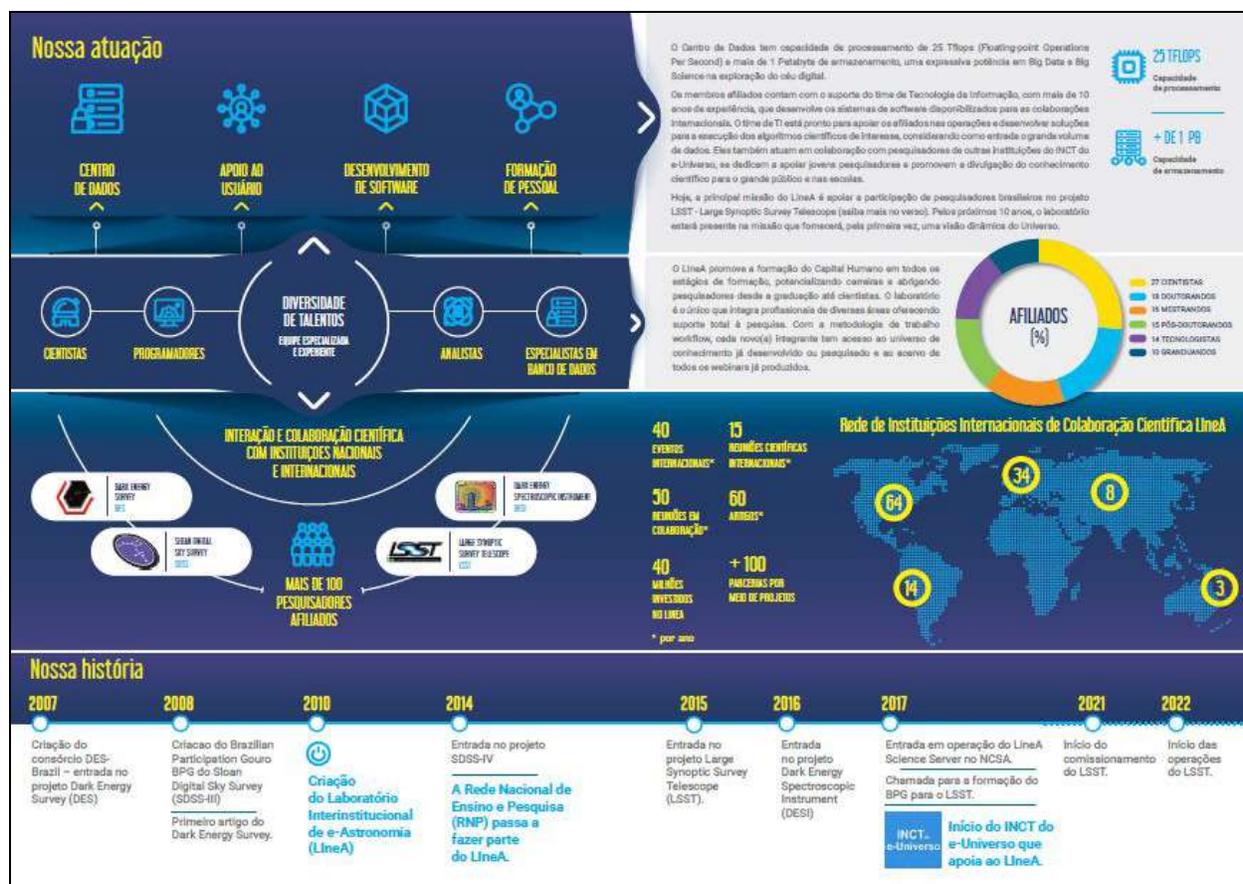


Figura 12 - Encarte produzido pelo LIneA para divulgar o trabalho sendo realizado pelo laboratório.

Durante o ano o uso das redes sociais para a divulgação das atividades do LIneA, anúncio de eventos e outros assuntos de interesse geral foi reestruturado dando mais ênfase à tweets e facebook para comunicações rápidas deixando os blogs para comunicados de imprensa ou assuntos que requerem maior profundidade. Em 2020 foram da ordem de mais de 170 tweets e 18 notícias na forma de blogs, resumidos num número mais reduzido de edições do LIneA News (figura 13) também de cara nova, enviado para mais de 1000 recipientes e acessível pelo site.



Figura 13 - Exemplo do resumo das principais notícias e atividades do LineA que são divulgadas para a comunidade periodicamente.

Outro esforço de divulgação foi a publicação do livro [***The Dark Energy Survey: The Story of a Cosmological Experiment.***](#) (figura 14) sobre a experiência adquirida na realização do projeto DES. O livro mostra o esforço envolvido em conduzir este tipo de experimento e tem também um valor educativo demonstrando a todos que grandes avanços na ciência não se fazem apenas com grandes idéias mas exigem a formação de equipes com multidisciplinares trabalhando em conjunto e uma organização para que mais de 500 cientistas e tecnólogos consigam trabalhar harmoniosamente por mais de uma década para um bem comum.

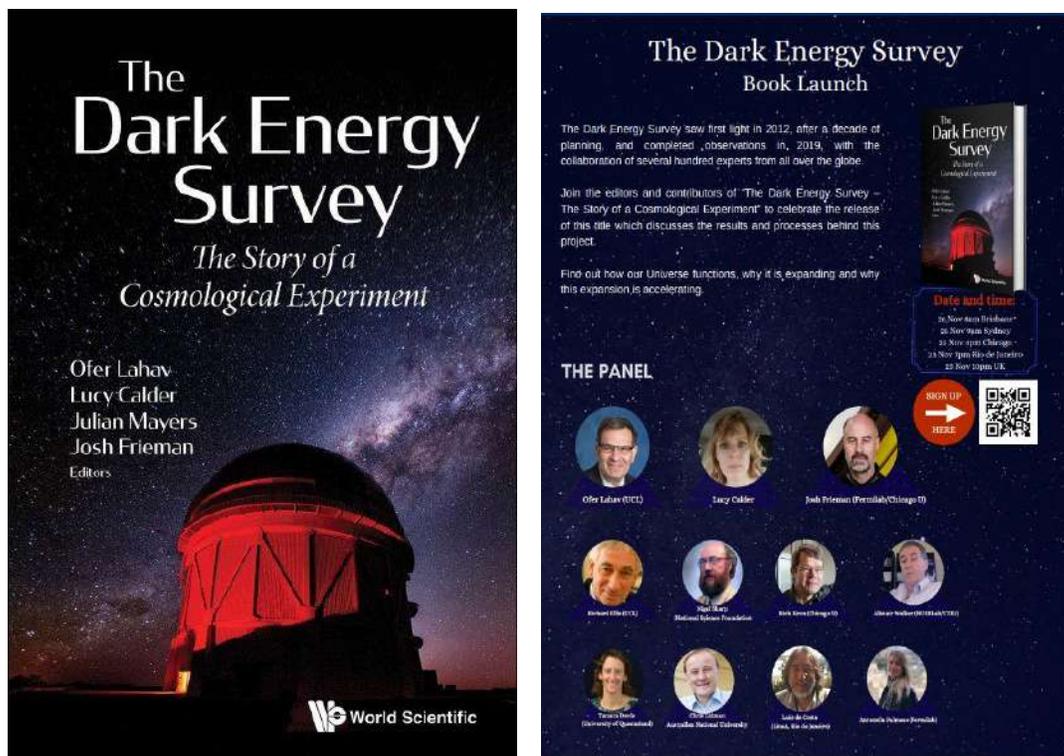


Figura 14 - A figura mostra à esquerda a capa do livro publicado descrevendo a experiência na realização desse experimento em andamento a quase duas décadas e à direita o painel organizado pela editora para o lançamento remoto do livro a partir de Londres com a participação do diretor do LIneA.

Como acontece todo o ano, o INCT/LIneA promove atividades na Semana Nacional de Ciência e Tecnologia. A exemplo do que foi feito em 2019 o INCT/LIneA promoveu em colaboração com o planetário do Rio de Janeiro uma série de palestras remotas divulgadas pelas redes sociais do INCT/LIneA e do Planetário. A série intitulada Astronomia ao Entardecer contou com a participação de palestrantes convidados do LIneA, INPE, UFSM,, UNESP, UnB. Estas palestras foram gravadas e estão disponíveis no [canal do LIneA](#) no youtube. A figura 15 mostra o cartaz digital usado para a divulgação do evento.

Astronomia ao Entardecer

Horário: **19 horas**
Como acessar: **Via YouTube**

A Fundação Planetário do Rio de Janeiro, o Instituto Nacional de Ciência e Tecnologia do e-Universo (INCT do e-Universo) e o Laboratório Interinstitucional de e-Astronomia (LIneA) realizarão durante a Semana Nacional de Ciência e Tecnologia de 2020 (19-23 de outubro), um ciclo de palestras virtuais sobre atualidades astronômicas intitulado Astronomia ao Entardecer.

Os temas estão relacionados com os estudos efetuados em diversos programas observacionais internacionais que são apoiados pelo INCT do e-Universo. As palestras duram de 30-40 minutos e após as mesmas, haverá possibilidade de colocação de perguntas.

Programa

<p>Dia 19 (Segunda-feira)</p> <p>Jaderson Schimola (IFSM) Buracos Negros Supermassivos - os raios cósmicos</p>	<p>Dia 20 (Terça-feira)</p> <p>Altair Gomes (UNESP) Mapeando os confins do sistema solar</p>	<p>Dia 21 (Quarta-feira)</p> <p>Alex Wuensche (INPE) Que destino nos reservam os astros?</p>
<p>Dia 22 (Quinta-feira)</p> <p>Adriano Pieres (LIneA) O nosso universo-ilha - a Via Láctea</p>	<p>Dia 23 (Sexta-feira)</p> <p>Mariana Penna Lima (UNB) Questões cosmológicas atuais</p>	

Planetário 50 anos | Prefeitura Rio de Janeiro | INCT e-Universo | LIneA

Figura 15 - Chamada para a série de palestras para a Semana Nacional de Ciência e Tecnologia patrocinada pela Planetário do Rio de Janeiro, INCT do e-Universo e o LIneA.

Nos dois últimos meses do ano foi possível contratar uma especialista em educação e divulgação que já teve uma grande impacto organizando as reuniões de avaliação para o acompanhamento dos grupos interessados nos produtos sendo desenvolvidos pelos time de TI, ou para a entrega dos mesmos, organizando a frente de comunicações e dessa forma liberando outros membros do time, realizando pesquisas junto a comunidade sobre o interesse e opiniões sobre diferentes iniciativas e atividades do LIneA, retomando a discussão sobre o novo site paralisada ao longo do ano e na organização de cursos de treinamento previstos para 2021.

III.4 Serviços

Apoio aos Usuários

O apoio direto ao usuários se dá via um sistema de helpdesk com mais de 700 tickets no ano de 2020 que inclui pedidos para:

- Abertura de contas no LIneA.
- Transferência de dados e ingestão no banco de dados.
- Disponibilização de datasets no sistema HDFS e Lustre.
- Instalação de bibliotecas e pacotes de software.
- Instalação de imagens para o Jupyter notebooks.
- Abertura de contas e apoio no uso do supercomputador Santos Dumont.
- Instalação de bibliotecas sob demanda no Santos Dumont.
- Atendimento de incidentes na infraestrutura computacional.
- Deployment de novas aplicações.

Recentemente, uma pesquisa foi feita com todos os membros do LIneA e apesar de um número reduzido de respostas a nota média dada ao serviço na escala de 1-10 foi de 8.5, excelente considerando o tamanho do time.

Operação

A operação foi uma das áreas com a perda de dois profissionais e a firma responsável pela manutenção dos mais de 100 equipamentos que compõem a infraestrutura computacional do LIneA, deixando o time com apenas dois profissionais. As atividades incluem monitoramento de dois banco de dados (MS SQL, PostgreSQL), compras, monitoramento dos equipamentos no LNCC que inclui três clusters de processamento, uma cluster de virtualização, sistemas de armazenamentos e máquina de transferência de dados, e rede.

Finalmente, como parte do compromisso com a colaboração SDSS, da ordem 20 TB correspondendo aos dados do DR16 foram transferidos e disponibilizados no [Skyserver/Casjobs](#) e o time do LIneA continua dando suporte ao sistema de computação associado ao espectrógrafo APOGEE-2 montado no telescópio Dupont do Observatório de Las Campanas, no Chile, com mais de 20 intervenções ao longo do ano consumindo da ordem de 40 horas. Além disso, um dos analistas contribui com apoio administrativo quando necessário ajudando na elaboração de relatórios técnicos e financeiros.

Desenvolvimento de software

O trabalho nesta área foi o mais intenso e se não foi possível expandir o time como imaginado originalmente, com pelo menos mais dois programadores, a equipe manteve seu pequeno mas excelente e experiente quadro. O desenvolvimento é feito usando a metodologia Agile em sprints normalmente de 10 dias úteis. Durante 2019 reuniões de *stand-up* eram diárias como recomendado pela metodologia, mas com a saída da *Scrum Master*, os *stand-ups* são

atualmente feitos às 4as e às 6as. Em 2020 foram realizados da ordem de 24 sprints, procurando avançar em paralelo nos projetos:

- MaNGa portal
- SSO portal
- LIneA Science server
- DES portal
- LSST Portal

Todos esses projetos são complexos e pelo menos quatro deles de longa duração, como discutido na seção II. Em particular, o desenvolvimento do LSST portal exige um longo período de P&D dado o grande volume de dados a ser processado.

III.5 Infraestrutura

Os mesmos profissionais responsáveis pela operação também estão envolvidos na manutenção e desenvolvimento da infraestrutura crítica para as atuais prioridades do LIneA. Um dos projetos é implantar um sistema de autenticação de usuários integrado à Comunidade Acadêmica Federada (CAFe), uma federação de identidade que reúne instituições de ensino e pesquisa brasileiras. Através da CAFe, um usuário mantém todas as suas informações na instituição de origem e pode acessar serviços oferecidos pelas instituições que participam da federação. Este trabalho foi realizado em colaboração com um grupo de trabalho da RNP (GidLab) ao longo de 2020 e a associação faz agora parte desta federação desde Julho de 2020.

A segunda parte deste esforço é a instalação de um sistema para gerenciar a autorização de usuários para acessar diferentes serviços com recursos (número de núcleos, espaço em disco, memória) atribuídos dependendo de credenciais pré-estabelecidas. Para isto, o sistema *Comanage* está sendo investigado em colaboração com a equipe do GidLab visando sua implantação no LIneA. A proposta é ter diferentes categorias de usuários:

- Platina - pesquisadores e estudantes vinculados aos projetos de colaboração científicas apoiados pelo LIneA, sem qualquer restrição.
- Ouro - pesquisadores da comunidade com demandas de processamento específicas que requerem os serviços do LIneA e para isto devem submeter propostas
- Prata - pesquisadores e alunos com demandas moderadas de pesquisa que exige acesso aos repositórios de dados com subquente necessidades de processamento.
- Bronze - para estudantes, professores de ensino médio e o público em geral que podem usar a infraestrutura para explorar os dados de forma exploratória. A Figura 16 mostra o esquema de registro sendo imaginado para estar disponível no novo site do LIneA.

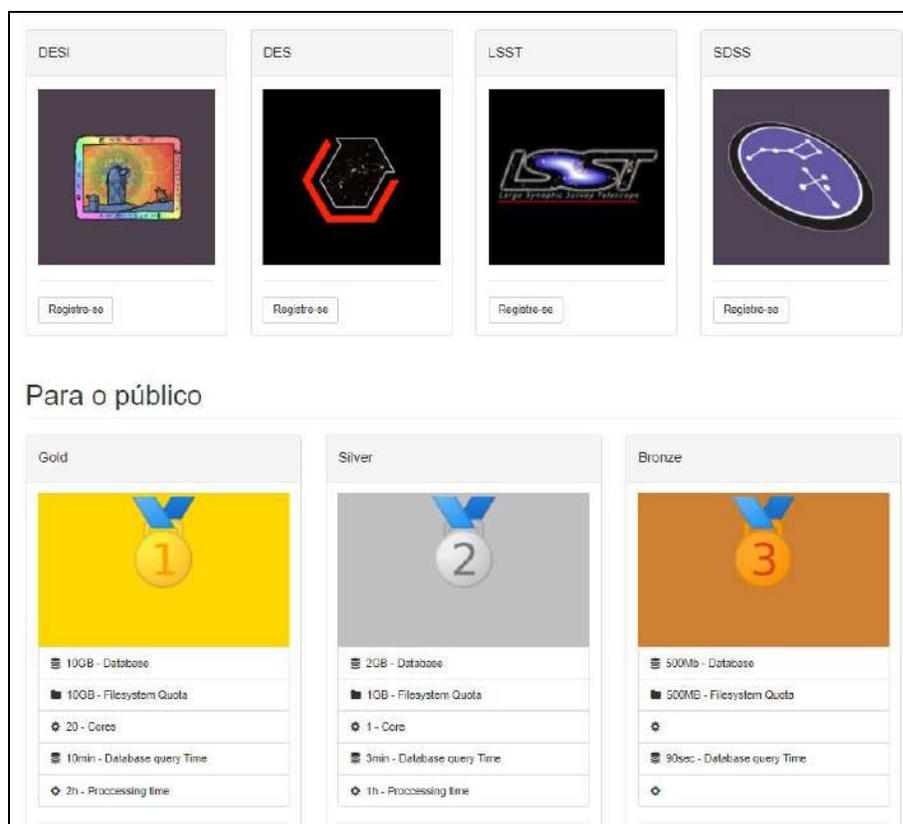


Figura 16 - Protótipo da interface de cadastramento de usuários prevendo diferentes categorias as quais serão atribuídos recursos de máquinas de acordo com suas necessidades e uma política de uso que garante a sustentabilidade da operação.

Os detalhes da proposta deverão ser definidos em discussão com o Comitê de Usuários do LIneA em consulta com a comunidade de potenciais usuários. No momento, testes estão sendo feitos para a integração deste sistema com o Ldap do LIneA. Finalmente, mudanças foram feitas a nível da federação CaFE para integrar o nosso sistema àquele sendo usado pelo LSST em preparação para o IDAC.

Outro importante projeto desenvolvido ao longo do ano foi a implantação de um *Kubernetes cluster* (K8S) sendo feito em colaboração com o NCSA e com a RNP. O objetivo é gerenciar os recursos de máquinas disponíveis para diferentes usuários de acordo com suas credenciais. É uma componente fundamental para a liberação do *LIneA Science Platform* e é um requisito para a implantação do IDAC e a disponibilização do *Rubin Science Platform* para acessar os dados do DES. Uma instalação de testes foi completada mas a entrada em produção mesmo que em fase de testes depende da disponibilização de máquinas que vai acontecer a partir da reorganização da infraestrutura do LIneA que teve início em Setembro com a transferência de conhecimento no encerramento do contrato com a firma responsável. Esta tarefa está quase completa mas foi interrompida com a chegada do novo sistema de armazenamento da DELL adquirido pelo LIneA, e por sérios problemas de suprimento de energia que ocorreram no

LNCC em Dezembro que impactou o processo de instalação sendo feito pela DELL mas com o envolvimento direto da equipe do LIneA. O repositório de dados em várias máquinas ainda está sendo avaliado.

O *Lustre Filesystem* sendo instalado permite alto desempenho computacional e foi projetado para ter vazão de 100 Gbps com armazenamento em discos SSD (70 TB) e HDD (500 TB), num total de 570 Terabytes de capacidade de armazenamento. A alta capacidade de armazenamento do *Lustre Filesystem* foi desenvolvida para atender a grande demanda de dados que serão produzidos pelo LSST. Além disso, o novo *Lustre* foi projetado para ser modular, e caso houver a necessidade de expandi-lo futuramente, a arquitetura do sistema permite que a equipe de TI do LIneA aumente sua capacidade de armazenamento e vazão (throughput) de forma mais simples. Este é um elemento crítico na proposta feita para hospedar o IDAC do LSST no Brasil. A instalação física já foi completada (figura 17), restando agora sua integração ao resto do sistema.



Figura 17 - Sistema de armazenamento Lustre de 570 TB recentemente adquirido pelo LIneA.

IV. Discussão

Os principais resultados do planejamento realizado no início de 2020 foram a elaboração do cronograma apresentado e da equipe necessária para cumprir o cronograma de maneira satisfatória a partir de 2021-2022, como mostram as figuras 18 e 19, respectivamente. A equipe foi concebida a partir de uma avaliação do esforço necessário por perfil profissional para cada projeto, levando em conta as datas críticas estabelecidas: 1) pela linha de tempo dos projetos internacionais; 2) pela necessidade de atender as demandas internas dos pesquisadores brasileiros atuando nos projetos apoiados; e 3) para suprir o mais rapidamente possível as necessidades da comunidade como um todo. O ponto de partida foi o time de 2019 que contava com 16 FTEs, além de uma firma responsável pela manutenção de toda infraestrutura computacional do LIneA e do serviço do site do SDSS e pela gestão de usuários.

Originalmente, a proposta foi a de manter este time e expandir em 2021 para 21 FTES e, a partir de 2022 para 25 FTEs. com cada perfil contribuindo as frações de tempo listadas na tabela para as tarefas de administração, ciência, desenvolvimento de software e operação e serviços em geral, onde ciência corresponde ao trabalho fundamental da supervisão científica dos produtos e serviços criados e oferecidos pelo LIneA.

Para 2020 a intenção era aliviar a pressão nos times de desenvolvimento de software com vários projetos em andamento simultaneamente (*DES Science Portal*, *LIneA Science Server*, *MaNGA Portal*, *SSO Portal*, *LIneA Science Server*), de analistas, para que estes pudessem se concentrar no desenvolvimento da infraestrutura necessária para expandir os serviços do LIneA para a comunidade (Kubernetes cluster, sistema de Autenticação e Autorização, Jupyter Lab) e ter alguém trabalhando na área de Comunicação, Divulgação e Educação em particular para desenvolver material para treinar jovens pesquisadores no uso dos dados e ferramentas dos projetos SDSS e DES, e futuramente do LSST.

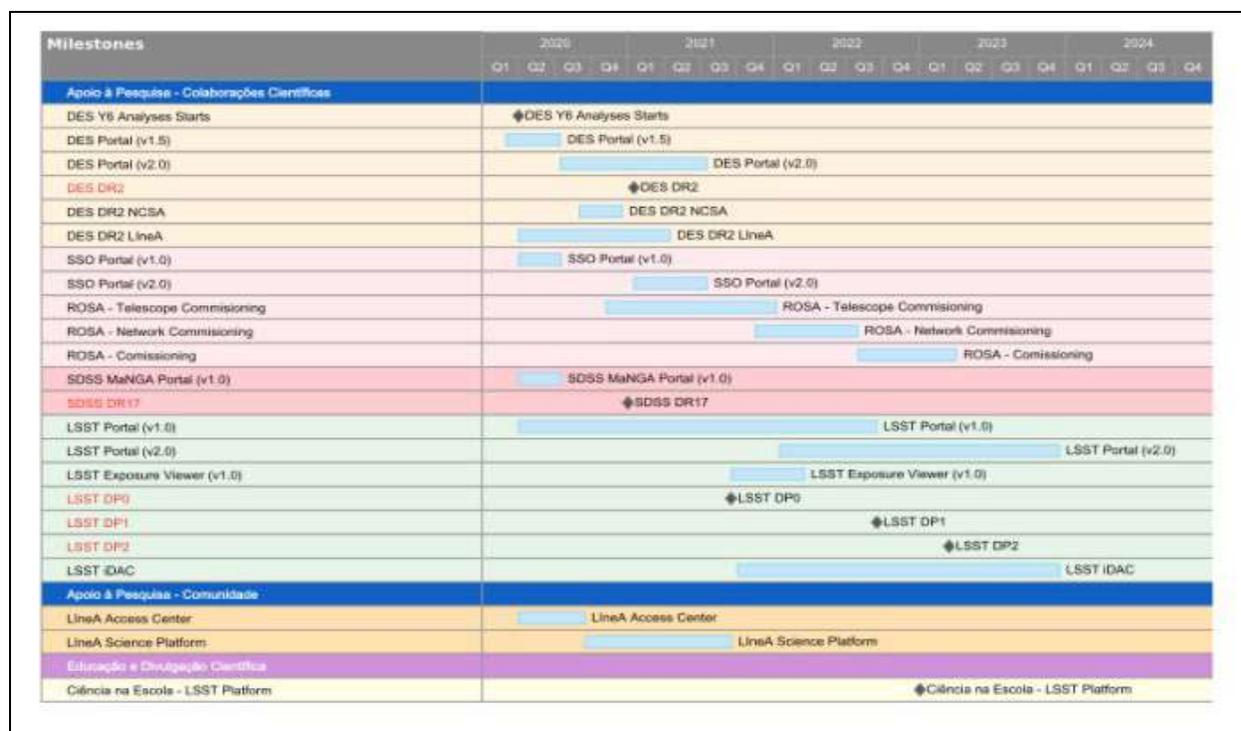


Figura 18 - Cronograma para o período de 2020-2024 preparado no início de 2020 para o plano diretor da associação LineA. A principal mudança foi o adiamento no início do projeto ROSA. Uma avaliação recente mostra um atraso em média de 6 meses causado pela abrupta redução do time em 2020.

Cargo	#	Administração	Ciencia	TI	Operacao & Serviços
Diretor-Geral	1	0.6	0.2	0.2	0
Coordenador Operacao & Serviços	1	0.25	0	0	0.75
Coordenador Administrativo	1	1	0	0	0
Coordenador de Projetos	1	0.25	0.25	0.5	0
Engenheiro de Software	1	0	0	1	0
Cientistas de Dados	4	0	0.25	0.75	0
Programador Cientifico senior	1	0	0	0.75	0.25
Programador Cientifico pleno	1	0	0	0.75	0.25
Arquiteto senior	2	0	0	0.75	0.25
Arquiteto pleno	2	0	0	0.75	0.25
Programadores de Front-end	4	0	0	1	0
Analista de sistemas - Especialista	1	0	0	0.75	0.25
Analista de sistemas - senior	2	0	0	0.5	0.5
Analista de sistemas - junior	2	0	0	0	1
Especialista Divulgação Científica	0.5	0.2	0	0	0.3
Assistente Administrativo	1	1	0	0	0
Total	25.5				

Figura 19 - Equipe considerada necessária para levar adiante o cronograma da figura 18. A figura mostra o número de profissionais para cada perfil e a fração de tempo dedicado a diferentes atividades.

Infelizmente, ao invés de expandir, o LIneA foi obrigado a reduzir a equipe devido às dificuldades financeiras, o que prejudicou o cumprimento do cronograma apresentado acima. Por exemplo, o desenvolvimento dos portais para os projetos SDSS (MaNGA Portal) e TON (SSO Portal) tiveram um atraso de 6 meses. Outros projetos profundamente prejudicados foram o *LIneA Access Center* e o *LIneA Science Platform*, apenas iniciados. O objetivo de ambos é permitir o acesso da comunidade aos dados e ferramentas do LIneA de maneira sustentável. No entanto, a área de atividade mais prejudicada foi a da infraestrutura com a perda de 2,5 FTEs além da firma responsável pela manutenção das máquinas, forçando os analistas a assumirem essas responsabilidades além da complexa instalação do sistema de armazenamento de dados Lustre ainda em andamento.

Outra área prejudicada foi o projeto *DES Portal 2.0* que envolve a revisão de workflows científicos já implementados no portal bem como a integração de novos que pudessem ser de interesse para analisar os dados finais do DES e para futuro uso na análise de dados do LSST já em 2022 usando dados do período de verificação científica com a diminuição do time e cientistas de dados de 2,5 para 1,25 e a perda da engenheira de software.

Por outro lado, apesar das inúmeras dificuldades encontradas durante o ano de 2020 as seguintes metas foram alcançadas:

- Cumprimento das obrigações internacionais do LIneA (apoio ao LCO, disponibilização do DR16 do SDSS, participação na liberação pública de dados do DES).
- A entrega de plataformas científicas para ajudar pesquisadores ligados a projetos apoiados (SSO Portal, Science Server v 2.0, MaNGA Portal).
- Importantes melhorias a infraestrutura computacional com a instalação do sistema Lustre, a reorganização das máquinas do LIneA, a preparação para a instalação do kubernetes cluster, jupyter lab e o sistema de autenticação/autorização necessário para liberar os serviços do LIneA para a comunidade.
- Apoio à análise utilizando o portal científico dos seis anos de dados do DES para diferentes grupos de pesquisa.
- Submissão e aprovação de proposta para quase dois milhões de CPU-horas no supercomputador Santos Dumont para os pesquisadores do LIneA.
- Submissão de proposta ao LSST para hospedar um centro de acesso a dados do LSST, para a criação de produtos científicos avançados em troca de novas vagas no projeto.
- Elaboração de acordo entre o LIneA, LNCC e RNP para a operação deste centro.
- A instalação do Rubin Science Platform que é um pré-requisito para o futuro Centro de Acesso a dados do LSST.
- Elaboração de documento de requisitos para o desenvolvimento do portal LSST a partir das lições aprendidas com o portal DES.

- Requisitos para a implantação do LIneA Science Platform, integrando os serviços do Science Portal, Science Server, Jupyter Lab e uma interface de acesso a produtos científicos avançados.
- Desenho preliminar do novo site do LIneA ainda em discussão, que será a base do *LIneA Access Center*.
- Importantes avanços na preparação de cursos de treinamento no uso dos serviços do LIneA planejados para 2021.

Todas essas realizações foram possíveis graças ao enorme esforço e dedicação da equipe do LIneA que multiplicou seus esforços para cobrir a lacuna deixada por aqueles que tiveram que ser dispensados em difíceis decisões tomadas ao longo do ano.

V. Conclusão

Apesar das grandes dificuldades enfrentadas pelo LIneA durante o ano de 2020 não deixa de surpreender o grande número de metas alcançadas e a intensidade do trabalho e de atividades realizadas pela pequena equipe em seu primeiro ano como associação. Sem dúvida, o apoio indispensável do INCT do e-Universo e a fundação da associação, e a motivação que isto trouxe, foram fundamentais para o sucesso alcançado.

O LIneA que celebrou este ano 10 anos de sua criação vem ao longo dos anos evoluindo em seus objetivos e como estabelecido por seu planejamento estratégico apresentado aos seus conselhos internos reconhece que sua missão deve ser não só dar suporte aos pesquisadores envolvidos nas colocação científicas apoiada pelo LIneA e o INCT do e-Universo, mas deve servir de facilitador para que a comunidade de astronomia e cosmologia brasileiras tenham os meios de explorar cientificamente o grande acervo de dados disponibilizados por esses levantamentos de diferentes maneiras, por exemplo: 1) usando grandes amostras estatísticas que requerem uma infraestrutura de HPC adequada; 2) estudos que combinam dados provenientes de diferentes instrumentos e comprimentos de onda; 3) mineração de dados em busca de fenômenos singulares; 4) análises de séries temporais que serão possíveis com os dados do LSST.

Com este objetivo o plano para 2021 o LIneA terá como prioridades:

1. A conclusão dos projetos *LIneA Access Center*, que envolve o novo site e o registro de usuários que irá permitir a atribuição de recurso de acordo com uma política de uso sustentável, construída com o apoio da comunidade.
2. Atividades para o engajamento da comunidade através de cursos de treinamento, documentação, palestras, mesas redondas e pesquisas de opiniões.

3. A implantação e liberação dos serviços que compõem o LIneA Science Platform para acesso aos dados e plataformas científicas para processamento usando Jupyter notebooks.
4. Participação no teste a ser feito pelo LSST na liberação de simulações.
5. O desenvolvimento do portal científico para o LSST dando continuidade ao escordo de P&D sedo feito no uso de novas técnicas para o processamento do grande volume de dados.
6. Caso aprovado a preparação de um memorando de entendimentos com o Projeto LSST, DOE e NSF para a implantação do IDAC brasileiro.
7. Iniciar a implantação do IDAC.

A velocidade com que estes projetos serão levados adiante dependerá do apoio que o LIneA terá ao longo do ano do INCT do e-Universo e do MCTI.

Data: 20/02/2021

Preparação

Luiz Nicolaci da Costa
Coordenador do INCT do e-Universo
Diretor-Geral do LIneA

Revisão: Márcio Maia (ON/LIneA)

Contribuíram para o documento:

Adriano Pieres (LIneA/INCT do e-Universo)
Altair Gomes (UNESP)
Anderson Souza (UNICAMP)
Antonino Troja (UNESP)
Beatriz Barbuy (IAG-USP)
Bruno Morgado (ON)
Carlos Adean (LIneA)
Cristiano Singulani (LIneA)
Felipe Oliveira (UNESP)
Felipe Ribas (UTFPR)
Fernando Silva (LIneA)
Glauber Villa-Verde (LIneA)
Hugo Camacho (UNESP)
Julia Gschwend (LIneA)
Julio Camargo (ON)
Marcelo Assafin (OV-UFRJ)

Marcos Lima (IF-USP)
Mariana Fernandes (LIneA)
Mariana Penna-Lima (UnB)
Michel Agüena (LAPP-CNRS)
Rodrigo Bouffleur (LIneA/INCT do e-Universo)
Rogério Riffel (UFRGS)
Rogério Rosenfeld (UNESP)
Sandro Vitenti (UEL)
Tassia Ferreira (UFES)
Thaisa Storchi-Bergmann (UFRGS)
Valerio Marra (UFES)
Valter Junior (LIneA/