ADASS XXXIV

ASTRONOMICAL DATA ANALYSIS SOFTWARE & SYSTEMS XXXIV

L-Università ta' Malta

10–14 November 2024 University of Malta, Valletta Campus, Malta um.edu.mt/events/adass2024

Invited Talks 8

How do you use yours? The Evolution of Proposal and Observing Preparation Tools. 8

From Daniel Dennett to Transformers: The Computational Evolution of Human Intelligence in AI 8

Declarative Data Management with DaCHS and the VO 9

Empowering Science with Good Design 9

The Challenges of Astronomical Data Systems 9

Beyond the Data: challenges and triumphs in data reduction and analysis 10

SDHDF: A new file format for spectral-domain radio astronomy data 10

The MeerKAT Science Data Processing Pipeline 11

Celebrating SAOImageDS9 11

Talks 12

Proposal and Observation Preparation Tools 12

How did we build ours? A modern proposal tool for a modern telescope 12

The proposal evaluation process: A unified user experience supporting different workflows. 12

STARS: A scheduling software for Space Missions and Ground-Based Observatories 13

Observation Scheduling Software Framework for Distributed Telescope Arrays in Time-Domain Surveys 14

The Rise of AI for Science and Data Center Operations 15

Machine Learning Enhancements for Real-Time Scientific Analysis of Cherenkov Telescope Data 15

AI Agents for Ground-Based Gamma Astronomy 15

Classification of HI Galaxy Profiles Using Unsupervised Learning and Convolutional Neural Networks: A Comparative Analysis and Methodological Cases of Studies 16

An AI-driven system for enhancing Astronomical Research workflows 17

Goal-Oriented Stacking: an novel approach to statistical image-domain inference below the noise threshold 18

Self-supervised learning of radio data for source detection, classification and peculiar object searches 19

Transforming Data into Insights: AI-Driven X-Ray Source Classification within the NADC Framework 19

Extending the Life of Software and Data 20

Synergies Unleashed: Bridging the Gap Between Science and Computing teams in the ALMA Observatory software deployments 20

The Chandra Data System at 25 years — What can it teach us? 21
DASCH: Bringing 100+ Years of Photographic Data into the 21st Century and Beyond 22
Curating a 20th century observation log in the 21st century 22
Longevity of a treasured database service 23
User Experience 23
Processing LISA's data as a human: The GlobalFit Framework user experience 23
User facing tutorials as code: reproducible and reliable tutorials with CI/CD 24
Accessible visualization of Astropy objects and Multi-Order Coverage Maps (MOCs) with the iPyAladin Jupyter widget 25
DARTS Timescape: Exploring 50 Years of Space Science Data Through Interactive Visualization 26
Integrating UX Design in Astronomical Software Development: A Case Study 27
Data Management and Trusted Repository in the Open Data Era 27
Securing Space Science: Advanced Data Protection in the HREDA Archive 27
Insights from a 30-Year international Partnership on Astronomical Archives 28
Roadblocks in Astronomical Data Analysis 29
Optimized Open-Source Tools for Scalable Solar System Science 29
Lowering in-memory footprint of antenna beams via polynomial approximation 29
A Reproducible Science Workflow System: DALiuGE in Action 30
Migrating Heterodyne Data Reduction to High-Performance Computing 30
The time-series visualization tool in ESASky 31
J-PAS early data release: unique processing challenges of an imaging sky survey in 57 optical filters 32
PyKOALA, a multi-instrument tool for reducing IFS data 33
Bridging the Gap: Enhancing Astronomical Data Analysis with Software Engineering Best Practices 33
Metadata, Semantics and Data Models Applicable to Data Formats 34
A data model to connect the ESO Data Processing System (EDPS) to ELT data archives 34
Using Felis to Represent the Semantics and Metadata of Astronomical Data Catalogs 35
Real-time and Near Real-time Processing Pipelines 35
NEOCC’s Aegis pipeline in asteroid orbit determination and impact monitoring. 35
Finding Fireballs in Lightning: A Daily Pipeline to Find Meteors in Weather Satellite Data 37
Sub-arcsecond degree-scale imaging pipelines with LOFAR 37
High-Performance Computing in Astronomy: Triumphs and Tribulations of Pipeline Processing on Supercomputers 38
High-Performance Pipeline Processing for the Australian Square Kilometre Array Pathfinder 38
Enhancing Keck Observatory Operations: The Data Services Initiative's Journey 39
Leveraging FPGAs as accelerators in real-time astronomical data-processing pipelines 40
Dynamic Imaging With MeerKAT: The Time Axis As The Final Frontier 40
Other 41

Astronomy Data and Computing Services: Changing the way research software is developed, supported and maintained 41

A Multi-Wavelength Data Viewer Realized through the Enhancement of hscMap 42

Data processing and preservation for CTAO 42

Asteroid Discovery with THOR on the Noirlab Source Catalog: An Engineering Perspective 43

FRELLED : An Astronomical Data Visualisation Package for Blender 43

Posters 45

Proposal and Observation Preparation Tools 45

ACROSS: Enabling Time Domain and Multimessenger Astrophysics 45

Simulation Tools to Help Prepare for Science with the Roman Space Telescope 45

Polaris: a new open source proposal tool 46

Scheduling the world’s largest observatory for gamma-ray astronomy 46

SPOT: Site Preparation and Observation Tool 47

The ScopeSim ecosystem 47

Continuation Requests in the 'Hedwig' Proposal System 48

ESO’s New Generation of Exposure Time Calculators 48

CPGS: a tool for preparing observations with the Coronagraph Instrument on the Roman Space Telescope 48

Usage and Management of TRDS files for JWST and HST 49

Delphi-Crew: An innovative methodology for proposals evaluation 49

Pandeia Exposure Time Calculator: A Cloud based Multi-mission Wizard to L2 and Beyond! 50

The Roman Telescope Proposal System 50

The Proposal Data Model as the basis for POLARIS a Proposal Submission Toolkit 51

Open Time Proposal Review System for the MeerKAT Radio Telescope 51

Managing target of opportunity (ToO) observations at Observatorio Astrofísico de Javalambre (OAJ) 52

Optimizing the Scheduling of the Cerro Chajnantor Atacama Telescope (CCAT) Surveys 52

The Rise of AI for Science and Data Center Operations 53

Machine-Learning Workflow for Morphology Classification of Galaxies 53

Machine Learning Mismatchings and Catalogues Creation: A Path to Finding the Milky Way Galaxies-Analogues 54

Open Science and Artificial Intelligence for supporting the sustainability of the SRC Network: The espSRC case. 54

Large Language Models: New Opportunities for Access to Science? 55

Maximising scientific return with NLP & ESA Datalabs at the Data Science and Archives division at ESAC55

ML Methods For Space Debris Detection in Bistatic Radar Data 56

Overcoming machine learning training data imbalance by simulating exoplanet transits 56

Leveraging Deep Learning for Efficient Galactic Characterization 57

From data to scientific breakthroughs using tools powered by Generative Deep Learning 58

To clean or not to clean? Influence of pixel removal on event reconstruction using deep learning in CTAO 58
Vision-Language Models for Spiral Galaxy Identification in SDSS: A Path to Finding Milky Way Analog Galaxies 59
Enhanced Exoplanet Detection Using Gradient Boosting Decision Tree Model with Multi-Characteristic Representations of Light Signals 59
Photometric redshift estimation of galaxies by machine learning 60
Stereograph: stereoscopic event reconstruction using graph neural networks applied to CTAO 60
Reinforcement Learning for Calibration in Radio Interferometry 61
Infrared Bubble Recognition in the Milky Way and Beyond Using Deep Learning 61
Extending the Life of Software and Data 62
A science platform for the SOFIA Data Center 62
Montage and Radio Astronomy 63
GDL version 1.1 63
Using OpenAPI to describe complex web services 64
PipeCat: A Flexible and Modular Framework To Go From Images to Catalogs in Just a Few Steps 64
The IRAM 30-meter control system upgrade and approach 65
What’s new with astropy? Updates on a Community Python Library for Astronomy 65
Setting Sail with ULLYSES 66
Rewriting and Modernization of C++ Code for Gain and Atmospheric Calibration of Far-Infrared Heterodyne Astronomical Data: A Python-Based Approach 66
Database of Markarian galaxies and classification of low-dispersion spectra 67
Renewing the 2dFdr legacy codebase as a sustainable Python package 68
Ziggy, Improving the Pipeline Infrastructure of Kepler and TESS 68
Container-based pre-pipeline data processing on HPC for XRISM 69
User Experience 69
Online SIMBAD TAP query interface 69
Leveraging the experience of chatbot development in the AI "explosion" era to improve the astronomical data services 70
Results from the 2023 ESA Astronomy Space Science Archives User Survey 70
VisiOmatic 3: Remote Image Visualization with New Python-Based Features 71
ALMA Wideband Sensitivity Upgrade: Science Archive impact 71
Jasmine - JAvaScript Multimodal INformation Explorer 71
FitsImageVOY: A FITS Image Viewer for Visual Studio Code and Modern Successor to SAOImage 72
dysh: A Python Package for Calibration of Single-Dish Spectral-Line Data 72
TOPCAT/STILTS Interoperability 73
Building DAST: Data analysis and support for PLATO scientists 73
The PADC ObsTAP Portal: a user-friendly gateway to archived observations in the Virtual Observatory 74
New Python-based Architecture for the Keck Observatory Archive 74

Data Management and Trusted Repository in the Open Data Era 75

The ASTRON Data Archive 75

Accessing Data from the Near-Earth Objects Coordination Centre 76

S3AI: Spoke 3 Archive Infrastructure 77

Astronomical Visualizers for Enhanced Data Analysis in ESA's HST and JWST Science Archives 77

AstroDB Toolkit: A collaborative data management tool 78

Mini-SiTian Data Release 1 79

Improving the findability of data in the European VLBI Network archive 79

The New Architecture of the Online Observation Quality System for the ASTRI Mini-Array Project 80

Hipsgen: news features for news HiPS 81

A New VO Publishing Registry Framework at NOIRLab 81

Introduction of China-VO New Paper Data Repository System 82

CDS Open Data commitment for data published by authors 82

Roadblocks in Astronomical Data Analysis 83

Distribution of the Högbom CLEAN Algorithm Using Tiled Images with Feedback 83

The OPS4 GAIA legacy project: transitioning towards a new data management system 83

RCSEDv2: algorithm for automatic analysis of individual spectra from large surveys 84

SpectralFitting: a fresh approach to fitting spectra 85

Milky Way galaxies-analogues as the isolated AGNs: multiwavelength data incompleteness 85

Cataloging the first year of Euclid data 86

Galaxy Detection Energy-Efficient Computation: Deep Learning With Tensor Methods To Speed-Up Astronomical Imaging 86

Integrating the Pegasus workflow management system into the CCAT Observatory data center 87

Overcoming barriers: simplifying accessibility and visualization of gravitational wave data with GW Data Plotter 88

Unix philosophy in astronomical data analysis 88

The Fornax Initiative: A NASA Astrophysics Science Platform 88

GPU Accelerated Image Quality Assessment-Based Software for Source Detection 89

Imbalance Learning in Astronomy 89

Metadata, Semantics and Data Models Applicable to Data Formats 90

Classification of Equatorial Plasma Bubbles (EPB) Using Convolution Neural Network (CNN) 90

Metadata-Driven Navigation System of DARTS and AI-Powered Data Exploration 90

Optimizing data models and data management solutions to address Gaia data exploitation challenges . 91

Real-time and Near Real-time Processing Pipelines 92

ALMA Fast Lane: the architecture of real-time data delivery 92

The real-time data processor framework for data handling and analysis of high-energy instruments 92

Breizorro: Image masking and cataloguing suite 94

Accelerating Source Finding in 3D HI-Datacubes: GPU Parallelization of the SoFiA-2 Smooth and Clip Finder 94
Progress in ML-Based Spectral Kurtosis RFI Detector 94
Statistics recording and processing in LOFAR2 using modern, cloud-native and web technologies 95
The BurstCube Data Analysis Pipeline 96
PypeIt: The Python Spectroscopic Data Reduction Pipeline 96
Connecting Zenodo and Analysis Facilities: a JupyterLab Extension Middleware 96
In-Situ High Performance Visualization for Astronomy & Cosmology 97
The Roman Data Monitoring Tool 98
Automated Deployment of Radio Astronomy Pipeline on CPU-GPU Processing Systems: DiFX as a Case Study 98
Evaluation of effectiveness of RFI removal algorithms for searching for fast radio transients with the SKA 99
Automating testing pipelines using Notebooks 99
Design and Implementation of a High-Throughput Data Transfer System for the CCAT Observatory 100
PyMerger: Detecting Binary Black Hole merger from Einstein Telescope Using Deep Learning 100
PyMerger: ADARI: Visualizing the quality of VLT data 101

Other 102

Ten (or more) reasons to register your astronomy research software with the Astrophysics Source Code Library (ASCL, ascl.net) 102
A software system for supporting untargeted HI surveys - from observations to public data release 102
The Asteroid Detection Polar and ESA’s Synodic Orbit Visualisation Tool 103
Navigating Astrophysics Literature: Harnessing AstroBERT and UAT 104
Developing the optimal cross-matching algorithm for the Gaia and BDB data 104
"Astronomy and Computing": evolution, current status and future plans 105
Automating the configuration of INAF radio telescopes' control system using Ansible 105
Science Data Centre Architecture for LOFAR2.0 106
Nonlinear fitting of undersampled discrete datasets in astronomy 107
The Gaia DR4 bulk-download repository: when worlds collide 107
ASPIS prototype: goals, organisation and results from the CAESAR project 108
Integral Legacy Archive: a new approach to explore High-Energy Space Radiation 108
Two Fundamental Julia Packages: Astrometry.jl and FITS.jl 109
A crude but efficient pipeline for JWST MIRI imager : the case of SN1987A 109
DAEPO Projects of National Astronomical Data Center (NADC) 110
Exploring New Instrument deGradation Models and Algorithms (ENIGMA) 110
The Merged Pan-STARRS/Gaia Photometric Reference Grid 111
Acceleration of identification process time of small solar system bodies(SSSBs) on the wide field imaging survey data 111
Generating VPHAS+ HiPS with Montage 112

Implementing SIAv2 Over Rubin Observatory's Data Butler 112
Dynamic Planetary Orbital Diagrams with Microsoft Word VBA 113
A Content-Driven Strategy for Roman's Science Platform: Enabling Low-Barrier Access and Collaboration 113
Communication Strategies and Achievements : Data-Driven Astronomy Education and Public Outreach (DAEPO) Projects of NADC 114
Searching for Fireballs in Lighting, but Finding Other Things 114
MWA Data Archive: A Decade of Radio Astronomy, 54 PB of Data 115

Birds of a Feathers 116

Built To Last? 116
Usability and User Experience in astronomical Software 117
Strategies for heterogeneous processing and archiving. 117
What New Data Formats is the Community Using? What New Data Models does the Community Need? 118
General-Purpose Spectroscopic Data Reduction and Analysis Tools 118
Software doesn't write itself: Prioritising Equity, Diversity, & Belonging to improve software output ... 119

Tutorials 120

Processing and analyzing XMM-Newton data with ESA Datalabs: A collaborative approach. 120
The first step when thinking about User Experience: Set-up an UX Vision 120
The Advanced Scientific Data Format (ASDF) 121
Programming the GPU on your laptop - is it easy, is it useful? 121

Focus Demos 122

JupyterLab extension: FireFly 122
The ESA Near-Earth Objects Coordination Centre Python Interface 122
Using LSDB to enable large-scale catalog distribution, cross-matching, and analytics 123
XRADIO: Xarray Radio Astronomy Data Input Output 123
Exploring Space Weather connections with the ASPIS prototype archive 124

Invited Talks

How do you use yours? The Evolution of Proposal and Observing Preparation Tools.

Author: Alan Bridger
Affiliation: Square Kilometre Array Observatory / UK Astronomy Technology Centre
ID: I101

In the past few decades the way in which astronomers use telescopes (or observatories) has changed dramatically. These changes have brought significant efficiency upgrades (science per hour) and have generally improved the professional lives of both users and observatory staff. In this review I will describe a personal view of this history highlighting some significant developments and their impacts along the way.

I will also briefly consider other, wider, impacts, and offer some thoughts on the future evolution of how we use our telescopes.

From Daniel Dennett to Transformers: The Computational Evolution of Human Intelligence in AI

Author: John Abela
Affiliation: University of Malta
ID: I201

In this talk we explore the philosophical and technological advancements that shape our understanding of artificial intelligence. The discussion begins with an examination of the late philosopher Daniel Dennett's views, particularly his assertion that human intelligence is Turing computable and can be replicated through computational procedures. Dennett's perspective finds potential vindication in the capabilities of large language models (LLMs), such as ChatGPT and Gemini, which exhibit 'emergent properties' — complex behaviors arising from simpler underlying processes. While the mathematical foundations of these models are well-understood, the sheer scale, involving trillions of parameters, challenges our ability to predict or even explain their behaviors.

We also consider contrasting views from other philosophers, notably David Chalmers, who offers alternative insights into the nature of intelligence and consciousness. The talk culminates with a brief discussion on the applications of Transformer models in fields beyond traditional AI, such as cosmology and astronomy. These models, through their sophisticated use of the attention mechanism and deep architectures, open new avenues for understanding and exploring the universe. This talk aims to bridge philosophical theories and cutting-edge AI technologies, illustrating the computational evolution of human-like intelligence in machines.

Declarative Data Management with DaCHS and the VO
Author: Markus Demleitner
Affiliation: Universität Heidelberg, Astronomisches Rechen-Institut
ID: I301

Publishing data seems easy: Put it on a web page, obtain a DOI, and you are done. In practice, this kind of "dead" data generally is hard to find, access and hence to reuse, not to mention interoperability. Hence, the Virtual Observatory defines "active" interfaces to the data: standard protocols enable uniform querying and access, rich metadata in standard formats on standard interfaces ensure discoverability. This means that data publishers need to run non-trivial software. Software, however, has a fairly short half-life, in particular because of changing platforms, but also because the standards occasionally evolve. In this talk, I discuss how the DaCHS data publication package tries to mitigate this specific sort of bitrot, first and foremost by introducing a declarative layer ("state the problem, not the solution") in data publishing from ingestion to service operation to registration. I will show some examples for how this has enabled us and others to run data centres over many years with low to moderate effort, while staying up to date with the evolving VO. I will also delineate where no suitable declarative approaches have been found and what that meant during major platform changes like the move from Python 2 to Python 3.

Empowering Science with Good Design
Author: Jenn Kotler
Affiliation: STScl
ID: I401

Scientists rely on software to make discoveries. As the creators of science tools, our design decisions can create gigantic challenges out of the simplest tasks or enable scientists to do more science with ease. This talk will cover practical design methods for building astronomy tools, best practices for user experience and accessibility, and examples of design patterns especially relevant to science workflows.

The Challenges of Astronomical Data Systems
Author: Ben Rusholme
Affiliation: Caltech/IPAC
ID: I501

Astronomical data systems are the critical link between telescopes and the scientific community. Creating and maintaining these systems requires a wide range of valuable skills, not just scientific, as unfortunately highlighted by recent security breaches. Is ADASS the correct venue for such extended topics?

Beyond the Data: challenges and triumphs in data reduction and analysis
Author: Nuria Lorente
Affiliation: AAO, Macquarie University
ID: I601

While astronomy data reduction and analysis software faces significant challenges, it has also seen major successes in standardisation, automation and collaboration, ensuring that data are processed efficiently and in ways that are accessible to ever larger fractions of the community.

The future of this field lies in ever greater automation, the incorporation of machine learning techniques for near-real-time data analysis, and more seamless integration of heterogeneous datasets. As the volume of data continues to grow, our challenge is to ensure that pipelines remain scalable, robust, and flexible enough to handle both routine and unusual datasets. Of course advances in algorithms and technology are only part of the solution: collaboration across disciplines - software engineering, astronomy, data engineering and computer science - is key to the success of this field.

In this presentation I will highlight some of the successes, and a few failures, of this field and explore how as a community we are preparing to tackle the challenges of the next generation of projects.

SDHDF: A new file format for spectral-domain radio astronomy data
Author: Lawrence Toomey
Affiliation: CSIRO
ID: I701

Radio astronomy file formats are now required to store wide frequency bandwidths and multiple simultaneous receiver beams and must be able to account for versatile observing modes and numerous calibration strategies. The need to capture and archive high-time and high frequency-resolution data, along with the comprehensive metadata that fully describe the data, implies that a new data format and new processing software are required. This requirement is suited to a well-defined, hierarchically-structured and flexible file format. In this paper we present the Spectral-Domain Hierarchical Data Format ('SDHDF') — a new file format for radio astronomy data, in particular for single dish or beam-formed data streams. Since 2018, SDHDF has been the primary format for data products from the spectral-line and continuum observing modes at Murriyang, the CSIRO Parkes 64-metre radio telescope, and we demonstrate that this data format can also be used to store observations of pulsars and fast radio bursts.

The MeerKAT Science Data Processing Pipeline

Author: Ludwig Schwardt
Affiliation: SARAO
ID: I801

The MeerKAT radio telescope has been producing high-quality scientific data for more than six years. Its Science Data Processor (SDP) subsystem produces signal displays, calibration products, continuum images and spectral cubes in a fully automated and pipelined fashion from the outputs of the correlator and tied-array beamformer. I review the major parts of the MeerKAT SDP pipeline with a focus on the software architecture, design choices and lessons learnt.

Celebrating SAOImageDS9

Author: Kenny Glotfelty
Affiliation: Center for Astrophysics | Harvard & Smithsonian
ID: I901

In this talk we will recognize SAOImageDS9 (DS9) and its creators William Joye and Eric Mandel for their vital role in astronomy and visualization. We will start with a brief history of DS9 and its predecessors. We will review some impressive metrics that demonstrate the breadth of DS9 usage and showcase some unexpected, yet wholly remarkable, applications of DS9. (Did you know DS9 was used to support Covid-19 research?) Along the way we will provide some little know facts and helpful tips that can change the way users interact with the application. Finally we will discuss current work on DS9 and discuss the uncertain prospects for future development.

Talks

Proposal and Observation Preparation Tools

How did we build ours? A modern proposal tool for a modern telescope
Author: Alissa Cheng
Affiliation: ASTRON
ID: C101

The LOFAR (Low Frequency Array) telescope, one of the largest low frequency telescopes in the world, is currently undergoing a major upgrade dubbed LOFAR2. This includes new hardware capabilities and improved software to use it with.

However, the end-of-life status of the telescope's old proposal tool, NorthStar, means a new one is needed for LOFAR2. Applying the lessons learned from NorthStar and other existing tools like HEDWIG, a conscious choice was made to develop the new proposal tool in a way that is also (relatively) new to the world of research software development. So that we can build it with emphasis on, and balance between, the aspects of user experience, software maintainability, and development speed.

This presentation will demonstrate how the above is achieved by means of:

- stakeholder involvement in short iterations
- design system implemented as an in-house component library
- code analysis and software monitoring tools
- integration with downstream software (scheduling tool, data explorer)

Last but not least, in order to showcase the resulting product, there will be a walkthrough of the proposal tool from the perspective of our users – both proposal writers and telescope operators.

The proposal evaluation process: A unified user experience supporting different workflows.
Author: Dario Dorigo
Affiliation: ESO
ID: C102

At the beginning of 2020 ESO approved a project to create a new tool supporting its proposals evaluation processes, called P1Flow. The main idea was to upgrade a series of old and standalone tools that were used to handle ESO peer reviews, into one flexible interface that could offer a user-friendly experience to the reviewers and a complete overview of the process to the ESO operator.

Due to the pandemic and the impossibility to hold panel meetings in-person, the reference model of ESO peer review, i.e., face-to-face panel meetings with about 80 people invited every six months to work for some days altogether in the same location became unfeasible. The sudden need to hold online panel discussions put the project under significant time pressure to quickly provide basic interfaces so that the ESO peer review system could continue. This implied a staggered implementation and release of interfaces supporting different phases of the process. However, the old infrastructure had to continue working in parallel with the new system, to cover for the remaining phases of the review.

The first version of the tool was released in April 2021. The project will reach its completion by the end of 2024.

In the article we will describe the project, how we kept the legacy system and the new tools working together by migrating data from one domain to the other, the strategies we used to keep complexity low, how we introduced new workflows in the system (e.g. Distributed Peer Review), the current status of the project and its possible evolutions in the future.

The description of the technologies, the project plan, the team evolution with time according to the different needs will also be provided.

STARS: A scheduling software for Space Missions and Ground-Based Observatories

Author: Néstor Campos
Affiliation: Institut d'Estudis Espacials de Catalunya
ID: C103

The efficient automatic planning and scheduling of astronomical observations from space and ground-based observatories has become essential for large astronomical surveys. It facilitates the coordination of multiple instruments and observatories located at different sites and enables a fast reaction to changes of the environmental conditions while maximizing the scientific return.

The main challenge for astronomical planning tools is to maximize the number of observed targets, taking into account available resources, time, and observation constraints. These constraints are mainly related to target visibility, observation cadence, the ephemerides of astronomical events, and the phase of the moon, among others. The resulting observational plans shall fulfill all the constraints and shall be optimal in terms of the telescope time used for science observations, that is to say, minimize the idle time of the instrument, and maximize the scientific return.

In this talk, we will present the STARS library (Scheduling Telescopes as Autonomous Robotic Systems), which provides the tools needed to generate optimal schedules for space and ground-based observatories. It includes the means to define the tasks to be planned, their constraints and the observational resources, the scheduling algorithms based on Artificial

Intelligence techniques, and their optimization goals (figures of merit). We also provide a web-based graphical user interface to easily visualize and analyze the generated schedules. The future development plans include, among other things, the ability to generate observation schedules directly from the browser, thus providing a graphical interface throughout the scheduling process.

STARS is being successfully used in several ground and space-based observatories. The CARMENES instrument at the Calar Alto Observatory and the TJO robotic telescope at the Montsec Observatory are already in operation and using the STARS library. Additionally, the library will be employed for the operations of the CTA Observatory, the scheduler of the Ground Observations Program of the PLATO M3-ESA mission, and the planning tool for the ARIEL M4-ESA mission, all of which are currently in development.

Observation Scheduling Software Framework for Distributed Telescope Arrays in Time-Domain Surveys

Author: Yajie Zhang
Affiliation: Tianjin University
ID: C104

Telescope arrays are increasingly valued for their higher resource utilization, broader survey areas, and more frequent space-time monitoring compared to single telescopes. This new observation mode poses a challenging demand for efficient coordination of distributed telescopes while coherently modeling abstract environmental constraints to achieve scientific goals. We propose a multilevel scheduling model and a flexible software framework for distributed time-domain survey telescope array. This framework is constructed from both global and site levels, successfully solves the telescope array scheduling problem considering the projected volumes of constraints and objectives. A remarkable feature of the framework is its ability to achieve global control of generic large-scale surveys through multi-level scheduling, dynamically responding to unexpected interruptions with robustness and scalability. Also, a Python simulator is built to model telescope array observations, including the creation of scheduling blocks obtained from global scheduler, observation conditions, telescope equipment status, and observation fields, enabling the evaluation of scheduling algorithms under various settings. This paper provides examples of telescope array scheduling algorithms and time-domain survey evaluation metrics implemented within the proposed framework. We envision this prototype framework being used to develop automated scheduling schema that support multi-telescope, multi-site coordinated observations. By integrating novel artificial intelligence techniques and solvers, further performance optimizations can be easily supported.

The Rise of AI for Science and Data Center Operations

Machine Learning Enhancements for Real-Time Scientific Analysis of Cherenkov Telescope Data

Author: Ambra Di Piano
Affiliation: Università di Modena e Reggio Emilia, Istituto Nazionale di Astrofisica - Osservatorio di Astrofisica e Scienza dello Spazio di Bologna
ID: C201

The Cherenkov Telescope Array Observatory (CTAO) will provide incredible opportunities for the future of ground-based very-high-energy gamma-ray astronomy. To optimise its scientific output, the CTAO will have a Science Alert Generation (SAG) system, which as part of the Array Control and Acquisition (ACADA) system will perform reconstruction, data quality monitoring and scientific analysis in real-time to detect and issue candidate science alerts. As part of the continuous research and development activity for improvements of future versions of the ACADA/SAG product, this work aims at implementing machine learning enhancements for the scientific analysis. In real-time technical and observational variability, as well as performance requirements, can highly impact the overall sensitivity of the automated pipelines. We developed two prototypes of Convolutional Neural Network based models with the aim of removing any a priori knowledge requirements that standard scientific tools have. The first model is an autoencoder trained to remove background noise from counts maps of a given observation, without requiring inputs on target position, background templates or instrument response functions (IRFs). The second model is a 2-dimensional regressor that extracts hotspots for the localisation of candidate sources in the field of view, without requiring inputs on the background template or IRFs. To verify both models we use the current version of ACADA/SAG (rel1), finding that they achieve comparable results with the additional benefit of not requiring a priori knowledge.

AI Agents for Ground-Based Gamma Astronomy

Author: Dmitriy Kostunin
Affiliation: DESY
ID: C202

The Cherenkov Telescope Array Observatory (CTAO) represents the next generation in ground-based gamma astronomy, marked by a substantial increase in complexity with dozens of telescopes. This leap in scale introduces significant challenges in managing system operations and offline data analysis. Traditional methods, which depend on advanced personnel training and sophisticated software, become increasingly strained as the system's complexity grows, making it more challenging to effectively support users in such a multifaceted environment.

To address these challenges, we propose the development of AI agents based on instruction-finetuned large language models (LLMs). These agents align with specific documentation and codebases, understand the environmental context, operate with external APIs, and communicate with humans in natural language. Leveraging the advanced capabilities of modern LLMs, which can process and retain vast amounts of information, these AI agents offer a transformative approach to system management and data analysis by automating complex tasks and providing intelligent assistance.

We present prototypes aimed at integrating with CTAO pipelines for operation and offline data analysis. The first prototype is a plugin that implements chatting and function calling for the Configuration Database of Array Configuration and Data Acquisition (ACADA). This plugin enhances operational efficiency by providing intuitive, context-aware assistance and automating routine tasks. The second prototype is an open-access custom ChatGPT tailored for the gammapy-based data analysis, which offers interactive support for researchers.

Classification of HI Galaxy Profiles Using Unsupervised Learning and Convolutional Neural Networks: A Comparative Analysis and Methodological Cases of Studies

Author: Gabriel Andres Jaimes Illanes
Affiliation: Instituto de Astrofísica de Andalucía (IAA)
ID: C203

Hydrogen is the most abundant element in the universe, making it essential to the formation and evolution of galaxies. The 21 cm radio wavelength neutral atomic hydrogen (HI) line maps the distribution and dynamics of gas within galaxies. The emission from this spectral line is an important tracer for galaxy interaction studies and understanding galactic structure, star formation processes and general behavior of the Interstellar Medium. The application of Machine Learning (ML) and BigData tools algorithms are assets to tackle the enhancement of the quality and efficiency of scientific analysis in this field, especially when it involves large radio astronomy databases and the study of spectrum classification.

Within this context, our work aims to propose a framework for the classification of HI spectral profiles using ML techniques. Several methodologies integrating unsupervised ML techniques and Convolutional Neural Networks (CNN) have been implemented. To carry out this approach, we have focused on HI datasets used in the AMIGA (Analysis of the interstellar Medium in Isolated GAlaxies) research group with a sample of 318 CIG (Catalog of Isolated Galaxies) spectral profiles and 30780 profiles from the ALFALFA (Arecibo Legacy Fast ALFA) survey.

To design this classification framework the first step was data preprocessing, using the Busyfit package (Westmeier et al, 2014) for HI spectrum profile fitting. A second data set was generated using iterative fitting with polynomial, Gaussian, and double-Lorentzian models. This approach also involved a multi-faceted strategy for profile clustering based on temporal shapelet transformation for features detection algorithms: K-means, spectral clustering, DBSCAN, agglomerative clustering, among others, as bootstrap for the extraction of features. Furthermore, we considered a series of classification techniques that include K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Random Forest classifiers. In order to optimize the performance of such models, CNN model was probed, where we made an in-depth evaluation for various configurations of the model with regard to their impact on classification accuracy.

The second part of this work is focused on the generation of an additional dimension to the profiles in order to improve the classification. This 2D analysis is based on the application of CNN techniques to determine the degree of asymmetry by carrying out the classification of the sample of CIG galaxies. The original data was modified by adding a new dimension to the profiles in order to improve the classification. Three distinct 2D image models were generated for the symmetry study: the first is a rotation of the fitted spectrum, the second involves rotating the spectrum after subtracting its right and left profiles to accentuate asymmetry features, and the third is a normalized version of the previous image, with pixel intensity adjusted to further emphasize specific image features. We explain the methodology with current ML techniques and discuss the extrapolation to the ALFALFA survey. The resulting classification was compared with a profile classification previously made by the AMIGA scientific group (Espada, 2011).

The study presents the application of ML techniques for classifying HI profiles, including an approach to extract profile asymmetries classification with HI profiles transformation into 2D images, to improve the accuracy and depth of future analyses. With this, we also have the intention to build and verify a minimal methodology that could potentially be applied to the ongoing Square Kilometre Array (SKA) precursor surveys such as MeerKAT (MIGHTEE HI) or Apertif, where the number of detections will be higher, thus laying the foundation for building a full-scope methodology in the SKA era. All material, code and models have been produced following the FAIR principles and have been published in an open access public repository.

An AI-driven system for enhancing Astronomical Research workflows

Author: Karthik Mahesh Rathod
Affiliation: Instel
ID: C204

The rapid expansion of astronomical literature has created significant challenges for researchers aiming to increase the visibility and impact of their work. To address this challenge, we developed Stellar Forge, an AI-powered platform that optimizes the research publication workflow by integrating the NASA Astrophysics Data System (ADS) and the Unified Astronomy Thesaurus (UAT).

Stellar Forge provides several key services. First, it includes a concept tagger that uses a large language model with a retrieval pipeline built upon a novel hierarchical representation of UAT to automatically tag papers with relevant UAT concepts, improving discoverability across platforms and reducing human-induced errors. The tool also provides proper justifications, hierarchies, and concept positions within a branch, and it also suggests new concepts with branch positions if they emerge, ensuring that research remains accessible and up-to-date. The platform also includes a performance prediction tool that estimates the potential readership of papers within the ADS platform. This feature analyses abstract content and previous trends of similar publications to probabilistically forecast readership for the upcoming months, helping researchers anticipate the impact of their work. In addition, Stellar Forge supports content creation with title and abstract optimization tools that improve the visibility of research, allowing researchers to cater their writing for specific audiences and purposes with precise controls. These tools use task-specific fine-tuned large language models and dynamic one-shot prompting to craft clear, effective titles and abstracts.

Built on a modular architecture, Stellar Forge integrates with existing research systems to provide continuous support and improvements to researchers, from content creation to maximizing research visibility. By bridging AI with astronomical resources and researcher needs, Stellar Forge opens the door to more efficient and impactful advancements in astronomical research.

Goal-Oriented Stacking: an novel approach to statistical image-domain inference below the noise threshold

Author: Roger Deane
Affiliation: University of the Witwatersrand / University of Pretoria
ID: C205

A commonly used approach to explore astrophysical sources below the detection threshold is image-domain stacking or co-adding. This uses known positions of a source population sample identified at one observing wavelength to make statistical measurements of the sample at a different wavelength, where the images are not sufficiently deep for direct detections of the individual objects. These samples are typically selected through intrinsic or observed properties such as stellar mass or optical colours in an attempt to limit biases, maximise completeness, or separate out sub-populations of interest. We explore the utility of an alternative approach by designing an algorithm (using a non-linear neutral controller) to select subsets of the input parent sample of galaxies based on what we refer to as “goal-orientated stacking” objectives. In this case, we set the goal as identifying a subset of galaxies with physically correlated properties (e.g. stellar mass, redshift, star formation rate) that maximise the radio continuum signal-to-noise level. We explore a few applications of this alternative approach and discuss possible extensions.

Self-supervised learning of radio data for source detection, classification and peculiar object searches

Author: Simone Riggi
Affiliation: INAF
ID: C206

New advancements in radio data post-processing are underway within the SKA precursor community, aiming to facilitate the extraction of scientific results from survey images through a semi-automated approach. Several of these developments leverage deep learning (DL) methodologies for diverse tasks, including source detection, object or morphology classification, and anomaly detection. Despite substantial progress, the full potential of these methods often remains untapped due to challenges associated with training large supervised models, particularly in the presence of small and class-unbalanced labeled datasets.

Self-supervised learning has recently established itself as a powerful methodology to deal with some of the aforementioned challenges, by directly learning a lower-dimensional representation from large samples of unlabeled data. The resulting model and data representation can then be used for data inspection and various downstream tasks if a small subset of labeled data is available.

In this study, we explored contrastive learning methods to learn suitable radio data representation from unlabeled images taken from the ASKAP EMU and MeerKAT GPS surveys. We evaluated trained models and the obtained data representation over smaller labeled datasets, also taken from different radio surveys, in selected analysis tasks: source detection and classification, and search for objects with peculiar morphologies.

Transforming Data into Insights: AI-Driven X-Ray Source Classification within the NADC Framework

Author: Xiaoxiong Zuo
Affiliation: National Astronomical Observatories, Chinese Academy of Sciences
ID: C207

The advent of AI has revolutionized the field of astronomy, particularly in the realm of time-domain astronomy. This talk focuses on the application of AI within the framework of the National Astronomical Data Center of China (NADC), which encompasses its data infrastructure and science platform. The NADC framework plays a pivotal role in converting raw astronomical data into valuable scientific insights. The Einstein Probe (EP) serves as a case study, exemplifying the integration of AI with the NADC framework to enhance the discovery and analysis of transients and variable sources. The Time Domain Information Center (TDIC) science platform within the NADC facilitates the application of AI for science and enabling the efficient handling and interpretation of vast datasets generated by astronomical satellites like the EP.

The core of this talk focuses on the development and implementation of a classification algorithm within the NADC framework. The algorithm, a Random Forest classifier, leverages features extracted from light curves, energy spectra, and spatial information to autonomously classify observed X-ray sources. Demonstrating remarkable accuracy rates of approximately 95% on EP simulation data and an impressive 98% on observational data from the EP pathfinder Lobster Eye Imager for Astronomy (LEIA). The integration of this AI classifier into the data processing pipeline not only accelerates the manual validation process but also serves as a testament to the NADC's commitment to advancing scientific research through technological innovation. The talk concludes with an exploration of the implications of the most effective features for X-ray source classification and the broader application of these AI techniques to other X-ray telescope data, thereby setting the stage for future advancements in time-domain astronomy. By showcasing the successful application of AI within the NADC framework, this talk aims to inspire further integration of technologies in astronomical research, paving the way for new discoveries and a deeper understanding of the universe.

Extending the Life of Software and Data

Synergies Unleashed: Bridging the Gap Between Science and Computing teams in the ALMA Observatory software deployments

Author: Jose Gallardo
Affiliation: ALMA Observatory
ID: C301

The ALMA Observatory has been collecting science data for more than 10 years. During the first years of operations (up to the end of Cycle 4, i.e., September 2017) the focus of the ALMA Integrated Computing Team (ICT) and the Department of Science Operation (DSO) was mainly on the data acquisition part (using what is known in the ALMA jargon online software).

Thanks to the stability of the online software, during Cycle 5 (October 2017 – September 2018), the Observatory reached a very high data acquisition performance, but with a negative impact on Data Processing (DP) and Data Delivery (DD), one of the reasons being that the software downstream data acquisition (known as offline software) was not mature enough to cope with such a large amount of incoming data.

One of the major contributors to the immaturity of the offline software was that ALMA underestimated the importance of an integration procedure. The applications were working as expected individually, but not as part of an entity which contains software components with interdependencies.

In the last years, significant resources have been allocated to consolidate the performance of the offline software. The situation was reversed, thanks to an efficient, coordinated, and collaborative plan between ICT (especially the EU part) and DSO. The outcome of this effort is a suite of regression and integration tests for the ALMA offline deployment process. This paper (a) describes in more detail the complete history behind this effort, (b) show the current regression and integration tests in place for the most important ALMA scenarios and (c) presents the cutting-edge technology behind the automation approach. We also discuss how this innovative approach in looking at science operations from the software perspective, with an enhanced and coordinated collaboration between the two mentioned teams (ICT and DSO), can become a game-changer in the improvement of an Observatory's performance. The statistics collected demonstrated that the offline software has become much more robust, as both the occurrence of bugs and the need for patches have significantly diminished during the past four cycles.

The Chandra Data System at 25 years — What can it teach us?

Author: Janet Evans
Affiliation: Center for Astrophysics | Harvard & Smithsonian
ID: C302

The Chandra X-ray Center Data System (CXCDS) software provides end-to-end software support for Chandra mission operations. The CXCDS Software Team develops, maintains, and supports the software that drive the CXC-based forward (proposer to spacecraft) and return (spacecraft to observer) threads necessary to perform the Chandra observing program. The Data System also includes the CIAO data analysis package and the Chandra Source Catalog (CSC) processing system that recently completed CSC 2.1. The software system consists of ~2 million logical lines of code, including C/C++, Python, SQL, Java, Perl, and a few stray algorithms written in Fortran. The Chandra Data Archive manages the operational threads and serves all data of the system. The data products are written in FITS format and are OGIP and IVOA compliant.

For twenty-five years the CXCDS has served Chandra science operations and the user community well. The up-front planning and detailed design, conceived in the mid-1990's and implemented for the operational system by launch in 1999, has paid off. The Software Team has managed operational changes to algorithms and processing, operating systems, compilers, scripting languages, and most recently a configuration management system migration to Git. All of these upgrades and many more have been possible given the structured architectural design that enabled, among other things, modularity and flexibility to manage change.

In this talk, I'll provide insight on the longevity of the Chandra Data system, our ability to introduce change successfully, and what a new project can take away from our experience going forward.

DASCH: Bringing 100+ Years of Photographic Data into the 21st Century and Beyond
Author: Peter K. G. Williams
Affiliation: Center for Astrophysics | Harvard & Smithsonian
ID: C303

The Harvard College Observatory was the preeminent astronomical data center of the early 20th century: it gathered and archived an enormous collection of glass photographic plates that became, and remains, the largest in the world. For nearly twenty years DASCH (Digital Access to a Sky Century @ Harvard) has been actively digitizing this library using a one-of-a-kind plate scanner. Earlier this year, after 470,000 scans, the DASCH project finished. Now, this unique analog dataset can be integrated into 21st-century, digital analyses. The key DASCH data products include ~350 TB of plate images, ~50 TB of calibrated light curves, and a variety of supporting metadata and calibration outputs. Virtually every part of the sky is covered by thousands of DASCH images with a time baseline spanning more than 100 years; most stars brighter than B ~ 15 have hundreds or thousands of detections. I will present the DASCH data release and discuss some of the lessons learned while trying to make data from the previous century accessible in the next century and beyond.

Curating a 20th century observation log in the 21st century
Author: Sébastien Derriere
Affiliation: CDS, Observatoire astronomique de Strasbourg
ID: C304

The International Ultraviolet Explorer (IUE) was a space mission which operated between 1978 and 1996. The final merged log of IUE observations, published in 2000, contains a vast collection of 110033 spectra, which are still of scientific value decades later.

A special operation was performed in 2000 at the Strasbourg astronomical data center (CDS) to provide links between astronomical objects from the SIMBAD database and the IUE Newly Extracted Spectra hosted at VILSPA by ESA. This resulted in 65872 spectra being linked from 7392 distinct SIMBAD objects.

Despite the considerable growth of the number of objects referenced in SIMBAD between 2000 and 2024 (from less than 3 to more than 18 millions), the links to the IUE archive remained unchanged. We therefore decided to attempt a significant update, trying to provide as many links as possible between SIMBAD objects and the IUE spectra.

We describe in this paper the challenges in trying to improve the discoverability of archival data several decades after the end of the mission. The most time-consuming part is the recovery of information hindered by the use of improper object identifiers, concealed implicit information, and human errors (typos).

Ultimately, we were able to link 99.7\% of the eligible IUE spectra to SIMBAD objects (calibration spectra, and spectra of solar-system objects are not relevant for SIMBAD), compared to only 67.7\% in the 2000 operation. We also generated the space-time coverage of the corresponding spectra using the ST-MOC IVOA standard.

Our work strongly advocates for the Best Practices for Data Publication in the Astronomical Literature (Chen et al., 2022), and also stresses the needs for extra care in the observer's proposal tools and metadata management, in order to facilitate the long-term use and optimal scientific return of astronomical observations.

Longevity of a treasured database service

Author: Xiuqin Wu
Affiliation: Caltech, IPAC-NED
ID: C305

The NASA/IPAC Extragalactic Database (NED) is a comprehensive database of multiwavelength data for extragalactic objects, providing a systematic, ongoing fusion of information integrated from hundreds of large sky surveys and tens of thousands of research publications. The contents and services span the entire observed spectrum from gamma rays through radio frequencies. NED has been serving the public for 34 years. Over this period, database and software technologies have advanced by leaps and bounds. To keep providing the invaluable services the community has come to rely on, NED team continues to integrate the science data as they were published, serve the data to community via user interface and APIs, while expand and upgrade the software and database systems. This talk will give an overview of what we have done and what we are doing now to ensure the longevity of this treasured system, both database and software.

User Experience

Processing LISA's data as a human: The GlobalFit Framework user experience

Author: Antoine Basset
Affiliation: CNES
ID: C401

Space missions have always recorded electromagnetic signals, from infrared light to gamma rays. Expected to launch in 2037, ESA's large-class mission LISA (Laser Interferometer Space Antenna) will survey gravitational wave signals from space. As the world's first in-orbit instrument to probe space-time itself, this is one of the most ambitious science missions ever. LISA promises a wealth of new science, allowing us to test our understanding of general relativity and to open a new window for astrophysics and cosmology. The data analysis for this mission will have to disentangle superposed signals from a variety of astrophysical sources, as well as modeling the instrumental noise. At the heart of a distributed data analysis system lies a gigantic Bayesian inference pipeline: the Global Fit. The computational challenge will be massive, expected to be about an order of magnitude heavier than the data processing of the recent ESA mission Euclid, in optimistic scenarios.

The inference of the parameters of each source will require source separation, complicating the estimation of their posterior distributions which is already challenging for isolated gravitational events. When separation is not possible, the number of superimposed sources becomes an unknown and the signals themselves form a confusion background comparable to noise; trans-dimensional analysis is then required, which yields additional complexity. To tackle the challenge of the Global Fit, the currently envisioned approach relies on a Markov chain Monte Carlo (MCMC) strategy, with block Gibbs sampling across the classes of sources (and the noise level) to reduce the complexity. Even using this trick, existing pipeline prototypes are computationally expensive and scale badly. As a consequence, the scientific community is looking for technological and algorithmic breakthroughs, e.g. relying on GPUs, sparsity-based modeling or artificial intelligence.

The so-called GlobalFit Framework will provide an abstraction layer between the distributed system components in charge of pipeline orchestration and execution on one hand, and the scientific modules on the other hand. This layer will offer to the scientific development team a convenient way to interact with the underlying components and hence focus on the algorithm development. Among others, the Framework will guarantee that the scientific modules can be integrated with low coupling, thus allowing dozens of labs to contribute with various languages and technologies. It will also handle the module scheduling, including the iteration logic, scalability and adaptation to available resources. It will feature check-pointing and resuming capabilities, and communicate in real time with concurrent pipelines running in different computing centers across the world. All of those technical requirements already make the Framework development an engineering challenge. Yet, the most complex features to be implemented relate to the user experience: Given such an algorithmic and computational beast, how to unleash its scientific potential? To do so, our Framework prototype is equipped with two dashboards which will be presented in more details: an Operation Dashboard for operators and a Monitoring Dashboard for science experts.

The proposed talk will address the following questions: How to support debugging and investigation at runtime, when thousands of sources are being processed in parallel? How to actively monitor the progress of the estimators and how to display a relevant synthesis of the source catalog in live? How to provide interactivity to the operators without wasting resources?

User facing tutorials as code: reproducible and reliable tutorials with CI/CD

Author: Brigitta Sipőcz
Affiliation: Caltech
ID: C402

User-facing tutorials typically combine code, narrative text, execution results, and visualization. However, the target audience of these tutorials can differ significantly, tutorials are often served to be part of the documentation to be accessed by individual users as part of their asynchronous learning, while at other times tutorials are presented at one-off workshops while they are deployed on specific science platforms.

This talk presents the best practices for reliable and reproducible tutorials assembled as part of a Scientific Python Ecosystem Coordination (SPEC) document. These practices distinguish between the different flavours of tutorials and offer guidance for each of them.

The talk specifically showcases how we have implemented and used these best practices at IRSA, and focuses on the ecosystem we adopted for maintaining, testing and deploying our tutorials to the scientific user community. In our approach, we treat these tutorials as library code, and test them in an automated and regular CI/CD setup we serve them in an aesthetically pleasing, user-friendly way.

Accessible visualization of Astropy objects and Multi-Order Coverage Maps (MOCs) with the iPyAladin Jupyter widget

Author: Manon Marchand
Affiliation: Strasbourg Astronomical Data Centre
ID: C403

iPyAladin allows to view astronomical data interactively within Jupyter notebook. In this presentation, we will demonstrate the new capabilities of `ipyaladin` and its new compatibility with the Astropy ecosystem.

We will highlight the new features of the latest versions of the `ipyaladin` widget. A few lines of code allows one to visualize any catalog downloaded thanks to `astroquery`. It is then possible to make a visual sub-selection of this catalog by drawing a circle or a rectangle in the interactive widget. This sub-table can then be retrieved in a new `astropy Table`. `ipyaladin` also allows to display FITS files from disk or from `astropy`'s objects. The other way around also works: the image survey currently displayed in the widget can be cutout into a new FITS file. An other new functionality is the display of sky regions -- with `astropy-regions` support. Approximated sky regions (Multi-Order Coverage) can also be overlayed onto the view by the supporting the Astropy-affiliated module `MOCpy`.

We will showcase how these new methods enable a workflow in which the programmatic approach within Python benefits from constant visual checks in a widget on the side of the notebook's cells.

DARTS Timescape: Exploring 50 Years of Space Science Data Through Interactive Visualization

Author: Miriam Sawczuck
Affiliation: JAXA ISAS
ID: C404

DARTS, the data archive operated by JAXA's Institute of Space and Astronautical Science, provides a wide range of time series data obtained from space science missions over a period of 50 years. In this paper, we present the development of "DARTS Timescape", a system designed to provide efficient access to these time series data and allow users to explore the vast temporal landscape of missions archived in DARTS.

To build this system, we utilized InfluxDB, an open-source time series database designed for fast, high-availability storage and retrieval of time series data. InfluxDB can be easily used as a data source in the likewise open-source tool Grafana to build an interactive web application for data visualization.

Users of DARTS Timescape can compare data from various space missions side by side, analyse long-term trends with data aggregation such as calculating averages, and more.

For demonstration purposes, we will present various datasets providing insights into lunar seismology, magnetospheric phenomena, and high-energy astrophysics, specifically:

1. Apollo Lunar Seismic Data: This dataset includes seismic measurements from the Moon's surface, captured by instruments deployed during the Apollo missions from 1969 to 1977. These data illustrate the motion and geologic characteristics of the lunar surface.

2. Arase, MAXI and CALET Magnetosphere Data: Observations from the Arase satellite (launched in December 2016), the Monitor of All-sky X-ray Image (MAXI), and the Calorimetric Electron Telescope (CALET), the latter two being mounted on the International Space Station (ISS) since July 2009 and August 2015, respectively, provide data on relativistic electrons and plasma waves trapped in the Earth's magnetosphere. This data is crucial for advancing space weather research, enhancing our understanding of space environment dynamics and their impacts.

3. MAXI X-ray Data: MAXI furthermore offers X-ray data from various celestial bodies, enhancing our understanding of high-energy phenomena in the universe.

By integrating these diverse datasets into a unified visualization platform, DARTS Timescape aims to facilitate comprehensive research and discovery across multiple domains of space science.

Integrating UX Design in Astronomical Software Development: A Case Study
Author: Yan Grange
Affiliation: ASTRON
ID: C405

In 2023, ASTRON took the step of incorporating a dedicated User Experience (UX) designer into its software development process. This decision aimed to enhance the accessibility and usability of services providing access to our data holdings, as well as to optimize the design of services within the SKA Regional Centres Network.

The field of astronomical software development has historically underemphasized UX design. ASTRON's initiative represents a shift, not only in improving our own tools but also in demonstrating to the broader community the value of integrating UX expertise into development teams.

This presentation will explore the impact of embedding a UX designer within our organisation. We will discuss:

- The rationale behind hiring a dedicated UX professional
- The integration of UX methodologies into our software development lifecycle
- Challenges and lessons learned in this integration process
- The potential for wider adoption of UX-focused approaches in astronomical software development

By sharing our experiences, we aim to contribute to the ongoing dialogue about best practices in software engineering within astronomy and astrophysics, emphasizing the critical role of user-centered design in creating more effective and accessible tools for the scientific community.

Data Management and Trusted Repository in the Open Data Era

Securing Space Science: Advanced Data Protection in the HREDA Archive
Author: Anastasia Andres
Affiliation: Starion for ESA
ID: C501

The HREDA (Human and Robotic Exploration Data Archive) is a data archive and information portal that contains ESA funded or co-funded investigations and experiments since 1972. These experiments are performed in the International Space Experiments from different investigation fields such as growing vegetables and fluids physics studies in the space.

The archive is a joint effort by ESA’s Directorate for Human and Robotic Exploration, the Directorate of Science, and the Science Data Centre (SDC) Madrid. It became operational in 2020 and supersedes the former Erasmus Experiment Archive (EEA) and the ESA Microgravity Database (MGDB).

HREDA is developed by the ESAC Science Data Centre (ESDC). The ESDC provides services and tools to access and retrieve observations and data from ESA's space science missions (astronomy, planetary science, heliophysics and human robotic exploration).

The data archive within the system is highly heterogeneous, requiring the management of different security levels. Some data is always public, some becomes public after an initial prior access, whereas sensitive data requires special permissions for accessing like medical analysis from astronauts. This paper focuses on the advanced mechanisms developed for accessing sensitive data.

Our archive incorporates an advanced security framework for managing sensitive data, ensuring compliance with data protection standards. The system ensures that all incoming data is received in an encrypted format, safeguarding it from unauthorized access from the moment it enters the network. Each dataset is assigned an unique certificate, adding an additional layer of security and traceability. Decryption keys are securely stored in a robust Key Management Service (KMS) server, further protecting the data from breaches.

Our solution integrates two-factor authentication (2FA) with HRE-IC Internet Secured Services (HISS) to provide an extra layer of security, ensuring that only verified users can access the data. Additionally, access authorization is meticulously managed on an individual basis, with each user requiring explicit approval. This personalized authorization process guarantees that only the right personnel have access to the sensitive information.

This implementation enables the end user to securely download decrypted data without having to manage the complexities of certification and key management.

Insights from a 30-Year international Partnership on Astronomical Archives

Author: David Rodriguez
Affiliation: Space Telescope Science Institute
ID: C502

In an era where astronomical data is expanding at an unprecedented rate, the importance of data sharing and accessibility among astronomy archives cannot be overstated. Since the 1990s, an international partnership between the Space Telescope Science Institute (STScI), the European Space Astronomy Centre (ESAC), and the Canadian Astronomy Data Centre (CADC) has been focused on this endeavor, facilitating the exchange of data from the Hubble and James Webb Space Telescopes.

We will present how this collaboration has evolved over time, highlighting key milestones and innovations in decision-making, communication, and technology. Additionally, we will discuss some of the challenges we have encountered and the strategies we employed to overcome them, offering insights that could benefit future archive collaborations.

Roadblocks in Astronomical Data Analysis

Optimized Open-Source Tools for Scalable Solar System Science
Author: Alec Koumjian
Affiliation: Asteroid Institute
ID: C601

Asteroid Institute is developing a suite of open-source tools tailored for solar system science, with a focus on precision, accuracy, and scalability. Key functionalities include orbit propagation, ephemeris generation, orbit fitting, residual calculation, coordinate transformations, arc extension, Monte Carlo simulations, and impact analysis. The libraries are all rigorously unit-tested and benchmarked as well as compatible with existing Python tools like astropy and pandas.

Our science workloads run compute-intensive tasks on billions of data points across thousands of VMs with a low memory footprint. We outline the design decisions of our libraries that significantly reduce memory consumption while enabling seamless multiprocessing. Additionally, we highlight the advantages of our type-safe dataframes and integer-based datetimes over traditional approaches. We also share our Python data packages for convenient access to SPICE kernels and more.

Lowering in-memory footprint of antenna beams via polynomial approximation
Author: Ali Taqi
Affiliation: University of Oxford, UK
ID: C602

With the emergence of new radio telescopes promising larger fields-of-view at lower observation frequencies (e.g., SKA), addressing direction-dependent effects (DDE) (e.g., direction-specific beam responses; sector-based ionosphere corruptions) has become all the more important. Be it through A-projection or major-cycle calibration strategies, addressing DDE often requires reliable representations of antenna/station beams; yet, these require significant amounts of computational memory as they are baseline-, frequency-, time- and polarisation-dependent. A novel prototype is reported here to approximate antenna beams suitable to SKA-MID using Zernike polynomials. It is shown that, beam kernels can be approximated up to 3 lobes with sufficiently few

coefficients, thereby replacing the memory-intensive sampled beams. It is hoped that these results facilitate more efficient beam-dependent solutions and approaches to tackling polarisation leakage; all of which are essential for large-scale radio telescopes.

A Reproducible Science Workflow System: DALiuGE in Action

Author: Andreas Wicenec
Affiliation: ICRAR, University of Western Australia
ID: C603

The DALiuGE science workflow system has been introduced to the ADASS audience in 2022 and since then it has evolved into a sophisticated tool allowing the construction, scheduling and execution of arbitrarily complex workflows on single machines and clusters with thousands of compute nodes. Almost any software package exposing a Python binding can be automatically introspected, including the in-line documentation and argument types. The extracted individual components, classes, methods and functions, can then be used to construct workflows in a graphical editor. Unlike most other workflow systems, in DALiuGE application _and_ data components are represented as nodes on a workflow graph. Fundamentally, this concept enables the extreme scalability as well as the separation of I/O from the algorithms. Data components can reside in memory, even across a compute cluster. Application components can be as complex as full MPI applications or as small as a single line function call. Along the whole workflow design, scheduling and execution chain, DALiuGE is recording hash codes of components and data artefacts into a Merkle tree and enables complex comparisons of the equivalence of graphs, software components, data artefacts and complete execution runs. Workflows and component descriptions are stored in user configurable GitHub or GitLab repositories and are thus fully version controlled and can be shared with collaborators or the world. DALiuGE also supports workflows, containing sub-workflows. These sub-workflows can be scheduled and executed at run-time, either on the same platform as the main workflow or somewhere else. When using existing software packages, users don't need to write any at all and can fully concentrate on the workflow design. The parameterisation of existing, established graphs to run on different datasets or re-run with slightly changed configuration of the individual components has been streamlined into a single table interface for entire graphs, exposing pre-selected so-called graph configuration parameters.

Migrating Heterodyne Data Reduction to High-Performance Computing

Author: Christof Buchbender
Affiliation: I. Physikalisches Institute University of Cologne
ID: C604

The data volume produced by astronomical experiments continues to grow with each new generation of instrumentation. This is particularly true for heterodyne receivers transitioning from single-pixel to multi-pixel arrays (hundreds of pixels), as we are doing with the CCAT Heterodyne Array Instrument (CHAI) at the upcoming CCAT observatory. Previous-generation receivers, like GREAT aboard SOFIA, with 21 pixels, generated up to 50-70 GB of data per flight. While challenging, these data volumes could still be

reduced and analyzed on local computers using traditional methods, such as the GILDAS software from IRAM. However, CHAI is expected to produce a peak data rate of 8TB per day. This volume crosses the threshold where traditional single-computer pipelines are insufficient, necessitating a migration to an automated high-performance computing (HPC) environment.

CHAI is one of two instruments at the CCAT observatory. The other instrument, Prime-Cam, a modular receiver with up to seven science modules, will yield a similar data rate. To manage these large data volumes from the CCAT observatory, we are developing the CCAT Data Center at the University of Cologne.

In this presentation, I will discuss the limitations of our traditional in-house single-dish heterodyne data reduction pipelines, such as those used at SOFIA and the NANTEN2 telescope, and how these limitations hinder migration to a distributed, fully automated computational environment. I will also present our approach for the CCAT Data Center to overcome these challenges. Specifically, we are transitioning to a Python-based pipeline optimized for distributed computing and HPC environments where we aim to to use existing solutions where possible. By employing a central database to track data from planning through observation, data transfer, reduction, and analysis, and by using a workflow management system to orchestrate the data reduction process, we aim to minimize manual interaction and increase efficiency.

However, implementing these solutions is not without challenges. One significant challenge is that existing solutions from other groups often meet 90% of our needs on paper, but the specifics of our data formats and processing requirements often prevent easy integration or native use. My hope is that by sharing our experiences, we can foster discussions with other groups to make our solutions more general and to learn from our respective experiences.

The time-series visualization tool in ESASky

Author: Elena Puga

ID: C605

The ESASky web interface primary goal is to enable users to access European Space Agency's (ESA) space astronomy mission data. In addition to the searching and downloading functionalities, the ESASky user community has expressed an eagerness for visualization tools that aid in the data inspection process. They are meant to help scientists assess whether datasets are ultimately useful to their science case (e.g. source variability, exoplanet transit, transient events) to then download and analyse them with specific research domain tools.

To that end, we have developed a photometry and spectroscopy time series data viewer API within ESASky. The ESASky time-series viewer scientific use case is the capacity to, from a tabulated catalogue or spectra/imaging view within ESASky, aggregate multi-mission or multi-object time series data and interact with them for detailed inspection. In this talk, I will tease the ESASky time-series viewer and present the challenges that such tool entailed, as well as those that its future development, scaling and exploitation will likely bring.

J-PAS early data release: unique processing challenges of an imaging sky survey in 57 optical filters

Author: Héctor Vázquez Ramió
Affiliation: Centro de Estudios de Física del Cosmos de Aragón (CEFCA)
ID: C606

The Javalambre-Physics of the Accelerating Universe Astrophysical Survey (J-PAS; https://j-pas.org) is a multifilter photometric sky survey using a unique set of 56 narrow-band filters (145Å FWHM) and one broad-band, i, covering the visible range at a magnitude depth of AB~22. It is being carried out at the Observatorio Astrofísico de Javalambre (OAJ), Spain, since 2023. It aims to cover thousands of square degrees of the the observable sky from the OAJ and to determine precise photometric redshifts for around 1.3×10^8 galaxies. The survey is conducted with the Javalambre Panoramic Camera (JPCam) attached to the 2.5m-class telescope, Javalambre Survey Telescope (JST250). JPCam consists of a mosaic of 14 identical 9.2k x 9.2k CCDs which cover $3.4 deg^2$ in total on the sky. On top of each CCD, a different narrow-band filter is positioned through a set of 5 filter trays of 14 slots each.

The J-PAS survey strategy and the instrumentation employed pose important challenges in the data processing, some of them listed hereafter.

Large FoV systems are prone to suffer from stray light that is not properly corrected with the classical flat-fielding. The needed illumination correction is obtained from the 2D residuals of the photometric calibration performed taking as reference Gaia DR3. Other remarkable difficulty is dealing with the variety of background patterns observed through the set of J-PAS narrow-band filters together with other ambient illumination gradients under the presence of the Moon. Another significant challenge is the two-dimensional mixture of PSFs in the stacked images for each filter, due to the fact that the observations are often made on different dates and seeing conditions. This has an important impact on the proper estimation of the photometric aperture corrections on those combined frames. On the electronics side, a specific procedure has to be implemented in order to correct from a time-varying bias signal of the CCDs.

In this talk I will discuss these and other challenges and the solutions we have successfully implemented. I will also announce the upcoming public early data release of J-PAS, scheduled for the end of 2024.

PyKOALA, a multi-instrument tool for reducing IFS data

Author: Pablo Corcho-Caballero
Affiliation: Kapteyn Astronomical Institute, University of Groningen, Netherlands
ID: C607

Over the past two decades, the advent of Integral Field Spectroscopy (IFS) has revolutionized the field of astronomy, enabling comprehensive analysis of both the spatial and spectroscopic properties of extended objects. However, this technological leap has also introduced significant challenges in the reduction and processing of IFS data. The complex nature of IFS datasets requires meticulous correction and calibration steps, often demanding customized pipelines tailored to specific instruments.

To address these challenges, we present PyKOALA: a cutting-edge, open-source Python package designed to streamline the reduction of IFS data. Originally conceived as an expert pipeline component to complement the outputs of 2dfdr and enhance the data reduction process for the Kilofibre Optical AAT Lenslet Array (KOALA) Integral Field Unit (IFU), PyKOALA's vision has expanded over the past few years from a single-instrument focus to a versatile, multi-instrument framework. It now provides a modular and flexible framework, that allows astronomers to customize their reduction sequences and apply an arbitrary number of corrections across various IFS instruments. PyKOALA offers a streamlined interface that facilitates the ingestion of data from different IFUs, standardizing the fundamental properties of IFS data for consistent processing.

The first official release of PyKOALA is expected during early 2025, though its current beta version already features comprehensive documentation and a suite of Jupyter notebook tutorials to ease the learning curve. In this talk, I will showcase PyKOALA’s powerful capabilities, highlighting key features such as multi-instrument support and advanced correction modules, with examples from the initial results of the HI-KOALA IFS Dwarf Galaxy Survey (HI-KIDS).

Bridging the Gap: Enhancing Astronomical Data Analysis with Software Engineering Best Practices

Author: Shvetha Chynoweth
Affiliation: San Francisco State University
ID: C608

Extracting meaningful information from large amounts of astronomical data requires not only a clear understanding of how it was collected and the physics involved, but also a disciplined approach to the analysis. As an example, in my current research I work with EPRV (Extreme Precision Radial Velocity) data, which - like other areas of astronomy - requires disentangling instrumental and physical effects. Unfortunately, this process is often hindered by the absence of best practices which makes subsequent analyses and models inherit these complications, compromising their accuracy. As a part of my research, I apply my background in software engineering by building a standard framework for processing data that can be reused across research projects. In this talk, I will discuss recommendations for bridging the gap between software engineering and astronomical data analysis, how I'm applying these practices to my research, and methods for integrating these enhancements in other works.

Metadata, Semantics and Data Models Applicable to Data Formats

A data model to connect the ESO Data Processing System (EDPS) to ELT data archives

Author: Hugo Buddelmeijer
Affiliation: Kapteyn Astronomical Institute / University of Groningen
ID: C701

Context

The METIS and MICADO instruments for ESO's Extremely Large Telescope passed their Final Design Review mid 2024. That heralds an era of new data flow challenges for optical to mid-infrared astronomy. Reasons: (1) the instrumental complexity of the Adaptive Optics assisted ELT and its PSF Reconstruction, (2) nightly data rates up to several Terabytes, already during the laboratory testing phase. To address the laboratory testing phase challenge for METIS we will build an METIS AIT (Assembly Integration Testing) Archive database. The database contains the metadata of data items plus a pointer to the bulk FITS files. An interface between this database and the ESO Data Processing System would allow it to initiate processing laboratory data by querying the database. A similar approach could be adopted for MICADO.

Talk

In this talk we present a prototype AIT Archive populated with simulated data. The archive is based on a data model. This model contains formalised descriptions of data items, of data processing recipes and pipelines (see [Buddelmeijer et al 2020]). The database tables are generated automatically from this data model.

We specify an interface between the ESO Data Processing System (EDPS) and the database. The interface is currently one-way: processed data can be automatically ingested into the database. We outline the plan to make the interface two-way such that the EDPS can also automatically retrieve necessary data.

Similarly, we specify an interface between the EDPS and the simulator. This ensures **an end-user on the EDPS can restrict itself to specifying the desired end data product and leave it to the system to decide itself how to acquire the necessary input data**: **read from local disk**, **retrieved from the database**, **or generated through the simulator**.

We outline the plan for these two interfaces, present the progress towards achieving it, and discuss how they will allow this AIT setup to evolve into a platform to support ELT data flow during ELT science operations (i.e., a "ELT Research Data Platform").

Using Felis to Represent the Semantics and Metadata of Astronomical Data Catalogs
Author: Jeremy McCormick
Affiliation: SLAC/Rubin Observatory
ID: C702

Data catalogs are a fundamental part of modern astronomical research, allowing scientists to view, search, and filter data according to their requirements. Tabular data models described by SQL Data Definition Language (DDL) are a common way to represent such catalogs. However, DDL does not provide a way to describe the semantics of the data, such as the meaning of a data column, units of measurement, or the relationships between columns. The International Virtual Observatory Alliance (IVOA) has developed several standards in this area, including VOTable and Table Access Protocol (TAP), which are widely used within astronomy for representing such information.

The Data Engineering group of the Vera C. Rubin Observatory has developed a data description language and toolset, Felis, for defining the semantics of its Science Data Model schemas, which represent its public-facing data catalogs. Felis uses a rich Pydantic data model for describing and validating catalog metadata, represented as a human-readable and -editable YAML format. Felis provides a Python library and application for working with these data models. The metadata is used to populate the TAP_SCHEMA tables for the IVOA TAP services that power the table UI of the Rubin Science Platform (RSP). The toolset is also being used to assist in data migrations and will be utilized in testing the conformance of LSST data products to the data model. Felis's current capabilities will be discussed, as well as recent developments and future plans.

Real-time and Near Real-time Processing Pipelines

NEOCC's Aegis pipeline in asteroid orbit determination and impact monitoring.
Author: Francesco Gianotto
Affiliation: Starion for ESA
ID: C801

The Near-Earth Objects Coordination Centre (NEOCC) is the main component of the Planetary Defence Office (PDO) within ESA's Space Safety Programme (S2P). Its mission is to support and coordinate the observations of small Solar System bodies and assess and track the threats they may pose to Earth. Central to this mission is Aegis (1), an automated orbit determination and impact monitoring system developed by SpaceDyS s.r.l. under ESA contract and operated by NEOCC. Aegis plays a critical role in the daily operations of the NEOCC, providing up-to-date information on orbital properties, impact probabilities, and risk analysis for near-Earth objects.

Aegis operates on an hourly basis, continuously downloading new astrometric data from the Minor Planet Center. It is primarily based on two components: orbit determination and impact monitoring. The orbit determination component maintains a dynamic catalogue of near-Earth asteroids, which includes orbital parameters with associated uncertainties, physical properties such as visual and absolute magnitude, observation details, residuals, close approaches, and ephemerides. The impact monitoring component computes the impact probabilities of near-Earth asteroids over the next 100 years. Objects with non-zero impact probabilities are listed in the NEOCC Risk List (2). When an object's impact probability exceeds a certain threshold, Aegis also computes the associated impact corridor, further refining the risk assessment.

This presentation will focus on the Aegis processing pipeline, which leverages Docker services for the automated download and integration of observational data. Aegis begins by downloading observational data from the Minor Planet Center, after which it employs a dedicated weighting scheme to prioritize observations. This scheme considers the observatory, technology, and program codes associated with each observation. This allows us to apply different rules based on the historical relevance of each observatory. The system's ability to adjust the significance of data inputs ensures robust and accurate orbit determination.

A crucial output of this pipeline is the "Residual Weights Observations" (rwo) format, an ad-hoc data structure that encapsulates not only the astrometric information, but also detailed data used in the orbit computation, such as weights and observational residuals. This data is accessible through the NEOCC webportal and its APIs (3).

This presentation will also cover the Aegis processing pipeline for the impact monitoring component. This method relies on sampling the Line of Variation, a 1-dimensional differentiable curve in the orbital elements space that identifies the direction with the largest uncertainty. The output of this sampling is then propagated for 100 years from the current epoch, searching for close approaches and potential impacts with Earth. Both sets of information are published in the NEOCC webportal on a daily basis.

By leveraging real-time processing capabilities and advanced data analysis through the rwo format, Aegis provides an essential service in the continuous monitoring of NEOs, contributing to the safety and security of our planet. This presentation will delve into the technical aspects of the Aegis pipeline, emphasizing its real-time functionality and the use of the rwo format and weighting scheme in astronomical data analysis.

(1) https://ui.adsabs.harvard.edu/abs/2023sndd.confE..73F/abstract
(2) https://neo.ssa.esa.int/risk-list
(3) https://neo.ssa.esa.int/computer-access

Finding Fireballs in Lightning: A Daily Pipeline to Find Meteors in Weather Satellite Data
Author: Jeffrey Smith
Affiliation: SETI Institute
ID: C802

Weather satellite data contains a wealth of information well beyond its application to meteorology. The GOES weather satellite lightning mapper instruments detect millions of lightning strikes per day. Within these haystacks are a handful of bolides (exploding meteors). Through a combination of hard manual work, advanced machine learning techniques, statistical analysis and supercomputers, our multi-disciplinary team has succeeded in creating an efficient pipeline to identify the bolides. Our algorithms are also sensitive to other interesting phenomena in the data. Funded by NASA's Planetary Defense Coordination Office (PDCO), our goal is to create a rich, calibrated, and statistically consistent data set of bolide light curves to inform the planetary defense community of the risks associated with large asteroidal impacts. We utilize a three-stage detection pipeline, with successively more computationally expensive algorithms: 1) simple Hierarchical Clustering, 2) Random Forests and then 3) Convolutional Neural Networks. Detections are promptly published on a NASA hosted publicly available website, https://neo-bolide.ndc.nasa.gov. We present the evolution of our pipeline, the ML techniques utilized and how we continue to incorporate new information to improve detection performance.

Sub-arcsecond degree-scale imaging pipelines with LOFAR
Author: Jurjen de Jong
Affiliation: Leiden Observatory
ID: C803

In recent years, significant efforts have been made to automatically calibrate and image observations conducted with the Dutch high-band antennas from the Low Frequency Array (LOFAR) observing the universe at 150 MHz. These efforts have led to the LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al. 2017, 2019, 2022) and the LoTSS-deep fields (Kondapally et al. 2021; Duncan et al. 2021; Tasse et al. 2021; Sabater et al. 2021), providing wide-field images of the northern sky at 144 MHz and 6” resolutions. However, 90% of the radio sources at 6” remain unresolved at 144 MHz. This necessitates higher resolutions by using data from all of LOFAR’s international stations, extending the maximum baselines to about 2000 km and resulting in sub-arcsecond resolutions.

We now present a calibration and imaging pipeline capable of producing deep sub-arcsecond resolution images, achieving the highest sensitivities and highest resolutions at the lowest frequencies radio frequencies (Morabito et al. 2022; Sweijen et al. 2022; de Jong et al. 2024). Given the challenge of working with hundreds of terabytes of data, we are now focused on reducing computational costs such that we enable the possibility of a near real-time calibration and imaging pipeline. This advancement will allow for LoTSS-type surveys with data from all of LOFAR's international stations, facilitating further research of radio sources at 150 MHz and sub-arcsecond angular scales.

High-Performance Computing in Astronomy: Triumphs and Tribulations of Pipeline Processing on Supercomputers

Author: Kevin Vinsen
Affiliation: The International Centre for Radio Astronomy Research - UWA
ID: C804

The advent of supercomputing has the potential to revolutionise astronomical data processing and is essential for the analysis of massive Radio Astronomy datasets of the SKA era at unprecedented speeds. This talk presents our experiences implementing a complex data reduction pipeline on a number of state-of-the-art supercomputer facilities. We demonstrate how, when optimised, our pipeline achieves remarkable throughput, reducing processing times from weeks to mere hours for large-scale surveys.

However, the transition from traditional computing environments to supercomputing infrastructures is not without challenges. We discuss several critical issues encountered, including:

1) Scratch file management in distributed file systems like Lustre
2) File handle limitations in massively parallel operations
3) Scheduler conflicts and queue optimisation with measurement sets being very different in size depending on observation time and flagging
4) Wall time constraints and job segmentation strategies

We offer some solutions for astronomers looking to leverage high-performance computing resources and tools, such as DALiuGE, to mitigate many of these issues. Our findings highlight both the transformative potential and the practical considerations of supercomputing for modern astronomical research.

High-Performance Pipeline Processing for the Australian Square Kilometre Array Pathfinder

Author: Matthew Whiting
Affiliation: CSIRO
ID: C805

The Australian Square Kilometre Array Pathfinder (ASKAP) is a new-technology radio telescope operated by CSIRO at Inyarrimanha Ilgari Bundara, the CSIRO Murchison Radio-astronomy Observatory in the Western Australian outback. Its innovative receivers, with their wide field-of-view, generate very large data rates, necessitating high-performance computing to create the required calibrated images and catalogues, and deposit them in the CSIRO ASKAP Science Data Archive (CASDA) for use by astronomers.

The processing is orchestrated by the ASKAP pipeline, a scripted workflow that interfaces with the Slurm workload manager to run all necessary data preparation, calibration, imaging, and source-extraction tasks. The computationally-intensive processing is done using a custom-written imaging package called ASKAPsoft, specially designed to handle the scale of data produced by ASKAP. Crucially, the pipeline must run in near-real-time to keep up with the incoming data rate, allowing the telescope to efficiently survey the entire sky.

The ASKAP pipeline is operational, with regular survey observing resulting in large amounts of data (currently >3.8PB since full-surveys started late 2022) being made publicly available through CASDA. ASKAP processing is a demonstration of what can be possible through a large and complex nearly-autonomous supercomputing workflow, and provides important lessons for planning of even larger workflows anticipated for future instruments.

This talk will describe the design decisions that went into creating and scaling up the workflow, and describe how it has been set up to work on the supercomputers at the Pawsey Supercomputing Centre. This will include the range of different types of processing jobs and their contrasting requirements, the impact of the high I/O on overall processing efficiency, and lessons learned from both developing and running the pipeline at scale. We look ahead also to planned upgrades, as well as considerations for implementing processing for future facilities such as the SKA.

Enhancing Keck Observatory Operations: The Data Services Initiative's Journey

Author: Max Brodheim
Affiliation: W. M. Keck Observatory
ID: C806

In 2021, the W. M. Keck Observatory and the NASA Exoplanet Science Institute enhanced their collaboration through the Data Services Initiative aiming to produce and archive science-ready data. This endeavor focused on refining two pivotal aspects of observatory operations: the planning and execution of observations, and the processes of data reduction and archiving. To achieve these, the team developed new tools, including web-based GUIs, to capture observer intentions into a machine-readable format and an execution engine to carry out these observation plans. A revamped data handling architecture was also introduced to expedite the creation of reduced data products and their subsequent near-real time ingestion into the Keck Observatory Archive. As a result, raw and reduced data are now made available to users well under the requirement of 5 minutes, significantly enhancing the efficiency of data access. Additionally, the number of community-supported data reduction pipelines for Observatory instrumentation expanded from 2 to 9 (across 11 instruments), exceeding the initial project goals. Despite these advancements, the team navigated significant challenges, such as integrating automated workflows within a classically scheduled observing model and managing with limited personnel and budgetary resources. This presentation will explore the project journey, highlighting the successes achieved, obstacles faced, and lessons learned throughout the development cycle that can be applied to your future work.

Leveraging FPGAs as accelerators in real-time astronomical data-processing pipelines

Author: Mitchell Mickaliger
Affiliation: The University of Manchester
ID: C807

Given the big-data regime in which we work today, real-time processing is a strict requirement to get the most out of data while it's available. Over the years, there have been many advances in processing power, from multi-core CPUs to off-the-shelf GPUs. However, both of these examples are best suited for certain situations, given their hardware architectures. FPGAs (field-programmable gate arrays), on the other hand, have an open, almost undefined architecture, allowing the user to define it, through the creation of image files that are flashed to the device. As well as having a highly-configurable architecture, FPGAs also use much less power than other accelerator boards like GPUs, while still delivering very good performance (e.g. 38 TFLOPS for an Intel Agilex 7), making FPGAs a useful resource when power is a consideration, or to reduce cooling requirements in high-packing-density situations. While these devices are becoming more commonplace and off-the-shelf versions are available, there is still some effort required to successfully integrate these seamlessly into a pipeline. In this talk, I will describe our efforts to integrate FPGAs into our real-time pipeline for pulsar and fast transient searching for the SKA (named cheetah), and show the performance benefits we have gained from doing so.

Dynamic Imaging With MeerKAT: The Time Axis As The Final Frontier

Author: Oleg Smirnov
Affiliation: Rhodes University & South African Radio Astronomy Observatory
ID: C808

With the increased sensitivity and field of view of SKA pathfinders, dynamic radio imaging (that is, imaging the time axis) is becoming a burgeoning field, yielding rich new discoveries of transients and variable sources. MeerKAT is capable of reaching sub-150 uJy image rms in an 8s integration, which opens up studies of variability on much shorter timescales than was possible with previous radio interferometers. This also has important implications for interferometric SETI, since any potential technosignatures would be a subset of such transient events.

At the same time, imaging at such short timescales introduces its own substantial challenges. Instrumental effects that tend to average out in a traditional long synthesis observation can become limiting for dynamic imaging if not addressed correctly. I will discuss these challenges and present MeerKAT dynamic imaging of Jupiter's radiation belts, which have led to the serendipitous discovery of a pulsar-class object named the PARROT (pulsar with abnormal refraction recurring on odd timescales).

This work has led to the development of (and given the name to) a more general dynamic imaging pipeline, developed in collaboration with the Breakthrough Listen initiative. The PARROT pipeline is capable of detecting short-duration transients in imaging data, and yielding light curves and dynamic spectra for thousands of field sources en masse. We are already starting to use it to "mine" existing archival MeerKAT data, yielding a couple of new discoveries. The longer-term plan is to develop the PARROT pipeline to a state where it can be run in real-time, commensally with any MeerKAT imaging observation. This would open the door to transient event triggers -- something that has never been done with a radio interferometer before. With 6 years of observational data in the MeerKAT archive ready to be mined, and new observations arriving daily, this has the potential to turn MeerKAT into a transient and variability discovery machine, opening up new frontiers in astrophysics and SETI.

Other

Astronomy Data and Computing Services: Changing the way research software is developed, supported and maintained

Author: Gregory Brian Poole
Affiliation: Swinburne University of Technology
ID: C901

Astronomy Data and Computing Services (ADACS) was established in 2017 by the Australian Astronomy community - with the leadership of Astronomy Australia Ltd. (AAL) - to provide astronomy-focused training, support and expertise to maximise scientific return on Australia's investments in astronomical data & computing infrastructure. One of our flagship services is the ADACS Merit Allocation Program (MAP). This program provides an opportunity for any Australian astronomer to compete (in the same way they compete for computing or telescope resources) for the time of dedicated software professionals to deliver:

1. Bespoke training in skills related to software development, &/or

2. Dedicated software development or design effort for new or established codebases.

To date we have successfully delivered over 120 projects involving over 100 unique applicants across all areas of astronomy through this program. The ADACS MAP is solving problems for researchers that otherwise could not have been solved; accelerating old science and enabling new science. More broadly however, we are working to shift the culture of computing in our community; moving to a more collaborative model, where wider skill-sets of multifunctional teams can be exploited.

A Multi-Wavelength Data Viewer Realized through the Enhancement of hscMap

Author: Hiyo Toriumi
Affiliation: Shibaura Institute of Technology
ID: C902

hscMap serves as the imaging data viewer for the HSC-SSP (Hyper Suprime-Cam Subaru Strategic Program). The SSP data is obtained using five broad-band filters and several narrow-band filters. In hscMap, users can assign any of these filters to the RGB channels to display pseudo-color images. This study aims to enhance hscMap to create a more flexible and multifunctional viewer.

To achieve this enhancement, two primary functionalities were added. Firstly, support for HiPS (Hierarchical Progressive Surveys) data in FITS (Flexible Image Transport System) format allows data obtained by instruments other than HSC to be easily managed through hscMap's simple interface. Secondly, the ability to freely increase or decrease the data to be displayed (in the case of HSC-SSP, data per band) facilitates experimentation under diverse conditions. Additionally, automating part of the color synthesis process makes it easier to generate beautiful images.

These enhancements enable hscMap to flexibly synthesize and display various wavelength data, contributing to efficiency improvements in astronomical research and data analysis. The new features provided by this study facilitate the visualization of multi-wavelength data, serving as a valuable tool for researchers to make new discoveries.

Data processing and preservation for CTAO

Author: Karl Kosack
Affiliation: CEA Paris-Saclay and CTAO
ID: C903

We present the design and status of the data processing and data preservation software and infrastructure for CTAO, the next-generation very-high-energy (VHE) gamma-ray observatory under construction in Chile and the Canary Islands. The unprecedented size and complexity of CTAO and the need for open and reproducible data present new challenges for the storage and treatment of data products. We have developed open pipeline software to process petabytes of data generated by the observatory after on-site data volume reduction, including the treatment of tens of petabytes of simulated data required to model the instrumental response. This software produces reduced, standard data products usable by observers, in a format aimed to be compatible with other high-energy observatories. The data will be processed in parallel, data-driven workflows across six CTAO data centers (two on-site and four off-site in Europe). We will describe the big-data infrastructure we use for ensuring the preservation of the raw data and for executing the processing workflows. The technologies include Rucio, FTS, CVMFS, and significant contributions to expand the capabilities of the DIRAC middleware for complex workflows. We will also discuss our efforts to ensure both open software, standardized data models and formats, and FAIR data products. This includes also the ongoing efforts to create standard data models and formats for high-level data products that allow interoperability of CTAO data with existing X-ray, gamma-ray and neutrino observatories (i.e. GADF and VODF), as well as extending IVOA standards to meet the needs of this waveband for discoverability.

Asteroid Discovery with THOR on the Noirlab Source Catalog: An Engineering Perspective

Author: Nate Tellis
Affiliation: Asteroid Institute
ID: C904

In this presentation, we discuss the deployment of the THOR asteroid discovery algorithm on the NOIRLab Source Catalog Data Release 2 (NSC DR2). We will focus on the technical challenges encountered and the software, architecture and infrastructure used to address them. We discuss the data extraction and filtering challenges associated with the NSC DR2 dataset. We cover memory management and handling the computational demands of THOR. We will also cover the execution strategy using Dagster for orchestration, Kubernetes spot instances, result validation, and our open-source tools: quivr, adam_core, and mpcq. The talk will conclude with future work implications, notably in relation to the Vera C. Rubin Observatory's LSST.

FRELLED : An Astronomical Data Visualisation Package for Blender

Author: Rhys Taylor
Affiliation: Astronomical Institute of the Czech Academy of Sciences
ID: C905

I present FRELLED, the FITS Realtime Explorer of Low Latency in Every Dimension. This is a data visualisation package specifically designed for examining 3D FITS files, primarily (but not exclusively) intended for HI and higher-frequency radio data sets such as those from ALMA. It provides a number of different visualisation techniques to maximise the scientific returns from the data. Users can view their data volumetrically, as isosurfaces, or as a traditional series of 2D images with the option to use displacement maps, or even in virtual reality. The display can be rapidly toggled between different viewing methods. Multi-volume rendering is possible, both by overlaying two volumetric data sets directly, or plotting contours or isosurfaces from an unlimited number of data cubes over one volumetric display. FRELLED incorporates tools to allow rapid visual cataloguing of data sets of up to 1500^3 voxels, as well as performing basic analysis tasks : comparing their data with the SDSS, querying NED, plotting integrated flux and velocity maps, and measuring the spectral features of the data. Users can create different virtual objects to mask and catalogue data, interactively rescaling and positioning them and using different colours to indicate different object types. These objects can be exported either to produce simple catalogues or for direct processing of the data. In an era where automatic techniques are increasingly dominant, I will demonstrate that the correct cools can still ensure that visual examination can still play an important role even in cataloguing large data sets.

Posters

Proposal and Observation Preparation Tools

ACROSS: Enabling Time Domain and Multimessenger Astrophysics
Author: Brian Humensky
Affiliation: NASA Goddard Space Flight Center
ID: P101

The U.S. Astro2020 Decadal Survey recommended an investment in Time Domain and Multi-Messenger Astrophysics (TDAMM) as the top-priority sustaining activity in space for the coming decade. One aspect of NASA's response to this recommendation is a pilot project, the Astrophysics Cross-Observatory Science Support (ACROSS) Initiative, designed to provide support to both missions and observers as they pursue TDAMM science. In this talk, we present our observations of needs in the community and initial plans for ACROSS activities, including services to facilitate and improve cross-mission follow-up planning and execution; a multi-messenger web portal with links to existing mission resources, community tools, and information targeted for TDAMM General Observers; development of "Smart Target of Opportunity submission page" proof-of-concepts; and ongoing development of a potential TDAMM general observing competitive grant solicitation. While the initial focus has been to enhance coordination between NASA missions, we are eager to work with ground-based and international partners as well. We invite discussion with both missions and observers to better understand their needs and concerns as ACROSS progresses. Here we present our efforts on the web-portal and API, along with our development to support NASA's BurstCube mission.

Simulation Tools to Help Prepare for Science with the Roman Space Telescope
Author: Charles-Philippe Lajoie
Affiliation: Space Telescope Science Institute
ID: P102

The Nancy Grace Roman Space Telescope (Roman) is set to launch in late 2026 and will survey large regions of the sky to help answer key questions in many areas of astrophysics, including dark energy and dark matter, galaxy formation, and exoplanets. In an effort to provide the community with tools to better understand and prepare for the large amounts of data from Roman’s Wide Field Instrument (WFI), the Science Operations Center (SOC) at the Space Telescope Science Institute has developed a suite of simulations tools. These tools range from Python notebooks with graphical interface to more advanced, full-scale simulations of WFI observations that include known instrumental effects. The different tools enable users to model with high fidelity the WFI point-spread function, perform exposure time and signal-to-noise calculations, as well as perform image simulations of complex astronomical scenes. These tools also include comprehensive user documentation and tutorials that allow users to make the most out of the tools and be ready for Roman's groundbreaking science. Here, we give an overview of the various SOC simulation tools currently available to the astronomical community.

Polaris: a new open source proposal tool

Author: Darren Walker
Affiliation: University of Manchester, Jodrell Bank Observatory
ID: P103

We are developing a new, open source proposal tool called Polaris as a deliverable of the Horizon 2020 OPTICON RadioNet Pilot project

This tool has been built on top of the Proposal Data Model, defined in the IVOA standard VO-DML. Although this is an important part of the tool, here we will be specifically focussing on the technologies used build and develop Polaris itself. We strove to keep the code as lean as possible by utilising code generation and by using existing libraries and frameworks. The functionality of the tool is exposed as a RESTful API in a micro-services architecture deployed on Kubernetes.

We have developed a web-based GUI frontend to access our API. It has been written in Typescript using the React framework, specifically leveraging the Mantine library. We are also developing a command line interface in Java to access the API. This is not yet as fully fledged as the GUI, but we are looking to add more functionality to it in the future.

The API, GUI and CLI are maintained in open-source repositories on GitHub (https://github.com/orppst). This allows others to contribute to the development of Polaris or fork the projects to change the tool to suit their mode of operation.

Scheduling the world's largest observatory for gamma-ray astronomy

Author: David Soldevila
Affiliation: ICE-CSIC
ID: P104

The Cherenkov Telescope Array Observatory is the world's largest and most powerful ground-based observatory for gamma-ray astronomy at very-high energies. [...] The main challenge of the scheduler is to generate the optimal schedule given a series of approved scientific proposals and the resources to observe them with, namely, up to 8 sub-arrays (logic group of telescopes) per hemisphere.

SPOT: Site Preparation and Observation Tool

Author: Eric Jeschke, Russell Kackley, Takeshi Inagaki, Jonathan Merchant, Erin Dailey
Affiliation: Subaru Telescope, National Astronomical Observatory of Japan
ID: P105

We present a graphical open-source tool for planning and assisting astronomical observations called SPOT (Site Planning and Observation Tool). SPOT is based on various open-source packages including Astropy, Skyfield, Astroquery and Ginga. Here are few of the things it can do:

- You can select a site and date/time when you plan to observe (it's also easy to add your own custom site).
- It can show you an astronomical almanac of information about a particular date (sunrise, sunset, moonrise, moonset, twilights, etc).
- It can load lists of targets and plot them on a polar plot for their position in the sky at the current time or any given time.
- It can show you the various targets' visibility as a plot of altitude vs. time.
- It can overlay fisheye-type sky camera images on the polar plot so that you can monitor for cloud coverage.
- With the right customization it can show you where your telescope target is, the current telescope position and the slew that it will take to get there on the polar plot.
- It can look up catalog images from various sources for a given target and show instrument detector overlays on top, with adjustable position angle.

Because SPOT is written in pure Python, it can run on any platform that has the required Python modules. In this paper we will describe the capabilities of SPOT and how to install, run and access documentation.

The ScopeSim ecosystem

Author: Fabian Haberhauer
Affiliation: University of Vienna
ID: P106

ScopeSim is a general-purpose observation data simulation ecosystem for astronomical instruments. It allows users to simulate observations with multiple instruments for the same (often custom built) target description using a common software platform, thus enabling "apples-to-apples" comparisons of the outputs. The simulation engine has been described in a previous proceedings paper,1 however behind the scenes a vast infrastructure has been built to support the ScopeSim engine. The supporting elements are in some cases major projects in their own right, with multiple additional use cases and user groups. For example, the Instrument Reference Database (IRDB) provides a public and open-source platform for instrument consortia to distribute a coherent picture of the optical properties and characteristics of their instrument(s).

Continuation Requests in the 'Hedwig' Proposal System
Author: Graham Bell
Affiliation: East Asian Observatory
ID: P107

'Hedwig' is an observing proposal handling system developed by the East Asian Observatory for use at the JCMT. It has been used successfully for the last 9 years (semesters 16A - 24B). For the most recent call for proposals, the system was extended to offer the option to submit 'continuation requests' as an alternative to full proposals.

Continuation requests allow the authors of a successful proposal to request an extension of their time allocation, for example to continue their project in the new semester if the observations have not been completed. The aim is to allow such requests to be processed more efficiently, with a simplified proposal structure and abbreviated review process. Continuation requests are then presented to the time allocation committee alongside regular proposals to allow them to allocate time as normal.

ESO's New Generation of Exposure Time Calculators
Author: Jakob Vinther
Affiliation: European Southern Observatory ESO
ID: P108

In recent years, ESO has initiated a project modernising its Exposure Time Calculators (ETC) to ETC 2.0, with a Python backend and a dynamic interactive web frontend, as well as a programmatic interface.

ETC 2.0 has been deployed for a range of new instruments at Paranal and La Silla observatories, including CRIRES, ERIS, FORS2, HARPS/NIRPS, and 4MOST. The development of the MOONS, UVES, and MUSE ETC 2.0 are currently underway.

The ETC 2.0 configurations are obtained from Instrument Packages to ensure compliance and seamless integration with the Phase 2 observation preparation tools. The Angular web frontend uses modern store technologies to efficiently manage, preserve, and share state across sessions.

CPGS: a tool for preparing observations with the Coronagraph Instrument on the Roman Space Telescope
Author: James Ingalls
Affiliation: Caltech
ID: P109

The Coronagraph Instrument (CGI) on board the Roman Space Telescope (set to launch in Fall 2026) will be the first in-flight coronagraph with active wavefront sensing and control. In this poster we describe Command Product Generation Software (CPGS), a web-based tool hosted at the Roman Science Support Center at Caltech/IPAC to plan and prepare CGI observation inputs to the Roman Planning and Scheduling Subsystem. The CPGS interface breaks up into 9 main sections, each of which takes input that governs a different aspect of a CGI observation; the poster is primarily made up of a "tour" through each of the sections. On the output side, the observation specifications created with CPGS contain all the information needed to schedule a CGI observing campaign, using XML to encode control variables and arguments to command scripts in Roman observatory visit files. Because the Roman ground system is distributed among several institutions, the creation of flight products that execute pre-scheduled commands on CGI occurs via a series of information exchanges between partner organizations, with CPGS as the first step in the process. We give an overview of the infrastructure to support this chain of exchanges from observer to flight hardware, and describe how CPGS accesses information from further down the chain. We also note that, as CPGS is the furthest link in the software chain from the observer to flight hardware, it has been designed to be the most flexible to changes in implementation.

Usage and Management of TRDS files for JWST and HST
Author: Meaghan McDonald
Affiliation: Space Telescope Science Institute
ID: P110

The Tools Reference Data System (TRDS) supports the management of reference data used with tools that facilitate the planning and analysis of JWST and HST observations. These tools rely heavily on data that combine the light transmission of the telescopes with models or astronomical objects in order to make predictions about observations. The Reference Data for Calibration and Tools (ReDCaT) Team is responsible for the documentation and validation of mission specific data, as well as community derived astronomical catalogs, to ensure successful observations and science results from general observers.

Delphi-Crew: An innovative methodology for proposals evaluation
Author: Miguel Doctor Yuste
Affiliation: Telespazio UK for ESA
ID: P111

Double-blind peer review has been the de facto standard technique for evaluating scientific proposals since its inception in the 18th century. However, with the significant rise in the number of scientists submitting proposals to compete for observing time on a wide variety of scientific missions, and the existing challenges in identifying experts with the necessary skills

and availability to effectively evaluate these proposals, there is a pressing need to consider new methodologies and techniques for more effective evaluation.

In an attempt to address this situation, we present Delphi-Crew, an innovative methodology for evaluating scientific proposals that combines the well-structured workflow and flexibility of the Delphi method—a strategic consulting technique used to obtain long-term forecasts from groups of experts—with the advanced text-processing capabilities of generative AI multi-agent systems. First, we provide a brief description of the Delphi method, along with the adjustments made to adapt it for proposal evaluation. Next, we present a proof-of-concept implementation using CrewAI, a cutting-edge framework for orchestrating role-playing autonomous AI agents. Finally, we discuss the challenges and limitations of our approach that must be overcome to apply this technique to real-world cases.

Pandeia Exposure Time Calculator: A Cloud based Multi-mission Wizard to L2 and Beyond!

Author: Oi In Tam
Affiliation: Space Telescope Science Institute
ID: P112

Space Telescope Science Institute (STScI) provides Exposure Time Calculators (ETC) for James Webb Space Telescope (JWST) and Roman Space Telescope's Wide Field Instrument (WFI).

It provides many useful features for the principle investigators and co-investigators to help on their proposals and collaborations since 2017 for JWST and September 2024 for Roman.

The ETC offers unique features such as sets of example and science workbooks and the ability to share and organize workbooks with calculations for principle investigators and co-investigators.

Due to high demand for JWST (Cycle 4 deadline: 3036 calculation per hour at peak), the ETC utilizes AWS Cloud architecture to scale in and out the resources based on the usage.

Both JWST and Roman ETCs are built on the Pandeia engine, which is a data driven python code based software that is designed for multi-mission support.

The Roman Telescope Proposal System

Author: Patrick Lowrance
Affiliation: Caltech/IPAC
ID: P113

The Nancy Grace Roman Space Telescope’s Wide Field Instrument’s observing program will include General Astrophysics Surveys (GAS), for which a minimum of 25% of Roman’s observing time will be reserved in the first five years and which will primarily be selected via traditional peer-reviewed calls for proposals. The Roman Telescope Proposal System (RTPS) is currently being designed for proposal submission and panel review for the GAS. This poster will highlight the current infrastructure built to use a web-based single-stage proposal form, allow for multiple re-submissions of proposals, and notify users of a successful submission. The current design also supports the panel review process, flexible panel creation and management, and user-friendly proposal review, all with Dual-Anonymous Peer Review (DAPR).

The Proposal Data Model as the basis for POLARIS a Proposal Submission Toolkit

Author: Paul Harrison
Affiliation: University of Manchester
ID: P114

Polaris is an open source proposal toolkit, developed as a deliverable of the Horizon 2020 OPTICON RadioNet Pilot project

This toolkit has been built on top of the Proposal Data Model (https://github.com/ivoa/ProposalDM), defined in the IVOA standard VO-DML. The model design will be discussed as although the Polaris toolkit uses the model internally, the Proposal Data Model can also be used as an interchange format for other proposal preparation systems.

The Polaris API, GUI and CLI are maintained in open-source repositories on GitHub (https://github.com/orppst). The development of the Polaris tool itself is described in a poster.

Open Time Proposal Review System for the MeerKAT Radio Telescope

Author: Rohan Schwartz
Affiliation: SARAO
ID: P115

Through periodic Calls for Proposals, the South African Radio Astronomy Observatory (SARAO) allocates time on the MeerKAT Radio Telescope to the international community to maximise the scientific impact through radio astronomy; while contributing to South African scientific leadership and human capital development.

Proposals are submitted through the proposal submission system, followed by a stringent review process where they are graded based on specific criteria. These proposal reviews are managed through the proposal review tool, after which, time on the telescope is awarded based on the grade achieved during the review process.

Managing target of opportunity (ToO) observations at Observatorio Astrofísico de Javalambre (OAJ)

Author: Tamara Civera
Affiliation: CEFCA (Centro de Estudios de Física del Cosmos de Aragón)
ID: P116

The Observatorio Astrofísico de Javalambre (OAJ) is a Spanish astronomical ICTS (Unique Scientific and Technical Infrastructures) located at the Sierra de Javalambre in Teruel (Spain). It has been particularly conceived for carrying out large-sky multi-filter surveys and consists of two main telescopes of large field of view (FoV): JST250, a 2.5m with 3.4deg² FoV and JAST80 with 2deg² FoV. The Unit for Data Processing and Archiving (UPAD), as the needed data center to deal with the images from both telescopes, is considered an essential infrastructure of the ICTS too.

As an ICTS, the OAJ is committed to offer at least 20% of the observing time. It actually offers more than 25% of Open Time to the astronomical community through Legacy Surveys, Regular Programs (RP) and Director discretionary time (DDT). Regarding the RP, a new call for proposals is made public each semester accepting only proposals under the modality of Target of Opportunity (ToO). These projects are defined as proposals for which the target and/or observation epoch is not known at proposal submission. Moreover, depending on the scientific requirements and merit they could require a fast response after a trigger and they could be granted with "override" status over other existing proposals. Another aspect is that they may need a quick access to the data products and, in those cases, these are provided in a matter of minutes after the observations are made.

This contribution summarizes how ToOs have been integrated into OAJ operations presenting the different applications designed and implemented at the observatory to deal with ToOs and the data flow between them focusing on the proposal preparation portal and the tools for triggering the observations.

Optimizing the Scheduling of the Cerro Chajnantor Atacama Telescope (CCAT) Surveys

Author: Yoko Okada
Affiliation: I. Physikalisches Institut, Universität zu Köln
ID: P117

In this poster contribution, we present our strategies for optimizing the scheduling of surveys at the Cerro Chajnantor Atacama Telescope (CCAT) Observatory. Scheduling observations is required at every observatory and is therefore not a new topic. However, each observatory has unique requirements and challenges. The CCAT Observatory will operate the Fred Young Submillimeter Telescope (FYST), the 6 m diameter, high surface accuracy telescope at 5600 m altitude in the Atacama Desert, Chile. The first generation instruments (the heterodyne receiver CHAI and the direct detection instrument Prime-Cam) have their own observing campaign periods, targeting submillimeter frequencies up to 850 GHz or even potentially up to 1.5 THz. The main scientific objectives range from the Cosmic Microwave Background (CMB) and the line-intensity mapping in the Epoch of Reionization (EoR) to magnetic fields, dust properties, and cycling of matter in the interstellar medium in the Milky Way. The observing plans are a mixture of large area long-term surveys and deep small field observations on a collaborative basis in the consortium. Under these conditions, our main optimization goals are to have the highest signal-to-noise data set for a given weather condition (mainly precipitable water vapor), and to obtain reasonable sets of early science observations in the short term for various science goals in the consortium. We decided to use a simple sequential scheduling with the scoring algorithm based on the astroplan package, and are optimizing the relative weights of different constraints. We will also show how the scheduler will be integrated into the observatory operation system.

The Rise of AI for Science and Data Center Operations

Machine-Learning Workflow for Morphology Classification of Galaxies
Author: Bernd Doser
Affiliation: Heidelberg Institute for Theoretical Studies
ID: P201

As part of the EU project SPACE Center of Excellence, the seven most commonly used astrophysical simulation codes are being improved to prepare them for exascale computing. This will result in a large amount of data (several petabytes) that will soon be waiting to be analyzed. The potential insights and discoveries these data sets could reveal are fascinating. Spherinator enables the reduction of these extensive data sets to a low-dimensional space using an unbiased artificial neural network with a variational autoencoder and the learning of the morphology structure of the galaxies. The latent space of the variational autoencoder represents a morphological classification of the galaxies of the entire data set ordered by their similarity. Using a spherical latent space allows us to use the Aladin software to visualize Hierarchical Progressive Surveys (HiPS) produced with the HiPSter module.

This work presents a machine-learning workflow covering all stages, from data collection to preprocessing, training, prediction, and final deployment. This workflow ensures full reproducibility by keeping track of the code, data, and environment. Additionally, the workflow allows for scalability in managing a large amount of data and complex pipelines. We use only open software and standards that align with the FAIR principles, which allow us to distribute the workflow to enable collaborations between teams by sharing code, data, and results.

Machine Learning Mismatchings and Catalogues Creation: A Path to Finding the Milky Way Galaxies-Analogues

Author: Daria Dobrycheva
Affiliation: Main Astronomical Observatory of the NAS of Ukraine
ID: P202

Our work aims to study the evolutionary properties of galaxies at redshifts z<0.1. We have created catalogs of galaxies at z<0.1 with specific morphological features, alongside estimations of the basic evolutionary parameters of the galaxies contained within these catalogs (Vavilova+, 2022). We developed this catalog using photometry- and image-based machine learning for classification. This catalog provides information on 32 features, such as bar, ring, merger, dust lane, edge-on, and irregular etc. It also includes attributes related to CNN accuracy in identifying these features. To ensure higher CNN accuracy in estimating the evolutionary properties of galaxies, we conducted a thorough visual inspection of the galaxy catalogs to remove misclassified objects. This meticulous process enhances the reliability and precision of our evolutionary analyses. We present the results of this visual inspection and identify the leading machine learning errors uncovered during the review.

By creating catalogs of galaxies with detailed internal features, we provide a foundation for more accurate searches for the Milky Way Galaxies-Analogues. These catalogs, containing information on key characteristics of galaxies, are a basis for our studies of structures similar to our galaxy. This approach enables a deeper understanding of the Universe's large-scale structures and their evolutionary properties through machine learning techniques.

Open Science and Artificial Intelligence for supporting the sustainability of the SRC Network: The espSRC case.

Author: Julián Garrido Sánchez
Affiliation: Instituto de Astrofísica de Andalucía (CSIC)
ID: P203

One of the major challenges in Radio astronomy is to build the SKA Observatory (SKAO, an ESFRI landmark), which aims to answer fundamental questions in astrophysics, fundamental physics and astrobiology. A network of SKA Regional Centres (SRCNet) will be established to process the scientific data products generated by the SKAO, bringing them to the final state required for scientific analysis. The SKAO will deliver approximately 700 PB/year of calibrated data to the SRCs. The SRCNet will provide science support and the resources for scientific exploitation. This includes access to SKA science data products via a portal, irrespective of the geographical location of the user, as well as to the analysis tools and the processing power necessary to fully exploit the science potential of these products.

In this talk, we will explain how the TED4SKA project, led by the Spanish SRC, aims to contribute to creating a sustainable SRC network. Its main goals are: a) minimising the environmental footprint of SRCs; b) incorporating a green perspective into the platforms for

SKA data storage, analysis and visualisation within the SRC Network prototypes; and c) developing Open Science technologies supporting end-to-end research reproducibility.

We will also discuss the deployment of a network of sensors in the IAA (Instituto de Astrofísica de Andalucía) computing centre, which is designed to gather detailed environmental and operational data in real-time. These sensors will facilitate the use of advanced Artificial Intelligence and Machine Learning (AI/ML) techniques to explore solutions for mitigating the environmental impact and improving Data Centre operations. By analyzing the data collected, these AI/ML tools will not only help identify areas where energy efficiency can be improved but will also enable predictive maintenance and provide a deeper understanding of the facilities.

Large Language Models: New Opportunities for Access to Science?

Author: Jutta Schnabel
Affiliation: ECAP, FAU Erlangen-Nürnberg
ID: P204

The adaptation of Large Language Models (LLMs) like ChatGPT for information retrieval from scientific data, software and publications is offering new opportunities to simplify access to and understanding of science for persons from all levels of expertise. They can become tools to both enhance the usability of the open science environment we are building as well as help to provide systematic insight to a long-built corpus of scientific publications.

In this talk, the uptake of Retrieval Augmented Generation (RAG)-enhanced chat applications in the construction of the open science environment of the KM3NeT neutrino detectors will be used as a focus point to explore and exemplify prospects for the wider application of LLMs for our science.

Maximising scientific return with NLP & ESA Datalabs at the Data Science and Archives division at ESAC

Author: Marcos López-Caniego
Affiliation: Aurora Technology for ESA
ID: P205

The ESA Science Directorate is committed to advancing its mission by leveraging data science and AI, in line with the organization's broader digital transformation strategy.

The Data Science and Archives division at ESAC contributes to this strategy supporting the ESA astronomy, planetary and heliophysics missions and their communities with new tools and services to facilitate scientific analyses of ESA’s archival data.

An essential aspect of this strategy is the ability for scientists to access space mission data directly through interfaces that support in-depth scientific analysis without the need to transfer data to local devices. With this purpose in mind, the ESA Datalabs Science platform was created. This platform extends mission archives with interactive analysis and collaboration functionality.

The Data Science team at ESAC is also looking to support various machine-learning use-cases and add corresponding hardware and frameworks to the science platform. This includes several projects dealing with Natural Language Processing techniques that use Large Language Models such as investigating the use of ChatGPT in scientific publications in astronomy, building a chatbot to explore the user manual's from XMM-Newton archive using open source LLM's enhanced with Retrieval Augmented Generation (RAG) techniques or using open source LLM's+RAG to infer object type information combining CDS' SIMBAD's object data base and ESASky's chatbot.

ML Methods For Space Debris Detection in Bistatic Radar Data

Author: Miguel Zammit
Affiliation: Institute of Space Sciences and Astronomy
ID: P206

The exponential increase of satellite launches in the past few decades has led to a significantly large population of debris objects in Earth's orbit, particularly in the Low Earth Orbit (LEO) regime. Their subsequent detection and monitoring have thus become ever more pertinent, with facilities such as the BIRALES space debris radar regularly used for detecting new NORAD objects and tracking confirmed bodies in the LEO regime. The current detection pipeline installed at BIRALES employs a Multi-beam streak detection strategy (MSDS) algorithm to identify and segment radar streaks in spectrogram data indicative of debris objects. However, the relatively slow computation time will pose a problem once BIRALES is scaled to a larger array. We aim to leverage the power of ML to develop a streak detection and segmentation model able to improve on the performance seen by the MSDS algorithm, both in terms of the recall and false positive rates, as well as being able to maintain a level of solid performance at lower SNR levels. These preliminary tests aim to fully constrain the prospective model's performance by testing a number of different architectures across different SNR levels, beam locations and detection bandwidths. These results serve as a springboard for future work in training a deployable model to rival and replace the MSDS detector.

Overcoming machine learning training data imbalance by simulating exoplanet transits

Author: Nika Gorchakova
Affiliation: Dublin City University
ID: P207

Simulations of exoplanet transits have the capacity to improve outcomes when training Machine Learning models. Machine learning has huge potential in exoplanet detection but faces challenges due to data imbalance and lack of ground truth in observational data.

Most stars do not show transits, leading to datasets being skewed towards non-transit light curves, which can result in overfitting and poor recall. Furthermore, the absence of ground truth complicates understanding the effects of noise and errors on detection outcomes.

To address these issues, we simulate exoplanet transits using key astrophysical parameters and diverse noise profiles to create balanced training datasets. This simulation-based approach will improve machine learning models, enhancing their outcomes in detecting exoplanets in real-world data.

Leveraging Deep Learning for Efficient Galactic Characterization

Author: Omar Anwar
Affiliation: The University of Western Australia
ID: P208

The rise of AI in scientific research and data centre operations is transforming how we process and analyse astronomical data. This talk presents a proof-of-concept study utilising deep learning to estimate the age, metallicity, and dust content of galaxies, significantly enhancing the efficiency of processing these critical parameters. Leveraging the MILES synthetic dataset, our approach begins with 636 spectral templates and expands the dataset to 111,936 diverse templates by incorporating varying extinction levels and creating combination templates from up to five different MILES spectra with random weights. These templates are further modified to account for different types of noise, ensuring robust and comprehensive model training over the spectral range from 4749.37 Å to 7100.49 Å with a resolution of 2.5 Å.

The model architecture employs four parallel attention-based encoders with varying kernel sizes to capture different spectral features. Model generates a 12 × 53 grid of metallicity and age predictions, and a single value of dust for the input spectrum. The model demonstrates a mean squared error (MSE) of 0.27% with a standard deviation of 0.10% between the input spectra and the model generated spectra for the test set. To validate the model's performance on real data, it was tested on two galaxies from the PHANGS survey, NGC4254 and NGC5068, captured with MUSE. The average spectral MSEs for these two galaxies are 0.394% and 0.389%, with standard deviations of 0.337% and 0.282%, respectively. The pPXF-generated maps of age, metallicity, and dust for both galaxies serve as a reference for comparison, and show strong correlation with GalProTE generated maps. Notably, while pPXF takes around 11 seconds per spectral fit, proposed model achieves this in 8 milliseconds. This approach not only accelerates analysis but also maintains a reasonable accuracy, offering a significant advancement for efficient galactic characterisation and astronomical studies.

From data to scientific breakthroughs using tools powered by Generative Deep Learning
Author: Sebastian Trujillo-Gomez
Affiliation: Heidelberg Institute for Theoretical Studies (HITS)
ID: P209

Current applications of machine learning to astrophysics focus on teaching machines to perform domain-expert tasks accurately and efficiently across enormous datasets. Although essential in the big data era, this approach is limited by our own intuitions and expectations, and provides at most only answers to the 'known unknowns'. To address this, we are developing a new conceptual framework and tools to help astronomers maximize scientific breakthroughs by letting the machine learn unbiased interpretable representations of complex data ranging from observational surveys to simulations. Our tools automatically learn low-dimensional representations of complex objects such as galaxies in multimodal data (e.g. images, spectra, datacubes, simulated point clouds, etc.), and provide interactive explorative access to arbitrarily large datasets using a simple graphical interface. Our framework is designed to be interpretable, work seamlessly across datasets regardless of their origin, and provide a path towards discovering the 'unknown unknowns'.

To clean or not to clean? Influence of pixel removal on event reconstruction using deep learning in CTAO
Author: Tom François
Affiliation: LAPP, Univ. Savoie Mont-Blanc, CNRS, Annecy, France
ID: P210

The Cherenkov Telescope Array Observatory (CTAO) is the next generation of ground-based observatories employing the imaging air Cherenkov technique for the study of very high energy gamma rays.

The software Gammalearn proposes to apply Deep Learning as a part of the CTAO data analysis to reconstruct event parameters directly from images captured by the telescopes with minimal pre-processing to maximize the information conserved.

In CTAO, the data analysis will involve a data volume reduction that will definitely remove pixels. This step is necessary for data transfer and storage but could also involve information loss that could be used by sensitive algorithms such as neural networks (NN).

In this work, we evaluate the performance of the gamma-PhysNet NN when applying different cleaning masks on the input images from simulated as well as real data from the first Large-Sized Telescope.

This study is critical to assess the impact of previous steps in the data processing, mainly motivated by data compression.

Vision-Language Models for Spiral Galaxy Identification in SDSS: A Path to Finding Milky Way Analog Galaxies

Author: Vladyslav Khramtsov
Affiliation: V. N. Karazin Kharkiv National University, Kharkiv, Ukraine
ID: P211

In this study, we explore the application of Vision-Language Models for spiral galaxy identification within the Sloan Digital Sky Survey (SDSS) at redshifts $z < 0.1$. Specifically, we leverage the CLIP model (ViT-B-32_laion2b_s34b_b79k) to classify ~170k galaxies into two categories: spiral galaxies and others (which include round, rounded in-between, cigar-shaped, and edge-on galaxies). Ground truth labels are derived from the Galaxy Zoo 2 dataset, offering reliable citizen-science-driven classifications. Using CLIP, we generate embeddings of RGB images constructed from the g, r, and i photometry bands of the SDSS. These embeddings are then employed to train a logistic regression classifier for binary classification, using 67% of the data for training. Our approach yields an F1 score of 89.4% on test dataset, demonstrating strong performance with minimal training requirements. The simplicity of the method underscores its potential for efficient large-scale galaxy classification. Ultimately, our goal is to refine the identification of spiral galaxies to facilitate the search for Milky Way analogs.

Enhanced Exoplanet Detection Using Gradient Boosting Decision Tree Model with Multi-Characteristic Representations of Light Signals

Author: Yutao He
Affiliation: University of California at Los Angeles (UCLA)
ID: P212

The detection of exoplanets, especially those located in habitable zones, has become a significant focus of astronomical research. Machine learning has played a critical role in advancing exoplanet detection efforts. This paper introduces a novel approach for exoplanet detection using a stacked Gradient Boosting Decision Tree (GBDT) model that learns from multi-characteristic representations of light signals. Our method integrates key transit characteristics from the time domain, such as depth, duration, period, and shape, with frequency components that may indicate the presence of exoplanets, capturing a comprehensive view of transit signals.

We apply careful data preprocessing techniques to mitigate systematic errors, instrumental noise, and stellar variability, ensuring the extracted features accurately reflect the underlying transit signals. The GBDT model is trained on these features with hyperparameters optimization and rigorously tested on a separate data set of Kepler light curves, based on performance metrics that provide a detailed evaluation of the model's accuracy in identifying potential exoplanet transits.

The results of our research show that the stacked GBDT machine learning model, combined with multi-characteristic representations, offers a robust and efficient approach for exoplanet detection.

Photometric redshift estimation of galaxies by machine learning

Author: Changhua Li
Affiliation: National Astronomical Observatories, Chinese Academy of Sciences
ID: P213

The accurate estimation of photometric redshifts plays a crucial role in accomplishing science objectives of the large survey projects. By utilizing machine learning, we can swiftly obtain redshift for large-scale galaxies. Based on the training set obtained by cross-correlating the BASS DR3 QSO candidate catalogue, DESI Legacy Imaging Surveys DR9 galaxy catalogue and many spectroscopic galaxy catalogue, serval photometric redshift estimators with different machine learning methods are trained. Last, we used the optimal model to predict the redshifts of galaxies in the DESI Legacy Image Survey.

Stereograph: stereoscopic event reconstruction using graph neural networks applied to CTAO

Author: Thomas Vuillaume
Affiliation: LAPP, Univ. Savoie Mont-Blanc, CNRS
ID: P214

The CTAO (Cherenkov Telescope Array Observatory) is an international observatory currently under construction. With more than sixty telescopes, it will eventually be the largest and most sensitive ground-based gamma-ray observatory.

CTAO studies the high-energy universe by observing gamma rays emitted by violent phenomena (supernovae, black hole environments, etc.). These gamma rays produce an atmospheric shower upon entering the atmosphere, which emits faint blue light, observed by CTAO's highly sensitive cameras. The event reconstruction consists of analyzing the images produced by the telescopes to retrieve the physical properties of the incident particle (mainly direction, energy, and type).

A standard method for performing this reconstruction consists of combining traditional image parameter calculations with machine learning algorithms, such as random forests, to estimate the particle's energy and class for each telescope. A second step, called stereoscopy, combines these monoscopic reconstructions into a global one using engineered weighted averages.

In this work, we explore the possibility of using Graph Neural Networks (GNNs) as a suitable solution for combining information from each telescope. The "graph" approach aims to link observations from different telescopes, allowing analysis of the shower from multiple angles and producing a stereoscopic reconstruction of the events. We apply GNNs to CTAO-simulated data from the Northern hemisphere and show that they are a very promising approach to improving event reconstruction, providing a more performant stereoscopic reconstruction. In particular, we observe better energy and angular resolutions and enhanced separation between gamma photons and protons compared to the Random Forest method.

Reinforcement Learning for Calibration in Radio Interferometry
Author: Brian Kirk
Affiliation: New Mexico Institute of Mining and Technology / NRAO
ID: P215

Data gathered by an interferometer requires substantial processing before astronomers can extract the scientific information from it. The selection and sequencing of calibration and analysis actions is a complex decision-making task. This decision process depends on the expertise of astronomers, who consider data characteristics, instrument knowledge, compute cost, and best practices to guide each choice. We apply reinforcement learning (RL) to this task, where an agent can autonomously explore and identify optimal decisions based on an objective function with metrics that quantify best practices. By framing data processing as a pathfinding and cost minimization problem, we can use this data-driven approach to learn effective sequences of actions. This has implications for improving the accuracy and efficiency of high-throughput automated interferometric data processing.

Infrared Bubble Recognition in the Milky Way and Beyond Using Deep Learning
Author: Shimpei Nishimoto
Affiliation: Osaka Metropolitan University
ID: P216

The application of artificial intelligence in astronomy is rapidly expanding, offering new possibilities for identifying celestial structures within vast observational datasets. In this study, we propose a deep learning model that can detect Spitzer bubbles accurately using two-wavelength near-infrared data acquired by the Spitzer Space Telescope and JWST. The model is based on the Single Shot MultiBox Detector as an object detection model, trained and validated using Spitzer bubbles identified by the Milky Way Project (MWP-Bubble). We found that using only MWP-Bubbles with clear structures, along with normalization and data augmentation, significantly improved performance. To reduce the dataset bias, we also use the data without bubbles in the dataset selected by combining two techniques: negative sampling and clustering. The model was optimized by hyperparameter tuning using Bayesian optimization. Applying this model to a test region of the Galactic plane resulted in a 98 %

detection rate for MWP-Bubbles with 8 µm emission clearly encompassing 24 µm emission. Additionally, we applied the model to a broader area of 1°≤ |l| ≤ 65°, |b| ≤ 1°, including both training and validation regions, and the model detected 3,006 bubbles, of which 1,413 were newly detected. We also attempted to detect bubbles in the high-mass star-forming region Cygnus X, as well as in the external galaxies Large Magellanic Cloud (LMC) and NGC 628. The model successfully detected Spitzer bubbles in these external galaxies, though it also detected Mira-type variable stars and other compact sources that can be difficult to distinguish from Spitzer bubbles.

The detection process takes only a few hours, demonstrating the efficiency in detecting bubble structures. Furthermore, the method used for detecting Spitzer bubbles was applied to detect shell-like structures observable only in the 8 µm emission band, leading to the detection of 469 shell-like structures in the LMC and 143 in NGC 628.

Extending the Life of Software and Data

A science platform for the SOFIA Data Center

Author: Benjamin Greiner
Affiliation: SOFIA Data Center at the Institute for Space Systems, Universität Stuttgart
ID: P301

The Stratospheric Observatory for Infrared Astronomy (SOFIA) has gathered a considerable amount of scientific data between first light in May 2010 to the final observing flight in September 2022. The joint mission by NASA and DLR produced a diverse set of astronomical data from 8 science instruments along with imaging data from the tracker cameras and a large amount of observatory housekeeping datasets. After the end of flight operations, NASA has handed over all datasets to IPAC, where they were published in the Infrared Science Archive (IRSA). Additionally, NASA supported a post operations phase for a bit over a year, which allowed for a limited data reprocessing of SOFIA Observation Cycles 5 to 9. The DLR funded SOFIA Data Center (SDC) at the University of Stuttgart aims to perform a more comprehensive reprocessing of all available SOFIA data using upgraded data processing pipelines. Along with the technical and operational engineering data of the telescope, the science data - re-ordered by astronomical criteria - will be stored in a publicly accessible archive. We plan to provide a science platform, enabling users to explore and work with the rich legacy SOFIA has generated. This will include a web portal accommodating the intricacies of the observatory, Jupyter notebooks to access and process SOFIA data within a Jupyter Lab environment, and comprehensive access via Virtual Observatory (VO) APIs. We expect to work with the German and international astronomical community to promote the exploitation and publication of rare Mid- and Far-Infrared data, establish a powerful scientific resource for the time after the mission, and bridge the gap until the next space- or stratospheric infrared observatory becomes available.

Montage and Radio Astronomy
Author: Bruce Berriman
Affiliation: Caltech/IPAC-NExScI
ID: P302

The Montage Image Mosaic Engine, first released in 2002, was developed for processing FITS images from wide-field surveys at visible and infrared wavelengths. It has since found applicability across the electromagnetic spectrum to support data processing and visualization, from gamma-ray astronomy to low frequency radio astronomy. This broad applicability has come about through its design as an Open Source ANSI-C toolkit (and Python binary extensions), with independent components to perform each step in the creation of a mosaic and with support for all WCS extensions. This design enables easy integration into custom environments, workflows and pipelines, and is the principal reason for its long lifetime. Here we emphasize the growing use of Montage in radio astronomy (36 peer-reviewed papers since 2020), and will focus on three high-profile applications:

(1) Analysis of observations made with SKA-precursor experiments, such as MeerKAT and the Murchison Wide-field array;
(2) Identification of fast-radio bursts;
(3) Faraday tomography of LOFAR Two-Metre Sky Survey data (LoTSS-DR2), which explores the structure of the local interstellar medium.

GDL version 1.1
Author: Alain Coulais
Affiliation: LERMA, Observatoire de Paris, CNRS, France & DAp, CEA, France
ID: P303

GDL, a free interpreter of the IDL language --- the second most used programming language in astronomy --- is continuing smoothly its development, driven by feedback's and requests from end-users who are more numerous and active now that GDL is hosted on GitHub. Among the most interesting features introduced during the last years are the large test on M1/M2/M3 processors, the very good performance for computation (OpenMP included) demonstrated on a wide benchmark, a simplification of compilation and installation, the availability of SHMMAP and Bridge functions that are great for deploying concurrent GDL runs on the same RAM on HPC.

As developers of GDL we believe that this language has a useful place in the modern world where efficiency and low power computing is a must. Indeed, GDL (not to mention IDL), written in C/C++ shows excellent efficiency in "real world" benchmarks, making it one of a few interpreted language that are really "green". And probably the only that comes with such a huge collection of free, multitested and proven astronomical procedures written by colleagues in the past. GDL is also interesting for long life code since it is based on a very stable syntax.

Using OpenAPI to describe complex web services

Author: Dave Morris
Affiliation: University of Manchester
ID: P304

Our team at the SKA Observatory (SKAO) [1] are developing a specification for the International Virtual Observatory Alliance IVOA [2] ExecutionBroker [3] service that will help to facilitate cross-platform code execution across a heterogeneous federation of different types of science platforms.

As part of this we have been participating in the IVOA Protocols Transition Team (P3T) [4] looking at ways to update the IVOA standards to be compatible with modern software development environments and current best practice.

We have been evaluating the OpenAPI [5] service description language as a tool for describing the web service interface in a concise and standardised machine readable format.

Our overall experience of using OpenAPI has been positive, and we believe that providing a machine readable definition of our service interfaces alongside the service specification documents will extend the usability and lifetime of the IVOA standards. Making it easier to port our implementations to new platforms and frameworks in the future.

In this poster we will describe our experience of using OpenAPI and we will present some examples of things that worked well and some that didn't work so well.

[1] https://www.skao.int/en
[2] https://ivoa.net/
[3] https://github.com/ivoa-std/ExecutionBroker
[4] https://wiki.ivoa.net/twiki/bin/view/IVOA/DALGWSTigerTeam
[5] https://swagger.io/specification/

PipeCat: A Flexible and Modular Framework To Go From Images to Catalogs in Just a Few Steps

Author: Fernando Caro
Affiliation: INAF - OAR
ID: P305

We introduce PipeCat, a flexible and modular software aimed at simplifying how image processing pipelines are set up to produce source catalogs. The main idea behind this project is to offer users a clean interface to quickly configure a pipeline, where the main astronomical tools can be easily employed in a sequential manner to carry out the usual image processing steps (such as mosaicing, detection, deblending or photometry) and generate a custom source catalog as a result in a straightforward and fully reproducible way.

We believe that a tool with these characteristics may greatly help astronomers improve their scientific results, for instance, when working with big imaging datasets. In such cases, the official catalogs might not be enough to comply with their specific science requirements and, therefore, access to custom catalogs, produced with a precise set of configuration parameters for each processing step, represents a crucial need to be fulfilled. We will also discuss practical cases where generating a custom catalog has proven particularly helpful in enhancing the scientific outcome derived from analyzing a given imaging dataset.

The IRAM 30-meter control system upgrade and approach

Author: Jean Tedros
Affiliation: Insituto de Radioastronomia Millimetrica
ID: P306

The IRAM 30-meter telescope has known a number of major upgrades since its inception in 1985. The generational effort cumulatively preserved its position at the frontier of millimeter astronomy during its 40 years of operations. In 2023, in an effort to further improve its telescope, IRAM embarked on its most major refurbishment endeavor to upgrade the telescope's servo control system and enable robust tracking, fast scanning, and agile wobbling. Moreover, this project inscribes in the context of extending the life of software and hardware at the observatory and enhancing the system's reliability for future use. This was achieved by replacing the central VME-based component of the control system with a real-time PLC; thus creating a two layer approach separating the low-level real-time hardware control from the existing high-level observational logic. The project not only updated the system's core component but also ensured the compatibility with existing control components through the development of an middleware control layer. This middleware layer acts as asynchronous message-passing layer ensuring the cohesion of the low-level PLC to the high-level VME control components. This layer makes extensive use of the async Python library to orchestrate timely communication throughout the systems at the observatory. The system was tested, integrated and successfully commissioned at the 30-meter in February 2024 and has been operational since.

What’s new with astropy? Updates on a Community Python Library for Astronomy

Author: Jeff Jennings
Affiliation: Flatiron Institute
ID: P307

The Astropy Project is a community effort to develop a core package for astronomy using the Python programming language and improve usability, interoperability, and collaboration between astronomy Python packages. The core astropy package contains functionality aimed at professional astronomers and astrophysicists, but may be useful to anyone developing astronomy software. This poster summarizes recent, near-future and long-term changes to the core astropy package, including the codebase and documentation. It also examines the overall health of the project and details how individuals can contribute.

Setting Sail with ULLYSES

Author: Jo Taylor
Affiliation: Space Telescope Science Institute
ID: P308

Specifically selected to leverage Hubble's unique UV capabilities, the Hubble Ultraviolet Legacy Library of Young Stars as Essential Standards (ULLYSES) program is a large Director's Discretionary program of approximately 1,000 orbits- the largest ever executed. The ULLYSES program recently completed a UV spectroscopic library of young high- and low-mass stars in the local universe, leveraging a considerable amount of archival observations. Half of the program is dedicated to the study of accretion physics through UV-optical-NIR spectroscopy of young, low-mass accreting (T Tauri) stars in star-forming regions in the Milky Way. The other half is focused on the characterization of the winds and photospheres of massive O and B stars in nearby galaxies.

While the program concluded in 2024, the data products and software tools have been designed to streamline data exploration and analysis in the long term, and maximize the impact and legacy of the ULLYSES observations. Here we present a description of the program deliverables such as data products, software, and documentation, as well as the efforts to maximize the longevity of the science return.

Rewriting and Modernization of C++ Code for Gain and Atmospheric Calibration of Far-Infrared Heterodyne Astronomical Data: A Python-Based Approach

Author: Juan Luis Verbena
Affiliation: Universität zu Köln
ID: P309

The analysis of far-infrared heterodyne astronomical data often requires sophisticated calibration techniques to account for atmospheric effects. In the case of the GREAT receiver on board SOFIA, this process has been handled using a complex C++ codebase that integrates FORTRAN and Python components. However, this approach presents several challenges, including difficulties in parallelization, maintenance complexity, and limited adaptability to new instrumentation.

In this work, we present the rewriting and modernization of this C++ code into a streamlined Python-based framework designed specifically for a new generation of far-infrared instruments, i.e., the CHAI heterodyne receiver for the CCAT telescope. The Python implementation leverages modern computational libraries and parallel processing techniques, enabling more efficient and scalable data calibration. By reducing the code complexity, we enhance the maintainability of the calibration pipeline. Additionally, the simplified observing modes of the target instrument will be fully integrated into the new framework, providing a more focused and optimized solution for its specific operational requirements. To further enhance performance, the framework will incorporate C++ native functions, accessed through Python wrappers, where needed. This approach ensures compatibility with existing code while leveraging the efficiency of lower-level implementations. This transition is designed to position the framework for future advancements in data analysis techniques, while maintaining the overall efficiency of the calibration process by utilizing established libraries and custom C++ integrations.

Our work aims to result in significant improvements in processing speed, code clarity, and ease of use. This work outlines the key aspects of the transition from C++ to Python, the challenges encountered, and the benefits obtained in the context of far-infrared astronomical data calibration.

Database of Markarian galaxies and classification of low-dispersion spectra

Author: Areg M. Mickaelian
Affiliation: Byurakan Astrophysical Observatory (BAO)
ID: P310

A database of 1515 Markarian (Mrk) galaxies has been created at the Byurakan Astrophysical Observatory (BAO) webpage at https://www.bao.am/TEMP/mg/index.php. Mrk lists, sky distributions, individual page for each object with all alternative names and various available data, DSS1/2 and SDSS images, the Digitized First Byurakan Survey (DFBS) low-dispersion and SDSS medium resolution spectra, Spectral Energy Distribution (SED), etc. are given. A new fine classification for activity types for 779 Mrk galaxies having SDSS spectra was carried out with the new classification scheme (https://www.bao.am/activities/projects/21AG-1C053/mickaelian/). Trials for classification of the objective prism low-dispersion spectra were carried out based on the DFBS for a number of types of objects: Mrk (UV-excess) galaxies, QSOs and Seyferts, white dwarfs (WD) and hot subdwarfs (sb), carbon (C) and late-type M stars, etc. DFBS comprises low-dispersion spectra (the resolution is between 22 and 60 A) for approximately twenty million objects. Considering the distinct spectral characteristics within subgroups, sub-object classification becomes crucial for a more detailed understanding of the dataset. We have proposed a comprehensive cloud-based service for classifying objects into spectral classes and subtypes. The service employs advanced ML algorithms trained on labeled data to classify objects into their respective spectral types and subtypes. The service can be accessed and utilized through a user-friendly interface, making it accessible to a wide range of users in the astronomical community.

Renewing the 2dFdr legacy codebase as a sustainable Python package

Author: Miguel A. Gonzalez-Bolivar
Affiliation: Australian Astronomical Optics
ID: P311

For over two decades, astronomers have relied on the AAO's 2dFdr package to process data from various multi-fibre spectroscopy instruments on the AAT, such as AAOmega and KOALA. While the original codebase, written in Fortran, C, and Tcl/Tk, was suitable at the time, few astronomers today are proficient in these languages, making ongoing maintenance and development challenging. With the astronomical community's widespread adoption of Python for data reduction and analysis, it is essential to develop a Python interface for 2dFdr to ensure its longevity. This presentation discusses the successful development of the twodfdr Python module. Using the F2PY package from NumPy, we have created Python bindings for the low-level Fortran and C functions in 2dFdr. These functions are now accessible through user-friendly, high-level Python classes such as Tdfio (file input and output) and various other methods to call key data reduction routines. We employ a taest-driven development approach, requiring unit tests for each function to ensure robust performance, and provide API and user-facing documentation. It is crucial to ensure that the new bindings do not introduce regression errors compared to the original well-characterised Fortran code. Parameters and execution of data reduction routines are managed in a standardized and reliable way using the AAO-developed Python Language Bindings for CPL (PyCPL), enabling 2dFdr to function like any ESO pipeline. Our overall strategy is to implement line-by-line Python replacements for key components of 2dFdr. This will enable developers and astronomers to incrementally transition 2dFdr's legacy code to sustainable Python native implementations.

Ziggy, Improving the Pipeline Infrastructure of Kepler and TESS

Author: Peter Tenenbaum
Affiliation: SETI Institute
ID: P312

We report on Ziggy, a portable, scalable infrastructure for science data processing pipelines.

Ziggy is based on the Pipeline Infrastructure (PI) software that was originally written to support the Kepler exoplanet search mission, and on the extension of that software for use by the Transiting Exoplanet Survey Satellite (TESS) mission. Ziggy maintains and extends the key features of its predecessors, including: strong integration with the NASA High End Computing Capability (HECC) supercomputer, execution sequencing, a well-organized logging system, generation and preservation of data accountability and traceability records, execution flow control, and exception handling. Improvements and upgrades to the software include simplified mechanisms for users to define pipelines and integrate algorithms developed by subject matter experts, a robust messaging system that links Ziggy's components, an intuitive tool for selecting collections of data for processing and for choosing

whether to process only data that has not yet been processed versus processing all data regardless of its processing history, and an event-driven processing manager that allows Ziggy to initiate processing upon detection of a signal from an external source. Ziggy has been released as open-source software as part of the NASA Open Source Science Initiative and is freely available at github.com/nasa/ziggy. Ziggy is currently used as the pipeline manager for a data analysis pipeline that operates on GOES satellite data, and the TESS data analysis pipeline is currently undergoing modifications to run under Ziggy rather than under one of its illustrious predecessors.

Container-based pre-pipeline data processing on HPC for XRISM

Author: Satoshi Eguchi
Affiliation: Kumamoto Gakuen University
ID: P313

The X-Ray Imaging and Spectroscopy Mission (XRISM) is the 7th Japanese X-ray observatory launched on 2023 September 7, whose development and operation are in collaboration with universities and research institutes in Japan, U.S., and Europe including JAXA, NASA, and ESA. The telemetry data downlinked from the satellite are accumulated into the SIRIUS database for abstraction of the tracking and control stations on the ground and for deduplication of telemetry packets, then are converted to scientific products such as cleaned event FITS files with the pre-pipeline (PPL) and pipeline (PL) software running on standard Linux VMs on the JAXA and NASA sides, respectively. Since both PPL and PL are still under improvement, whole (> 100) products must be reprocessed at once with the latest version of PPL and PL at some points to homogenize the quality. To boost the throughput of PPL, we decided to port our PPL to the JAXA "TOKI-RURI" HPC system capable to complete ~ 100 PPL processes within 24 hours. By utilizing the container platform of Singularity and its "--bind" option to create arbitrary mappings between a real filesystem on HPC and that inside a container instance, we could prepare the HPC version of PPL in two weeks. In this paper, we briefly show the data processing in XRISM and present our porting strategy of PPL to the HPC environment in detail.

User Experience

Online SIMBAD TAP query interface

Author: Anaïs Oberto
Affiliation: CNRS - Observatoire astronomique de Strasbourg
ID: P401

SIMBAD is a service to provide identifications, bibliography and measurements on astronomical objects. It is used mainly through the web site, with forms to query by coordinates, identifications, papers, and more.

It provides also a Table Access Protocol access (TAP, promoted by the IVOA) to allow the users to write their own query, using any type of selections and criteria.

SIMBAD is working on a new web interface built as an IDE (Integrated Development Environment), to facilitate the user to write its SQL-like query.

We propose a code completion function giving access to the 326 columns of the SIMBAD schema, user-defined functions, and if any, uploaded columns.

Beside that, the interface gives the possibility to edit multiple queries in multiple opened tabs, and an access to the user's history of queries and their results.

Leveraging the experience of chatbot development in the AI "explosion" era to improve the astronomical data services

Author: André Schaaff
Affiliation: ObAS, CNRS, CDS
ID: P402

Over the past years the CDS has undertaken R&D work on Natural Language Processing applied to the querying of astronomical data services. The motivation for this is to enable new ways of interaction as an alternative to the traditional forms exposing parameter fields, check boxes, etc. Our aim is to answer the fundamental question: is it possible to reach query results satisfying professional astronomers? The Virtual Observatory (VO) brings us standards like TAP, UCDs, ..., implemented in the CDS services. The VO enables the interoperability which is a mandatory backbone, helping us to query our services in NL and which will be useful in a further step to query the whole VO. We will present our pragmatic approach exploiting all the resources we have (authors in SIMBAD, missions and wavelengths in VizieR, UCDs, ADQL/TAP, ...) using a chatbot interface. Initially based successively on several libraries and APIs (Stanford NLP Java, Google Dialogflow, RASA) we started in 2023 to explore how to improve it with the OpenAI API. We are now forking from this initial work to study how to apply it to the improving of our services, in a wider AI use.

Results from the 2023 ESA Astronomy Space Science Archives User Survey

Author: Deborah Baines
Affiliation: ESA
ID: P403

The majority of ESA's space science mission archives are developed and maintained by the ESAC Science Data Centre (ESDC), in coordination with the science operations centres, instrument teams, and mission consortia. On the astronomy side, these archives include the Gaia archive, ESA Hubble Science Archive, ESA JWST Science Archive, XMM-Newton Science

Archive, Herschel Science Archive, Planck Legacy Archive, ISO Data Archive, EXOSAT Science Archive and Lisa Pathfinder Legacy Archive. ESDC also develops ESASky, an all-sky science exploration interface that provides easy access to high-level science data from both ESA archives and major ground- and space-based observatories. In mid-2023, the ESA Astronomy Archives User Group initiated a survey to gather feedback from the astronomy community on ESA's archives and ESASky. This poster highlights the key findings of the survey, providing an overview of how these archives are being used, user satisfaction, and opportunities for future improvements across the archives.

VisiOmatic 3: Remote Image Visualization with New Python-Based Features

Author: Emmanuel Bertin
Affiliation: CFHT and CEA/AIM/Université Paris-Saclay
ID: P404

The VisiOmatic package is a comprehensive remote visualization system designed for large, multispectral astronomical image datasets. Version 3 is primarily developed for advanced 'quicklook' image visualization of imaging data at CFHT, but it is versatile enough for a wide range of applications, including use as a stand-alone viewer. The server-side code has been completely rewritten in Python, introducing new features such as support for hyperspectral datacubes and multi-extension files from mosaic cameras, just-in-time caching of FITS data, and temporal image sequence playback. `VisiOmatic` 3 is released under the MIT license.

ALMA Wideband Sensitivity Upgrade: Science Archive impact

Author: Felix Stoehr
Affiliation: ESO/ALMA
ID: P405

The Atacama Large Millimeter/Submillimeter Array (ALMA) will be substantially upgraded. The upgrade contains a new correlator, upgraded receiver bands and signal chain as well as updates to the existing software systems like the data processing system and the ALMA Science Archive (ASA). We briefly describe the current usage of the ASA, the short-term plans as well as the impact and possible long-term plans to allow users to deal with the increased yearly data volume of 11 times (2029) and 46 times (2032?) compared to today.

Jasmine - JAvaScript Multimodal INformation Explorer

Author: Fenja Schweder
Affiliation: University of Bremen, HITS gGmbH
ID: P406

Astronomical data is rich in volume, information, and facets. While this offers countless research perspectives, the inspection and exploration of the data remains a challenge. A single visualization of a data point is not capable of displaying all contained information without being overloaded. In addition, some aspects require different visualization techniques than others, due to their dimensionality or behavior.

To address this issue, we introduce Jasmine, the JAvaScript Multimodal INformation Explorer. Jasmine allows users to open different data viewer modals that show a specific data point from a set. The viewer currently supports image data, as well as point cloud objects. Users can decide on which information about the data point they like to have displayed. Point clouds are interactive and allow zooming, tossing, and turning. Picking a data point is enabled by providing a structured view of the set, arranged by a key property. The arrangement is achieved by autoencoding. This offers the user to browse the data set with similar points in the same region.

FitsImageVOY: A FITS Image Viewer for Visual Studio Code and Modern Successor to SAOImage

Author: Guillem Megias
Affiliation: Stanford University
ID: P407

The visualization of astronomical data is pivotal to the process of discovery, and tools that integrate seamlessly with development environments are increasingly in demand. FitsImageVOY, a FITS image viewer extension for Visual Studio Code, is devised as the natural successor to SAOImage, offering a scaled-down yet robust tool for viewing and analyzing FITS files. This extension enables astronomers to interact with FITS images directly within their coding workspace, providing an intuitive and efficient way to inspect and manipulate astronomical data. FitsImageVOY incorporates essential features such as customizable image scaling, color mapping, and data inspection, maintaining the core functionalities of SAOImage while taking advantage of the versatile Visual Studio Code environment. This presentation will highlight the capabilities of FitsImageVOY, its integration into the astronomer's workflow, and its role in modernizing the tools available for FITS data visualization. Future developments aimed at enhancing the extension's utility will also be discussed, underscoring its potential to become a central tool in the astronomer's toolkit.

dysh: A Python Package for Calibration of Single-Dish Spectral-Line Data

Author: Marc Pound
Affiliation: University of Maryland
ID: P408

The Laboratory for Millimeter-Wave Astronomy (LMA) at The University of Maryland is working with The Green Bank Observatory (GBO) to create dysh, an open-source and portable Python-based spectral line reduction and analysis package for Singledish FITS (SDFITS) data (https://dysh.readthedocs.io/). dysh provides a new, rich Python software environment for interaction with Green Bank Telescope (GBT) spectral line data by observers, post-observation scientists, and GBO staff. It is intended to replace GBO's current data reduction package, GBTIDL. Features include simultaneous calibration of multiple scans using improved algorithms, a user-friendly object oriented interface, an interactive plotter, and out-of-the-box calibration for all GBT receivers. To lower barriers to adoption, dysh is built on packages familiar to astronomers such as Astropy, specutils, NumPy, Matplotlib, and pandas. It can be run through IPython, notebooks, or scripts. Our development paradigm is Agile-like, with frequent communication between core LMA and GBO project members as well as regular feedback-gathering from user tests. dysh is scheduled for completion by the end of 2025 with several intermediate releases currently available, and it has the flexibility to be expanded to work with SDFITS files from other single dish radio telescopes, such as the Large Millimeter Telescope (LMT).

TOPCAT/STILTS Interoperability

Author: Mark Taylor
Affiliation: Space Telescope Science Institute
ID: P409

TOPCAT is an established interactive GUI desktop tool for data analysis of tables, offering visualisation, crossmatching, data manipulation, and access to VO services among other capabilities. STILTS is a suite of command-line tools providing a scriptable interface to much of the same functionality. While scripting is a powerful way to approach many data analysis tasks, the learning curve for STILTS is rather steeper than for its point'n'click counterpart, with the result that some TOPCAT users may be reluctant to exploit the scripting capabilities on offer because of the (perceived?) difficulty of learning to use them.

To address this, many windows in TOPCAT can now display to users the STILTS invocation equivalent to the operation being performed by the GUI. It is hoped that this will encourage more users to benefit from the scripting interface.

Building DAST: Data analysis and support for PLATO scientists

Author: Martin Schäfer
Affiliation: Max Planck Institute for Solar System Research
ID: P410

ESA's PLATO mission (Planetary Transits and Oscillations of stars) is a multinational space mission aimed at discovering Earth-like planets around Sun-like stars and advancing our understanding of stellar physics. To facilitate human analysis and validation of the large-scale observational data, a suite of integrated, powerful, and user-friendly tools is essential. Since no commercial off-the-shelf solutions exist, the Data Analysis and Support Tools (DAST) are being developed within the PLATO mission to address this need. DAST provides web-based access to scientific pipeline data products for visual inspection, along with various tools for on-the-spot analysis, enabling the evaluation of alternative pipeline parameters. Additionally, DAST allows PLATO scientists to offer feedback and rank planetary candidates.

Additionally, the development of DAST presents the typical challenges of creating maintainable, reusable software in an academic setting. These challenges include evolving institutional responsibilities, personnel changes, and the ongoing refinement of requirements and algorithms.

This poster will highlight the features currently implemented in DAST and provide an outlook on tools planned for future releases. Further we'll outline the approach taken by the MPSSR in developing the DAST software using the SCRUM methodology.

The PADC ObsTAP Portal: a user-friendly gateway to archived observations in the Virtual Observatory

Author: Renaud Savalle
Affiliation: PADC / Observatoire de Paris - PSL
ID: P411

In the context of Open Science and FAIR practices, the IVOA plays a key role in developing standardized protocols to access astronomical data within the Virtual Observatory. Many data centers worldwide now publish archived observations using the ObsCore data model and the Table Access Protocol. We have developed an ObsTAP web-based portal to empower astronomers to effortlessly query this wealth of information through a user-friendly interface. Contents of the web form for space, time, and energy constraints are used to generate an ADQL query which can be edited before being sent in parallel to about 40 IVOA-registered servers. Results are summarized in a tabular format with footprints and previews, and can be further analyzed using dedicated applications connected with the Simple Application Messaging Protocol, such as Aladin, TOPCAT or CASSIS. Our portal has become the primary client at Paris Observatory for accessing PADC data collections such as digitalized plates from the new NAROO scanner, and students observations.

New Python-based Architecture for the Keck Observatory Archive

Author: Rosemary Moseley
Affiliation: Caltech/IPAC-NExScI
ID: P412

We describe the development of the Keck Observatory Archive (KOA) Data Discovery Service, a web-based dashboard that returns metadata for wide-area queries of the entire KOA archive within seconds. Currently in beta, this dashboard will support exploration, visualization, and data access across multiple instruments. This effort is underpinned by open-source VO compliant query infrastructure, and will offer services that can be hosted on web pages or in Jupyter notebooks. The effort also informs the design of a new modern landing page that meets expectations of accessibility and ease of use.

The new query infrastructure is based on nexsciTAP, a component-based, DBMS-agnostic Python implementation of the IVOA Table Access Protocol, developed at NExScI and integrated into the NASA Exoplanet Archive and the NEID archive, and into the PyKOA Python client. This infrastructure incorporates R-tree spatial indexing, built as memory-mapped files as part of Montage, a software toolkit used to create composite astronomical images. R-trees are used most often in geospatial analysis, here they enable searches of the entire KOA archive, an eclectic collection of 100 million records of imaging and spectroscopic data, in 2 seconds, and speeds up spatial searches by x20. The front end is built on the open-source Plotly-Dash framework, which allows users to build an interactive user interface based on a single python file.

Data Management and Trusted Repository in the Open Data Era

The ASTRON Data Archive
Author: Aida Ahmadi
Affiliation: ASTRON
ID: P501

This poster presents an overview of the ASTRON Data Archive (ADA) and its ingest and access statistics over the past two years. The ADA provides access to the instrument data archives for the Low Frequency Array (LOFAR) and APERTIF. LOFAR consists of an interferometric array of 52 dipole antenna stations distributed throughout the Netherlands and Europe operating in the frequency range 10-90 MHz and 110-240 MHz. After more than a decade, LOFAR ceased operations at the end of August 2024 and is ongoing a major upgrade to come back on sky in 2026. With more than 70% operational efficiency, more than 60 PB of data in mixed state of reduction and science readiness (mostly uncalibrated) have been ingested in the LOFAR Long Term Archive (LTA), the largest astronomical data collection to date and accessible to the world wide scientific community. LTA's content includes raw telescope measurement of correlated visibilities for calibrators and science targets, raw beam forming observations, interferometric averaging pipeline and pulsar pipeline data products. These are distributed over three data centers: SURF (Netherlands), Forschungszentrum Juelich (Germany) and the Poznan Supercomputing and Networking Center (Poland). Over the past two years, on average ~250,000 data products have been ingested in the LTA with ongoing efforts to reduce the size of data products and retire unused files. APERTIF is a phased-array feed installed on the focal plane of 12 of the 14 25-m dishes of the Westerbork Synthesis Radio Telescope (WSRT), operating at the wavelength of 20 cm. APERTIF is a survey instrument, carrying observations up to March 2022. The APERTIF Long Term Archive (ALTA) consists of two connected data repositories, one disk-based, hosted at ASTRON, and one tape-based, hosted at SURF. ALTA hosts data at different levels of calibration and processing, from unprocessed or compressed observation timeseries and visibility data to calibrated (imaging) visibility data, to quality assessed and curated image FITS files. APERTIF data processing is done by dedicated science teams and low-level science data products are stored on tape media to reduce costs. Several LOFAR and APERTIF community-generated data releases are hosted and/or published through the ASTRON VO service (vo.astron.nl), making science-ready data products available to the wider astronomical community. The upgraded LOFAR2.0 will generate significantly more data than its predecessor, posing a challenge in data processing and storage. As we move to this new era, advanced processing pipelines will need to run in the operational environment to generate science-ready data products automatically. ASTRON envisions a new data portal, currently named ADEX (ASTRON Data Explorer), that will provide unified access to LOFAR and APERTIF data in the future.

Accessing Data from the Near-Earth Objects Coordination Centre

Author: Dario Oliviero
Affiliation: ESA / PDO / NEOCC
ID: P502

The Near-Earth Objects Coordination Centre (NEOCC), is the main component of the Planetary Defence Office (PDO) within ESA's Space Safety Programme. Located in ESRIN, ESA’s establishment near Frascati (Italy), the NEOCC serves as the ESA’s central hub for a variety of data related to near-Earth object (NEO) and planetary defence activities. Every day, the NEOCC uses the data gathered from across the globe, to provide orbital information, impact monitoring and risk analysis and to distribute this information. Its mission is to support and coordinate the observation of small Solar System bodies to assess and track the threats they may pose to Earth, and to recommend mitigation strategies when necessary. The NEOCC also provides the community with a web portal dedicated to the distribution and visualization of orbital data of NEOs, including FITS images, raw and processed observations, asteroids’ orbits and ephemerides, the list of objects with a non-zero probability of impacting Earth (“risk list”), along with lists of recent and upcoming asteroids' close approaches.

In this presentation, we will focus on the NEOCC API system, which is a key component of our service. The API service provides a comprehensive access to our most current data (detailed at https://neo.ssa.esa.int/computer-access). This system allows developers to seamlessly integrate and utilize data from "Aegis", NEOCC's asteroid orbit determination and impact monitoring system (1). The presentation will also include a preview of the upcoming release of the Python interface to NEOCC's data, designed to provide a more user-friendly way of accessing, manipulating and analysing our data.

The NEOCC is also integrating its data into the VESPA service, a web-based data search interface for Planetary Science and Heliophysics data (2). The data integration process is quite advanced, and a preview of the data is already publicly available (3).

By presenting our work, we aim to increase the visibility of the data publicly available via our services, as well as to highlight the significance of the NEO hazard and the role of NEOCC in monitoring and assessing it. We also hope to foster collaboration and communication with other NEO-related entities, such as observatories, space agencies, and scientific institutions. Our work is relevant not only to the scientific community but also to software engineers and data analysts, who can benefit from the constantly updated and freely accessible data we maintain on our portal. Moreover, by sharing and presenting our data and methods, we contribute to the advancement of knowledge and technology in the field of Planetary Defense.

(1) https://ui.adsabs.harvard.edu/abs/2023sndd.confE..73F/abstract
(2) https://vespa.obspm.fr/planetary/data/
(3) https://dachs.neo.s2p.esa.int/tap

S3AI: Spoke 3 Archive Infrastructure
Author: Giacomo Coran
Affiliation: INAF OATS
ID: P503

An innovative archiving system has been developed for the Spoke3 WP4 project. A Python based software manages the ingestion of incoming files, storing extracted metadata in a PostgreSQL database and separating actual data in a Rucio infrastructure. Metadata can be searched enabling the data download using a custom web application called jPortal, providing additional "Cut & Merge" and "Transits" services to address big data challenges. Further services are currently under development.

Astronomical Visualizers for Enhanced Data Analysis in ESA's HST and JWST Science Archives
Author: Javier Espinosa Aranda
Affiliation: Starion for ESA
ID: P504

Efficient data access and analysis are crucial in the rapidly evolving field of astronomy. This poster introduces powerful data visualization and analysis tools for conducting research with the European Space Agency mirror of the Science Archives of the Hubble Space Telescope (HST, https://hst.esac.esa.int/ehst) and James Webb Space Telescope (JWST, https://jwst.esac.esa.int/archive). Researchers will be guided through accessing and utilizing various visual tools for exploring and cross-checking extensive astronomical data repositories.

Starting with HST, participants will discover how to use an image viewer built on JS9 for a detailed exploration of astronomical images. They will also explore a spectra viewer that supports cross-comparison across multiple observations. And once the new version is released, the HCV (Hubble Catalog of Variables) Explorer application will be migrated to this new interface, providing enhanced capabilities for source analysis.

In the JWST Science Archive, users will be introduced to a similarly powerful image viewer and a data cube viewer, ideal for exploring the three-dimensional structure of astronomical data. Additionally, ongoing developments for a new spectra viewer will be highlighted, which promises further capabilities for visualizing and comparing spectral data. Additionally, the integration of automated HiPS (Hierarchical Progressive Surveys) generation for JWST will be featured, showcasing how astronomical images are rendered interactively in ESASky.

As a final step, a detailed walkthrough will demonstrate how to seamlessly integrate these searches into automatically generated Jupyter Notebooks within the User Interfaces, allowing users to execute their own scripts for visualizing data.

By focusing on these advanced visualization tools, the workflow accelerates the research process, offering astronomers new ways to visualize, cross-check, and analyze vast datasets with ease. This approach facilitates deeper insights and supports a more intuitive exploration of archival data.

AstroDB Toolkit: A collaborative data management tool

Author: Kelle Cruz
Affiliation: CUNY Hunter College
ID: P505

We present the AstroDB Toolkit, a collaborative database management tool, intended for small-to-medium sized projects to manage their living datasets. The database management relies on a review using GitHub, thus integrating itself into the existing workflows of many Astronomers. The Toolkit relies on the astrodbkit and astrodb_utils Python packages to convert back and forth between a non-relational and relational database that can be used with established packages like SQLAlchemy. This facilitates a novel way of reviewing, updating, and version-controlling the contents of medium-scale databases while still integrating with existing tools. We have also developed a web app which presents the database contents as an interactive, queryable website. The proof-of-concept application is the SIMPLE Archive of low mass stars, brown dwarfs, and exoplanets (https://simple-bd-archive.org/). In this poster, we present the architecture of the AstroDB Toolkit, the document-store model, and the GitHub workflow for reviewing and approving database modifications.

Mini-SiTian Data Release 1
Author: Linying Mi
Affiliation: National Astronomical Observatories, CAS
ID: P506

The SiTian project is a next-generation large-scale time-domain survey, aiming to establish an array of 54 optical telescopes across multiple observatory sites. This array will enable 30-minute single-exposure observations of the entire northern sky, capturing true color (gri) time-series data down to about 21 mag. As the pathfinder for the SiTian project, the mini SiTian project utilizes an array of three 30 cm telescopes to simulate a single node of the full SiTian array. Over the past two years, mini SiTian has achieved a series of significant progresses, including the successful testing and validation of control systems, data processing pipelines, and design of observational strategies. It also developed image processing systems (both static and real-time), enabling the rapid detection of transient events and the generation of real-time alerts, effectively addressing the challenges of observing transient events in modern astronomy. During its pilot operation, the mini SiTian collected a substantial amount of observational data, capturing light curves of supernovae and tidal disruption events (TDEs), providing community with crucial data for a deeper understanding of these transient phenomena.

This presentation will focus on the upcoming first data release mini SiTian survey, highlighting the potential of this data to advance research in time-domain astronomy. It will provide a detailed overview of the observational sky coverage, the generated star catalogs, and the light curves. The data release system has been developed by the China-VO team, and the presentation will demonstrate how users can efficiently browse, search, and download data on the platform, as well as introduce the scientific tools available on the platform. Additionally, the presentation will delve into the development and application of the VO-based data management and release platform, which not only ensures the standardization of data management processes but also aims to establish a trusted repository. This platform will provide a robust infrastructure that meets the highest standards of the open data era, supporting the future release of SiTian data to the global scientific community.

Improving the findability of data in the European VLBI Network archive
Author: Mark Kettenis
Affiliation: Joint Institute for VLBI ERIC
ID: P507

While many radio telescopes have an open skies policy and make data public some time after observation, it isn't always easy for astronomers to find the data relevant to their research. In recent years some of the larger radio facilities have made their data archives accessible through Virtual Observatory (VO) protocols. However the existing protocols are not always a good fit for the wide variety of radio data products available. That is why the Radio Astronomy Interest Group within the International Virtual Observatory Alliance (IVOA) has been working on an extension of the ObsCore model that allows astronomers to search for observations with parameters that better describe observations made by radio interferometers as well as single dish radio telescopes. An implementation of this proposed extension is now available in the popular DaCHs software that implements a large range of VO protocols. We have used this implementation to offer a TAP service that provides the ObsCore extension for the data archive of the European VLBI Network (EVN). This poster describes the Python software that was built to extract the relevant parameters from data files in the archive. It also showcases some use cases that use the new parameters to do targeted searches in the data archive as well of some interesting statistics that can be done using these new parameters to characterise the data in the EVN archive.

The New Architecture of the Online Observation Quality System for the ASTRI Mini-Array Project

Author: Nicolò Parmiggiani
Affiliation: INAF/OAS Bologna
ID: P508

The ASTRI Mini-Array is an international collaboration led by the Italian National Institute for Astrophysics. This project aims to construct and operate an array of nine Imaging Atmospheric Cherenkov Telescopes to study gamma-ray sources at very high energy (TeV) and perform stellar intensity interferometry observations.

We describe the updated software architecture of the Online Observation Quality System (OOQS) for the ASTRI Mini-Array project. The OOQS is one of the subsystems of the Supervisory Control and Data Acquisition (SCADA) system. The OOQS aims to execute real-time data quality checks on the data acquired by the Cherenkov cameras and intensity interferometry instruments and provide feedback to both SCADA and the Operator about abnormal conditions detected. When OOQS detects anomalies, it is possible to correct these anomalies and optimizie the duty cycle of the observations.

The OOQS is designed to analyze different data packets received from the Array Data Acquisition System through the Kafka service. The data are serialized and deserialized for transmission using Avro. Internally the OOQS uses ZeroMQ to transfer data between processes, monitor the execution of the analyses, and receive commands from the OOQS Manager. The OOQS Manager is an Alma Common Software (ACS) component that is interfaced with other SCADA subsystems. The OOQS Manager sends commands to the OOQS pipeline which is deployed as a docker container in a Kubernetes cluster. There are nine instances of the OOQS, one for each telescope of the array.

The results generated by the OOQS are quality check statuses, camera plots, histograms, tables, and value versus time plots. The results are stored in the Quality Archive for further investigation and sent to the Operator Human Machine Interface (HMI) through Kafka.

Hipsgen: news features for news HiPS
Author: Pierre Fernique
Affiliation: Centre de Données astronomiques de Strasbourg
ID: P509

The HiPS standard was standardized by IVOA in 2019 to meet the challenge of visualizing and exploiting large astronomical surveys. Its hierarchical access method, based on HEALpix resampling, makes a sky survey accessible, displayable and even processable, whatever the size of the survey, the quality of the network and the computing power available to the astronomer. It's a response to the challenge of big data, particularly important with the advent of the next big surveys such as the Rubin or SKA. HiPS - Hierarchical Progressive Survey - has been implemented for several years now in a number of scientific, amateur and popular visualization tools, including Aladin, hips2fits, ESAsky, ESO portal, Firefly, DIGISTAR and Stellarium. In contrast, most of the 1400 HiPS currently available have been generated using a unique tool: Hipsgen, developed and maintained by CDS since the invention of HiPS. In this poster, we present recent developments and improvements to Hipsgen: iterative generation, contrast adjustment, alternative storage, etc. We'll also mention promising future developments notably HiPS-3D, as well as alternatives to Hipsgen.

A New VO Publishing Registry Framework at NOIRLab
Author: Robert Nikutta
Affiliation: NSF NOIRLab
ID: P510

With a growing number of survey datasets and data services offered at NOIRLab the need to support their discoverability was long overdue. With initial support by the IVOA VORegistry working group,

NOIRLab set out to write a lean publishing registry framework to facilitate the publishing of records to the VO Registry through a simple plug-and-play architecture. No knowledge of the publishing

mechanics is required to run an instance, and only the VO resource records need to be generated and pushed up via PRs; they can be in XML, JSON, or YAML formats. Our development implements the OAI-PMH protocol, is fully open-source, written in Python, documented, and can be easily deployed on GCP or other cloud providers using Terraform. It specifically can run on temporary instances that only incur costs when requests are being processed. This is ideal for teams, and even individual surveys, who have no internal or national VO support. The deployment of the registry is automated, and the templating allows for quick customization to the needs of the records. Various projects within an organization can simply contribute their VOResource records to the registry instance, whichever way they generate them. In the first deployment at NOIRLab, the new registry contains datasets and

services from the Astro Data Lab science platform and the Astro Data Archive, with plans to include other projects. In this talk we will discuss the motivation behind our implementation, and also outline how other organizations can simply clone and adopt NOIRLab's registry code for their use cases.

Introduction of China-VO New Paper Data Repository System

Author: Youfen Wang
Affiliation: National Astronomical Observatories, Chinese Academy of Sciences
ID: P511

The China-VO Paper Data Repository offers long-term storage and open access to paper data, encompassing a wide range of materials such as tables, figures, images, videos, source codes, models, and software packages referenced in scientific publications. Each dataset is assigned a permanent URL and a Digital Object Identifier (DOI), ensuring easy citation and reference. It is important to note that the copyrights of these materials remain with the authors. The repository system has been updated since 2022, and we are pleased to introduce the updated system in this overview.

CDS Open Data commitment for data published by authors

Author: Gilles Landais
Affiliation: Université de Strasbourg
ID: P512

In this presentation, we will show how the CDS infrastructure meets the challenges of open science for data published in articles. Open science requires trusted Data Centers that meet criteria for both data preservation and distribution. The preservation principles embodied by the acronym TRUST (Transparency-Responsibility-User focus-Sustainability-Technology) are reflected in the CDS by a strong partnership with data providers such as space agencies, the scientific journals or national institutes and by an architecture that includes redundancy and the use of mirrors. Applying FAIR principles to data distribution means implementing standards. The standards in astronomy have been implemented for history in CDS, in particular those provided by the Virtual Observatory - These are completed by interdisciplinary norms which has grown significantly the recent years. For instance, the persistent identification, which is both a unique identifier and a means of extending the distribution of authors' products across different interconnected networks. Applying these principles allowed the CDS to be awarded by the Core Trust Seal certification. Finally, we will discuss the challenges and investments involved in continually renewing our hosting offering.

Roadblocks in Astronomical Data Analysis

Distribution of the Högbom CLEAN Algorithm Using Tiled Images with Feedback
Author: Daniel Wright
Affiliation: Oxford University
ID: P601

Data sizes for next generation radio telescopes, such as the Square Kilometer Array (SKA), are far above that of their predecessors. The CLEAN algorithm was originally developed by Högbom in 1974, long before such data sizes were thought possible and is still the most popular tool used for deconvolution in interferometric imaging. In order to facilitate these new large data sizes and reduce computation time a distributed approach to the algorithm has been investigated. The serial nature of the CLEAN algorithm, due to its matching pursuit design, makes this challenging. Splitting the image into a number of tiles which can be individually deconvolved has been investigated, but this creates discontinuities in the deconvolved image and makes it difficult to deconvolve faint sources in the presence of a point spread function associated with bright sources in other tiles. A method of feedback between each of the tiles has been developed to deal with these problems. This new approach has been tested on a simulated dataset containing multiple point sources of known intensity. When compared to a standard Högbom deconvolution the tiled feedback version produced a reconstructed image, containing sources up to 2.1 Jy, which agreed to between -0.1 Jy and +0.04 Jy of the standard method across the whole deconvolved image.

The OPS4 GAIA legacy project: transitioning towards a new data management system
Author: Enrico Licata
Affiliation: INAF
ID: P602

Gaia Legacy idea is born to enhance the platform system of Gaia's Big Data science data center in Turin, and to be a center for the management, visualization, processing, manipulation and analysis of large amounts of data that require the development and implementation of innovative systems with exascale approach, guaranteeing high performances. The system responds to the scientific needs of the Gaia INAF community, beyond the core science of the Gaia mission itself, under a multimessenger approach such as characterization of cosmological gravitational waves and degenerate binary systems in the Milky Way.

Moving from an approach where data is stored and processed in real time (OLTP) to an online analytical processing (OLAP), which uses complex queries to analyze aggregated historical data from the OLTP system, we need fast queries and analysis of data from different perspectives.

The system will extend its capability to engineering data collected by space instrumentation for studies of future missions, observation, calibration and qualification of instrumental models. This will be achieved using a hybrid data management system, dynamic data modeling and dedicated data reduction softwares.
We present the Gaia Legacy repository project which goal is the generation of a deep and complete sky, on 4π sterad , as a reference tool and therefore interoperable for the integration of multiband data (from radio to high energies) and multimessenger data (e.g. sources of gravitational waves, neutrinos,...) for efficient data mining aimed at fast multidimensional scientific data exploitation.

The investigation regarding possible different solutions such as, but not limited to, an alternative DBMS to the one already in use, will be conducted in synergy with the activities currently planned within the “Italian National Center in HPC, Big Data and Quantum Computing” and in particular under the IGUC project.

RCSEDv2: algorithm for automatic analysis of individual spectra from large surveys
Author: Evgenii Rubtsov
Affiliation: Sternberg Astronomical Institute, M.V. Lomonosov Moscow State University
ID: P603

RCSEDv2 is the second version of the RCSED catalog (Chilingarian et al, 2017) – A Value-added Reference Catalog of Spectral Energy Distributions https://rcsed2.voxastro.org/. It is based on spectral and photometric data for more than 4 million extragalactic sources collected from several publicly available spectral (SDSS, LAMOST, Hectospec, 2dF, etc.) and photometric (SDSS, UKIDSS, GALEX, WISE, etc.) all-sky surveys in a wide wavelength range. To analyze them, we used the NBursts full spectrum fitting technique developed by the authors (Chilingarian et al, 2007), with additional capabilities for analyzing emission lines. Analysis of such large samples of objects is only possible in automatic mode, for which procedures have been developed for detection, classification and selection of emission lines templates with several possible action scenarios. The properties of the stellar components (age and metallicity) were determined using stellar population models (PEGASE, MILES, X-SHOOTER, etc.). If emission lines were detected in the spectrum, the following scenarios were checked: a single set of lines, a set of narrow lines with a broad lines component (NLR+BLR) and a double set of narrow lines (DP/outflow). The emission lines profiles were described both parametrically and non-parametrically. From the resulting fluxes, gas properties (metallicity and ionization parameter) were assessed using the IZI procedure (Blanc et al, 2015). As a result, for a large number of extragalactic sources, we obtained a complete and homogeneous set of analysis results, which can be used for further statistical analysis and study of the evolution of galaxies.

SpectralFitting: a fresh approach to fitting spectra
Author: Fergus Baker
Affiliation: University of Bristol
ID: P604

Reducing observational data and fitting astrophysical models to the resulting spectra is a multistage process involving a number of different software programs. Many of these programs were designed in an era where spectral resolution was low, and machines were predominantly single threaded. Now, we are in the era of ever higher resolution and high volume astronomy, with machines capable of massively parallel compute, and where data science is increasingly augmented with machine learning techniques: the processes of fitting spectra should be taking full advantage of this to maximise useful scientific output and minimize the technical overhead on the analyst. We present SpectralFitting, an open-source, work-in-progress package for the Julia programming language that aims to modernise the spectral fitting process. Our approach is to create modular, lazily loaded, strictly reproducible and easily distributable software that can fully utilize available computational resources. We aim to unite existing analysis tooling across the electromagnetic spectrum, to provide a single tool that can fit large volumes of multi-wavelength data, with a choice of optimization algorithms and hardware acceleration. This talk will outline the software architecture we are implementing to try and achieve this goal. I will touch upon how we are making it easier to distribute and use already existing C and Fortran spectral models, and how you can interactively define your own. I will also discuss how we are using machine learning methods to augment the fitting procedure, and our plan to build a tool ready for future challenges in astronomical data analysis. We welcome collaboration and comments.
https://github.com/fjebaker/SpectralFitting.jl Universität Stuttgart

Milky Way galaxies-analogues as the isolated AGNs: multiwavelength data incompleteness
Author: Iryna Vavilova
Affiliation: Main Astronomical Observatory, NAS of Ukraine, Kyiv, Ukraine
ID: P605

The Sgr A* is currently inactive. However, several billions years ago a central engine of the Milky Way was typical active nuclei. The proof is that the well-known Fermi bubbles, observable in X-ray and gamma-ray spectra, are thought to be a result of past activity of the supermassive black hole (SMBH). The Milky Way evolved as an isolated galaxy during 10 Gyr: last major merger with the Gaia-Sausage-Enceladuswas at $z \approx 2$. We can observe Milky Way galaxies-analogous, for example, NGC 3521, which is a twinin terms of baryon mass, rotation velocity, scaled disc length, metallicity. Multiwavelength analysis via the CIGALE software reveals ongoing star formation inNGC 3521, with a star formation rate SFR = 1.26 $MSun$ per year, along side substantial dust mass $Mdust$ = 8.36 · 10^7 $MSun$ and neutral hydrogen mass MHI = 1.3 · 10^10 $MSun$. The BPT diagram analysis indicates that it hosts a LINER-type active nucleus. Our 2MIG sample of 61 isolated galaxies at z < 0.05 shows that most AGNs in these galaxies are faint sources with SMBH masses comparable to Sgr A*. However, some isolated galaxies have a changing-look AGN, for example, Mrk 42. Other isolated galaxies exhibit powerful AGNs with clumped dust torus, leading to variability in their X-ray emission. Examples of such galaxies include NGC 6300, ESO 116-018, MCG 02-09-040, NGC 3081. We consider that these variations in nuclear properties within the sample of isolated AGNs are due to various evolutionary stages. So, studying them will enhance our understanding of the Milky Way's past.

Cataloging the first year of Euclid data

Author: Martin Kuemmel
Affiliation: LMU Faculty of Physics
ID: P606

The Euclid satellite is an ESA mission that launched in July 2023. Euclid is working in its regular observing mode with the target to observe an area of 14,000 deg^2 with two instruments, the Visible Imaging Channel (VIS) and the Near IR Spectrometer and imaging Photometer (NISP) down to VIS=24.5mag (10 sigma). Ground based imaging data in griz from surveys such as the Dark Energy Survey and Pan-STARRS complement the Euclid data to enable photo-z determination. Euclid investigates the distance-redshift relationship and the evolution of cosmic structures by measuring shapes and redshifts of galaxies and clusters of galaxies out to redshifts ~2

Generating the multi-wavelength catalogs of Euclid and ground based data is a central part of the Euclid data processing system. In this contribution we discuss the results from the pipeline we implemented to detect objects in the Euclid data VIS and NIR and measure their properties. We show the four different photometric measurements we offer to our users and show that these fulfill the tight requirements on photometric accuracy. We also present the other measurements such e.g. morphology that are included in our catalogs. We list all output products of the cataloging pipeline which go far beyond the object catalogs.

With the first internal and public data releases approaching fast the automatic validation of the results from the various on-the-fly and re-processing campaigns has become a central issue, and we show the software and procedures we developed and implemented to achieve this difficult task with a minimum of interactive work.

Galaxy Detection Energy-Efficient Computation: Deep Learning With Tensor Methods To Speed-Up Astronomical Imaging

Author: Mauricio Solar
Affiliation: Universidad Tecnica Federico Santa Maria
ID: P607

Addressing the "astronomical triad" of volume-AI-energy represents a critical frontier in contemporary astronomy. Mitigating the computational and energy consumption for galaxy classification is important in astroinformatics, given the large amount of data obtained from astronomical surveys. In this research, we focus on the application of tensor factorization to enhance the energy efficiency of deep learning models in astroinformatics, specifically for the task of galaxy classification. Our study introduces a refined implementation of the Faster R-CNN framework, optimized through tensor factorization techniques, to reduce the computational load and energy consumption inherent in processing vast astronomical datasets. By integrating tensor factorization, we systematically decrease the model's complexity and memory footprint, while preserving, and in certain aspects, improving the precision of galaxy morphological classification. The efficacy of this optimized model is evaluated using data from the Galaxy Zoo DECaLS, demonstrating that tensor factorization can significantly enhance computational efficiency without sacrificing accuracy. This research contributes to the field of astroinformatics by offering a methodologically advanced, resource-efficient approach to the analysis of large-scale astronomical data, aligning with the growing demand for sustainable computational practices in the analysis of cosmic phenomena.

Integrating the Pegasus workflow management system into the CCAT Observatory data center

Author: Michael Commer
Affiliation: Institute of Physics I, University of Cologne, Germany
ID: P608

The CCAT Observatory is a facility currently under construction at over 5,600 meters on Cerro Chajnantor (Atacama Desert, Chile). It will host the Fred Young Submillimeter Telescope, a 6-meter diameter submillimeter to millimeter wavelength telescope. Regular operation will lead to approximately 10 PB of raw sky-mapping data in the first five operational years. The CCAT project exemplifies how modern astrophysical research has been driven by the analysis of ever-increasing data amounts. Becoming a data-intensive discipline translates to high-performance-computing demands and an automated-processing infrastructure in order to ensure the streamlining of analysis tasks. Discipline-specific workflow systems (WFS) try to respond to such demands, specifically, pipeline construction, task automation, reproducibility of results, and fault-tolerance. However, WFS can also incur a certain development and administration overhead. This poster presents first experiences to date and a prototype that utilizes the Pegasus WFS. The prototype will embed Pegasus in conjunction with the operations database of the CCAT data center infrastructure.

Overcoming barriers: simplifying accessibility and visualization of gravitational wave data with GW Data Plotter

Author: Panagiotis Iosif
Affiliation: University of Trieste / INFN Trieste
ID: P609

Accessing data released by a particular scientific community can be challenging for researchers from different disciplines. Such a task typically has a significant learning curve and involves knowing how to use specialized software packages. Entering the field of gravitational wave (GW) astronomy for the first time, one will undoubtedly face this obstacle.

GWs are distortions in spacetime caused when massive objects accelerate. Some of the most violent and energetic phenomena in the Universe, e.g. the merging of two black holes or neutron stars, are responsible for producing them. Data from GW events are recorded by the international network of GW detectors, i.e. the LIGO-Virgo-KAGRA collaboration, and released publicly via the Gravitational Wave Open Science Center after a set proprietary period.

Even basic interaction with a public GW data file requires programming skills and adequate familiarization with the tools used by the GW community. To break down this barrier, we present GW Data Plotter, a graphical interface, desktop application, developed to facilitate visualization and basic analysis of GW data without the need to write a single line of code.

Unix philosophy in astronomical data analysis

Author: Peter Teuben
Affiliation: University of Maryland
ID: P610

In this poster we discuss the virtues of using the Unix philosophy in a CLI toolset, with the intent to build flexible and powerful pipelines. The poster will start with reminding the audience of the Unix philosophy and continue with a few short presentations summarizing this approach in a few well known packages, after which we open the floor to the audience.

See also: https://en.wikipedia.org/wiki/Unix_philosophy

The Fornax Initiative: A NASA Astrophysics Science Platform

Author: Vandana Desai
Affiliation: Caltech/IPAC
ID: P611

In the era of big data collected from worldwide telescopes, astronomers face many roadblocks in data analysis. These roadblocks include limited access to storage and compute resources, a lack of expertise in data management and scalable techniques, and insufficient time for the overhead of data, software, and compute management. We will discuss how the Fornax Initiative will address these roadblocks in the context of Open Science, which encourages transparency, reproducibility, and inclusivity in the research endeavor. Fornax is a joint project between the NASA Astrophysics Archives and Goddard to collaboratively create systems, software, and standards for cloud-based astronomical data analysis. We will introduce the three pillars of the effort: (1) Fornax Scientific Components, including client-side libraries, tutorial notebooks, curated astrophysics software environments, and open deployment astronomy cloud services; (2) the Fornax Science Console, a cloud-hosted web-based application that users log into for access to cloud computing, data storage, and interactive data analysis in JupyterLab; and (3) Science Support Systems, including a help desk and training opportunities for researchers interested in using all aspects of the Fornax system.

GPU Accelerated Image Quality Assessment-Based Software for Source Detection

Author: Xiaotong Li
Affiliation: University of Oxford
ID: P612

Fast imaging localises celestial transients using source finders in the image domain. The need for high computational throughput in this process is driven by next-generation telescopes such as Square Kilometre Array (SKA), which, upon completion, will be the world's largest aperture synthesis radio telescope. It will collect data at unprecedented velocity and volume. Due to the vast amounts of data the SKA will produce, current source finders based on source extraction may be inefficient in a wide-field search. In this paper, we focus on the software development of GPU-accelerated source finders based on our Image Quality Assessment (IQA) methods, Low-Information Similarity Index (LISI) and augmented LISI (augLISI). These methods serve as precursors to our transient-oriented source finding method, tLISI. We accelerate the algorithms using GPUs, achieving kernel time of approximately 0.1 milliseconds for source finding in 2048 x 2048 images. The data transfer time can be eliminated when applying the software within a pipeline, as the data is already on the device.

Imbalance Learning in Astronomy

Author: Yanxia Zhang
Affiliation: National Astronomical Observatories, , Chinese Academy of Sciences
ID: P613

Searching for rare objects or phenomena is a crucial task in astronomy. This task falls under the category of imbalanced learning, which presents a significant challenge in the field, particularly because astronomical datasets often contain a large number of examples from one class (such as background noise or common objects like stars) and very few examples from another class (such as rare phenomena like supernovae or black holes). This imbalance can severely distort the performance of machine learning models, resulting in poor generalization for the rare but important classes. In this paper, we summarize various methods for addressing imbalanced learning, along with their advantages and disadvantages. Additionally, we focus on the application of imbalanced learning techniques within the realm of astronomy.

Metadata, Semantics and Data Models Applicable to Data Formats

Classification of Equatorial Plasma Bubbles (EPB) Using Convolution Neural Network (CNN)

Author: Ameh James Adah
Affiliation: University of Jos, Nigeria
ID: P701

Equatorial plasma bubbles (EPBs) are irregular structures of ionosphere plasma density depletions that can range in size from tens to hundreds of kilometers. They arise as a result of depletion of localized electron density in the F-region ionosphere at low and Equatorial latitudes. The study of EPB has become an increasing area of research as it is responsible for disturbance in geomagnetic field, increase risk in navigation system, changes in atmospheric density and severe scintillation in radio wave from Global Navigation Satellite System (GNSS), This study focused on using convolution neural network (CNN) to classify images with the presence or absence of plasma bubbles in the ionosphere. 3800 images collected from the All-sky air glow imager in the Space Environment Research laboratory of the National Space Research and Development Agency, FCT-Abuja, Nigeria, was used for this study. These images were preprocessed and a Convolution Neural Network was used to classify them. The model achieved an accuracy of 98%. Notably, the images with plasma bubbles has a precision score of 98% and a recall score of 96%, while the images without plasma bubbles has score of 97% and a recall score of 98%. These results reveal that the model is highly effective for the classification of images with plasma bubbles in the ionosphere. These findings have significant benefits for space weather monitoring, navigation system and radio communication.

Metadata-Driven Navigation System of DARTS and AI-Powered Data Exploration

Author: Satoshi Nakahira
Affiliation: Japan Aerospace Exploration Agency
ID: P702

DARTS, managed by the Institute of Space and Astronautical Science (ISAS), serves as a space science data archive. It offers over 300 datasets from approximately 40 missions, some dating back to before 1970, covering a wide range of scientific fields. However, over time, variations

in data archiving standards and differences in field-specific policies have led to the data becoming somewhat fragmented, making the archive challenging to navigate for users without specific goals in mind.

To resolve this, we consolidated the information into machine-readable metadata and developed a website and web applications that autonomously function using this metadata. The metadata adheres to the schema.org Dataset format, drawing on various text and HTML sources, with a unique metadata entry created for each dataset and observation mission.

The website, built using a static site generator based on nuxt.js, generates pages directly from the data, including the metadata. We have also developed dynamic web applications, such as a "Dataset Catalog Display" tool that presents the dataset contents in a user-friendly format and enables search functionality, and an "Observation Time Coverage Display" tool that visualizes the temporal coverage of numerous datasets. These tools are seamlessly integrated across the site, allowing users to explore data relationships through various keywords.

In addition, we are developing a data search tool that indexes individual data files and uses faceted navigation to facilitate filtered searches. This tool is also designed for API access, enabling detailed table retrieval by connecting to a database using specific filtering criteria.

With much of the site’s data now machine-readable, we are also exploring AI integration. We are currently working on an interactive assistant to aid users in exploring data within DARTS. This AI system interprets user queries, utilizing language models for search and summarization, and makes API calls to provide features such as: 1) explanations of observation missions, 2) dataset descriptions and suggestions, 3) information on data usage (e.g., code examples), and 4) data search and retrieval, all through agent-based responses.

Through these developments, DARTS is becoming a more user-friendly data archive for a broad audience, promoting new discoveries while also enabling improved transparency in data management.

Optimizing data models and data management solutions to address Gaia data exploitation challenges

Author: Sara Gelsumini
Affiliation: INAF - OATo
ID: P703

Gaia's data for astrophysical research poses significant challenges due to its size and complexity, demanding tailored data models and management solutions. This project - within the activities foreseen by the INAF involvement in the Italian National Center for HPC, Big Data, and Quantum Computing - focuses on the identification of database and filesystem archive solutions, intending to support requests from the astrophysical Italian community on particular scientific cases.

Gaia's dataset includes objects such as CompleteSource (CS), which stores a comprehensive set of data related to a specific source, like a catalog entry, and AstroElementary (AE), which holds calibration data for each observation of each source, along with ancillary tables. Additionally, raw and intermediate Gaia data are stored in GBIN files, which are compressed, serialized Java objects. While this format was chosen for operational constraints, it is challenging to handle directly for the broader scientific community. To improve both data access and analysis outside the data reduction pipelines, we identified the HDF5 to be a suitable format, since it is more structured and human-readable. Furthermore, The HDF5 format allows the nesting of primitive fields, arrays, and complex objects, thereby enhancing accessibility and computational efficiency.

We focus our attention on defining a data model and its related metadata satisfying the most challenging Gaia use cases, and on the choice of the most suitable and performant database for the metadata descriptors. For example, a critical aspect is the ability to pair the source information from CS with the transit data from AE relying on the CrossMatch table, given that each of the billion astronomical objects is observed an average of 100 times over the 10 years of the mission's duration.

This approach represents a crucial step towards creating scalable and efficient systems capable of managing Gaia's datasets, ultimately unlocking the potential to discover new astrophysical phenomena.

Real-time and Near Real-time Processing Pipelines

ALMA Fast Lane: the architecture of real-time data delivery
Author: Alexandros Chalevin
Affiliation: FEV etamax GmbH, ESO
ID: P801

ALMA's Fast Lane is a workflow mode in which the software automatically performs quality assurance on processed observational data and delivers it directly to end users. This functionality accelerates the process and prepares ALMA for future upgrades. In this discussion, we outline all stages of the ALMA data delivery workflow, address remaining challenges, and explore solutions to enable a fully automated data flow for ALMA.

The real-time data processor framework for data handling and analysis of high-energy instruments
Author: Andrea Bulgarelli
Affiliation: INAF/OAS Bologna
ID: P802

Context:

In the multi-messenger and multi-waveform era, space and ground-based facilities implement real-time data acquisition and analysis software to analyze the acquired data as soon as possible and generate scientific results.

Aims:

We implemented the real-time data processor (rta-dp) framework. This framework can be used to implement real-time analysis pipelines and data handling systems to manage high-throughput data streams with distributed applications.

Method:

The rta-dp is based on the ZeroMQ in-memory communication framework to receive input data, share data between distributed processes, and send or receive commands and pipeline configuration.

It is possible to implement a custom data receiver, for instance, using Kafka, or to connect directly rta-dp to a data acquisition system. The rta-dp is available in both Python and C++ implementations. Each pipeline is configured through a JSON file. A Supervisor manages the pipeline and receives the input data, managing low and high-priority data streams. Each Supervisor instantiates one or more WorkerManager following the configuration file and transfers to them the input data. The WorkerManager instantiates one or more Workers using multithreading or multiprocessing features to analyze the data in parallel. The Workers of the same WorkerManager execute the same analysis. They can send the analysis results in output using ZeroMQ and this can become the input for a new rta-dp pipeline. With this flexible architecture, it is possible to create a distributed analysis system customized to the requirements of each scenario.

The rta-dp framework also provides monitoring capabilities of the running processes and sends housekeeping, logging, alarm and informative messages that can be acquired by a monitoring process.

Results:

We are using the rta-dp in several contexts such as the reprocessing of archiving data, exploiting them not only in a real-time context, and in the Online Observation Quality System of the ASTRI Mini-Array. We will use this system for the real-time data quality pipeline of the Cherenkov Telescope Array Observatory. In addition, we are developing data handling systems for gamma-ray scintillator detectors using the rta-dp framework to exploit its flexibility and performance.

Breizorro: Image masking and cataloguing suite
Author: Athanaseus Ramaila
Affiliation: Rhodes University
ID: P803

Breizorro is a versatile software tool designed to streamline image analysis tasks, particularly isolating islands of emission and creating and manipulating image masks (a common application in radio interferometric images). Using customisable parameters, users can create precise binary masks from FITS images. The software offers a range of features for mask manipulation, including thresholding, inversion, gap filling, dilation, and combination. Beyond masking, Breizorro facilitates region creation and manipulation based on the masked image. Users can generate region lists, combine or subtract sections, and effectively catalogue and visualise region attributes. This comprehensive package is a great resource for astronomers looking for an effective image masking tool for data processing pipelines.

Accelerating Source Finding in 3D HI-Datacubes: GPU Parallelization of the SoFiA-2 Smooth and Clip Finder
Author: David Bernhardt
Affiliation: Hochschule Bonn-Rhein-Sieg Institute of Visual Computing
ID: P804

The identification of sources in HI-datacubes is a critical task in astronomical data analysis. In the light of ever growing data sets recorded by large telescope arrays like the Very Large Array or, in the future, the SKA, fast processing of HI-datacubes is more important than ever. We present our work on enhancing the SoFiA-2 pipeline, which employs a parallelized algorithm to significantly improve the efficiency of the smooth and clip finder (S+C finder).

By leveraging modern GPU architectures, we have successfully parallelized the S+C finder, achieving performance enhancements of up to 38 times compared to the traditional CPU implementation. Our approach not only accelerates the processing of large-scale HI-datacubes but also shows, that real-time analysis of mediocre sized HI-datacubbes is possible. Therefore a transition from batch processing to interactive analysis is now possible.

We will discuss the technical details of our parallelization strategy, including optimization techniques and how noise calculations have been ported to GPU code. Additionally, we will present benchmarking results that demonstrate the effectiveness of our GPU-based solution.

Progress in ML-Based Spectral Kurtosis RFI Detector
Author: Dustin Lagoy
Affiliation: National Research Council Canada
ID: P805

Radio astronomy telescopes at the Dominion Radio Astrophysical Observatory (DRAO) are being upgraded with new capabilities to enable cutting-edge science through increased frequency coverage, frequency resolution and time fidelity. Radio frequency interference (RFI) is prevalent in all recorded telescope data and must be removed prior to scientific use, a process historically involving archiving large volumes of high-fidelity data and significant effort by scientists performing offline manual RFI removal. This process would be prohibitively expensive for modern telescopes, including the recently upgraded John A. Galt telescope which produces spectral data at a rate of up to 400 MB/s.

A spectral kurtosis (SK) estimator is a well studied technique that can be used for the detection of RFI in radio telescope data, but traditional approaches suffer from limited accuracy and its use with production telescopes is limited (partly due to the additional power squared products required, which the upgraded Galt telescope now produces). We aim to significantly improve the accuracy of the SK estimator by using machine learning techniques to train an adaptive multi-scale estimator on a telescope's unique RFI environment and to implement the developed estimator into the Galt telescope's real-time data processing pipeline.

Statistics recording and processing in LOFAR2 using modern, cloud-native and web technologies

Author: Hannes Feldt
Affiliation: ASTRON
ID: P806

In the upgraded Low Frequency Array (LOFAR2) telescope, modern, cloud-native object storage solutions, combined with simple data formats are used to record and process statistics. This allows easy access for scientists and operators as well as increased overall robustness.

In LOFAR, various statistics are produced on each of the over 50 stations across Europe. While in LOFAR1, the statistics were only manually recorded on an individual station, they will be recorded continuously in LOFAR2. Afterward, the statistics will be collected and automatically processed centrally for quality control or to be provided as data products to scientists and operators.

The distributed nature of LOFAR and the owner structure of international stations require a robust and flexible way to record, collect, and access the statistics. In LOFAR2 the MinIO Object Store is used to manage data storage on the station and centrally, providing a robust way to replicate the data between multiple locations. This setup not only decreases the risk of data loss during network or system failures over multiple hours but also provides an easy and flexible way to access and replicate the statistics in various places.

The BurstCube Data Analysis Pipeline
Author: Joseph Asercion
Affiliation: NASA/GSFC
ID: P807

BurstCube is a 6U CubeSat (10 x 20 x 30 cm), operated from April to September of 2024, designed to rapidly detect and disseminate information about gamma-ray bursts (GRBs) to both expand sky coverage and supplement existing observatories. To do this, the mission required a lightweight yet robust data processing pipeline system which could be easily deployed, require minimal maintenance, could perform standard gamma-ray burst analysis to convert raw instrument data to FITS standard data files, and could handle transfers to science archiving and alert systems. In addition, the software was designed to be easily reusable by similar scale missions performing gamma-ray analysis in the future. This poster details the multi-level software architecture of the BurstCube Data Analysis Pipeline, pipeline deployment instructions for future use, and provides an end-of-mission "lessons learned" retrospective in light of adaptations that needed to be made to accommodate changing in-flight mission parameters.

PypeIt: The Python Spectroscopic Data Reduction Pipeline
Author: Kyle Westfall
Affiliation: University of California Observatories
ID: P808

PypeIt (pronounced "pipe it") is a python-based software package that provides a series of end-to-end scripts that process raw astronomical spectroscopic observations to yield science-grade data products with minimal user interaction. PypeIt addresses a significant need in the astronomy community, providing a comprehensive data-processing product that leverages modern, industry-driven support infrastructure for open-source code development. Although PypeIt implements a fixed processing structure, it provides critical flexibility via built-in entry points that enable users to introduce new data from new spectrographs and manipulate the performance and control flow via an exhaustive set of runtime parameters. Its core development team is centered at the University of California Observatories, but it has been an open-source project since its inception in 2018. Its various pipelines now support processing data from more than 50 spectrographs at 18 different observatories. In this talk, I will briefly summarize PypeIt's capabilities and its near-term development goals.

Connecting Zenodo and Analysis Facilities: a JupyterLab Extension Middleware
Author: Michael Zengel
Affiliation: The University of Alabama, CERN
ID: P809

As a part of the ESCAPE Project, the Virtual Research Environment was created to improve the technical experience of physicists by aggregating infrastructure and software into a single, cloud-based platform, following Open Source and FAIR best practices. Pursuing the goal of developing a fully-scoped analysis facility, we aim to integrate a general purpose, public repository into the Virtual Research Environment, for the purpose of preserving any digital object with a unique identifier. Zenodo is a software developed by CERN as a catch-all repository for European Council funded research, as motivated by Open Source guidelines. For the scope of this project, the implementation of a Zenodo feature into a JupyterHub-based system has been the focus. More specifically, this has been done via the development of a Jupyter Lab extension, designed for simple user experience and interaction with the Zenodo software from any Jupyter-based environment.

The development of a visual interface with the Zenodo REST API Command Line Interface increases the ease with which it can be used. Furthermore, its integration into a Jupyter framework significantly reduces the need for local storage use in uploading or downloading various data to Zenodo, catalyzing a fully cloud-based interface.

In-Situ High Performance Visualization for Astronomy & Cosmology

Author: Nicola Tuccari
Affiliation: University of Catania
ID: P810

The Astronomy & Cosmology (A&C) community is presently witnessing an unprecedented growth in the quality and quantity of data coming from simulations and observations. Writing results of numerical simulations to disk files has long been a bottleneck in high-performance computing. To access effectively and extract the scientific content of such large-scale data sets appropriate tools and techniques are needed. This is especially true for visualization tools, where petascale data size problems cannot be visualized without some data filtering, which reduces either the resolution or the amount of data volume managed by the visualization tool.

A solution to this problem is to run the analysis and visualization concurrently (in-situ) with the simulation and bypass the storage of the full results. In particular we use Hecuba (https://github.com/bsc-dd/hecuba), a framework offering a highly distributed database to stream A&C simulation data for on-line visualization. We will demonstrate the Hecuba platform integration with the ChaNGa (Charm N-body GrAvity solver) high performant cosmological simulator (https://github.com/N-BodyShop/changa/wiki/ChaNGa) and the in-situ visualization of its N-body results with the ParaView (https://www.paraview.org/) and the VisIVO tools (https://visivo.readthedocs.io/en/latest/).

Acknowledgements: This work is funded by the European High Performance Computing Joint Undertaking (JU) and Belgium, Czech Republic, France, Germany, Greece, Italy, Norway, and Spain under grant agreement No 101093441 and it is supported by the spoke "FutureHPC &

BigData" of the ICSC – Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing – and hosting entity, funded by European Union – NextGenerationEU"

The Roman Data Monitoring Tool

Author: O. Justin Otor
Affiliation: Space Telescope Science Institute
ID: P811

The Nancy Grace Roman Space Telescope (Roman) is a NASA flagship mission planned to launch in late 2026. Roman's large-scale surveys will produce 30 PB of science data over its five-year primary mission, an order of magnitude more than the combined yields of all active NASA flagship missions in astrophysics. To keep pace with this volume of data, Roman's Science Operations Center (SOC) at the Space Telescope Science Institute is developing the Roman Data Monitoring Tool (RDMT). The RDMT is an automated software system that will assess the scientific accuracy of calibrated, detector-level data from the Roman Wide Field Instrument and the reference files used to calibrate them in real time. Because these products are stored in different computing environments, separate arms of the tool operate in the Amazon Web Services commercial cloud and on-premises at the SOC. Each will track data quality metrics in a database and alert the SOC to any anomalous updates. Select results and algorithms will be made publicly available so the community can stay up to date with the SOC's calibration analysis.

Automated Deployment of Radio Astronomy Pipeline on CPU-GPU Processing Systems: DiFX as a Case Study

Author: Ophélie Renaud
Affiliation: IETR, INSA Rennes
ID: P812

The next generation of radio telescopes, such as the Square Kilometre Array (SKA), will require peta-FLOPS processing power to handle the massive amount of data acquired. A new generation of computing pipelines are required to address the SKA challenges leading to the integration of the pipelines on a dedicated heterogeneous High-Performance Computing (HPC) system. The tight real-time and energy constraints are driving the community to study the use of hardware accelerators like GPUs in the computing system. Allocating resources, such as processor times, memory, or communication bandwidth, to support complex algorithms in such systems is known as an NP-complete problem. Existing tools such as Dask and Data Activated 流 Graph Engine (DALiuGE) rely on dataflow Model of Computation (MoC) and have proven to be an efficient solution to specify parallel algorithms and automate their deployment. These models are efficient programming paradigms for expressing the parallelism of an application. However, state-of-the-art dataflow resource allocation only

targets CPUs and usually relies on complex graph transformations resulting in a time-consuming process. This paper introduces an automated dataflow resource allocation method and code generation for heterogeneous CPU-GPU systems. Our method efficiently and quickly manages pre-scheduling graph complexity, and optimizes the dataflow model to the target architecture. Experimental results show that the proposed method improves resource allocation and speeds up the process by a factor of 13 compared to the best existing method on a basic architecture. Moreover, the execution times of the obtained implementations are comparable to those of manual implementations.

Evaluation of effectiveness of RFI removal algorithms for searching for fast radio transients with the SKA

Author: Raghuttam
Affiliation: University of Manchester
ID: P813

Radio Frequency Interference (RFI) has always been a problem for radio observations but is increasing with the growing number of telecommunication satellites, terrestrial transmitters, active Radar systems, mobile phones, satellite televisions, etc. In time-domain observations, this could potentially lead to a large number of false positives, cause issues during periodicity search, completely obscure the signal of interest and much more. RFI Mitigation (RFIM) therefore becomes an important process in any kind of pulsar or transient searches. Generally, choosing the right algorithm for RFIM is done by inspection by the observer and cleaning of RFI stops when it is simply ‘good enough' with no metrics used or based on any statistics. Currently, most of the astronomers perform the data analysis offline which provides discretion to try and decide on a combination of techniques which works the best for a particular scenario. With modern systems planned with the SKA telescopes, such as SKA-PSS (Pulsar Search Subsystem), which would run in real-time, it will not be scientifically rigorous just to choose any algorithm without strong reasoning. We have tried to quantitatively define the metrics that can be used to decide the best mitigation practice.

Automating testing pipelines using Notebooks

Author: Rosa Diaz
Affiliation: Space Telescope Science Institute
ID: P814

We developed a testing pipeline using Notebooks and CI infrastructure. This setup has several advantages for small non-critical projects compared to other testing pipelines, including low-cost infrastructure, ease of setup, ease of understanding by collaborators, straightforward organization of stages of processing and log outputs, and easy-to-identify issues with the run. Possible disadvantages include sharing runners with other projects that limit resources, significant network bandwidth when downloading the same data to multiple runners, and

individual runner resource constraints. Here, we will describe our pipeline and how it can complement standard software testing, particularly when input from non-developers is required to address problems.

Design and Implementation of a High-Throughput Data Transfer System for the CCAT Observatory

Author: Thanh Liem Ngo
Affiliation: Institute of Physics I - University of Cologne
ID: P815

The CCAT Observatory hosts the Fred Young Submillimeter Telescope (FYST) at 5,612 meters in the Chilean Atacama Desert and is set to generate approximately 10 PB of raw data over five years, originating from its main instruments, Prime-Cam and CHAI. The CCAT infrastructure currently under development involves an elaborate system that synchronizes data-transfer between the FYST storage buffer and downstream archives, the latter primarily located at the Regional Computing Center (RRZK) of the University of Cologne.

To achieve a data transfer rate of up to 8 terabytes per day, we plan to use the bbcp protocol. Drawing on the ALMA Observatory's experience—transferring data from the Atacama Desert to Germany via parallel streams of ~50-gigabyte files—we will assemble raw data into DataTransferPackages of similar size. These packages are sent downstream in parallel streams, optimizing bandwidth and ensuring data integrity through checksum verification upon arrival before they are archived on tape.

To streamline this task chain, an Operations Database (ops-db) will track all the necessary auxiliary observatory information. The data transfer process is modular, with services for packaging, transfer, integrity-verification, archiving, and deletion, scheduling Python-Celery tasks on the computers that hold that data at the respective stage of the task chain. Initial testing on local infrastructure at RRZK is underway. Access to a data-generating telescope emulator, located in Chile at 22 milliseconds in latency time from FYST, will allow for the simulation of long-distance transfers, so that our system can be optimized before prime time in late 2025. We adhere to best software-development practices, including unit testing and continuous integration. Established monitoring tools supporting our efforts are Prometheus, Loki, and Grafana. This poster outlines the implementation of the module ensemble in terms of architecture and interfaces.

PyMerger: Detecting Binary Black Hole merger from Einstein Telescope Using Deep Learning

Author: Wathela Alhassan
Affiliation: Nicolaus Copernicus Astronomical Center, Warsaw, Poland
ID: P816

We present PyMerger, a Python tool for detecting binary black hole (BBH) mergers from the Einstein Telescope (ET), based on a Deep Residual Neural Network model (ResNet). ResNet was trained on data combined from all three proposed sub-detectors of ET (TSDCD) to detect BBH mergers. Five different lower frequency cutoffs (F_{low}): 5 Hz, 10 Hz, 15 Hz, 20 Hz, and 30 Hz, with match-filter Signal-to-Noise Ratio ($MSNR$) ranges: 4-5, 5-6, 6-7, 7-8, and >8, were employed in the data simulation. Compared to previous work that utilized data from single sub-detector data (SSDD), the detection accuracy from TSDCD has shown substantial improvements, increasing from 60%, 60.5%, 84.5%, 94.5% to 78.5%, 84%, 99.5%, 100%, and 100% for sources with $MSNR$ of 4-5, 5-6, 6-7, 7-8, and >8, respectively. The ResNet model was evaluated on the first Einstein Telescope mock Data Challenge (ET-MDC1) dataset, where the model demonstrated strong performance in detecting BBH mergers, identifying 5,566 out of 6,578 BBH events, with optimal SNR starting from 1.2, and a minimum and maximum D_{L} of 0.5 Gpc and 148.95 Gpc, respectively. Despite being trained only on BBH mergers without overlapping sources, the model achieved high BBH detection rates. Notably, even though the model was not trained on BNS and BHNS mergers, it successfully detected 11,477 BNS and 323 BHNS mergers in ET-MDC1, with optimal SNR starting from 0.2 and 1, respectively, indicating its potential for broader applicability.

PyMerger: ADARI: Visualizing the quality of VLT data

Author: Cesar Enrique Garcia Dabo
Affiliation: ESO
ID: P817

ADARI (Astronomical DAta Reporting Infrastructure) is a system designed for creating graphical reports of astronomical data so that the quality of these products can be assessed. It has been designed from the ground up to be backend-agnostic, meaning the same ADARI code can be sent to a web plotting API, or a code-based plotting API, with no alteration.

Quick data inspection is an important feature in data reduction systems. The use cases range from quality control at the telescope, advance quality checks prior to delivering data to the scientists as well as data inspection for users running the pipelines at their home institutes. The goal of ADARI is to deliver the same experience and code for data visualization for all the environments, either running automatically in the Paranal Observatory environment or at the PI premises.

ADARI contains a library of pre-packaged Panels for astronomical instruments with similar data types. The goal is to minimize the amount of code need to support the reports for a new instrument and have a uniform system that is easier to maintain.

Most of the time the reports are generated as part of the execution of a data reduction workflow implemented with EDPS, the new ESO system for automatically organising data from ESO instruments and for running the reduction pipelines on them.

Other

Ten (or more) reasons to register your astronomy research software with the Astrophysics Source Code Library (ASCL, ascl.net)
Author: Alice Allen
Affiliation: Astrophysics Source Code Library/University of Maryland
ID: P901

This presentation covers the benefits of registering astronomy research software with the Astrophysics Source Code Library (ASCL, ascl.net), a free online registry for software used in astronomy research. Indexed by ADS and Clarivate's Web of Science, the ASCL currently contains over 3500 codes, and its entries have been cited over 16455 times. Registering your code with the ASCL is easy with our online submissions system. Making your software available for examination shows confidence in your research and makes your research more transparent, reproducible, and falsifiable. ASCL registration allows your software to be cited on its own merits and provides a citation that is trackable and accepted by all astronomy journals and journals such as Science and Nature. Adding your code to the ASCL also allows others to find your code more easily, as it can then be found not only in the ASCL itself, but also in ADS, Web of Science, and Google Scholar.

A software system for supporting untargeted HI surveys - from observations to public data release
Author: Austin Shen
Affiliation: AusSRC
ID: P902

The Australian SKA Regional Centre (AusSRC) has been supporting the Widefield ASKAP L-band Legacy All-sky Blind surveY (WALLABY) to produce advanced data products. This has involved the development and deployment of post-processing pipelines for observation quality checking, mosaicking, and source finding, back-end software for automating the deployment of such pipelines, and custom web applications to enable scientists to interactively construct catalogs from raw data. Through these software tools, we have helped to prepare the data products for two public data releases and three associated publications for the WALLABY survey in its Pilot Phase. In the poster we provide a high-level overview of the workflow, our chosen software stack, and other features that enable us to streamline the process of generating advanced data products from observations.

The Asteroid Detection Polar and ESA's Synodic Orbit Visualisation Tool

Author: Dario Oliviero
Affiliation: ESA / PDO / NEOCC
ID: P903

Asteroid observability from the ground depends on the object distance to the Sun and to the observer, on the object's Sun-Earth phase angle, on the object shape and on its surface reflectivity properties. Additionally, several magnitude systems have been proposed in the last decades to model the visual magnitude of the object based on the mentioned parameters, as for example the H, G system, the H, G1, G2 system or the H, G12 system. Independently of the magnitude system used, there is a three-dimensional representation of the geometrical locus of equal visual magnitude when this value is constrained, which we have called the detection polar. We analyse the shape of this geometrical locus for the H, G magnitude system and its applicability to the representation of the detectability of an asteroid in its trajectory. We also present ESA's Synodic Orbit Visualisation Tool and provide examples of its use to analyse the graphical representation of a NEO trajectory and the observability region for a given limiting visual magnitude defined by the detection polar.

Given the H, G magnitude system and a limiting visual magnitude Vlim the detection polar computation results from the intersection of two families of curves. The first set of curves is given by the phase angles, having the Sun and the observer as focal points and resulting in arcs of circumference passing by both focuses. The second family of curves is derived from constant values of the distance product from the asteroid to the two focuses as derived from the H, G system equation. These are represented by the so-called Cassini ovals. Defining the relative magnitude B as Vlim-H, the results of the intersections of both families are a new set of curves for varying values of the parameter B. Starting with a negative value of B two heart-like solutions are located at each of the focal points with the shape tip pointing in the direction of the opposition effect. As the value of the relative magnitude grows the shapes size increases accordingly. When B reaches a value of 1.6685 mag (for G=0.15) the two shapes make contact and they transform into two other shapes; an external one that keeps growing and an internal one that joins the two focal points and that decreases in size as the relative magnitude grows. The detectability region is described by the area between both shapes. At this point, no visibility constraints other than the mentioned magnitude system are discussed. However, it is expected that the derived solution close to the Sun will be affected by a low solar elongation as seen from the Earth, thus limiting the possibilities of observing that part of the detection polar.

All those constraints and shapes can be easily simulated in ESA's Synodic Orbit Visualisation Tool (SOVT), a software tool which allows the graphical representation of the mentioned detection polars extended to 3D shapes, in conjunction with the representation of the co-rotational motion of the object in space. Examples of such representation will be provided in the poster.

The SOVT thus allows analysing in a very direct and intuitive way when and how a NEO would become observable from Earth by a given telescope.

Navigating Astrophysics Literature: Harnessing AstroBERT and UAT

Author: Dhruv Sondhi
Affiliation: University of Western Ontario
ID: P904

The advent of LLMs or large language models such as ChatGPT has propelled research towards using these AI models for understanding and summarising research papers. There is a huge number of research papers being published daily and keeping them in check is becoming more and more difficult. Currently, there are state-of-the-art models such as AstroBERT, a large language model, using the BERT architecture, for astrophysics that is built and trained by NASA/ADS (NASA Astrophysics Data System). This model is being used in the astronomy domain to assist in automatically identifying astrophysical entities of interest and to classify research papers in different subdomains of astrophysics. In addition to this, there are tools like the Universal Astronomy Thesaurus (UAT for short) which is an open, interoperable and community-supported thesaurus which aims to unify the existing divergent and isolated Astronomy and Astrophysics thesauri into a single high-quality, open thesaurus formalising astronomical concepts and their inter-relationships.

Our goal is to integrate this high-quality knowledge about astrophysical entities from the UAT which contains specialized vocabulary, into the dataset. This developed dataset would be used to fine-tune and train AstroBERT. This helps us create a rich collection of documents that contains a single “root” word as well as other “synonyms” (or different ways of expressing) of this word, allowing the model to learn these different contextual embeddings and provide higher confident predictions.

Our expectations involve improving the AstroBERT model's current performance by modifying the dataset to include duplicated sentences which contain these “root” words. We also aim to keep the original sentences to further improve the confidence in the prediction of the correct type of astrophysical entity. This in turn will allow us to have a better model that can eventually be used to classify research papers in different subdomains of astrophysics.

Developing the optimal cross-matching algorithm for the Gaia and BDB data

Author: Ekaterina Malik
Affiliation: Institute of Astronomy RAS, Moscow, Russia
ID: P905

Binary and multiple systems are important parts of the stellar population. Depending on the observational method, they might have various sets of observational parameters, identification systems and designations in different catalogues. Thus, it is necessary to identify an object in different datasets. This could be done with a unified catalogue containing all binary and multiple systems with proper cross-identification. This catalogue is named an index catalogue of the Binary Star Database (BDB)—Identification List of Binaries (ILB). Though there still remains the necessity of updating the catalogue, in particular, the results of the Gaia mission could significantly augment its data. This work focuses on developing an optimal cross-match algorithm for the ILB and Gaia-based data.

"Astronomy and Computing": evolution, current status and future plans

Author: Fabio Pasian
Affiliation: INAF - O.A. Trieste
ID: P906

Elsevier's "Astronomy and Computing" (ASCOM) is a peer-reviewed journal publishing algorithms, computing methods, software and data relevant to astronomy. Established in 2013, it was groundbreaking as there were no alternatives to publish refereed papers on methods and software. Born thanks to efforts of active members of the ADASS community, ASCOM still represents a natural venue for papers that deserve a more in-depth discussion than what conference proceedings can afford.

During the past decade the journal has extended its coverage to solicit, besides standard and review papers, also notes on practice, white papers and software/data releases. Additionally, the journal scope has expanded to cover areas receiving increasing interest (e.g. Machine Learning, Applications of Artificial Intelligence to astronomy, etc.). From time to time, Special Issues on specific topics are published.

Other important improvements concern a more flexible and efficient policy towards Open Access (OA), including allowing to publish in ArXiv the final text of papers (Green OA), and defining agreements with the major scientific institutions to cover the publishing fees for Gold OA.

This poster provides a summary of the current ASCOM activities, statistics and access policies, and the plans for the near future.

Automating the configuration of INAF radio telescopes' control system using Ansible

Author: Giuseppe Carboni
Affiliation: INAF - Osservatorio Astronomico di Cagliari
ID: P907

The operation of large radio telescopes relies on stable control systems to ensure flawless observations and accurate data collection.

The three INAF radio telescopes' control systems, based on the Alma Common Software (ACS) framework, run on a dedicated cluster of Linux workstations.

These machines run the control software and provide an interface for observers to control the telescopes through commands and graphical interfaces, as well as some other tools that can be useful during the observing operations.

This poster presents an automated procedure that makes use of Ansible to configure the aforementioned control workstations.

Ansible, an open-source automation tool, enables a reliable and repeatable provisioning process, ensuring uniform configuration across all workstations, according to each telescope specific requirements.

This procedure can also be used to create virtual machines that can be distributed to the members of the development team, simplifying the writing of new features of the control software by replicating the operational environment.

Moreover, the provisioning procedure is continuously validated through a set of automated tasks on GitHub Actions.

This process ensures that any update to the Ansible configuration procedure is automatically tested online, providing a safe and reliable way to restore or replace a control software machine in case of eventual failure.

Science Data Centre Architecture for LOFAR2.0
Author: Hanno Holties
Affiliation: NWO-I ASTRON
ID: P908

This contribution presents the architecture and roadmap for the Science Data Centre system for LOFAR2.0. The upgrade of the LOFAR instrument is accompanied by a program for realizing a Science Data Centre that supports astronomers, providing access to data as well as generating advanced data products in the LOFAR2.0 era. The Science Data Centre system is an evolution from the services that have been in use since the start of LOFAR science operations in 2012. Existing capabilities are modernized and in many cases replaced, while new capabilities are added to address the challenges of analyzing data from the even more versatile and data-intensive future telescope. These include a new application for submitting proposals for observing time and data processing services, an automated data processing environment, a new generation of data processing pipelines, support for the community to return advanced data products to the central archive, and VO compatible data discovery services. Technologies and operational activities are adopted to lower the boundaries for astronomers to work with LOFAR2.0 data and keep cost and environmental impact within bounds while stimulating science output. User experience design and full chain integration of services aim to take away roadblocks for astrophysicists of all levels of expertise as well as for the LOFAR observatory for working with a complex instrument, the demanding data it generates, and to contribute to the data processing capabilities.

Nonlinear fitting of undersampled discrete datasets in astronomy

Author: Igor Chilingarian
Affiliation: Center for Astrophysics -- Harvard and Smithsonian
ID: P909

Data analysis and interpretation often relies on an approximation of an empirical dataset by some analytic functions or models. Actual implementations usually rely on some kind of a non-linear multi-dimensional optimization algorithm, typically Levenberg-Marquardt or other flavors of Newton gradient methods. A vast majority of data in optical astronomy and infrared astronomy are represented by values on a discrete grid because the actual signal is sampled by regularly shaped pixels in the light detectors. Here we come to the main problem of nearly all widely used implementations of non-linear optimization methods: the function that is being fitted is evaluated at central pixel positions rather than integrated over the pixel areas. Therefore, the best-fitting set of parameters returned by the minimization routine might not be the best representation of the observed dataset, especially if a dataset is undersampled. For example, a central pixel of a 1D Gaussian with a dispersion of 1 pix (2.36 pix FWHM; so not too badly undersampled) will be about 4.2% lower than its central evaluated value if integrated. To handle this effect properly, one needs to perform numerical or analytic integration of a model within pixel boundaries. We will discuss possible computationally efficient solutions and present a preliminary implementation of a non-linear fitting using Levenberg-Marquardt minimization that correctly accounts for the discrete nature of the data. We also discuss potential astronomy-related applications such as PSF fitting of stars in undersampled images and emission line fitting in 1D spectra.

The Gaia DR4 bulk-download repository: when worlds collide

Author: Jos de Bruijne
Affiliation: European Space Agency
ID: P910

We present a science-driven proposal for the contents, data format, metadata, partitioning scheme, and delivery scheme of the Gaia DR4 bulk-download file repository. It successfully merges different sets of constraints from the point of view of efficient large-scale data processing and more traditional astronomy standards.

ASPIS prototype: goals, organisation and results from the CAESAR project
Author: Marco Molinaro
Affiliation: INAF-OATs
ID: P911

The CAESAR project (Comprehensive Space Weather Studies for the ASPIS Prototype Realisation) was funded by ASI (Italian Space Agency) and INAF (Italian National Institute for Astrophysics) to develop the prototype of ASPIS (ASI SPace weather InfraStructure). This contribution aims at reporting design considerations, challenges and final status of the ASPIS prototype, which allows for the study of the chain of phenomena from the Sun to Earth and planetary environments. The database within ASPIS is aimed at handling the heterogeneity of metadata and data while storing and managing the interconnections of various Space Weather events. On top of the database, user interfaces, including a graphical web interface and an advanced Python module (ASPIS.py), have been developed to facilitate data discovery, access, and analysis. The high-level metadata, to inform the discovery phase in the database, have been collected using an internally developed tool, ProSpecT (Product Specification Template). This tool utilises JSON Schema and JSONForms to create a web interface to guide the data providers in describing their "Products" and generate a JSON object with the necessary metadata. Approximately 100 product descriptions in JSON format have been collected and used to create wiki-like documentation pages, as well as to help examine formats and metadata details for the implementation of the database prototype. The graphical web interface helps the users discover, filter, and access the database content, while ASPISpy provides more advanced analysis tooling. Moreover, ASPISpy sits on top of commonly used Python packages like SunPy, scikit-learn, matplotlib to help integrate research analysis with other tools and research domains. The database has been built with the adherence to FAIR principles in mind and with the idea of making it easily interoperable with other research data infrastructures in the Space Weather or sibling research domains.

Integral Legacy Archive: a new approach to explore High-Energy Space Radiation
Author: Monica Fernandez
Affiliation: ESAC
ID: P912

The Science Archives, managed by the ESAC Science Data Centre (ESDC), are the ultimate repositories for ESA missions' data, ensuring the preservation, accessibility, and usability of invaluable scientific information for researchers worldwide.

In developing the Integral Science Legacy Archive (ISLA), significant efforts were made to involve mission scientists. The goal was to enhance the user experience by transitioning from traditional data forms to specialized Integral Science Portals. This change provides a more seamless and intuitive interface for conducting scientific research and facilitating advanced data analysis.

In the poster, we will showcase these specialized portals and demonstrate how this innovative approach has inspired improvements in other ESDC archives.

One portal features a sophisticated, grammar-based smart search tool, making it easier to explore INTEGRAL observations and time series around specific targets or coordinates.

Another section serves as a centralized access point for Multi-Mission Observer Data Archive (MMODA) products, which are generated by the ISDC through an advanced data processing pipeline. Users can now effortlessly access light curves, spectra, mosaics, and other essential scientific datasets through this section of ISLA.

We will conclude the poster by highlighting the importance of remaining technologically up-to-date to develop robust and maintainable software tools capable of addressing complex scientific requirements

Two Fundamental Julia Packages: Astrometry.jl and FITS.jl
Author: Paul Barrett
Affiliation: The George Washington University
ID: P913

Julia is a programming language designed for high performance scientific computing. This poster presents the current state of two Julia packages for astronomy, Astrometry and FITS. The Astrometry package implements basic astrometric algorithms that are fast and simple to use. It also contains a subpackage that implements a complete set of SOFA functions in Julia. The FITS package currently supports the standard compressed and uncompressed header-data units (HDUs). Its interface is similar to the original PyFITS module, where access to HDUs and cards is via array syntax.

A crude but efficient pipeline for JWST MIRI imager : the case of SN1987A
Author: René Gastaud
Affiliation: CEA/IRFU/DEDIP
ID: P914

Observations in imager mode of supernova sn1987A with MIRI have been made in July 2022, just after Performance Verification Phase of the instruments onboard space mission Webb (aka "James Webb Space Telescope" or JWST). We quickly realised that the official pipeline was too limited at that time and does not give the high quality level we need. So we developed in GDL a simple but efficient pipeline to process our data from scratch. We made significant improvements in data processing for :

- ramps to slopes (better Allan Variance)
- mosaicing (better resolution in maps)

- odd/even removal (less SNR on maps)

Even if those improvements are now known and published some are still not in the official pipeline. Thanks to our experience as data scientists, following KISS philosophy, we wrote the custom tools working on raw data, circumventing the constraints from official infrastructure, in a month. Being able to code with our favorite language was a way to quickly test new ideas and have high quality results, without waiting for official product.

DAEPO Projects of National Astronomical Data Center (NADC)

Author: ShanshanLi
Affiliation: National Astronomical Observatories, CAS
ID: P915

Mega-science astronomical projects along with large-scale astronomical simulations generate massive amounts of data daily. Currently, these data are primarily designed and utilized by astronomers, physicists, and computing experts. However, some individuals and groups are exploring ways to harness these data for education and public outreach (EPO), recognizing their value as an ideal resource for such endeavors. Among all astronomy EPO activities, data-driven astronomy education and public outreach—abbreviated as DAEPO—is particularly special and important. It benefits from the advancements in Big Data and Internet technology, offering flexibility and diversity. The National Astronomical Data Center (NADC) in China has been actively designing and operating such projects for over a decade. The projects are now mainly divided into two categories: one is the EPO ecosystem centered around the WorldWide Telescope (WWT), and the other is the China-VO citizen science platform, which includes a variety of citizen science projects based on astronomical data. I will present the achievements, experiences, and summarize these projects.

Exploring New Instrument deGradation Models and Algorithms (ENIGMA)

Author: Stephane Beland
Affiliation: LASP - University of Colorado Boulder
ID: P916

The Solar Radiation and Climate Experiment measured Total Solar Irradiance (TSI) and Solar Spectral Irradiance (SSI) from 2003 to 2020. The Solar Irradiance Monitor (SIM) instrument measured SSI from 200nm to 2400 nm on a daily basis. The current SORCE-SIM instrument degradation correction uses a measurement equation derived from accessible telemetry, the known instrument refraction geometry, inter-detector comparisons, and inter-spectrometer comparisons. While the current degradation model captures much of the long-term trending, some of the parameters were adjusted without well-defined physical interpretations.

The degradation model used for SORCE-SIM is not unique. We're reporting on work to Explore New Instrument degradation Models and Algorithms (ENIGMA) to address issues with the

current model and with the derived corrected irradiances. The development of enhanced SORCE-SIM measurement equations permits the evaluation, and potential inclusion, of degradation mechanisms not captured in the present model.

We present an initial updated degradation model which improves our ability to construct a composite irradiance time series in combination with TSIS-SIM, whose measurements overlap with SORCE for 2 years and has well-defined uncertainties. Combining the two datasets, allowed the construction of a consistent composite irradiance time series from 2003-2024.

The Merged Pan-STARRS/Gaia Photometric Reference Grid
Author: Stephen Gwyn
Affiliation: Canadian Astronomy Data Centre
ID: P917

Since its release in 2016, Pan-STARRS has become the definitive grizy photometric reference for the northern sky. The Pan-STARRS photometry forms the basis of the photometric calibration for CFHT-MegaCam and Subuaru-HSC. A comparison with the GAIA DR3 G/BP/RP photometry (transformed into the Pan-STARRS system) shows reassuringly small residuals between the two systems for the most part. However, it also indicates that there are small, isolated but significant flaws in the photometry of both catalogs. Further, the comparison breaks down near the galactic plane where stellar metallicity and reddening effects invalidate the GAIA to Pan-STARRS photometric transformations. The GAIA DR3 spectra can remove some of these problems, but only for the brighter stars. These stars tend to saturate in 4m and 8m class telescopes, limiting their utility in photometric calibration. The Merged Pan-STARRS/Gaia Photometric Reference Grid (MPGPRG), presented here, uses the GAIA spectra to re-calibrate the Pan-STARRS photometry. The GAIA continuous spectra are converted into sampled spectra using an MCMC method. Multiplying the sampled spectra by the well-understood Pan-STARRS passbands produces synthetic Pan-STARRS down to the depth of the GAIA spectra. This relatively shallow synthetic photometry catalog is then used to re-calibrate the Pan-STARRS catalog. Smoothing the re-calibration over suitable spatial scales minimizes the flaws in both the Pan-STARRS and GAIA photometry. The resulting catalog provides uniform photometry.

Acceleration of identification process time of small solar system bodies(SSSBs) on the wide field imaging survey data
Author: Tadafumi Takata
Affiliation: National Astronomical Observatory of Japan
ID: P918

Extracting and identifying small solar system bodies (SSSBs) is one of the important roles of data analysis of wide field and multiply imaged data taken in large imaging survey. We need

to develop and operate the efficient method to find out the SSSBs among a huge number of detected sources. As reported in Ootsubo et al.(2022), we use photometric and imaging data obtained with the Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP) survey for the SSSB extraction. For the next step, we develop the system based on the Apache Spark with photometric data of detected sources for all frames obtained with the HSC-SSP Wide survey taken between March 2014 and January 2020.

By this development, we achieve much faster process speed (5-10 times totally) for identifying more than 1.4 million detection(events) for 248,000 SSSBs from the HSCSSP data of 255 nights. The final goal of our project is to develop a system with faster hyper-speed database management system on a non-huge parallel computing system.

Generating VPHAS+ HiPS with Montage

Author: Thomas Boch
Affiliation: Observatoire Astronomique de Strasbourg, Université de Strasbourg, UMR 7550, 11 rue de l'Université, 67000 Strasbourg, France
ID: P919

We will present our process to generate VPHAS+ HiPS datasets using the Montage tool. While Hipsgen is heavily used to generate 99% of available public HiPS, it is not the most suitable solution for image surveys showing strong background variation. As HiPS tiles geometry can be described with a WCS, it makes usage of Montage possible to build a background-corrected HiPS. Our poster will detail each step involved in the generation of the 5 VPHAS+ HiPS datasets available.

Implementing SIAv2 Over Rubin Observatory's Data Butler

Author: Tim Jenness
Affiliation: Vera C. Rubin Observatory / AURA
ID: P920

The IVOA Simple Image Access version 2 protocol defines an easy way to provide community access to a collection of data. At the Vera C. Rubin Observatory we currently enable ObsTAP access to our data holdings via an ObsCore export or view of our Data Butler repositories. This approach does come with some deployment constraints, such as requiring pgsphere and compatibility with our CADC TAP implementation, so recently we decided to see whether we could instead provide an SIAv2 service that talks directly to our Data Butler. This poster describes our motivation, implementation strategies, and current deployment status, as well as discussing some metadata mismatches between the Butler data models and SIAv2.

Dynamic Planetary Orbital Diagrams with Microsoft Word VBA

Author: Travis Stenborg
Affiliation: University of Sydney
ID: P921

Microsoft Word allows insertion of shapes into word processing documents. Importantly, such shapes can be manipulated by Word's built-in programming language VBA (Visual Basic for Applications). Planetary orbital element data, their rates of change, and planetary dynamics equations can all be implemented in VBA. This approach was used to produce Word documents that display dynamic planetary orbital diagrams in documents. I.e., diagrams that render orbital paths and planetary positions that are accurate to the system date the document is opened on.

The VBA code combined a Julian date calculator with key orbital elements data (semi-major axis, eccentricity, inclination, longitude of perihelion, longitude of ascending node, mean longitude) and real-time derivation of associated anomalies (mean, eccentric, true). Kepler's equation was solved via a modified Newton-Raphson method. Solar system state was calculated and stored in custom Word document properties, then drawing objects' dimensions, multi-axis rotations and positions adjusted appropriately. Associated VBA macros and an example Word document are available publicly via GitHub.

A Content-Driven Strategy for Roman's Science Platform: Enabling Low-Barrier Access and Collaboration

Author: Tyler Desjardins
Affiliation: Space Telescope Science Institute
ID: P922

Science platforms offer a unified user experience for discovering, exploring, accessing, and analyzing data. As observational astronomy enters an era of petabyte-scale data, cloud-based science platforms will become essential interfaces between researchers and data. In this poster, we present our efforts to develop content specifically designed to provide resources and engage the community with NASA's next flagship astrophysics mission, the Nancy Grace Roman Space Telescope (Roman). With its wide field of view and unprecedented survey speed, Roman's Wide Field Instrument is expected to generate an estimated 30 PB of data during its five-year primary mission. To prepare for Roman's scientific discoveries, the Space Telescope Science Institute (Roman Science Operations Center) is developing a user-friendly science platform focused on low-barrier access and collaboration. We outline the creation of simulated observations, Jupyter tutorials, science workflows, and other tools for the cloud-based platform. These resources demonstrate the platform's capabilities, ease of use, and help prepare users for a new era in observational astronomy.

Communication Strategies and Achievements : Data-Driven Astronomy Education and Public Outreach (DAEPO) Projects of NADC

Author: 杨涵溪
Affiliation: National Astronomical Observatories, Chinese Academy of Sciences
ID: P923

NADC(National Astronomical Data Center) has launched a series of DAEPO(Data-Driven Astronomy Education and Public Outreach) projects, in virtue of the extensive reach of the internet and a wealth of astronomical data. This study explores how NADC uses a variety of media platforms and communication methods to promote the DAEPO projects, with a particular focus on the role of these methods in enhancing public interest and participation in astronomy. By analysising the data of users, the research confirms that tailored media strategies significantly enhance public participation, especially among younger amateurs. At the end of the study, suggestions were put forward to optimize the DAEPO project in China, such as strengthening cooperation and interaction with other media or institutions, and discovering more popular media tools.

Searching for Fireballs in Lighting, but Finding Other Things

Author: Joshua Grajales
Affiliation: Columbia University
ID: P924

The Geostationary Lightning Mapper (GLM) is an instrument onboard the Geostationary Operational Environmental Satellite (GOES) 16, 17, and 18 weather satellites and has been used to detect lightning, forecast weather, track severe storms, and study bolides (exploding meteors). GLM is of particular interest to NASA’s Asteroid Threat Assessment Project (ATAP) as they seek to leverage its data to investigate how incoming bolides interact with the Earth’s atmosphere. As part of ATAP’s asteroidal studies, a machine learning-based bolide detection and light curve generation pipeline has been developed. In addition to the bolides and meteorological phenomena detected by this pipeline, there exists an additional class of objects: Near Field Glints Streaks (NFGS). Given the recent development of the bolide detection pipeline and the relatively infrequent occurrence of these NFGS, little is known about these objects. This exploratory data analysis project aims to study NFGS by gathering a series of statistics on their motion and occurrence rates. This study can improve the bolide detection pipeline by helping better distinguish NFGS from other meteorological and asteroidal phenomena but may also lead to a novel method of observing Earth orbiting objects. To conduct this study, we developed a python-based identification pipeline for NFGS objects that leverages both image-processing and statistical algorithms to isolate streak-like objects in GOES GLM data. After development of the prototype, we then integrated the code into the bolide detection codebase and processed several years of GLM data using the computing resources of NASA’s Advanced Supercomputing Facility (NAS).

MWA Data Archive: A Decade of Radio Astronomy, 54 PB of Data

Author: Mouriyan Rajendran
Affiliation: Curtin university
ID: P925

The Murchison Widefield Array houses 54 PB of raw radio astronomy data in its archive. This poster illustrates how such a massive volume of data is stored, processed, and served to users worldwide. It also highlights how the MWA follows VO standards to ensure interoperability.

Birds of a Feathers

Built To Last?
Author: Mark Nicol
Affiliation: UK Astronomy Technology Centre (STFC, UKRI)
ID: B101

Proposal for a Birds of a Feather (BoF) discussion around the challenges facing building software for observatories, telescopes and instruments whose operational lifetime is often measured in decades.

Developing software for modern observatories brings its own challenges. Work on the design and architecture of the software may begin several years before the observatory or its instruments have been fully designed, and certainly long before they have been fully developed and commissioned.

Key parts of the observation control software may be expected to work reliably for the next 20 to 30 years or even longer.

How do experienced teams, who have worked with different telescopes and observatories, manage some of the issues that stem from the inherent unpredictability of technology, changing requirements, and the operational environment over such a long-time frame.

What lessons have been learnt from designing and building the software of the current generation of observatories, telescopes, and instruments and how are these being applied to the next generation of facilities currently in progress?

The list below reflects some of the possible shared challenges for discussion, it is expected that others may emerge during the other sessions:

Living with changing requirements - such as user needs, operational necessities, and the integration of modern technologies, while managing hardware and software obsolescence.

Evolving standards and regulations including a changing security threat landscape.

Long term maintenance and support – best practices for knowledge transfer and maintaining a healthy codebase.

Designing for the unknown – building software for a massively complex project that is itself evolving and may change radically between conceptualization and when it becomes operational.

External economic, financial, political, and environmental factors

Usability and User Experience in astronomical Software
Author: Kai Polsterer
Affiliation: HITS gGmbH
ID: B401

The most crucial requirement to scientific software is output correctness. Large amounts of development resources are necessary to achieve this goal, leading aspects like interface design, usability and user experience to be put in second place. Ironically, not paying attention to user-friendliness creates the monster that we swore to defeat: an error-prone system. A user that suffers from cognitive overload, confusion or being irritated due to bad software design is more likely to make careless mistakes or to misuse the system.

During this BoF session, we invite you to discuss the role of usability and user experience in astronomical software. We will point out aspects that affect the user-friendliness of a system in a positive or negative way, and review our current software portfolio regarding these aspects. Furthermore, we will speak about how to address usability and user experience in our development procedures to ensure future software pieces for astronomers to be user-friendly.

Strategies for heterogeneous processing and archiving.
Author: Yan Grange
Affiliation: ASTRON
ID: B601

By pushing the boundaries of discovery, several modern astronomical observatories are daunted with the challenge of processing and storing up to many petabytes of data per year. To tackle the challenges of processing and storing those massive amounts of data, organisations are increasingly adopting heterogeneous and distributed infrastructures for computing and archiving to leverage the opportunities for scalability and collaboration. Typically, such architectures can be built based on custom-built systems, national infrastructures, up to and including commercial clouds.

In this session, representatives from various facilities will share their experiences, plans and strategies for managing those complex processing and archiving systems. After a brief introduction, we will transition into a moderated discussion, encouraging participants to share their insights, ask questions, and engage in a collaborative exchange of ideas and solutions.

The goal is to foster a productive dialogue that can help the community better understand and address the unique challenges of managing heterogeneous processing systems and archives in the context of modern astronomical research.

Possible topics of the discussions could be

- Architectural design and implementation of distributed computing systems for data processing and analysis
- Effective strategies for integrating and orchestrating heterogeneous resources
- Best practices in data management and archiving approaches
- Emerging technologies and trends that may shape the future of heterogeneous computing in astronomy
- Operational concerns, like scheduling and provisioning

What New Data Formats is the Community Using? What New Data Models does the Community Need?

Author: James Tocknell
Affiliation: AAO, MQ
ID: B701

FITS is in its 5th decade, and ADASS has traditionally had a FITS BoF. That is not this BoF. This BoF aims to find out what other formats are out there: the good, the bad, and the ugly. Whether it is VOTable encoded in to Parquet or simulations written to HDF5, this BoF aims to discover what formats and models the community is using, thereby fostering collaborations and driving interoperability.

General-Purpose Spectroscopic Data Reduction and Analysis Tools

Author: Kyle Westfall
Affiliation: University of California Observatories
ID: B901

Motivated by recommendations from the Astro2020 Decadal Survey, developers across many software projects have met over the past year to discuss current and future needs for general-purpose spectroscopic data reduction and analysis tools. Specifically for data reduction software, the discussion has centered on the development of more general-purpose tools (e.g., the AstroPy-affiliated packages specutils and specreduce) to enable users to more easily build pipelines for unsupported and newly commissioned instruments. This is needed to address the fact that most available packages focus on specific instruments or surveys.

In this BoF, we will briefly review the recommendations from Astro2020, the discussions from follow-up meetings, and ongoing development efforts stemming from these discussions. We will then hold an open discussion to collect information about parallel efforts and potential new collaborations that can be formed to address specific, timely needs of the community.

Software doesn't write itself: Prioritising Equity, Diversity, & Belonging to improve software output

Author: Simon O'Toole
Affiliation: AAO
ID: B902

You might not think that discussions around diversity and inclusion are relevant to you. You might not feel that you are in a privileged position. Yet if you are white or male, or both, you have probably had more or better opportunities than many others. What we are talking about here is not just socioeconomics, but the invisible and institutional systems that extend advantages to some groups over others. This BoF will explore different perspectives of these privileges, with a particular focus on research and software. The fundamental question we aim to address is: "how do you make sure that you and everyone in your team has equal and fair access to opportunities?" Come along to a safe, non-judgemental discussion around how we build a community where ensuring everyone has the opportunity to contribute their creativity leads to more fulfilling careers, excellence in engineering, and better research outcomes.

Tutorials

Processing and analyzing XMM-Newton data with ESA Datalabs: A collaborative approach.
Author: Aitor Ibarra Ibaibarriaga
Affiliation: Telespazio UK for ESA, Madrid, Madrid, Spain
ID: T301

The XMM-Newton satellite is one of the most successful missions ever built for ESA. It has been operating as an open X-ray observatory since the beginning of 2000, producing high quality scientific results since then. The Science Analysis System (SAS) is the application used for processing the data obtained with the scientific instruments on board XMM-Newton, an indispensable tool that has been helping scientists in the publication of nearly all refereed scientific papers published up to date.

Recently, SAS has been adapted to be able to run within the ESA Datalab platform. A set of Jupyter notebook threads have been developed within the XMM-Newton team to allow newcomers and experienced scientists to explore the X-ray domain.

ESA Datalabs is the perfect platform to, in a collaborative way, exploit the vast amount of XMM-Newton data gathered during the last 24 years.

The goal for this tutorial will be the introduction to the participants to the analysis of the XMM-Newton data using ESA Datalabs as a collaborative platform.

The first step when thinking about User Experience: Set-up an UX Vision
Author: Kevin Tai
Affiliation: Astron
ID: T401

This tutorial is aimed at anybody who wants to implement High-Level UX (User Experience) thinking in their software development process. A glance at how we used to set up a UX vision group and UX templates at Astron. Gathering stakeholders and collaborating on:

- Persona(s)
- User story maps
- User journey maps

Establishing a clear UX vision is the foundational step in creating a successful User Experience. This vision serves as a guiding star for the design process, ensuring that all design decisions align with overarching goals and objectives. A well-defined UX vision helps create a cohesive and consistent user experience, considering user needs and objectives. By defining a strategic framework for designers and stakeholders, it enables teams to deliver meaningful and valuable experiences that meet user needs and organization goals.

The Advanced Scientific Data Format (ASDF)
Author: Nadia Dencheva
Affiliation: Space Telescope Science Institute
ID: T701

ASDF is a language-neutral file format for serializing scientific data, in use by JWST, DKIST and Roman. It is integrated in data analysis and visualization tools. It has human-readable hierarchical metadata structure, followed by binary blocks and is designed to be easily extensible and customizable. The tutorial will be a hands-on session focusing on reading, writing and creating ASDF files in Python. It aims to show the advantages of ASDF over other formats by working through an example challenging in other formats (WCS). Participants will gain understanding of how ASDF works at the format and library level.

Programming the GPU on your laptop - is it easy, is it useful?
Author: Keith Shortridge
Affiliation: K&V
ID: T901

This tutorial is aimed at ADASS attendees who may have sat through numerous talks about how GPUs make everything faster, wondered about making use of the GPUs to speed up compute tasks, and then somehow never found the time to actually try it. The aim is to give people who have no experience with GPU programming a kick-start towards trying it for themselves on their own laptops. The tutorial will be based around a small set of example C++ command line programs that perform calculations on 2D data, all of which run on MacOS, Linux, and Windows. Attendees will be able to build, run, modify, and experiment with these programs, seeing how the GPU performance compares with the CPU. The structure of the programs will be explained together with details of the code, with an emphasis on what is actually going on in the GPU when it runs. Metal and Vulkan versions of each program will be provided. Metal is Apple's current GPU infrastructure, and Vulkan - a descendant of OpenGL - will run on almost any recent GPU. (Those wondering about CUDA should note that CUDA and systems based on it need an Nvidia GPU, while most laptops use different GPUs.) As the underlying operations they perform are the same, seeing what these two quite different systems have in common will provide some insight into the internal workings of a modern GPU. Comparing the GPU and CPU code will show which programs gain from using the GPU and which may not. There will be a bit of fun stuff at the end with graphics, and everyone will take away an amount of potentially helpful example code.

The latest release of the example code, and the latest installation instructions, can be found at: https://github.com/KnaveAndVarlet/ADASS2024_GPU

Focus Demos

JupyterLab extension: FireFly
Author: Denis Zaytsev
Affiliation: Caltech, IPAC
ID: F201

Data visualization plays a crucial role in the everyday work of data scientists and ML engineers. It is important to present data in different visual forms at all stages of the ML process: when data scientists prepare the training data set, inspect how well the model converges during the training, and then when it is time to validate the trained model and analyze the results of inference. Existing data visualization tools available in JupyterLab are mostly limited to a static data representation and do not provide enough interactivity.

In this demo, we will look at how FireFly integrates with JupyterLab to enhance your AI/ML and astronomical data science processes by providing interactive visualization components integrated into the JupyterLab notebook. We will see how FireFly visualizations can be helpful for preliminary data inspection and cleaning, exploratory data analysis, feature engineering, and model inference results analysis. The demo will showcase JupyterLab FireFly extensions for displaying tabular data, FITS images, plotting charts, visualizing HiPS maps, or overlaying data on the reference images.

The ESA Near-Earth Objects Coordination Centre Python Interface
Author: Eduardo Peleato
Affiliation: ESA / PDO / NEOCC
ID: F501

The Near-Earth Objects Coordination Centre (NEOCC), part of the ESA's Planetary Defence Office, is dedicated to monitoring, tracking and assessing the risks associated with near-Earth objects (NEO).

The NEOCC offers a variety of services and tools, including public access to data through its portal and related APIs. The NEOCC Python Interface is a powerful tool designed to streamline access to critical data from the NEOCC API. The demonstration will highlight how this Python library simplifies the retrieval and analysis of near-Earth object data like orbital data and physical properties data, and how to efficiently integrate NEOCC data into third-party libraries like ESA's flight dynamics software GODOT. Furthermore, it will be demonstrated how to use it for retrieval and analysis of advanced information about asteroids, ephemeris and other ancillary data such as the risk list and the close approach list. Finally, a simulation of an Apophis 2029 intercept mission will be showcased.

Using LSDB to enable large-scale catalog distribution, cross-matching, and analytics
Author: Neven Caplar
Affiliation: Rubin Observatory/University of Washington
ID: F601

In recent years, the exponential growth of large survey catalogs has introduced new challenges in the joint analysis of astronomical datasets, particularly as we move towards handling petabytes of data. Our demo will showcase the latest advancements in our Large Survey DataBase (LSDB) framework. The framework utilizes a particular hierarchically sharded spatial partitioning of large datasets, using Parquet to store the data. This approach facilitates efficient and scalable cross-matching and analysis of big datasets.

In this demo, we will explore the new features in LSDB; such as support for nested Pandas/Dask, making it easier to work on time-domain data and spectral data by storing observations from the same astronomical objects in the same dataframe row. We will demonstrate how users can start their analysis on small subsection of sky and easily scale up their analysis after initial testing. We will showcase the cross-matching ability across large datasets and demonstrate real-world applications by applying analysis functions on complete wide-sky synoptic datasets. We will highlight our collaborations with our partners (such as STScl, IPAC, Rubin, and CDS) to provide various catalogs in this format and show how we can utilize the Fornax cloud platform to work with diverse datasets in a unified cloud-based framework.

XRADIO: Xarray Radio Astronomy Data Input Output
Author: Jan-Willem Steeb
Affiliation: NRAO
ID: F701

The advent of next-generation radio interferometers—ALMA-WSU (Atacama Large Millimeter Array Wide Band Sensitivity Upgrade), ngVLA (Next Generation Very Large Array), and SKA (Square Kilometre Array)—will increase astronomical data volumes by orders of magnitude. Current pipelines for ALMA and the VLA rely on CASA (Common Astronomy Software Applications) and store data in MSv2 (Measurement Sets v2). This approach, utilizing considerable custom software, faces maintenance challenges and scaling limitations. To address these issues, we present a new data schema, MSv4 (Measurement Set v4), implemented in the open-source Python package XRADIO (Xarray Radio Astronomy Data IO). This initiative represents a collaborative effort between the National Radio Astronomy Observatory (NRAO), European Southern Observatory (ESO), National Astronomical Observatory of Japan (NAOJ), and Square Kilometre Array Observatory (SKAO), combining expertise from leading astronomical institutions.

The MSv4 contains data for a single spectral window, polarization setup, and observation setup within a fully self-describing structure, allowing for finer partitioning as needed.

Collections of MSv4, termed PS (Processing Set), facilitate deployment across distributed computing environments. Departing from MSv2's relational tables, MSv4 employs labeled n-dimensional arrays.

XRADIO leverages off-the-shelf technology to ensure scalability and maintainability. It relies on Zarr for efficient storage and serialization, while Xarray provides in-memory data representation as NumPy arrays or lazy Dask arrays, complete with dimensions, coordinates, and attribute labels.

This focus demo, presented through interactive Jupyter Notebooks, will:

- Explore the PS and MSv4 schemas
- Demonstrate how to easily convert legacy MSv2 datasets to a PS
- Demonstrate efficient data selection techniques
- Showcase data visualization methods
- Illustrate parallel processing capabilities

By adopting these modern tools and approaches, we aim to equip the radio astronomy community with a robust framework capable of handling the data challenges of the next generation of interferometers.

Exploring Space Weather connections with the ASPIS prototype archive
Author: Mirko Stumpo
Affiliation: INAF
ID: F901

Advancing Space Weather Science and predictions requires the understanding of the linkage of the different physical processes that occur simultaneously or sequentially in many application domains (Sun, Interplanetary space, Earth’s environment, other planets, hazards). In order to meet the needs of the Space Weather community, the ASI SPace weather InfraStructure (ASPIS) prototype has been designed and set up by a partnership at the national level in Italy, with efforts from INAF, ASI, INGV, INFN, and seven Universities (Aquila, Calabria, Catania, Genova, Perugia, Rome Tor Vergata, Trento).

This infrastructure includes a database of many heterogeneous resources/products (calibrated data, derived and L3 data, models) pertaining to different domains, a user interface and a Python package that can be used to explore Space Weather phenomena and “chain” them into events that span from the Sun through the interplanetary space to the Earth’s atmosphere up to the ground (and other bodies in the solar system).

Connecting and exploring different kinds of data and model outputs related to different phenomena occurring in the Heliosphere and driven by solar activity is not trivial due to the different formats, scopes and resource collectors available. In this sense, the ASPIS prototype has homogenized data and metadata and provides a new way to explore and interact with Space Weather data.

The goal of this focus demo is to show the audience the various contents of ASPIS, guiding the attendees in exploring the resources and tools provided by the infrastructure.

The audience will be shown how to find, access and explore data and model resources in the Space Weather domains using a new infrastructure designed on purpose and the related tools.

Going from discovery through visualization and then access to the data, investigation of the Space Weather phenomena and events will be showcased using a graphical interface and Python scripting/coding through the dedicated ASPISpy package and general, interoperable tools.