



2015 ASTRONOMICAL DATA ANALYSIS SYSTEMS

AND SOFTWARE CONFERENCE

25th – 29th October 2015

Rydges World Square, Pitt Street, Sydney, NSW, Australia

POSTER ABSTRACT BOOKLET

P001 David Aikema

University of Calgary

Integrating Globally Distributed Resources into the CyberSKA Platform

The CyberSKA platform offers astronomers an online portal / gateway for data intensive research. Integrating a social network environment with e-science applications and data management systems, CyberSKA provides a platform for collaboration by distributed teams of researchers, for use of integrated applications, and for sharing and remote visual analytics of multi-dimensional data sets which may be of large size. The portal is used by over 500 astronomers distributed around the world working on SKA science pathfinder projects. Here we describe recent changes to the platform to improve the user experience and also enable the platform to function in a globally distributed and scalable manner. These changes, and experimentation with a global testbed separate from the production system, enable CyberSKA to serve as a platform to test ideas for global distribution of SKA data and services. Progress towards our next-generation remote visualization tool called CARTA is also described, with support for extensibility through a plugin-based architecture while retaining an HTML5-based interface.

P002 Anastasia Alexov

Space Telescope Science Institute (STScI)

Integrating Single Sign-On into the STScI Archive: Lessons Learned

The Space Telescope Science Institute (STScI) migrated a handful of web services to a Single Sign-On (SSO) solution in 2014, using the Shibboleth and CAS software solutions. SSO was implemented in order to consolidate usernames and passwords used for a variety of web services, to improve



security at STScI and to better the user experience. It took an additional year to integrate the STScI Archive into the SSO Framework, including multiple web services and a non-web service authentication back end which allowed for one set of user credentials. The SSO Framework was expanded to allow external users to register with STScI and use those SSO credentials to access multiple services at STScI, including the Archive. We took great care in migrating and informing over 13,000 STScI Archive users of their new SSO credentials and new user interfaces. We coordinated help desks between IT, the instrument science groups and the Archive to have a successful, seamless transition of the STScI Archive users to their new SSO credentials.

We outline our SSO architecture, hurdles, lessons learned and implementation solutions which we have taken in order to migrate STScI's Archive services to using SSO.

P003 Alice Allen

Astrophysics Source Code Library

Astrophysics Source Code Library ver 3.0

Co-authors: G. Bruce Berriman, Infrared Processing and Analysis Center, California Institute of Technology; Kimberly DuPrie, Space Telescope Science Institute; Robert J. Hanisch, National Institute of Standards and Technology; Jessica Mink, Harvard-Smithsonian Center for Astrophysics; Robert Nemiroff, Michigan Technological University; Judy Schmidt, Astrophysics Source Code Library; Lior Shamir, Lawrence Technological University; Keith Shortridge, Australian Astronomical Observatory; Mark Taylor, University of Bristol; Peter Teuben, Astronomy Department, University of Maryland; John Wallin, Middle Tennessee State University

The Astrophysics Source Code Library, started in 1999, moved to a new infrastructure in 2014 with enhancements developed in response to feedback from users and publishers. With one-click author search, flexible browsing options, and a simple form-based submission process, the ASCL offers a better experience for users. Since the introduction of the new platform in mid-2014, users have submitted nearly 100 codes, more than in all previous years combined. Data sharing options, including the ability to pull all of the resource's public data in JSON and XML, provide new ways to collaborate with the resource. The ASCL now houses information on more than 1000 codes and its entries are increasingly used for citation, with over 15% of its entries cited, up from 7.5% in January of last year. Exciting opportunities to collaborate have been presented to the ASCL, including participation in the 2nd Workshop on Sustainable Software for Science: Practice and Experiences and a request from the AAS to form a special interest group on software publishing. This presentation will demonstrate the new capabilities of the ASCL and discuss its growth and recent outreach and collaborations.



P004 James Allen

University of Sydney

The SAMI Data Reduction Pipeline

The SAMI Galaxy Survey is the first of a new generation of integral field spectroscopic surveys, using multi-object instruments to observe thousands of galaxies. Reducing the raw data from these observations presents a number of unique challenges, ranging from processing metadata to balancing the requirements of different science goals. I will present the pipeline used to process SAMI Galaxy Survey data, which builds on existing software packages to deliver a consistent high-quality data product for the survey team.

P005 Maria Arevalo-Sanchez

ESA/ESAC

The New ESA Hubble Science Archive

The ESA Hubble Science Archive moved from ESO to ESAC Science Data Centre (ESDC) in 2012. At this new home, the archive's front-end has been completely re-engineered, and embedded into the common framework used for other science archives at ESAC, to ensure long preservation and maintenance of the Hubble data: over 1.1 million observations from 10 different scientific instruments that conform a treasure of astronomical data.

As a consequence of the re-engineering work and new technologies applied, all the public HST data, Hubble Legacy Archive and high-level science data products are available to the user in a carefully designed user interface served from a robust system, with many extra key functionalities. Among others, the new ESA Hubble Science Archive includes an accurate search on sources in the field of view of observations, showing precise footprints projected onto the sky and also enhanced data previews of images, interactive spectral plotting and FITS header metadata inspection. The new archive offers connectivity to common astronomical tools over VO protocols and fully compliance as well with SIAP and SSAP, and has been designed to easily integrate the recently released Hubble Source Catalog.

We present here the new archive's architecture and complete set of new features that are currently available to the astronomy community interested in Hubble data.

P006 Ji-Hye Baek

Korea Astronomy and Space Science Institute

Difference Image Application for SDO

We have developed a difference image application for Solar Dynamics Observatory (SDO) data at Korea Data Center for SDO. This application provides functions such as search engine, data download, and difference images. It makes running difference images for all selected data and could



save them in local devices. It handles Flexible Image Transport System (FITS) to download and process difference images.

P007 Carlo Baffa

INAF - Osservatorio di Arcetri

SKA CSP Controls: Technological Challenges

C.Baffa, E.Giani, S.Vrcic, M.Mickaliger, C.Williams

The Square Kilometre Array (SKA) project is an international effort to build the world's largest radio telescope, with eventually over a square kilometre of collecting area. For SKA Phase 1, Australia will host the low-frequency instrument with more than 500 stations, each containing around 250 individual antennas, whilst South Africa will host an array of close to 200 dishes. The scale of the SKA represents a huge leap forward in both engineering and research & development towards building and delivering a unique instrument, with the detailed design and preparation now well under way. As one of the largest scientific endeavours in history, the SKA will bring together close to 100 organizations from 20 countries.

Every aspect of the design and development of such a large and complex instrument requires state-of-the-art technology and innovative approach. This poster (or paper) addresses some aspects of the SKA monitor and control system, and in particular a choice of the framework to be used as a communication infrastructure for SKA monitor and control system. At the SKA workshop in April 2015, SKA monitor and control community has chosen TANGO Control System as a technology of choice for the implementation of the SKA monitor and control.

Work is on the way to incorporate TANGO Control System in SKA. We have started this process for the SKA Central Signal Processor.

In particular we now have:

- a uniform class schema proposal for all sub-Element systems
- a timed command approach to reach a millisecond coordination for a subset of the SKA-CSP.

In this work we introduce our proposals and the early results of the prototype development

P008 Jorgo Bakker

ESAC/ESA

Herschel Data Processing System - Managing Long Term Software Development with Changing User Expectations and Development Environments

Starting in 2002, the Herschel Data Processing software is a long term and largely distributed development project. Up to now, it has successfully supported Instrument Level tests, pre-flight characterisation of instruments, systematic processing of observations as well as multiple bulk



reprocessing campaigns during operations, bulk processing campaigns during post-operations and it provides interactive analysis software for both astronomers and instrument scientists.

Currently the Herschel project is in post-operations phase and the ongoing software development is geared towards adjusting the software to ensure that we have the best possible legacy products in the Archive. We also work towards having our software preserved in a way that will survive end of the current post-operations phase.

Along the way, we had to manage significant changes in the user expectations, developer expectations, underlying software libraries and development tools. We will summarise what major changes we faced, the challenges that arose and how we successfully coped with them in the context of a distributed development.

J. Bakker, on behalf of the Herschel Science Ground Segment consortium
<http://herschel.esac.esa.int/twiki/bin/view/Public/DpHipeContributors>

P009 Pascal Ballester
ESO

Data Flow Tools at the Very Large Telescope

Authors: P.Ballester, T.Bierwirth, S.Castro, V.Forchi, Y.Jung, L.Lundin, S.Zampieri

Affiliation: European Southern Observatory, Karl-Schwarzschildstr. 2, D-85748 Garching, Germany

We present a suite of tools demonstrating the main phases of the Data Flow System at the ESO Very Large Telescope, which are used during observation preparation, observation execution, and for data processing. The selection of tools shows the recent developments of the VLT Data Flow System in view of unifying the operational support of the instrumentation set of the VLT.

The unified GUideCam Tool (GUCT) is a Java-based observation preparation tool currently offered for VISIR, HAWK-I and VIMOS. The tool makes use of the [Aladin desktop application](#) and allows the user to visualize the instrument Field-of-View on a sky image, define the central pointing of the telescope, plan observations that involve a sequence of telescope offsets, select telescope guide stars, and create ESO compliant finding charts. Furthermore, GUCT can directly interact with the ESO Phase 2 Proposal Preparation in order to attach finding charts or to propagate acquisition template parameters, such as guide star coordinates, to the Observing Block.

The Night Log Tool (NLT) is a web application to automatically generate detailed reports on the observatory operational activities during night observation. These reports are automatically generated, collecting information from various sources, and completed by the operator, reducing human intervention to the minimum. The NLT has been developed within the Grails framework, using javascript and jQuery in order to provide a modern and dynamic look and feel. Since 2013, all Paranal telescopes are equipped with the NLT. La Silla telescopes have been upgraded in 2015.

ESO-Reflex is an environment that provides an easy and flexible way to reduce VLT/VLTI science data using the ESO pipelines. ESO-Reflex is based on the concept of a scientific workflow, built upon the Kepler workflow system, and is used for the development of data reduction workflows based on the

[ESO Common Pipeline Library for all new VLT instruments](#). The data reduction cascade is rendered graphically and data seamlessly flow from one processing step to the next. The data organization necessary to reduce the data is built into the system and is fully automatic. It is distributed with a number of complete test datasets so that users can immediately start experimenting and familiarize themselves with the system.

P010 Ugo Becciani

Istituto Nazionale di Astrofisica

Advanced Environment for Knowledge Discovery in the ViaLactea Project

The Milky Way galaxy is a complex ecosystem where a cyclical transformation process brings diffuse baryonic matter into dense unstable condensations to form stars, which produce radiant energy for billions of years before releasing chemically enriched material back into the InterStellar Medium in their final stages of evolution. Although considerable progress has been made in the last two decades in the understanding of the evolution of isolated dense molecular clumps toward the onset of gravitational collapse and the formation of stars and planetary systems, a lot remains still hidden.

The aim of the European FP7 ViaLactea project is to exploit the combination of all new-generation surveys of the Galactic Plane to build and deliver a galaxy scale predictive model for star formation of the Milky Way. This model will be used as a template for external galaxies and studies of star formation across the cosmic time.

Usually the essential steps necessary to unveil the inner workings of the Galaxy as a star formation engine (such as the extraction of dust compact condensations or robust reconstruction of the spectral energy distribution of objects in star-forming regions) are often carried out manually by the astronomer, and necessarily over a limited number of sources or very restricted regions.

To explore very large regions a new framework is implemented using advanced visual analytics techniques, data mining methodologies, machine learning paradigms and VO-based data representation and retrieval standards. All such specialized tools are integrated into a virtualized computing environment, resulting as an efficient and easy-to-use gateway for the scientific stakeholder community. An overview of the methodologies and technologies, able to fulfil the scientific expectations of the project is summarized as follow:

- Database and Virtual Observatory Infrastructure. The ViaLactea Knowledge Base (VLKB) is implemented containing i) the Radio Datacubes with search and cutout services and ii) Hi-Gal catalogue sources and related band merged information. The VLKB allows easier searches and cross correlations between ViaLactea data using the software tools the user community will have at its disposal. Consuming the VLKB alongside VO available resources is possible through the implementation of the TAP service so that the project's community can exploit the data without the need to continuously retrieving and downloading external resources.
- Data Mining Systems. The main tools are related to: i) compact source extraction to obtain a more refined version of band-merged catalogues based on the positional cross-match

among sources at different wavelengths; ii) filamentary structure detection to refine and optimize the detection of the edges of filamentary structures; and iii) source kinematical distance estimation combining all available information from Galactic rotation curve, spectroscopic survey data in molecular gas lines or 3D extinction maps in the near and mid-Infrared.

- 3D Visual Analytics Systems. A 3D-aided visual analytics environment allows the astronomer to easily conduct research activities using methods for multidimensional data and information visualization. Real-time data interaction are performed to carry out complex tasks for multi-criteria data/metadata queries on the VLKB, subsample selection and further analysis processed over the Science Gateway, or real-time control of data fitting to theoretical models.
- Science Gateway. Science Gateway technologies offer a collaborative web-based environments to allow the astrophysicists to run the aforementioned applications with little concern for learning and managing the required infrastructures that execute them. Furthermore the gateway allows users to share tools among the members of the ViaLactea consortium and to the whole astrophysics community at the end of the project. The adoption of an open source framework (WSPGRADE/gUSE) employed by a large community and the membership within a gateway alliance (the STARnet federation) ensures the maintainability of the gateway.

These tools will be able to access data products as well as libraries of millions of radiative transfer models necessary for the science analysis in an integrated environment. The emerging field of visual analytics brings data analysis and visualization into a single human-in-the-loop process. New technologies such as new 3D images of our galaxy with interactive data manipulation capabilities will provide powerful analytical, educational and inspirational tools for the next generation of researchers.

P011 Thomas Boch

CDS - Observatoire de Strasbourg - CNRS/Unistra

Aladin Lite: Status and Perspectives

Thomas Boch, Pierre Fernique

Aladin Lite is a lightweight version of the Aladin tool, running in the browser and geared towards simple visualization of a sky region. Launched in 2013, it has quickly gained traction and interest from various projects. This poster will give an overview of the current status of Aladin Lite.

We will detail how Aladin Lite is used, both inside CDS and by external projects, in different contexts including visualization of results, center part of a discovery portal, visualizer for outreach images, etc. We will also discuss about some recent additions and new developments: revamped and improved documentation, support for progressive catalogues, visualization of MOC (Multi-Order Coverage maps), easy access to Simbad, VizieR and NED data, preview of FITS images. Eventually, we will detail planned developments for the near future, including visualization of cubes and HTTPS support.

P012 François Bonnarel

CDS

*IVOA Data Access Layer: Goals, Achievements and Current Trends**F. Bonnarel, P. Dowler, K. Noddle, D. Tod*

The IVOA Data Access Layer (DAL) working group was created in October 2002, during the first year of existence of the International Virtual Observatory Alliance. The intent was to define protocols to homogenize data discovery, data description, data retrieval, and data access processes, with data access providing direct client access to dynamically computed subsets of remote astronomical datasets. The early developments in the development of these protocols were a Cone Search protocol for astronomical catalogues, and Simple Image Access version 1.0 (SIIV1) for images. Later came the Simple Spectral Access protocol (SSA) with the data access protocol defined in terms of a comprehensive underlying Spectrum 1.0 data model, a precursor to the current IVOA data models used to characterize astronomical observations and datasets. A major development providing standardized access to remote relational database management systems and astronomical archives was later made with the Table Access Protocol (TAP). Currently a new generation of DAL protocols is emerging devoted to advanced, interoperable data services in the age of the big data and complex multidimensional datasets. The relationship of the DAL protocols with other IVOA protocols and lessons learnt from this long development process will be illuminated. Current trends for the future development of the protocols will be described.

P013 Stacey Bright

Space Telescope Science Institute

JWST/MIRI Data Reduction and Products

The Mid-Infrared Instrument (MIRI) is one of four science instruments to be flown aboard the James Webb Space Telescope (JWST). MIRI operates from 5 to 28.5 microns and provides a suite of versatile capabilities including imaging, low-resolution spectroscopy (LRS), medium-resolution spectroscopy (MRS) via an integral field unit, and coronagraphy.

The MIRI data reduction plan is based on extensive testing of the instrument. It is split into three stages: first, the raw data is processed into an uncalibrated slope image; second, each slope image is calibrated; and third, slope images from multiple exposures are combined and processed to produce the final data product. Calibrated mosaics and catalogs of point sources will be produced for imaging observations; extracted spectra will be produced for LRS observations; PSF subtracted images will be produced for coronagraph observations; and spectral cubes will be produced for MRS observations. The goal of the MIRI pipeline is to provide well-calibrated, high level data products that maximize the scientific return from the instrument.

P014 Nicolas Cardiel

Universidad Complutense de Madrid

Looking for a Slope Change in a Linear Relationship

The initial understanding of many scientific problems relies on the simplifying assumption that some data of interest can initially be modelled by a linear relationship. Classical examples in astronomy, among a great many others, include surface brightness profiles of elliptical galaxies following a de Vaucouleurs law (when plotting surface brightness as a function of $r^{1/4}$), or surface brightness profiles of galaxy discs exhibiting an exponential dependence with radius. However, the availability of high quality data can rapidly demonstrate that models based on linear relationships are just oversimplifications of reality.

In some cases, the simultaneous fit of two lines, joined at an intermediate point, constitutes the next natural step towards a more elaborated modelling. This approach, that mathematically is equivalent to a linear spline fit to two line segments, can be employed to determine a possible slope change at a particular point, the break radius. In this poster we show how it is possible to use the method of Lagrange multipliers to obtain such a fit with the break radius as a free parameter, with special emphasis on the sensibility of the method to other critical parameters such as the number of fitted points, the slope difference between the two line segments, or the relative position of the break point within the fitted interval.

This method was used by Cardiel et al. (1998) to determine the break radius where star formation was taking place in central dominant galaxies, and more recently (Marino et al. 2015, in preparation) is being used within the CALIFA survey to analyse the possible connection between the external flattening observed in the ionised-gas metallicity gradients and the presence of breaks in the surface brightness profiles of disk galaxies.

P015 Sandra Castro

European Southern Observatory (ESO)

CASA HPC and Parallelization Development for ALMA

Authors: S. Castro, S. Bhatnagar, M. Caillat, P. Ford, K. Golap, J. G. Villalba, J. Jacobs, J. Kern, S. Loveland, D. Mehringer, G. Moellenbrock, D. Petry, M. Pokorny, U. Rao, M. Rawlings, D. Schiebel, V. Suoranta, J. Taylor, T. Tsutsumi, K. Sugimoto, W. Kawasaki, M. Kuniyoshi, T. Nakazato, R. Miel

CASA, the Common Astronomy Software Applications, is the offline data reduction package for ALMA and the VLA. As these new telescopes transition to operations, data sets are increasing in size and data processing time is increasing proportionately. In addition, ALMA has the requirement to deliver “science ready” data products to users, which requires the use of the CASA pipeline in a production environment.

The scope of the CASA-HPC project is to parallelize CASA in order to improve performance on computing clusters as well as modern PC architectures. Although our approach is applicable to a

variety of hardware platforms, the target environments are the clusters operated by ALMA and the partner organizations (ESO, NRAO, NAOJ). This article will describe the technologies used in the data parallel implementation of CASA and the necessary steps to successfully parallelize the ALMA data processing pipeline.

P016 Emmanuel Caux

IRAP/UPS-CNRS Toulouse, France

CASSIS (Centre d'Analyse Scientifique de Spectres Instrumentaux et Synthétiques)

CASSIS (Centre d'Analyse Scientifique de Spectres Instrumentaux et Synthétiques, <http://cassis.irap.omp.eu>) is a software package aimed to speed-up the scientific analysis of high spectral resolution observations, particularly suited for broad-band spectral surveys analysis. CASSIS is written in Java and can be ran on any platform. It has been extensively tested on Mac OSX, Linux and Windows operating systems. CASSIS is regularly enhanced, and can be easily installed and updated on any modern laptop. It can be used either directly connected to the VAMDC (<http://www.vamdc.eu>) portal, or via a fast Sqlite access to a local spectroscopic database combining the JPL (<http://spec.jpl.nasa.gov>) and CDMS (<http://www.astro.uni-koeln.de/cdms/catalog>) molecular spectroscopic databases, the atomic spectroscopic database NIST (<http://physics.nist.gov/PhysRefData/ASD/>), and user proprietary spectroscopic databases. The tools available in the currently distributed version (3.9) include, among others, a powerful spectrum plotter, a LTE model, a RADEX (<http://www.strw.leidenuniv.nl/~moldata/radex.html>) model connected to the LAMDA (<http://www.strw.leidenuniv.nl/~moldata/>) and the basecol (<http://www.obspm.fr/basecol>) molecular collisional databases, a spectral toolbox allowing to perform various changes on the displayed spectra, a rotational diagram module, a SAMP and a SSAP module. A scripting interface allows to find the best LTE or RADEX model of a spectrum computing regular grids of models, or via a MCMC module.

P017 M. Teresa Ceballos

Instituto de Física de Cantabria (CSIC-UC)

SIRENA: a Software Package for the Energy Reconstruction of the Athena X-IFU Events

We describe the current status of the SIRENA software, developed to reconstruct the energy of the X-ray events detected by the TES calorimeter instrument X-IFU of the Athena observatory. SIRENA is currently integrated in a larger project for end-to-end simulations where data are simulated with the SIXTE simulator (<http://www.sternwarte.uni-erlangen.de/research/sixte/>), then triggered and finally ingested in the SIRENA package for energy reconstruction. While the baseline implementation of SIRENA is based on the optimal filtering technique, some other algorithms are also being considered (principal component analysis, resistance space analysis). The integration of SIRENA into SIXTE allows the study and comparison of different triggering algorithms, reconstruction techniques, and



instrument configurations, and their optimization for in-flight use. The best performance algorithms for event triggering and energy determination will be finally implemented in the X-IFU Digital Readout Electronics (DRE) unit

P018 Igor Chilingarian

Smithsonian Astrophysical Observatory

New Galaxy Properties and Evolutionary Mechanisms in the Data Discovery Era

We have entered the new era when forefront research can be done by mining the data in astronomical databases and archives publicly accessible via the Virtual Observatory framework. In this light, development of novel and efficient knowledge extraction technologies and tools pioneered by the ADASS community has become as important as the progress in astronomical instrumentation and observational techniques. We will review research results based solely on the analysis of publicly available data and obtained over the last several years. Two most recent examples, both originated from ADASS tutorials in 2010 and 2012 and published in refereed journals are: (a) the fundamental colour relation of normal galaxies (a.k.a. the generalization of the famous "red sequence to star-forming galaxies) and (b) the discovery of runaway compact galaxies ejected from clusters and groups by gravitational interactions similar to hypervelocity stars in the Milky Way but on a completely different scale. The latter finding became the first astrophysical research paper published in the interdisciplinary journal "Science" without a single photon of new observations and a single byte of new numerical simulations. We foresee the role of data mining and knowledge extraction techniques in astrophysics to grow substantially over the next decade and call for the ADASS community to lead the effort.

P019 Patrick Clearwater

University of Melbourne

A Virtual Laboratory for Gravitational Wave Science in the Advanced LIGO Era

The Advanced Laser Interferometer Gravitational wave Observatory (aLIGO) is expected to start its first observing run towards the end of this year, and from that will come a wealth of data, perhaps containing the first direct detection of gravitational waves. With these data also comes the challenge of making the existing data analysis software and pipelines accessible to the broad astronomical community. To address this challenge, we are building a web-based, open source Virtual Laboratory to expose the existing LIGO data processing tools and pipelines in a coherent and user-friendly way. We are using Apache Airavata as an underlying framework, and key features include re-use and sharing of mature and proven data processing workflows, automatic management of data provenance and metadata, and platform-agnostic use of a range of underlying computational infrastructure. The project is a collaboration between the Australian Consortium for Interferometric Gravitational Astronomy and the CSIRO Digital Productivity Flagship.

P020 Vito Conforti

INAF - IASF Bologna

The Camera Server of the ASTRI SST-2M Telescopes proposed for the Cherenkov Telescope Array

The Cherenkov Telescope Array (CTA) project is an international initiative to build the next generation of ground-based very high energy gamma-ray instrument. Three classes of telescopes with different mirror size will cover the full energy range from tens of GeV up to hundreds of TeV. The full sky coverage will be assured by two arrays, with one site located in each of the northern and southern hemispheres. In the current design scenario, the southern hemisphere array of CTA will include seventy of small size telescopes (SST, 4m diameter) covering the highest energy region. Their implementation includes proposed intermediate steps with the development of mini-arrays of telescope precursors like the ASTRI mini-array, led by the Italian National Institute for Astrophysics (INAF) in synergy with the Universidade de Sao Paulo (Brazil) and the North-West University (South Africa). The ASTRI mini-array will be composed of nine telescope units (ASTRI SST-2M) based on double-mirror configuration whose end-to-end prototype has been installed on Mt. Etna (Italy) and is currently undergoing engineering tests. In the ASTRI SST-2M prototype, operating in single telescope configuration, the basic Camera Server software is being deployed and tested; it acquires the data sent by the Camera Back End Electronics as a continuous stream of packets. In near real time, the bulk data of a given run are stored in one raw file. In parallel they are sorted by data type, converted in FITS format and stored in one file for data type. Upon closure, each file is transferred to the on-site archive. In addition, the Quick Look component allows the operator to display the camera data during the acquisition.

This contribution presents how the Camera Server software of the Prototype is being upgraded in order to fulfil the mini-array requirements, where it will be deployed on the Camera Server of each ASTRI SST-2M telescope. Particular emphasis will be devoted to the most challenging requirements that are related to the stereoscopy, when two or more telescopes have triggered simultaneously. To handle stereoscopy, each Camera Server has also to: (i) get the timestamp information from the Clock Distribution and Trigger time stamping System, and associate it to the related camera event; (ii) get from the Software Array Trigger the timestamp which passed the stereo trigger criteria; and (iii) forward to the Array Data Acquisition system the stereo trigger events, according to the required data format and communication protocol.

P021 Evan Crawford

Western Sydney University

WTF? Discovering the Unexpected in next-generation radio continuum surveys

The majority of discoveries in astronomy have come from unplanned discoveries made by surveying the Universe in a new way, rather than by testing a hypothesis or conducting an investigation with



planned outcomes. For example, of the 10 greatest discoveries by HST, only one was listed in its key science goals. Next generation radio continuum surveys such as the Evolutionary Map of the Universe (EMU: the radio continuum survey on the new Australian SKA Pathfinder telescope), will significantly expand the volume of observational phase space, so we can be reasonably confident that we will stumble across unexpected new phenomena or new types of object. However, the complexity of the instrument and the large data volumes mean that it may be non-trivial to identify them. On the other hand, if we don't, then we may be missing out on the most exciting science results from ASKAP. We have therefore started a project called "WTF", which explicitly aims to mine EMU data to discover unexpected science that is not part of our primary science goals, using a variety of machine-learning techniques and algorithms. Although targeted specifically at EMU, we expect this approach will have broad applicability to astronomical survey data.

P022 Steven Crawford
SAAO

PySALT: The SALT Data Reduction Framework

PySALT is the suite of data reduction tools for the Southern African Large Telescope (SALT), a modern 10m class telescope with a large user community consisting of 13 partner institutions. The current software supports all of the facility class instruments on the telescope. This includes SALTICAM (an imaging and acquisition camera with high speed modes), the Robert Stobie Spectrograph (a multi-mode medium resolution spectrograph), and the High Resolution Spectrograph. In addition to the data reduction tools, software for archiving, distribution, and assessing the data quality have also been developed. Most of the software has been developed in python and several new packages have been developed including ones for modelling spectrographs, reduction of high speed modes, and archiving and distribution of observations. The most recent addition to the package has been the tools for the reduction of observations from the High Resolution Spectrograph. These provide tools for reduction of all modes except the high stability mode and have been integrated into the daily pipeline to provide users with reduced data the day after the observations are made.

P023 Chenzhou Cui
National Astronomical Observatories, CAS

Update of the China-VO AstroCloud

As the cyber-infrastructure for Astronomical research from China-VO project, AstroCloud has been archived solid progresses during the last one year. Proposal management system and data access system are rebuilt. Several new sub-systems are developed, including paperdata system, statics system and public channel. More data sets and application environments are integrated into the platform. LAMOST DR1, the largest astronomical spectra archive was released to the public using the



platform. The latest progresses will be introduced. Experiences and lessons learned will be discussed.

P024 Zhao Cui

National Astronomical Observatories, CAS

HSOS Data Query Form based on MVC

Huairou Solar Observing Station (HSOS) is one of the key stations of the National Astronomical Observatories, Chinese Academy of Sciences which began to observe and obtain solar data from 1986. Large number of datas have been obtained, thus the data query form is essential to be built. Here we will introduce a system based on MVC, which can provide data management, query, download and analysis. That is helpful to data sharing in solar physics, space physics and geophysics.

P025 Catherine de Burgh-Day

The University of Melbourne

Direct Shear Mapping: Measuring Weak Gravitational Lensing Directly for the First Time

While there are other techniques to measure the fraction of dark matter in the Universe, weak gravitational lensing is the only tool available to measure the spatial distribution of dark matter relative baryonic matter. Understanding the distributions of baryonic and dark matter in galaxies is vital for forming a complete picture of galaxy formation and evolution. Current weak lensing techniques require hundreds of galaxies for a single weak lensing measurement. They are insensitive to the shape of the dark matter halo and are useless for analyses that require individual, direct shear measurements.

We have developed a new technique called Direct Shear Mapping (DSM) to measure gravitational lensing shear directly from observations of a single background source. DSM assumes the velocity map of an un-lensed, stably-rotating galaxy will be rotationally symmetric. Lensing distorts the velocity map, making it asymmetric. DSM uses a MCMC Maximum-Likelihood method to fit for the shear in the velocity maps of galaxies by attempting to restore symmetry.

We demonstrate that we are in principle able to measure shears as small as 0.01. In practice, we have fitted for the shear in very low redshift (and hence un-lensed) velocity maps, and have obtained null result with an error of ± 0.01 . This high sensitivity results from analysing spatially resolved spectroscopic images (i.e. 3D data cubes), including not just shape information (as in traditional weak lensing measurements) but velocity information as well.

Spirals and rotating ellipticals are ideal targets for this new technique. Data from any large IFU or radio telescope is suitable, or indeed any instrument with spatially resolved spectroscopy such as SAMI, ALMA, HETDEX and SKA.

P026 Siamak Dehghan

Victoria University of Wellington

Performance Comparison of Algorithms for Point-Source Detection

The generation of a sky model for calibration of Square Kilometre Array observations requires a fast method of automatic point-source detection and characterisation. In recent years, point-source detection in two-dimensional images has been implemented by using several thresholding approaches. In the first phase of the SKA we will need a fast implementation capable of dealing with very large images (80,000 x 80,000 pixels). While the underlying algorithms scale suitably with image size, the present implementations do not. In this study, we present a detailed analysis of the performance of these methods using different input parameters and data, and compare their features, specifications, computational costs, and resource requirements. We make some comments on the pertinent tradeoffs for scaling these implementations to SKA-levels.

P027 Markus Demleitner

Heidelberg University, GAVO

Datalink and DaCHS

Datalink is a new Virtual Observatory protocol designed to facilitate the publication of complex datasets. It provides a standard way to declare relationships between

- parts of data sets (e.g., separate data and error files),
- different representations (e.g., previews, derived plots, or data provided in multiple resolutions),
- ancillary calibration data,
- files in provenance trees,
- data services for server-side processing of such data sets (e.g., cutout and rebinning services).

Datalink services can run standalone, where access links are passed around directly, or embedded in VO DAL protocols like ObsTAP or SSAP, where a special annotation of the VOTable service responses allows Datalink-enabled clients advanced access patterns while legacy clients continue to work unchanged. Datalink is already in use at sites like the CADC or the GAVO data center, with first clients (e.g., SPLAT, TAPHandle) already providing interfaces.

This poster gives a brief overview of Datalink itself as well as its integration in the VO server suite DaCHS.



P028 Nadezhda Dencheva

GWCS - A General Approach to Computing World Coordinates

CS provides tools for managing World Coordinate Systems in a much more flexible way than is possible with the FITS standard. Transformations from the detector to a standard coordinate system are combined in a way which allows for easy manipulation of individual components. The framework handles discontinuous models (e.g. IFU data) and allows quantities that affect transforms to be treated as input coordinates (e.g. spectral order).

The package is written in python and is based on astropy. It is easy to extend it by adding new models and coordinate systems.

P029 Sebastien Derriere

CDS, Observatoire astronomique de Strasbourg

Building the ASTRODEEP Data Portal

ASTRODEEP is a European project aiming at exploiting the deepest multi-frequency astronomical survey data. The data validation, analysis and sharing involves heterogeneous data: multi-frequency deep mosaic images, spectra, and additional data products such as spectral energy distributions or fits of photometric redshifts.

We present the dedicated data portal being developed for ASTRODEEP, using HiPS for image access through AladinLite or Aladin Desktop, SAADA for linking heterogeneous datasets, and a combination of simple widgets to build custom data analysis portals.

P030 Rosa Diaz

Space Telescope Science Institute

Reference Data Management System for Calibration and Tools

The accuracy and correctness of the calibration software and tools like APT, ETC and pysynphot, used to support HST, and in the future JWST, depend greatly on the data they use. The large number of reference data delivered on regular basis requires careful validation and verification against the systems that will use them. These, in some cases, are also used in more than one of these systems. The Reference Data for Calibration and Tools Management Team (ReDCaT), makes use of the Calibration Reference Data System (CRDS) to track and validate reference data used by the HST calibration pipelines and will also be used by the JWST pipelines. In the case of other planning and analysis tools, the ReDCaT team is in the process to develop a different system, one that will address their particular needs. In this poster we describe both systems and how these can help users to identify the best and most current reference data for their calculations.

P031 Sofia Dimoudi.

University of Oxford

Pulsar Acceleration Searches on the GPU for the Square Kilometre Array

The study of binary pulsars enables science investigations such as tests of general relativity and the detection of gravitational waves, and is an important scientific objective for the planned Square Kilometre Array (SKA) radio telescope. Pulsar detection is commonly done using a Fourier transform over the signal observation length, and a search of the resulting power spectrum for frequencies with signal levels above noise. Orbital motion in binary pulsar systems causes the frequency of the pulsar radiation to drift, reducing the amount of signal power that can be recovered on the radio telescope with the Fourier transform. Acceleration searches are methods for recovering the additional signal power that is lost due to the effect of orbital acceleration. Existing methods are currently computationally expensive, and so enabling such searches in real-time - a vital requirement for the SKA due to data volumes produced - will be bound to strict processing time limits. Modern Graphics Processor Units (GPUs), are demonstrating much higher computational performance than current CPUs in a variety of scientific applications, and there is a high potential for accelerating pulsar signal searches by using GPUs. We present an overview of an implementation of the Fourier Domain Acceleration Search on the GPU in the context of the SKA specifications, as part of the Astro-Accelerate real-time time-domain data processing library, currently under development at the Oxford e-Research Centre (OeRC), University of Oxford.

P032 Geoff Duniam

University of Western Australia; International Centre for Radio Astronomy Research

Big Data Architecture in Radio Astronomy - The Sky Net and SKA

The overall aim of this study was to determine if a scalable data warehouse architecture comprising disparate database systems would provide an effective solution for large astronomical data analysis in the peta and potentially the exa scale. The focus was on the analysis and testing of available database architectures and data model designs appropriate for the storage and retrieval of catalogue data and metadata, as well as the capabilities of these architectures to integrate with standard astronomical analytic paradigms (for example, R, Python, Scala, C/C++ and Java).

The study investigated various technology platforms to assess the disparate capabilities of a number of technology solutions.

This study utilised synthetic catalogue data provided by The Sky Net as real data from the SKA and ASKAP was not available. Data products available after validation and assimilation were raw catalogue data files. Various catalogue formats were evaluated.



Initial investigations into raw data partitioning and compression using Hadoop/HDFS integrated with a RDBMS containing parameter and detection metadata are promising for a search and retrieval architecture. We have also investigated various frameworks to integrate R and Spark paradigms into the HDFS framework; these investigations are also promising for the analysis of very large data sets.

P033 Kimberly DuPrie

STScI

Designing an Archive

The James Webb Space Telescope will have five instruments and 18 detectors. It will generate more than 25 different types of FITS files whose metadata will be stored in the archive. The archive must be designed in such a way that makes it easy to ingest the files quickly. This poster explores the evolution of the archive design, discussing the challenges that led to each change.

P034 Timothy Dykes

University of Portsmouth

Accelerated 3D Visualization of Mock Galaxy Catalogues for the Dark Energy Survey

Nowadays numerical simulations (grid or particle based) constitute powerful instruments for describing, investigating and ultimately understanding, the dynamics and properties of a multitude of astrophysical objects. Due to rapidly evolving technological advances in High Performance Computing (HPC), HPC infrastructures are more and more employed to execute such simulations resulting in increasingly accurate datasets and contributing to dramatic growth in their sizes and complexity. Although a typical scenario for scientific discovery is comparison of simulation and observational datasets, this is often not a straightforward process, e.g. due to the fact that in general the simulation and observed quantities are different. Visualization algorithms can then provide suitable and effective ways for inspection and exploration to underpin data analysis, e.g. aiding in rapidly focusing on relevant features of interest and detecting non-obvious correlations or even intrinsic data characteristics.

This paper presents a novel implementation of our volume rendering code, Splotch, specifically designed to address the aforementioned challenges, and its application to the visualization of data simulating galaxy surveys, which can be directly compared to the corresponding observations. Splotch is a volume ray-casting algorithm for effectively visualizing large-scale, particle-based numerical simulations. The algorithm is optimized in terms of memory usage and exploitation of HPC architectures, e.g. multi-core, multi-node, multi-GPU heterogeneous supercomputing systems, by efficiently combining CUDA, the OpenMP and the MPI paradigms. This allows processing huge datasets, with size of tens or even hundreds of terabytes (unreachable by most of currently available visualization software packages), while keeping the computational time reasonably low. The new implementation exploits atomic operations on the GPU, which strongly simplify our previously

developed CUDA based algorithm [1], with increasing performance. While atomic operations were previously considered too slow to be of use in this type of application they are strongly optimized in the more recent NVIDIA cards, as on the Tesla K20-40-80 accelerators.

The optimized Splotch algorithm has been adopted for the visualization of mock galaxy catalogues for the Dark Energy Survey (DES) [2], which can be compared directly with corresponding DES observations. Such mock catalogues are calculated from the data generated by N-body simulations requiring tens of billions of particles to follow the dynamics of the matter on a volume of $(3h^{-1} \text{ Gpc})^3$, identical to that of the on-going DES survey. Halo and galaxy catalogues are built upon it [3]. By construction the galaxy catalogue matches observed luminosity functions, color distributions and clustering as a function of luminosity and color at low redshifts. Galaxy properties are then evolved into the past-lightcone using evolutionary models. The full procedure is highly computationally demanding. For this reason alternative approaches are being under development. An example is given by the L-PICOLA [4] software, which has the advantage of generating and evolving a set of initial conditions into a dark matter field much faster than a full non-linear N-Body simulation. Galaxies are then placed on L-PICOLA halos and matter fields using the same technique as in the N-body simulations. Additionally, L-PICOLA has the ability to include primordial non-Gaussianity in the simulation and simulate the past lightcone at run-time, with optional replication of the simulation volume.

The results of the L-PICOLA code must be validated against other models and visualization provides a first prompt and effective way to compare different datasets, addressing the solution of possible mismatches or even errors. To this end, Splotch proved to be particularly effective, manipulating datasets of any size and providing their full 3D rendering in a reasonable time by means of its multi-GPU accelerated implementation, enhancing the scientific discovery process. We thus present and discuss the potentiality of our optimized visualization algorithm as a fast and effective debugging tool for the aforementioned simulations, in particular focusing on gaining rapidly an insight into potential anomalies and aiding in planning of appropriate remediation.

[1] GPU Accelerated Particle Visualisation with Splotch, Journal of Astronomy and Computing, DOI: 10.1016/j.ascom.2014.03.001, [2] <http://www.darkenergysurvey.org>, [3] <http://arxiv.org/abs/1312.2013>, [4] <http://arxiv.org/abs/1506.03737>"

P035 Satoshi Eguchi
Fukuoka University

Blade Runner – What Kind Objects are there in the JVO ALMA Archive?

The JVO ALMA Archive provides users one of the easiest ways to access the ALMA archival data. The users can have a quick look at a 3 or 4-dimensional data cube without downloading multiple huge tarballs from a science portal of ALMA Regional Centers (ARCs). Since we just synchronize all datasets with those of ARCs, the metadata are identical to the upstream, including "target name" for each dataset. The name is not necessarily a common one like NGC numbers, but sometimes one of sequential numbers assigned in an observation proposal. Compilation of these artificial names into astronomical ones could provide users more flexible and powerful search interfaces; for instance,

with the knowledge of the redshift for each source, the users can easily find the datasets which observed their interested emission/absorption lines at not the observer frame but the rest frame, fitting well with theoretical studies. To implement this functionality, cross-identification of all the sources in our archive with those in some other astronomical databases such as NED and SIMBAD is required. We developed a tiny Java application named "Blade Runner" for this purpose. The program works as a crawler for both the JVO ALMA Archive and SIMBAD, storing all information onto a SQLite-based database file; this portable design enables us to communicate results to each other even under different computing environments. In this poster, we introduce its software design and our recent work on the application, and report a preliminary result on the source identification in our archive.

P036 Tony Farrell

Australian Astronomical Observatory

*DRAMA2 - DRAMA for the modern era.**Co-Author: Keith Shortridge*

The DRAMA Environment provides an API for distributed instrument software development. It originated at the Anglo-Australian Observatory (now Australian Astronomical Observatory) in the early 1990s, in response to the need for a software environment for a large distributed and heterogeneous systems, with some components requiring real-time performance. It was first used for the AAOs 2dF fibre positioner project for the Anglo-Australian Telescope. 2dF is still in use today, but has changed dramatically over time. DRAMA is used for other AAO systems and is or has been used at various other observatories looking for a similar solution. Whilst DRAMA has evolved and many features were added, there had been no big changes. It was still a largely C language based system, with some C++ wrappers. It did not provide good support for threading or exceptions. Ideas for proper thread support within DRAMA have been in development for some years, but C++11 has provided many features which allow a high quality implementation. We have taken the opportunity provided by C++11 to make significant changes to the DRAMA API, producing a modern and more reliable interface to DRAMA, known as DRAMA2.

P037 Pierre Fernique

Observatoire Astronomique de Strasbourg – CDS

*MocServer: What & Where in a few milliseconds**Pierre Fernique, Thomas Boch, Anais Oberto, Francois-Xavier Pineau [CDS]*

The MocServer is a new astronomical service dedicated to the manipulation of data set coverages. This server sets together about 15000 spatial footprints associated to catalogs, data bases and pixel surveys from the CDS and partners. Thanks to the Multi-Order Coverage map coding method (MOC1.0 IVOA standard), the MocServer is able to provide in a few milliseconds the list of data set identifiers intersecting any polygon on the sky. Moreover, it allows to solve some use cases difficult to realize before. For instance, it is now straightforward to retrieve the list of catalog identifiers



containing velocity measurements and for which the coverage overlays simultaneously HST and GALEX observations. Also, the generation of the global coverage of all tables published in a journal such as A&AS requires less than an half second.

The MOC server has been deployed in June 2015 by the Centre de Donnees astronomiques de Strasbourg. It is already used by Aladin Desktop and Aladin Lite prototype versions. It is freely queriable at the address:

<http://alasky.unistra.fr/MocServer/query>

P038 Francoise Genova

CDS, Observatoire astronomique de Strasbourg

The Research Data Alliance

The Research Data Alliance (<https://www.rd-alliance.org/>) aims at enabling research data sharing without barriers. It was founded in March 2013 by the Australian Government, the European Commission, and the USA NSF. It is a bottom-up organisation which after little more than 2 years of existence gathers around 3,000 members from 100 different countries.

Work in RDA is organised in a bottom-up way: members propose Interest Groups and Working Groups tackling any aspect of research data sharing, which means a huge diversity in the activities. The RDA works to implement functional infrastructure through Working Groups. Working Groups are comprised of experts from the international community that are engaged in creating deliverables that will directly enable data sharing, exchange, or interoperability. Working Groups conduct short-lived, 12-18 month efforts that implement specific tools, code, best practices, standards, etc. at multiple institutions. Interest Groups are comprised of experts from the community that are committed to directly or indirectly enabling data sharing, exchange, or interoperability. Interest Groups serve as a platform for communication and coordination among individuals, outside and within RDA, with shared interests. They produce important deliverables such as surveys, recommendations, reports, and Working Group case statements. There are currently about 15 Working Groups and 40 Interest Groups, tackling very different kinds of topics, from very technical ones to more sociological ones. Some scientific communities use the RDA as a neutral place to hold the discussions about their disciplinary interoperability framework.

Astronomy has the IVOA and the FITS Committee for that purpose, but many RDA topics are of interest for us, for instance data citation, including citation of dynamic data bases and data repositories, or certification of data repositories. Also lessons learnt in building the IVOA and data sharing in astronomy are injected in the discussion of RDA organisation and procedures. The RDA is a unique platform to meet data practitioners from many countries, disciplines and profiles, to grab ideas to improve our data practices, to identify topics of common interest and to raise interesting subjects for discussion involving specialists from many countries by proposing new Interest and Working Groups. The talk will present the current status of the RDA, and identify topics of interest



for the ADASS community and topics which would be interesting to promote by proposing new RDA Groups.

P039 Claudio Gheller
ETHZ

In-situ, Interactive Calculations of Faraday Rotation Maps for ENZO Cosmological Simulations
Vazza, F. et al. 2015, arXiv:1503.08983, A&A accepted, Vacca, V. et al. 2015, arXiv:1501.00415

The management and analysis of modern large-scale datasets related to scientific experiments/observations and numerical simulations is becoming an outstanding issue due to continuously increasing size and complexity. Storing data products as they are produced is becoming unfeasible, requiring more and more capable and performing storage facilities and fast and reliable networks, which, in the future, may be not available or affordable.

Interactive, in-situ data processing can represent a viable solution for a prompt data analysis and volume reduction, simultaneous to its acquisition or calculation. In this paper, we propose a solution which combines effective 3D visualization, efficient algorithms and High Performance Computing (HPC) in order to interact and manipulate huge data volumes while they are still in memory, with no need of saving them on a disk.

For this purpose, we have exploited the VisIT visualization and analysis framework (<https://visit.llnl.gov>). VisIT relies on the VTK graphic library (www.vtk.org/) and is designed for the efficient processing of large datasets through a combination of optimized algorithms, the exploitation of high-performance computing architectures and the support of client-server capabilities, allowing efficient remote visualization. It can be used interactively, through a graphical user interface, and it can be scripted using Python. Most interesting, VisIt supports in-situ visualization by instrumenting the (simulation or data analysis) code through a specific library, exposing relevant datasets that can be visualized and processed exploiting all available VisIT functionalities.

VisIt in-situ library has been coupled to the ENZO code (<http://enzo-project.org/>), an Adaptive Mesh Refinement code for astrophysical simulations, which uses an N-body method to follow the dynamics of the collisionless Dark Matter component and an adaptive mesh method for ideal fluid dynamics. ENZO is used for a programme of cosmological simulations finalized to the study of the evolutions and properties of the cosmic web, running on the Piz Daint system of the CSCS Swiss National Supercomputing Center. Piz Daint is a CRAY XC30 hybrid CPU+GPU system with a peak performance of 7.8 Petaflops. VisIt can run concurrently to a simulation on the nodes of Piz Daint supporting the interaction with properly instrumented running processes, exploiting parallel processing not only for the simulation but also for our on-the-fly data analysis and visualisation.

ENZO has been instrumented with the in-situ VisIt library, exposing the main simulated variables, which are distributed across computational nodes that can be accessed by VisIt. The AMR data structure, which provides high spatial resolution where this is required by the physics of the problem, describing uninteresting regions by a coarse mesh, is fully supported. In this way, all the

details of the simulated data are available. From such 3D data, observables can be calculated, which can be directly compared to images, maps and catalogues coming from telescopes and detectors. In particular we have focused on the generation of 2D Faraday Rotation maps at radio wavelengths, calculated as the projection of the electron density times the parallel component of the magnetic field (both calculated by ENZO) along line of sights. We focus on the Faraday Rotation measurement since it is candidate to provide one of the first extensive direct detections of the filamentary network characterizing the large scale distribution of matter in the universe by means of future radio telescopes, as, first of all, SKA (e.g. Vacca et al. 2015; Vazza et al. 2015).

Our implementation allows generating the maps on the fly, while the simulation is running, from any possible point of view and supports interaction with the data, setting the main parameters. An important side effect, is that only maps are saved as results of the simulation, with a dramatic drop in the saved data size and in the storage requirements. This will be crucial for the next generation of MHD cosmological runs that will drastically increase the amount of physical variables to account for in the study of cosmic magnetogenesis (e.g. chemical fields and star forming particles to account for the run-time seeding of magnetic fields in galaxy formation processes). Results will be presented together with performance and scalability analysis, highlighting the effectiveness and robustness of this approach for prompt data analysis and scientific discovery. It will also show how such approach use ENZO only as one of the possible use cases, but it is extensible to a large class of algorithms requiring HPC for processing large data volumes, as in the case of data analysis software for next generations of observations, making it possible interactive, real-time data processing on petabytes sized datasets.

P040 James Gilbert
University of Oxford

Learning from History: Adaptive Calibration of 'Tilting Spine' Fibre Positioners

This paper discusses a new approach for determining the calibration parameters of independently-actuated optical fibres in multi-object astronomical fibre positioning systems. This work comes from the development of a new type of piezoelectric motor intended to enhance the 'tilting spine' fibre positioning technology originally created by the Australian Astronomical Observatory. Testing has shown that the motor's performance can vary depending on the fibre's location within its accessible field, meaning that representative information on the behaviour of an individual fibre is difficult to obtain with a one-time calibration routine. Better performance has resulted from constantly updating calibration parameters based on the observed movements of the fibre during normal closed-loop positioning. This information is used to keep a recent history of an individual fibre's behaviour within different sub-sections of its accessible field. Over time, an array of location-specific data is built up that can be used to better predict the results of a fibre movement. Results from a prototype system are presented, showing a significant reduction in overall positioning error when using this new approach.



P041 Ranpal Gill

The Mercury Transit 2016: a World-Wide Outreach Opportunity

The Mercury Transit is a rare phenomenon that will occur on the 9th of May 2016 11:12-18:42 UT. Most of the world will be party to this observable event whereby the planet Mercury will pass in front of our sun (a solar transit), visually it will appear as a small black dot moving across the face of the sun. It will occur approximately ten years since the last transit with the next one due in 2019.

Occurring just a year after the end of the MESSENGER mission and a mere few months before the launch of BepiColombo it provides an ideal opportunity to engage the public by highlighting the fascinating science aspects of these missions to Mercury. Furthermore, a dual point observation campaign facilitating live web transmission (possibly in 3D) and world-wide media events will enable the public to feel involved and even observe.

The observation of the transit will produce tangible educational benefits by stimulating interest in reproduction of classical transit measurements, Earth-Sun distance calculations, parallax measurements and the production of science results.

This outreach project will involve and include a multitude of organisations, people, communications channels, conventional media papers, presentations, posters and of course the full plethora of social media. All these actors will need to communicate efficiently, to do so a central control point is planned. This paper will provide details of the plan and it will provide a channel for the community to get involved.

P042 Juan González Núñez

ESAC Science Data Centre

Open Archives for Open Science at the ESAC Science Data Centre (ESDC)

Can the discipline of Scientific Data Archiving provide answers to the challenge for a more reproducible astronomical research? Open Science initiatives such as of Open Data and Software provide keystones in the form of open licenses. But besides licensing, an open approach also requires the necessary dissemination systems to be developed - so that open data and software are actually openly available and more accessible to wider communities. The ESAC Science Data Centre is carrying out several efforts to maximise this.

Next generation archives for ESA astronomical missions such as Gaia and Euclid have had open data in mind since early stages of development. Usage of VO data sharing protocols such as VOSpace and extensions implemented over TAP and UWS not only allow for effective publication of space missions data, but they also provide mechanisms for the sharing and publication of intermediate products and research results that will be permanently available and may be safely referenced.

Effort is also being put into moving towards more open software; from the definition of an infrastructure to host analysis workflow software at the Gaia Added Value Interface to the

publication as libraries of the code used to expose VO services, which will revert back to the community as forks for several projects (TAP, UWS, VOSpace) published under GNU licensing.

Open accessibility will not be restricted to mission specialists or specialised scientific communities; efforts on visual interfaces to science products that may be equally appealing to specialists and wider communities are being put into place with projects such as the Multi Mission Interface to all of ESA astronomical missions, or the Archive Image Browser at ESAC (imagearchives.esac.esa.int)

The publication of Rosetta NAVCAM images through the Archive Image Browser has become a full loop test: openly licensed data (Creative Commons) being released through open source software, in a very openly accessible portal. Positive feedback from both science community on the easy to browse interface, and from the general public on the accessibility to all raw science data encourages extending this to other data and missions.

P043 Javier Graciá Carpio

Max-Planck-Institut für Extraterrestrische Physik

The JScanam Map-Maker Method Applied to PACS/Herschel Photometer Observations

The Herschel far-infrared (FIR) satellite was launched in May 2009 and during its almost 4 years of technical and scientific operations it observed up to ~10% of the sky, ranging from Galactic star-forming regions to galaxies in the high redshift Universe. One of its 3 key instruments was the Photodetector Array Camera and Spectrometer (PACS). In its photometric scanning mode, PACS was able to observe wide areas of the sky, revealing the filamentary structure of the cold interstellar medium and resolving ~80% of the FIR cosmic infrared background. Being a bolometer array, the PACS photometer signal is dominated by the so-called 1/f noise. This noise produces stronger signal at longer timescales and results in visible signal drifts that affect all spatial structures in the maps. High-pass filtering techniques can be applied to remove this kind of noise, leaving point sources untouched. Unfortunately, these techniques also remove real astronomical extended emission from the maps. In order to solve this problem, a large fraction of the PACS observations were done in a scan/cross-scan observing scheme. In theory, this observing mode provides enough information to separate the 1/f noise from the extended emission. Several map-making methods (MadMap, Unimap, Scanamorphos, Sanepic) have been applied to reduce PACS photometer observations, reaching different ranges of success. In this talk I will present the JScanam map-maker, a HIPE/Jython implementation of the Scanamorphos techniques initially developed by Helene Roussel. JScanam is currently the default mapper for reducing PACS photometer observations and its automatic pipeline results can be easily inspected and downloaded from the Herschel Science Archive. I will briefly discuss the theory behind the JScanam algorithms and dedicate the rest of the time to explain the key steps that are necessary to produce high quality maps, both from the astronomer and the software developer point of view. The application of these techniques to similar bolometer arrays will also be discussed.

P044 Andy Green
AAO

FIDIA: A Python Interface for Astronomical Data that Astronomers Actually Understand!

We present a Python package for interacting with data for astronomical objects in a consistent and intuitive manner, regardless of the source or format of the data. Considerable past work has focused on standardised formats and protocols for working with data. However, these tools are structured around images and tables, not stars and galaxies. We take such work a step further by defining an interface to data on disk, on the web, in a database, etc. that has a vocabulary (based in part on existing IVOA UCDS) for the physical properties of astronomical objects. In our package, the primary objects are astronomical objects, such as stars and galaxies, with data members such as spectra, magnitudes and velocities. By presenting the properties and measurements of objects in this way, we simplify analysis and visualisation by providing a uniform representation of an astrophysical object's characteristics. Scope is included for simple, standard analysis operations to be executed automatically in response to requests for data not present, such as k-correcting magnitudes, computing narrow band images from integral-field spectroscopy, or even inferring star-formation rates from H α luminosities, while maintaining the possibility of overriding the default behaviour. The package accepts that data will always be available in a wide variety of formats and places, and therefore it provides a formulaic way to connect to and import from new data sources such as on the user's computer. Cross matching is done between data sources as needed. Throughout, the package has been designed to deliver the most likely result (with a warning) in the face of ambiguity, such as multiple versions of data or multiple potential cross-matches, but allow for arbitrary, repeatable reconfiguration. The package takes care of mundane details such as memory management, local caching, unit conversions, etc. The astronomer can use this package to leverage existing data about their stars or galaxies in their analysis, or to rapidly apply existing analysis tools based on this package to new data.

P045 Fabien Grisé
Observatoire Astronomique de Strasbourg

Recent and future developments of the 3XMM XCatDB:

The 3XMM XCatDB (<http://xcatdb.unistra.fr>) is a database hosting the 3XMM catalogue of X-ray sources observed by the XMM-Newton observatory. It provides possible (multi-wavelength) identifications in archival catalogues, an online spectrum fitting module and gives access to other products that can be previewed (spectra, time series).

We present the recent inclusion of the XMM-Newton spectral-fit database (Corral et al. 2014) inside the XCatDB. Users can now access the spectral parameters for all X-ray sources that have been fitted (up to the 3XMM-DR4 release) for a variety of spectral models. We detail here the implementation of these products and give a preview on some new developments to come that will put some emphasis on an active use of these products. This will provide a means to search for X-ray sources



meeting certain criteria and therefore, allowing possible new research/discoveries within the 3XMM database.

P046 Anthony Gross
CNRS – LAM

SPHERE TDB: Catalogue Management, Observation Planification and Detections Visualization

SPHERE is the extreme adaptive optics system and coronagraphic facility at the VLT. Its goal is imaging, low-resolution spectroscopic, and polarimetric characterization of extra-solar planetary systems. We developed at CeSAM (Centre de Données Astrophysiques de Marseille) the SPHERE Target DataBase (TDB). It is a web based information system aiming at the management of the catalogue of NIRSUR guaranteed time survey dedicated to the search and the characterization of Giant Planets. For the observation preparation, the TDB makes it possible to select targets with multiple criteria. Results are formatted and sent to a set of tools in order to optimize the observation plan. It also keeps track of all NIRSUR observations and resulting detections. In a second step, TDB offers the possibility to visualize detections and to ease the determination of false positive.

P047 Stephen Gwyn
Canadian Astronomy Data Centre

An Empirical Exposure Time Calculator for CFHT/Megacam

The CFHT/MegaCam exposure time calculator, allows users to determine the appropriate exposure time to achieve a given magnitude limit or photometric accuracy in different filters under different atmospheric conditions. Like most such calculators, relies on assumptions about the throughput of the instrument, the transmission of the filters and so on. While generally close to reality, these assumptions are not always completely accurate. The MegaCam archive contains close to 200 000 images taken under a wide range of seeing conditions and transparencies, with exposure times spanning several orders of magnitude. The empirical exposure time calculator presented here uses this existing data collection to more accurately predict the relationship between exposure time and limiting magnitude for CFHT/MegaCam.

P048 Jonas Haase
ESAC/ESA



The Big Hubble Metadata Spring Cleaning

At all three Hubble Space Telescope Archive sites, projects were started to transfer the existing observation meta-data into new containers. This led to a joint concerted effort to clean, repair and document HST meta-data to ensure all three archives share a correct and uniform collection. In addition this provided the opportunity to augment existing metadata by, among other things, computing footprints for all data. The long history of the telescope and heterogenous nature of the HST Instruments provided many challenges and also illuminated the gaps in quality control and background knowledge that had opened up over time. In this talk/paper we would like to spread the message that housekeeping tasks like these are essential and need to be done while the knowledge is at hand to ensure that an archive can be used by the next generation of astronomers.

P049 Roger Hain

Harvard-Smithsonian Center for Astrophysics

The Chandra Source Catalog Release 2 Master Match Pipeline

The Chandra Source Catalog (CSC) is a catalog of pointed and serendipitous sources observed by the Chandra X-Ray Observatory since its 1999 launch (cxc.harvard.edu/csc). The second release of the CSC will feature additional sources detected in observations released publicly since the CSC first release, as well as additional fainter sources found by combining observations whose pointings differ by no more than one arcminute into "detect stacks". Algorithm improvements are expected to further increase the number of fainter sources detected, with a goal of identifying sources down to approximately five net counts on-axis. The CSC release 2 master match pipeline is tasked with matching observations of the same celestial sources from multiple detect stacks to identify "master sources" on the sky, and identifying the detections in all stack observations that contribute to those master sources. The problem of matching photons from sources in multiple detections to form master sources is conceptually similar to multi-catalog crossmatch problems. The core algorithm that will be used in the master pipeline is based on the Bayesian probability approach outlined by Heinis, Budavari and Szalay (2009) and takes into account the position errors in assigning match probabilities. Additional pipeline features will include the ability to incorporate the effect of Chandra's spatially variable Point Spread Function (PSF) to allow matching of multiple compact sources detected on-axis in some stacks with single off-axis detections (with significantly larger PSF) from other stacks. Although the match from an off-axis source may be ambiguous as to which on-axis compact source the detected photons belong, knowledge of this type of match is still useful when calculating certain source properties. The algorithm will also be extended to match Convex Hull Sources (CHS), which are areas of extended emission identified in the CSC by a larger polygonal region than the typical elliptical point source region. Individual source pipelines and a following master merge pipeline will derive source parameters for each detect stack detection and ultimately for each master source. Input data and matching information from each stack detection contributing to the master source will be identified. Additionally, the non-detection of a master source in a detect stack that included the location of the master source in the sky will also be used to provide upper limit flux information. Ultimately, master match data will allow release 2 of the CSC to maximize the detection and source property determination of fainter sources as all detect stack observations are



combined to provide information about master sources. This work has been supported by NASA under contract NAS 8-03060 to the Smithsonian Astrophysical Observatory for operation of the Chandra X-ray Center. The work depends critically on the services provided by the ADS.

P050 Paul Hancock
Curtin University

Source Finding in Wide-Field, Wide-Band Width, Radio Images.

In this poster I describe the difficulties associated with finding and characterising sources in wide-field and wide-bandwidth radio data sets, and how these difficulties have been overcome. I will focus on how the Aegean source finding algorithm has been adapted to be able to account for a direction dependent point spread function in the GLEAM survey images, and how we are able to catalogue ~300 000 sources at 20 frequencies without cross-matching.

P051 Boliang He
National Astronomical Observatory of the Chinese Academy of Sciences

The LAMOST Data Archive and Data Release

In last four years, the LAMOST telescope has published four edition data (pilot data release, data release 1, data release 2 and data release 3). To archive and release these data (raw data, catalog, spectrum etc), we have set up a data cycle management system, including the transfer of data, archiving, backup. And through the evolution of four software versions, mature established data release system.

P052 Samuel Hinton
University of Queensland

Marz: Utilising Web Application for Scientific Analysis

The Australian Dark Energy Survey (OzDES) is a 100-night spectroscopic survey underway on the Anglo-Australian Telescope using the fibre-fed 2-degree-field (2dF) spectrograph. We have developed a new redshifting application Marz with greater usability, flexibility, and capacity to identify a wider range of object types than the runz software package previously used for redshifting spectra from 2dF. Marz is an open-source, client-based web-application which provides an intuitive interface and powerful automatic matching capabilities to consume FITS files generated from the AAOmega spectrograph and produce high quality spectroscopic measurements. Behind the scenes, a cross-correlation algorithm is used to match input spectra against a variety of stellar and galactic



templates, and automatic matching performance for high quality spectra has increased from 57% (runz) to 94% (Marz). Spectra not matched correctly by the automatic algorithm can be easily redshifted manually by cycling automatic results, manual template comparison, or marking spectral features.

P053 Michael Hoenig
Gemini Obsesrvatory

Automated, adaptive reduction of calibration frames

At Gemini Observatory, observers typically take a series of calibration frames (bias frames and twilight flats for GMOS, an optical imager and spectrograph) at the end of each night. Once the required number of frames has been obtained for each band and instrument mode, they are then reduced by observatory staff and ingested into our archive on a monthly basis. This poster describes a new algorithm that performs these tasks fully automatically.

P054 Ecaterina Howard
Macquarie University

Machine Learning Algorithms in Astronomy

We analyze the current state and challenges of machine learning techniques in astronomy, in order to keep up to date with the exponentially increasing amount of observational data and future instrumentation of at least a few orders of magnitude higher than from current instruments. We present the latest cutting-edge methods and new algorithms for extracting knowledge from large and complex astronomical data sets, as a versatile and valuable tool in a new era of data-driven science. As telescopes and detectors become more powerful, the volume of available data will enter the petabyte regime, in need for new algorithms aimed at better modeling of sky brightness, requiring more computing and providing more promising data analysis for billions of sky objects. We emphasize the most important current trends and future directions in machine learning currently adapted to astronomy, pushing forward the frontiers of knowledge through better data collection, manipulation, analysis and visualizing. We also evaluate the emergent techniques and latest approaches for various types and sizes of data sets and show the vast potential and versatility of machine learning algorithms in front of the new challenge of the Fourth Paradigm.

P055 Aitor Ibarra
ESAC/ESA



Embedded XMM-Newton processing light services in the XMM-Newton archive.

Authors: A. Ibarra, M. Sarmiento, E. Colomo, N. Loiseau, J. Salgado and C. Gabriel

While the XMM-Newton satellite keeps taking invaluable X-ray data, within the ground segment, we keep upgrading and maintaining the Science Analysis System (SAS) software to keep pace with the newest IT technologies to provide users the best available services. As part of this process, few years ago we started the development of the Remote Interface to Science Analysis (RISA) software to explore the possibility of offering SAS functionalities encapsulated within web technologies and grid infrastructures. The re-engineered XSA is now at a level to offer the interoperability needed to continue developing RISA. In parallel with SAS evolution, we have kept RISA system active at a low level.

To help scientists from different disciplines and at the same time to offer X-ray users newer processing capabilities, we present in this poster, the first implementation of the XMM-Newton SAS light services accessible from the archive interface.

With these light SAS processing services, users exploring XMM-Newton data from the archive, will be able to generate on-the-fly images, light curves and spectra using the latest SAS version together with the latest calibration files without having to install any software.

P056 Tim Jenness
LSST Data Management Team

The LSST Data Processing Software Stack: Summer 2015 Release

The Large Synoptic Survey Telescope (LSST) is an 8-m optical ground-based telescope being constructed on Cerro Pachon in Chile. LSST will survey half the sky every few nights in six optical bands. The data will be transferred to NCSA and within 60 seconds they will be reduced using difference imaging techniques and an alert list, using VOEvent, will be issued to the community. Annual data releases will be made from all the data during the 10-year mission, producing catalogs and deep co-added images with unprecedented time resolution for such a large region of sky. In the paper we present the current status of the data processing software, describe how to obtain it, and provide a summary of the construction plan.



P057 Wolfgang Kausch

University of Vienna, Department of Astrophysics, Vienna, Austria

An Advanced Atmospheric Model for Astronomical Observing Sites

The Earth's atmosphere unavoidably affects any ground-based observations taken with astronomical telescopes by scattering, absorption, and emission of light. Due to the large number of various molecular species all wavelength regimes are affected. Within the framework of Austria's accession to ESO, we have developed an advanced sky background model for Cerro Paranal in the Chilean Atacama desert. As one component, the software incorporates an atmospheric model based on a static standard atmosphere. Created in 2001 and being representative for equatorial regions, it was the best one available at the time of the development of the Cerro Paranal model. In order to take the characteristic dry conditions of the Atacama desert and temporal variations of the water vapour into account, it is dynamically refined by on-site measurements and a global model containing humidity, temperature, and pressure information corresponding to Paranal.

In the last decades, several satellites dedicated to atmospheric and climate research were successfully launched, e.g. ENVISAT, Aura, and OCO2. In addition, significant progress in global modelling of atmospheric properties based on ground-based measurements has been achieved by the European Centre for Medium-Range Weather Forecasts (ECMWF). This led to a large amount of data providing a much more accurate insight into the chemical composition and dynamics of the Earth's atmosphere.

Within the framework of the development of the European "Extremely Large Telescope (E-ELT) instrument simulators and pipelines, we are currently refining the Cerro Paranal sky background model and adapting it to Cerro Armazones, the observing site of ESO's new 39m telescope located at a distance of about 22km from Paranal. Our software aims to assemble atmospheric data from various sources and to create atmospheric height profiles of various molecular species which are representative for ESO observing sites. These profiles are significantly more accurate than the previously used ones, taking into account the chemical composition and high temporal variations of the Earth's atmosphere introduced by diurnal, seasonal, and climate change effects. In addition, a recently started project aiming at determining the water vapour content above Cerro Armazones is currently being conducted with the help of archival data taken with the BESO spectrograph. These data span several years and will give a good coverage of the seasonal variations of H₂O and a good comparison with a similar, previously performed study for Paranal.

In addition, another study aimed at better understanding of the airglow emission is currently being conducted at the University of Innsbruck. This emission arises in the mesopause region at heights between 80 to 110 km, is caused by various chemiluminescence processes, and is not fully understood yet. Due to its strength and high variability on various time scales, it has a high impact on optical and infrared astronomy. Results from this project will also be part of the new atmospheric model.

Our model will provide a very good basis for several astronomical purposes. It is foreseen to be incorporated into future E-ELT instrument simulators, exposure time calculators, and the data pipelines to remove the influence of the Earth's atmosphere. Moreover, the parts which are based



on world-wide satellite measurements and models can be easily adapted to any observing site, leading to a very versatile tool for various astronomical software purposes. In this presentation, we describe the status of the project and the technical details. In addition, we discuss the impact of the model on the planning of observations due to a better estimate of the obtainable signal-to-noise ratio, and the quality of the data reduction due to a better removal of the influence of the Earth's atmosphere.

P058 JJ Kavelaars

Canadian Astronomy Data Centre

cadcVOFS: A File System View of Virtual Observatory Space.

The Canadian Astronomy Data Centre Virtual Observatory File System (cadcVOFS) provides a FUSE based POSIX FS view of VOSpace. Version 2.0 of cadcVOFS, just now released, is uses the Python 'requests' and 'FUSE' packages to provide the interaction into VOSpace and a File System View. cadcVOFS 2.0 also provide file system cache, partial file reads (from VOSpaces that support Range Requests), fail-over between multiple copies of a VOSpace exposed file, and improved exception logging. This poster will describe the underlying architecture of cadcVOFS and demonstrate the capacity of cadcVOFS.

P059 Wataru Kawasaki

National Astronomical Observatory of Japan

Vissage: Development in 2015

Wataru Kawasaki, Yuji Shirasaki, Christopher Zapart, Tsuyoshi Kobayashi, George Kosugi, Masatoshi Ohishi, Yoshihiko Mizumoto (NAOJ), Satoshi Eguchi (Fukuoka University), Yutaka Komiya (University of Tokyo) and Toshihiro Kawaguchi (Sapporo Medical University)

Vissage (VISualisation Software for Astronomical Gigantic data cubEs) is a standalone FITS browser. Being developed as a desktop tool to complement the capability of ALMA data service run by JVO (Japanese Virtual Observatory), its primary aim is to support astronomers utilising public ALMA data by furnishing easy visualisation of 3- or 4-dimensional FITS data cubes and seamless connection to JVO services, together with a flexible and intuitive GUI. Still in a preliminary shape, but development of Vissage is underway to add new functionalities and to make improvements in usability and performance. Recent development items to be put into the next update include (1) viewing spectrum of ALMA data cube and/or 1-dimensional FITS data as well, (2) supporting large size data, (3) exporting images in PostScript format and so on. In this poster, we report our current development status of Vissage and introduce the newly available functionalities, then outline our plans for the near future.



P060 Baerbel Koribalski
CSIRO

3D visualisation of gas and stars in galaxies

P061 Uwe Lammers
European Space Agency

Status of Gaia Science Operations One Year into the Nominal Mission

Gaia commenced its nominal mission phase in July 2014 and 10.5 months later has delivered almost a quarter of a Trillion (250E9) astrometric measurements, 45 Billion (45E9) low-resolution, and 4.5 Billion high-resolution spectra. We critically review the daily operations system which processes the incoming satellite telemetry in near-real time for dissemination to other external processing centres and feeding into the health-monitoring of the satellite. Trial runs of the Astrometric Global Iterative Solution system are taking place in preparation for a first public catalogue release in summer next year and we also report on the status of these activities.

P062 Yajuan Lei
National Astronomical Observatories, Chinese Academy Sciences

The X-ray spectrum of the X-ray binary 4U 1728-34 observed with SUZAKU

With SUZAKU data of 2010 October, we report the spectral results of the neutron-star X-ray binary 4U 1728-34. The continuum spectrum can be fitted with the model of a multicolor accretion disk for the soft energy and a power law for the hard energy. The X-ray spectrum shows a broad, iron K α fluorescence line with equivalent width ~ 0.4 keV. We discuss interpretations of the broad line, and compare our results with the previous work.

P063 Kieran Leschinski
University of Vienna, Institute for Astrophysics

The MICADO IDS - An Instrument Data Simulator for the E-ELT first-light instrument MICADO

Mocking the MICADO - Developing an End-to-End model of the E-ELT's First-Light Near-Infrared Wide-field Imager K. Leschinski [1], O. Czoske [1], W. Zeilinger [1], J. Alves [1], W. Kausch [1,2], R. Koehler [2,1], G. Verdoes Kleijn [3] [1] University of Vienna, Department of Astrophysics, Vienna, Austria, [2] Institute for Astro and Particle Physics, Universität Innsbruck, Innsbruck, Austria, [3] Kapteyn Astronomical Institute, University of Groningen, The Netherlands

The era of 40m class extremely large telescopes (ELTs) will soon be upon us. These telescopes will open up new frontiers in astronomy, from exploring the formation of the first galaxies to directly



imaging nearby super-Earths. Realistic end-to-end data simulations are an important tool not only for the design of an instrument but also for its efficient operation.

MICADO will be the first-light near-infrared imager for the European Extremely Large Telescope (E-ELT) enabling unprecedented levels of sensitivity and resolution. MICADO will provide diffraction limited imaging with an angular resolution of 4mas as well as a medium spectral resolution ($R > 4000$) long-slit spectrograph. When combined with the 978m^2 collecting area of the E-ELT it will be sensitive to objects as faint as ~ 30 mag in the J, H, and Ks filters.

As part of its development we are constructing an end-to-end model of the instrument in order to simulate the output of the focal plane array by modelling the optical train for the E-ELT and MICADO. Not only will the simulator and mock images aid the development of the data reduction pipelines, they will also help the community to test out the viability of their favoured science cases and observing strategies.

For the development of the instrument data simulator, we are following a two-phased approach. The first phase involves the construction of a “pre-simulator” for the purpose of evaluating the critical elements in the optical train. This will allow us to better quantify the effects that these elements have on the overall performance (Strehl ratio, total transmission, etc) of the instrument. The second phase will then refine the accuracy of these simulations by modelling the critical elements in more detail within the context of MICADO as a complete optical system.

Here we will present the current status of the pre-simulator and show how it can be used to generate mock data for science and calibration purposes. We will also show how this mock data is being used to fine tune the science drivers for the E-ELT and MICADO. Currently the details of the MICADO optical system are still being refined and so we have to be able to easily adapt the simulator to changes in the design of the instrument. Here we will also discuss our strategy for keeping the model framework as flexible as possible while still providing accurate simulations of the instrumental output.

P064 Changhua Li

National Astronomical Observatories, Chinese Academy of Sciences

The Design and Application of Astronomy Data Lake

With the coming of many large astronomy observatory device, such as LAMOST, TMT, FAST, LSST, SKA etc. the storage of data is confronted with greatness challenge. Astronomy Data Lake, a large storage system, is designed to meet the data storage requirement of astronomer under the big data environment. Astronomy Data Lake uses master-slave framework, integrate many geographic distribution data storage resource, then provider user a single mount point. It implements automatic data backup and disaster recovery, capacity expansion easily. Based on this system, we developed many data storage service, including database storage, private file storage, computing data and paper data service to meet the data storage requirement of astronomer.



P065 Yinbi Li

National Astronomical Observatories, Chinese Academy of Sciences

The metallicity distribution and unbound probability of 13 SDSS hyper-velocity stars

Hyper-velocity stars are believed to be ejected out from the Galactic center through dynamic interactions between (binary) stars and the central massive black hole(s). In this paper, we analyze 13 low mass F/G/K type hyper-velocity star candidates reported in our previous paper, which were from the seven data release of SDSS. We compare the metallicity distribution of our 13 candidates with the metallicity distribution functions of the Galactic bulge, disk, halo and globular cluster, and the result shows that the Galactic halo or globular cluster are likely the birth place for our candidates, which roughly consistent with our previous results obtained by kinematic analysis. In addition, we determine the unbound probability for each candidate using a Monte-Carlo simulation by assuming a non-Gaussian proper-motion error distribution, Gaussian heliocentric distance and radial velocity error distributions, which shows the probability that a HVS candidate can escape the Milky Way.

P066 Suo Liu

National Astronomical Observatories, Chinese Academy of Sciences

Toward Standard Data Production for Magnetic Field Observations at Huairou Solar Observing Station

Standard data to facilitate the use of scientific research, the productions of data with internationally agreed standards are the common pursuits for all astronomical observation instruments both ground and space-based undoubtedly. The routine solar observations are available at Huairou Solar Observing Station (HSOS) since 1987, which should be regarded as one main solar observations in the world. The data storage medium and format at HSOS experienced a few changes, so there exist some inconveniences for solar physicist. This paper shows that the observations data of HSOS are further processed both for storage medium and format toward international standards, in order to explore HSOS observations data for scientific research.

P067 Mireille Louys

Universite de Strasbourg/ CDS / ICUBE

Unique identifiers for facilities and instruments in astronomy and planetary science

Mireille Louys 1,2 Baptiste Cecconi 3 Sebastien Derriere 1, Pierre LeSidaner 4 CDS, Icube, Strasbourg, LERMA, GEPI, Paris, F

1 CDS, 2 ICUBE, Strasbourg, 3 LESIA, Observatoire de Paris-CNRS, 4 VOParis, Observatoire de Paris, Paris, France

The Virtual Observatory seeks a repository or thesaurus that could easily list the facilities and instruments used in the astronomical community at large, including solar physics and planetology.



We examined various existing lists, and propose a description scheme for all facilities and instruments, and a mechanism to define unique identifiers for the various items. Various use-cases for a name resolver application and examples are discussed.

P068 Lance Luvaul

Australian National University

Running the SkyMapper Science Data Pipeline: Better to be a big fish in a small pond or a small fish in a big ocean?

The SkyMapper Survey of the Southern sky was launched recently in 2014. Here we review the structure and frameworks behind the pipeline that will process the flood of new data, and consider the challenges of deploying on two disparate platforms: 1) a publicly shared, massively parallel, queue-scheduled compute fabric; and 2) a dedicated NUMA-based, multi-core, mini-supercomputer. Concepts reviewed include a) how to impose a layer of central operator control over hundreds of jobs of varying type and CPU/IO profile, all running at once and all at different stages in their logic, b) how to maintain configuration control in an ever-changing algorithmic environment while not giving up ease of build and deployment, and c) how to configure a NUMA-architected machine for optimal cache buffer usage, process-to-memory locality, and user/system CPU cycle ratio.

P069 Michael Mach

Institute for Astrophysics, University of Vienna

Mid-Infrared Astronomy with the E-ELT: Data Reduction Software for METIS

I will present a poster about the Data Reduction Software for the Mid-Infrared E-ELT Imager and Spectrograph (METIS) for the European Extremely Large Telescope (E-ELT). METIS will feature diffraction limited Imaging, low/mid resolution slit spectroscopy and high resolution Integral Field Spectroscopy at mid-infrared wavelengths, and will be equipped with state of the art adaptive optics. In the wavelength regime between 3-19 μm (L/M/N/Q1 Bands), data reduction is particularly challenging, as thermal radiation from the telescope / detector itself and the Earth's atmosphere is orders of magnitude higher than the flux from the science target, which greatly increases the detection noise. One of the contributions of the A* consortium (Universities of Vienna, Innsbruck, Graz and Linz) to the E-ELT is to provide an on- and off-line data reduction pipeline with a graphical user interface to produce science ready data for METIS. The pipeline will use ESO's highly specialized data reduction library named 'Common Pipeline Library' (CPL) and will contribute recipe additions to the new 'high level data reduction library' (HDRL) provided by ESO. The Poster will give an overview of the current status of our efforts in this project and will highlight the particular challenges of developing a robust and versatile data reduction environment for an instrument for a 40m-class telescope.



P070 Jeffrey Mader
Keck Observatory

The Design and Development of the NIRSPEC Data Reduction Pipeline for the Keck Observatory Archive

The Keck Observatory Archive (KOA), a collaboration between the NASA Exoplanet Science Institute and the W. M. Keck Observatory, serves science and calibration data for all current and retired instruments from the twin Keck Telescopes. In addition to the raw data, we publicly serve quick-look, reduced data products for four instruments (HIRES, LWS, NIRC2 and OSIRIS), so that KOA users can easily assess the quality and scientific content of the data. In this paper we present the design and implementation of the NIRSPEC data reduction pipeline (DRP) for KOA. We will discuss the publicly available reduction packages for NIRSPEC, the challenges encountered when designing this fully automated DRP and the algorithm used to determine wavelength calibration from sky lines. The reduced data products from the NIRSPEC DRP are expected to be available in KOA by mid-2016.

P071 Thomas McGlynn
NASA/GSFC HEASARC

The NASA Astronomical Virtual Observatories (NAVO): Coordinated and Comprehensive Access to NASA Mission Data Holdings

In October 2014, NASA commenced support for the NASA Astronomical Virtual Observatories (NAVO), a collaboration of four major NASA archives (IRSA, NED, MAST and the HEASARC) committed to the maintenance of the US Virtual Observatory (VO) infrastructure and the establishment of a coordinated and comprehensive VO interface to all major NASA-archived astronomical datasets. This includes nearly 100 mission/observatory datasets: the Great Observatories (Hubble, Spitzer, Chandra and Compton), other active missions (e.g., Fermi, Planck and Kepler) completed missions and ground-based datasets (e.g., 2MASS, the DSS and RXTE), and key high-level products, notably the holdings of the NASA Extragalactic Database. In this paper we describe our progress in the first year of the NAVO collaboration and plans for the future.

While NASA archives had separately provided some VO access to many of these datasets, under NAVO auspices we are working to ensure that all mission tables and data are available through a consistent set of VO protocols that users can easily discover and utilize. During our first year, powerful generic query capabilities have been provided for almost all holdings using the VO Table Access Protocol. All these holdings are described and discoverable in the Virtual Observatory Registry and a common Web presence has been developed to provide a consistent description of NASA's VO interfaces regardless of which NASA archive that a user may start at.

We are working to provide comprehensive access to mission data -- not just tables -- through VO protocols. In future years we will also be working to ensure that mission descriptions and metadata are consistent throughout NASA's holdings. Working within the framework of the International Virtual Observatory Alliance (IVOA), NAVO will help develop and implement the standards needed to provide access to modern astronomical data sets. The maturing of VO protocols will allow us to



provide VO access to our mission datasets at levels that begin to match our custom archive interfaces and enable the community to address their science needs through IVOA standards.

P072 Bruno Merin

European Space Astronomy Centre

ESA's Astronomy Multi-Mission Interface

ESA is working on a science-driven discovery portal for all its astronomy missions at ESAC with the provisional name Multi-Mission Interface. The first public release of this service will be demonstrated, featuring interfaces for sky exploration and for single and multiple targets. It requires no operational knowledge of any of the missions involved. From a technical point of view, the system offers all-sky projections of full mission datasets using a new-generation HEALPix projection called HiPS; detailed geometrical footprints to connect all-sky mosaics to individual observations; and direct access to the underlying mission-specific science archives.

A first public release is scheduled before the end of 2015 and will give users worldwide simplified access to high-level science-ready data products from all ESA Astronomy missions plus a number of ESA-produced source catalogues. A focus demo will accompany the presentation.

P074 Laurent Michel

SSC XMM-Newton - Observatoire de Strasbourg

Experience with Arches

ARCHES is a 3 years project funded by the FP7 European Community programme and ending in December 2015. It aims at producing well-characterised multi-wavelength data in the form of spectral energy distribution for large sets of objects extracted from the XMM-Newton source catalogue. These multi-wavelength data will help the scientific community in the exploration of a wide range of forefront astrophysical questions. Public outreach is an important activity of EC funded projects. After evaluating several ideas we have adopted the concept of an "Arches Walker" which provides an attractive way to display featured sky objects at different wavelengths. The description of the objects is stored in a simple corpus whereas the sky views are provided by HIPS maps published by the CDS. Arches Walker can be run from a WEB page as well as in a booth mode where the view of both images and data is remotely controlled by a dedicated application running on an Android tablet. The data corpus can easily be adapted to the targeted audience (object sample, language, education level) and any modification is automatically propagated to the Android controller. We are now in the dissemination process. The Arches Walker prototype has already been presented to a general audience and the first public release has been issued. The product is now available for the outreach community (planetarium, public center, etc).

P075 Marco Molinaro

INAF - Osservatorio Astronomico di Trieste

Taking Advantage of Cloud Solutions to Balance Requests in an Astrophysical Data Center

A complete astrophysical publishing environment, working as a helper system for an astrophysical data center, requires various components, from custom data back ends up to more or less standardized (e.g. Virtual Observatory driven) front end solutions. Combining this environment into one framework can lead to a potentially non scalable or hardly improvable system. In this contribution we describe what we are planning and developing to take advantage of cloud computing infrastructures and of a modular/distributed component architecture to provide a scalable and maintainable publishing environment at the Italian center for Astronomical Archives (IA2) at the INAF (Italian National Institute for Astrophysics) Astronomical Observatory of Trieste. Using a set of modular services, connected by registered interfaces, we are planning to use automated balancing at the front end to allocate services on demand on a cloud environment and allow generic data access on the back end archive solution.

P076 Chrystel Moreau

LAM/CeSAM

ASPIC: Public Spectroscopic Archives at CeSAM

Archive of Spectra Publicly available In Cesam (ASPIC) uses recognized scientific laboratory expertise and the technical expertise of CeSAM to make available to the scientific community data and tools for spectroscopic massive programs.

Many projects led at LAM or in which LAM is involved, such as VVDS, zCOSMOS, VIPERS, EUCLID, PFS, Athena, demonstrate the level of expertise and international acknowledgement of LAM in this area.

In each of these projects ASPIC has a major role: responsible for the development of the redshifts measurement and validation pipeline, 1D spectra production, spectra archiving.

ASPIC proposes to any similar program, even of smaller extent, to produce and make available the final spectroscopic data by providing tools for increasing their value and / or after the period of operation of the mission or of the observer program through a web application (ANIS) with high level services.



P077 Christian Motch

Observatoire de Strasbourg

The ARCHES Project

Authors: C. Motch, F. Carrera, F. Genova, L. Michel, A. Mints, A. Nebot, F.-X. Pineau, D. Reminiac, S. Rosen, A. Schwobe, E. Solano, M. Watson on behalf of the ARCHES consortium."

The Astronomical Resource Cross-matching for High Energy Studies (ARCHES) project is a three-year long FP7-Space funded programme started in 2013. The project involves the Observatoire Astronomique de Strasbourg including the CDS (France), the Leibniz- Institut for Astrophysik Potsdam (Germany), the University of Leicester (UK), the Universidad de Cantabria (IFCA, Spain) and the Instituto Nacional de Tecnica Aeroespacial (Spain). The project aims at providing the international community with well-characterised multi-wavelength data in the form of spectral energy distributions (SEDs) and catalogues of clusters of galaxies for large samples of both resolved and unresolved sources extracted from the 3XMM DR5 X-ray catalogue of serendipitous sources. SEDs are based on an enhanced version of the 3XMM catalogue and on a careful selection of the most relevant multi-wavelength archival catalogues (GALEX, SDSS, UCAC, 2MASS, GLIMPSE, WISE, etc). For that purpose ARCHES has developed advanced methods providing probabilistic cross-identification of several catalogues in one go as well as a multi-wavelength finder for clusters of galaxies. Importantly, these tools are not specific to the X-ray domain and are applicable to any combination of well-described multi-wavelength catalogues. Both tools and data will be made available through specific interfaces, through CDS services and through the Virtual Observatory. These enhanced resources are tested in the framework of several science cases involving both Galactic and extragalactic researches.

P078 Demitri Muna

Ohio State University

Introducing Nightlight: A New, Modern FITS Viewer

The field of astronomy distinguishes itself in having standardized on a single file format for the majority of data it produces. Visualization of FITS data, however, has not kept up with modern software design, user interfaces, or user interaction; the simple task of inspecting a file's structure or retrieving a particular value requires writing code. While the file format has its shortcomings, a significant reason many astronomers dislike FITS is not due to the organization of the bytes on disk, but rather the ease of use in accessing or visualizing them. Nightlight is a new desktop application whose aim is to bring a modern interface to FITS files with the polish and design people expect from applications like iTunes. By making it a native Macintosh application, one is able to leverage cutting edge frameworks (not available in cross-platform environments) that enable GPU acceleration, multithreading, interactive touch interfaces, real-time image processing, desktop metadata indexing, and more. Nightlight is designed to provide a common platform on top of which data or survey specific visualization needs can be built. The initial public release of Nightlight is expected in 2015.



P079 Yujin Nakagawa

Japan Aerospace Exploration Agency

User-Friendly Data Analysis Software and Archive for the Japanese X-ray All Sky Monitor MAXI on-board ISS

The Japanese X-ray astronomical instrument "Monitor of All-sky X-ray Image (MAXI) was launched on July 2009. It was equipped on the exposed facility of the Japanese experiment module "Kibo" on-board the International Space Station (ISS). The main scientific objective of the MAXI is to discover and promptly report new X-ray variable objects as well as steady monitoring of known X-ray sources in the whole sky. The MAXI science operation is approved by JAXA at least until the end of March 2018. Releases of the MAXI data mainly consist of three steps. The first step was releases of scientific products on the MAXI website. Following the second step of on-demand data release using a web interface started on November 2011 developments of data archives and softwares are performed by the MAXI archive team as the third step. We have reported requirements and plans of the developments at ADASS in 2014. We further report developed system of the MAXI data archives and softwares this time. Among the on-board detectors referred to as Gas Slit Camera (GSC) and Solid State Camera (SSC) the developments for the GSC are almost finished. On the other hand the developments for the SSC are in progress. We will finalize the developments and will release the MAXI data archives and softwares by the end of March 2016. We also have incorporated all-sky X-ray images of the GSC and the SSC into JAXA Universe Data Oriented 2 (JUDO2) with which users can easily browse MAXI X-ray images at any parts of the sky being superposed with other astronomical images. Now we create an automated X-ray image processing system for JUDO2. The release of data archives and softwares will allow scientist in the world to perform scientific analyses in a more flexible way and will enhance publications of scientific articles using the MAXI data by the scientists.

P080 Vicente Navarro Ferreruela

European Space Agency ES

GAVIP – Gaia AVI Portal, Collaborative PaaS for Data-Intensive Astronomical Science

Gaia was launched in December 2013 with a primary objective to determine the position, motion, brightness and colour of the stars in our galaxy, the Milky Way. In addition to performing the star survey, Gaia will be able to detect planets outside our solar system and asteroids inside our solar system. Gaia orbits around the L2 Lagrange point where it is shielded from our sun by the Earth. During the five years of its operational mission there is a series of planned releases of "products" that will culminate in a final catalogue at the end of operations. The Data Processing and Analysis Consortium (DPAC) is responsible for the systems that will produce these releases. Furthermore, the DPAC is responsible for the development of the Gaia science archive that will serve as the interface between the public and the DPAC products.

The Gaia archive, hosted at ESAC, will only become accessible to the public with the first data release from the DPAC that is expected not before Summer 2016. As Gaia data becomes available and



reaches a wider and wider audience, there is an increasing need to facilitate the further use of Gaia products. Considering the much richer data environment of the early 2020's, when the final Gaia archive will be available over the net, astronomers will want to connect with ground-based and other space-generated archives. As complex applications are developed, there will be a requirement to run a distributed application accessing one or more complete archives without pulling large datasets over the net. To address this need, the Gaia AVI portal (GAVIP) is to offer an innovative collaboration platform for scientists to deploy, close to the data, separately developed "Added Value Interfaces" (AVIs). These AVIs, packaged as Docker containers, define an ecosystem of processing elements that can be shared and combined into complex scientific workflows. The portal functionality is expected to include features such as:

- Collection of user specified requests and associated parameters.
- Search and retrieval of products from the Gaia archive.
- User specified processing tasks.
- Management and storage of processing outputs.
- Provision of services to integrate resources using VOSpace, TAP and SAMP.

The main objectives of this activity are:

- Design, develop, verify and deploy a web-based portal that will serve as a centralised access point for all Gaia AVIs.
- Develop an initial set of AVIs to be integrated in the deployed portal.
- Support the integration of other AVIs, which will be specifically developed to interface with the Gaia AVI portal.

Therefore GAVIP aims at supporting ESA's important role in the vanguard of data-intensive astronomical science for the next decade. This paper will describe GAVIP's work in progress with a focus on prototype implementation and major design decisions in areas like architecture, virtualisation, cloud computing and programming languages derived from the System Concept phase.

P081 Jon Nielsen

Mount Stromlo Observatory, Australian National University

Data Reduction for Integral Field Spectrographs: Past, Present, and Future.

Co-author: Rob Sharp

Integral Field Spectrographs have become mainstream on optical/NIR telescopes, both large and small. The Australian National University has a long history with these instruments. ANU designed and built NIFS for Gemini North and WiFeS for the 2.3m telescope at Siding Spring, both of which are in active use. Building on this success ANU is designing the GMTIFS instrument for the 25 meter Giant Magellan Telescope. This combined near-infrared AO-assisted IFS and Imager will be delivered in 2023. Past experience has shown that an accurate, robust, and responsive (i.e., near real-time) data reduction pipeline is essential for the successful early exploitation of IFS instruments.



Data reduction for NIFS (and initially also for WiFeS) was performed within the IRAF environment. While this approach provided high quality data products, it was found to be difficult to maintain and improve, and community support was limited. More recently the python-based PyWiFeS pipeline has provided a user-friendly data processing solution. Based on PyWiFeS, a prototype data simulator and reduction pipeline has been developed for GMTIFS.

We provide an overview of the three instruments, their data reduction pipelines, and discuss the future of IFS data reduction.

P082 Sara Nieto

ESAC Science Data Centre

ESA VOSpace: New Paradigm for Astronomical Data Sharing and Dissemination

In the context of the activities done by the ESAC Science Data Centre (ESDC), in charge of the development and operations of the ESA Science Archives, two main challenges are being faced; to guarantee the public access and preservation of data coming from ESA astronomical and planetary missions and to provide support to the community to new missions that produce big data outputs like, e.g. Euclid and Gaia. In order to fulfil the second requirement, sharing data mechanism is a needed to enable the new paradigm "bring the software to the data". This mechanism has been implemented through an ESA VOSpace service that it is expected to be made available to all ESA Science Archives users.

VOSpace, the IVOA standard protocol to persist and exchange data, has been implemented by the ESDC to provide data sharing capabilities from the ESA Science Archives to the users community in an interactive and easy to use way.

In case of the ESA Science Archives, and especially for missions producing big data outputs, VOSpace will allow the users to send the results from queries, data retrieval and crossmatch operations among others, to a common data sharing infrastructure, saving a huge amount of local storage resources. In that sense, ESDC VOSpace will connect the ESA Science Archives through a standard and common interface.

From an interoperability perspective, ESDC VOSpace, is intended to connect to other VOSpace services from different astronomical data centres, providing a common science network for the dissemination of astronomical data.

ESDC VOSpace provides two access mechanisms to achieve these goals: a simple and elegant Web based interface for human interaction and a RESTful interface for command-line access.

The Gaia and Euclid missions have triggered the use of ESDC VOSpace as the interface to exchange science data for the community. In this context, Gaia will be the first ESA mission to provide this service through its archive, allowing the users exchange their scientific results with the rest of the community.



We present the ESDC VOspace service for astronomical and planetary data exchange within the ESAC Science Archives and the ongoing work on its interconnection within the ESA Gaia Archive that would offer the Gaia scientific community a new way of sharing data.

P083 Maria Nieto-Santisteban

Space Telescope Science Institute

Engineering Data Services for the JWST Data Management System

The James Webb Space Telescope (JWST) data management subsystem (DMS) includes an innovative new engineering database (EngDB) system following a service-oriented architecture that ingests data, manages processes, and serves data to pipelines and the community. EngDB services are implemented using a RESTful architectural style to drive processes and access data via the well-defined HTTP protocol. JWST will generate roughly 5 TB of calibrated engineering data per year, with up to 300 million samples per day for 20000 parameters. Given the nature and volume of the data, separate databases are used to store approximately 2000 frequently-accessed parameters and many more rarely-accessed parameters from each instrument and other observatory components. This decomposition provides us with a natural way to provide operational robustness by duplicating the frequently-accessed database in operations to support the pipelines and in the Mikulski Archive for Space Telescopes (MAST) to support public access. For many previous missions, getting time series engineering data into the hands of observers was quite difficult. Access to the full time series of JWST engineering telemetry will enable scientific researchers and instrument scientists to identify and correct calibration artifacts that correlate with engineering parameters. This type of analysis can be very useful, for example when analyzing exoplanet transit data where the highest possible precision is needed.

P084 Simon O'Toole

Australian Astronomical Observatory

The All Sky Virtual Observatory AAT Node: Managing Heterogeneous Datasets

In this talk, we discuss the motivation and development of the Anglo-Australian Telescope node of the All-Sky Virtual Observatory. The project aims to provide IVOA-compliant access to all AAT and UKST data including legacy, ongoing and future surveys. The data to be included is in a wide range of formats, ranging from tabular, FITS and HDF5. The primary challenge of the project is therefore to store large amounts of heterogeneous data and then query it transparently. We also plan to develop the capability to cross-match data across different ASVO nodes (SkyMapper, TAO and MWA). Based on a survey of user requirements, we developed a simple data model for our system, which suggests that traditional relational database solutions are not sufficient for our requirements; greater flexibility is required than these systems can offer. We have settled on the Hadoop ecosystem, following an investigation of the various noSQL options (including MongoDB and Cassandra), as well

as SQL systems that allow more flexible schema, such as MariaDB and PostgreSQL. Hadoop gives us the flexibility to include relational data, as well as more document-driven data, and is therefore well matched to our needs. Finally, we discuss the advantages and disadvantages of using Hadoop for astronomical Big Data. A detailed discussion of our Hadoop implementation, including performance, will be presented by Harischandra et al.

P085 Sergio Pascual

Universidad Complutense de Madrid

PyASB, All-Sky Night Sky Brightness pipeline

All-sky cameras have proven to be powerful tools to continuously monitoring the sky in a wide range of fields in both Astrophysics and Meteorology. We have developed a complete software pipeline (pyASB) to analyze the night CCD images obtained with the ASTMN-OT all-sky camera. The pipeline provides parameters used assess the quality of the night sky, such as the Sky Brightness, the Cloud Coverage and the Atmospheric Extinction, how they evolve over the time and their variability.

Repository: <https://github.com/guaix-ucm/PyASB>**P086 Michal Pawlak**

Warsaw University Astronomical Observatory

Hierarchical Machine Learning in Large Photometric Surveys - Selection and Classification of Eclipsing Binaries in the OGLE-IV Data

With the rapidly increasing amount of data collected by large photometric surveys, the amortization of the process of selection and classification of variable object becomes more and more important. The OGLE-IV project monitoring more than billion stars is a good example of such case. For the purpose of effective selection a classification of eclipsing binaries, the system based on machine learning technique is proposed. Catalogs of binary stars from previous phase of the project are used as training sets. The classification process is hierarchical including various steps. Each of them focuses on one slightly different task i.e. preselection of candidates, rejection of false detections, and separation of eclipsing binaries from other variables etc. Such approach results to be significantly more effective than a single-step classification.



P087 Joshua Peek

Space Telescope Science Institute

The GALFA-HI Data Reduction Pipeline & DR2: Objectless Åœbercal with 7679 filters

GALFA-HI is a neutral hydrogen 21-cm line survey of the Galaxy that has taken over 10,000 hours of data on the Arecibo telescope, covering 1/3 of the sky. I will describe methods we developed to mitigate many of the strong systematics inherent in single dish, obstructed aperture radio telescopes, allowing us to push well beyond the standard systematic floor. I will discuss our version of the "Åœbercal" self-calibration method, issues related to pushing self-calibration into the hyperspectral domain, and relevance for pathfinders like ASKAP. I will also debut the GALFA-HI Data Release 2, and show a number of stunning scientific and visualization highlights.

P088 Maura Pilia

INAF - Osservatorio Astronomico di Cagliari

Data reduction of multi-wavelength observations

Authors: M. Pilia, A. Trois (INAF - Osservatorio Astronomico di Cagliari, Italy)

We are developing a software to combine gamma-ray data from multiple telescopes with the aim to cross calibrate different instruments, test their data quality and to allow the confirmation of transient events using different instruments. In particular, we present the first applications using pulsar data from the AGILE and Fermi satellites and show how we solved the issues relative to combining different datasets and rescaling the parameters of different telescopes. In this way we extend the energy range observed by a single telescope and we can detect fainter objects. As a second step, we apply the technique of pulsar gating to the imaging data of the satellites combined, to look for pulsar wind nebulae. The same procedure is adopted in the radio domain, using data from the Sardinia Radio Telescope. We aim to be able to use similar techniques for multifrequency datasets spanning a large range of the electromagnetic spectrum. We also present the work in progress to include the automatic search for gamma-ray counterparts within the pipeline for pulsar search in radio.

P089 Franois-Xavier Pineau

JNanocubes: on-the-fly generation of HiPS density maps for exploration and visualization of large datasets

We present a follow-up on JNanocubes, an astronomy oriented Java implementation of the nanocubes data structure.



Nanocubes is developed by AT&T Research who provide C++ open source code. We have adapted it so that it can be used to enable the exploration and visualization of large astronomical data sets, in the particular framework of Hierarchical Progressive Survey (HiPS). By using HiPS, the spatial indexation is based on HEALPix, and we make use of the Aladin Lite HiPS visualizer. JNanocubes allows on-the-fly generation of individual HiPS tiles where each tile is a density map that is constructed from the catalogue data with user-provided constraints. We report new major developments including: the serialization of the data structure; the support of multi-dimensional discretized parameters; linked access to the original records. We illustrate these new features with tests made on color-color diagrams of hundreds of millions of sources extracted from the SDSS catalogue.

We also assess how we could provide these Nanocubes features as a service that would allow easy exploration of any dataset.

P090 Nikolay Podorvanyuk

Sternberg Astronomical Institute, Moscow State University

Stellar atmosphere interpolator for empirical and synthetic spectra

We present a new stellar atmosphere interpolator which we will use to compute stellar population models based on empirical and/or synthetic spectra.

Empirical stellar spectra (e.g. ELODIE and MILES stellar libraries) are broadly used in stellar population synthesis codes, however they do not cover the temperature-gravity-metallicity (Teff-logg-[Fe/H]) parameter space uniformly that lead to artefacts in stellar population models computed with fully empirical grids. Those gaps can be filled using synthetic stellar atmospheres (e.g., PHOENIX), however, using fully synthetic grids is not recommended because they suffer from incomplete line lists

We propose a new flexible interpolation scheme based on: (a) a combination of empirical and synthetic spectra which cover the entire Teff-logg-[Fe/H] parameter space; (b) in each spectral pixel we use smoothing splines (b-spline) on Teff and a low-order two-dimensional polynomial surface on fit on logg and [Fe/H]. The resulting parametrization is then evaluated at desired point of the parameter space. The use of b-splines helps us to resolve a long-standing issue in the stellar population modelling regarding a mathematically correct way of stellar atmosphere interpolation that does not cause discontinuities in resulting stellar population models that later hamper stellar population analysis in real galaxies and star clusters. We present a semi-empirical stellar atmosphere grid that can be plugged into the PEGASE.HR stellar population synthesis code.

We computed a new grid of stellar population models using our stellar atmosphere grid and analysed spectral data for 3000 giant elliptical galaxies from the SDSS survey and other galaxies using full spectrum fitting. We compare our results (ages, metallicities) with those obtained using existing stellar population models (MIUSCAT and PEGASE.HR) and analyze the artefacts in the age and metallicity determination using different model grids.

P091 Kai Lars Polsterer

HITS gGmbH

*Virtual Observatory Virtual Reality**Authors: Kai L. Polsterer, Mark Taylor*

The virtual observatory (VO) and its standards have become a success story in providing uniform access to a huge amount of data sets. Those data sets contain correlations, distributions, and relations that have to be unveiled. Visualization has always been a key tool to understand complex structures. Typically high-dimensional information is projected to a two dimensional plane to create a diagnostic plot. Besides expensive stereoscopic visualization cubes, only stereoscopic displays provided an affordable tool to peek into a three dimensional data space.

We present a low-cost immersive visualization environment that makes use of a smart-phone, a game controllers and Google cardboard. This simple equipment allows you to explore your data more natively by flying through your data space. The presented software consists of a central server application running on a computer and a client implementation performing the rendering on multiple smart-phones, enabling users to inspect the data jointly. As the server application uses the VO simple application messaging protocol (SAMP), it is seamlessly integrated with other VO tools, like topcat or aladin. Access the data in the usual way and employ Virtual Observatory Virtual Reality (VOVR) to explore it.

P092 Anne Raugh

University of Maryland

Metadata Wrangling in the New PDS4 Standards

The Planetary Data System (PDS) archives, supports, and distributes data of diverse targets, from diverse sources, to diverse users. One of the core problems addressed by the PDS4 data standard redesign was that of metadata - how to accommodate the increasingly sophisticated demands of search interfaces, analytical software, and observational documentation into the PDS product labeling standards without imposing limits and constraints that would impinge on the quality or quantity of metadata that any particular observer or team could supply.

Both data suppliers and data users want detailed information in a predictable - and preferably programmatically accessible - structure, while retaining a high degree of flexibility to deal with variety, specificity, and new developments in observations and their associated documentation. PDS, as a multi-generational archive, wants explicit documentation for every aspect of intelligence included with the data as part of its preservation and usability mandate. But PDS also knows from decades of experience that attempts to predict or constrain the breadth and depth of metadata that come with new instruments and techniques are as ill-advised as they are doomed to failure. And yet, PDS must have detailed documentation for the metadata in the labels it supports, or the institutional knowledge encoded into those attributes will be lost - putting the data at risk.



The PDS4 metadata solution is based on a three-step approach: adoption of open standards to define the architectural foundation; a hierarchy of namespaces that allows modularization of the information model and delegation of stewardship to appropriate localized authorities; and a strict model-driven design approach, based on the object-oriented methodologies of the foundational standards, that accommodates extension of existing metadata models to adapt to changing circumstances without invalidating or obscuring existing metadata in the archive.

This poster illustrates the key features of the PDS4 approach to metadata and its implications for data preparers, for end users, and for interoperability between archives.

P093 Jennifer Riding

University of Melbourne

Shapelet Modelling for the MWA

The Murchison Widefield Array (MWA), like other low frequency or recently upgraded radio telescopes, has a large field of view. For the MWA, this means any observation will contain a mix of point-like sources, extended sources and diffuse emission. The variety makes for fascinating science and challenging calibrations and deconvolutions. For example, bright extended sources complicate the calibration process by corrupting the visibilities of point-like calibrator sources. These effects are more concerning when a bright extended source exists at the edges of the beam or in its side lobes.

There are a number of ways to handle bright extended sources and the method proposed here subtracts a shapelet model of the offending source from the raw uvdata. Shapelet models make use of smooth gauss-hermite polynomials to better model extended emission with fewer data points than the traditional clean-component maps. Due to their simple Fourier relation, shapelets can be used in the UV domain. It is this method that the MWA's Real Time System will employ to peel foregrounds from MWA fields in the search for the Epoch of Reionisation signal

P094 Carlos Rodrigo Blanco

Centro de Astrobiología, INTA-CSIC

VOSA, VO SED Analyzer.

VOSA (VO Sed Analyzer, <http://svo2.cab.inta-csic.es/theory/vosa/>) is a public web-tool developed by the Spanish Virtual Observatory (<http://svo.cab.inta-csic.es>) and designed to help users to (1) build Spectral Energy Distributions (SEDs) combining private photometric measurements with data available in VO services, (2) obtain relevant properties of these objects (distance, extinction, etc) from VO catalogues, (3) analyze them comparing observed photometry with synthetic photometry from different collections of theoretical models or observational templates, using different techniques (chi-square fit, Bayesian analysis) to estimate physical parameters of the observed objects (temperature, mass, luminosity, etc), and use these results to (4) estimate masses and ages using collections of isochrones and evolutionary tracks from the VO. In particular, VOSA offers the



advantage of deriving physical parameters using all the available photometric information instead of a restricted subset of colors. The results can be downloaded in different formats or sent to other VO tools using SAMP.

VOSA is in operation since 2008 (Bayo et al, 2008, A&A 492,277B). At the time of writing this proposal there are more than 500 active users in VOSA (~7.000 files uploaded by users and ~600.000 objects analysed), and more than 50 refereed papers have been published making use of this tool.

In the framework of the GENIUS (<https://gaia.am.ub.es/Twiki/bin/view/GENIUS>) project we have upgraded VOSA to provide access to Gaia photometry and give a reliable estimation of the physical parameters (effective temperatures, gravities, metallicities, masses and ages) of thousands of objects at a time. This upgrade has required, on the one hand, the implementation of a new computation paradigm (including a distributed environment, the capability of submitting and processing jobs in an asynchronous and the use of parallelized computing to speed up processes) and, on the other hand, a redefinition of the web interface to handle large lists of objects and the corresponding information. One of the advantages of this upgrade is that processes are, in average, ten times faster.

P095 Michèle Sanguillon
LUPM, France

Assessment of the IVOA Provenance Data Model Concepts for CTA Data Products

In order to describe final data products delivered by the Cherenkoc Telescop Array (CTA) project, as well as their dependency on their progenitors, we examine which concepts defined in the W3C data model can be adequately re-used.

We describe current use-cases for the computational Provenance in the CTA production pipeline and explore the proposed W3C notations like Prov-N formats for our usage.

P096 Andre Schaaff
CDS, Observatoire de Strasbourg, CNRS, UDS

3D-Visualization of astronomical data in a Web browser

Authors: Andre Schaaff, Nicolas Deparis, Nicolas Gillet, Pierre Ocvirk (CDS, Observatoire de Strasbourg, CNRS, UDS), Arnaud Steinmetz, Pierre Lespingal, Nicolas Buecher (ENSIIE Strasbourg)

We present an ongoing work started this year around 3D-Visualization of astronomical data in a simple Web browser, especially simulation data and data from Vizier catalogues. The development is based on Javascript / WebGL and offers a multi-view display, several ingestion formats of the data, highlighting of physical values, etc. After native developments during the last years with the Oculus Rift to navigate in simulation data we have also implemented the Oculus view as a capability of our tool. It works with nightly build browsers implementing MOZVR. As the tool can run in a smartphone



Web browser we provide also a Cardboard view but currently with limited interactions. The immersive capabilities are proposed as an added value without being mandatory.

P097 Christopher Schollar

SKA SA, University of Cape Town

Radio Frequency Interference Monitoring for the MeerKAT Radio Telescope

South Africa is currently building MeerKAT, a 64 dish radio telescope array, as a pre-cursor for the proposed Square Kilometre Array (SKA). Both telescopes will be located at a remote site in the Karoo with a low level of Radio Frequency Interference (RFI). It is important to maintain a low level of RFI to ensure that MeerKAT has an unobstructed view of the universe across its bandwidth. The only way to effectively manage the environment is with a record of RFI around the telescope. The RFI management team on the MeerKAT site has multiple tools for monitoring RFI. There is a 7 dish radio telescope array called KAT7 which is used for bi-weekly RFI scans on the horizon. The team has two RFI trailers which provided a mobile spectrum and transient measurement system.

They also have commercial handheld spectrum analysers. Most of these tools are only used sporadically during RFI measurement campaigns. None of the tools provide a continuous record of the environment and none of them perform automatic RFI detection. Here we design and implement an automatic, continuous RFI monitoring solution for MeerKAT. The monitor consists of an auxiliary antenna on site which continuously captures and stores radio spectra. The statistics of the spectra describe the radio frequency environment and identify potential RFI sources. All of the stored RFI data is accessible over the web. Users can view the data using interactive visualisations or download the raw data. The monitor thus provides a continuous record of the RF environment, automatically detects RFI and makes this information easily accessible. This RFI monitor functioned successfully for over a year with minimal human intervention. The quality of the data and visualisations has been tested by MeerKAT engineers and astronomers. The monitor represents a clear improvement over previous monitoring solutions used by MeerKAT and is an effective site management tool.

P098 Min-Su Shin

Korea Astronomy and Space Science Institute

Applications of Multiple Machine Learning Algorithms for Reliable Detection of Variable Sources in Time-Series Data Depending on Required Speed and Precision Requirements.

Due to increasing size of astronomical data and expected boom of survey projects, it becomes important to detect interesting objects reliably in the large amount of data. Focusing on application of clustering algorithms to detect groups in data, I introduce a non-parametric Bayesian clustering method and a consensus clustering method which improves reliability of detecting genuine variable sources in astronomical time-series data. I also present a new strategy of time-series data analysis to

identify variable sources quickly by using ensemble of clustering methods as the data size grows, allowing people to find important candidates for follow-up observations quickly.

P099 Yuji Shirasaki
NAOJ

Current Status of JVO Portal

JVO portal is an astronomical data discovery service utilizing the Virtual Observatory as a basic data search system, and it can be accessible at <http://ivo.nao.ac.jp/portal>. The main features of the JVO portal are:

- (1) quick search on multiple major big catalogs,
- (2) dedicated search interface for Subaru and ALMA dataset,
- (3) VO-enabled data search.

We started to redesign the VO search interface in 2013 to improve the usability of VO data search functionalities, and the redesigned interface is open to public as an experimental version of JVO portal v2. On this version, the most of the fundamental VO search interfaces such as ""VO basic search.

P100 Petr Skoda

Astronomical Institute of the Czech Academy of Sciences

The Distributed Cloud Based Engine for Knowledge Discovery in Massive Archives of Astronomical Spectra

The current archives of large-scale spectroscopic surveys, such as SDSS or LAMOST, contain millions of spectra of celestial objects as well as catalogues of their classification estimated by automatic pipelines based on global template fitting. Some interesting objects (e.g. emission line stars, cataclysmic variables, or quasars) can be identified, however, only by checking shapes of certain spectral lines. As this can be hardly done by visual preview of millions of spectra, machine learning techniques have to be applied, complemented by flexible visualization of results. The proper application of methods of Knowledge Discovery in such mega-surveys may also bring new discoveries of yet unknown types of objects which were not correctly classified by pipelines. This motivation initiated our development of VO-CLOUD, the distributed cloud-based engine, providing the user with the comfortable web-based environment for conducting machine learning experiments with different algorithms running on multiple nodes. It allows visual backtracking of the individual input spectra in different stages of preprocessing, which is important for checking the nature of outliers or precision of classification. The engine consists of a single master server, representing the user portal, and several workers, running various types of machine learning tasks. The master holds the database of users and their experiments, predefined configuration parameters



for individual machine learning models and a repository for a data to be preprocessed. The workers have different capabilities based on installed libraries and HW configuration of their host (e.g. number of CPU cores or GPU card type) and may be dynamically added to provide new machine learning methods. Spectra for experiments can be obtained by several ways, including local files upload, ftp and http recursive download and, namely, using Virtual Observatory SSAP, DataLink and SAMP protocols with advanced post-processing involved. We present the architecture of VO-CLOUD, identify features specific to handling of Big Data and demonstrate its functions and typical user interaction in a short demo.

P101 Enrique Solano

Centro de Astrobiología (INTA-CSIC)

Gran Telescopio Canarias OSIRIS Catalogues

The Gran Telescopio Canarias (GTC) (<http://gtc.sdc.cab.inta-csic.es/gtc/>) archive is the result of a collaboration agreement between the Centro de Astrobiología (CAB, INTA-CSIC) and GRANTECAN S.A. The archive, in operation since November 2011, has been developed in the framework of the Spanish Virtual Observatory and is maintained by the Data Archive Unit at CAB. The archive contains both raw and reduced data of two instruments: OSIRIS and CanariCam. It has been designed in compliance with the standards defined by the International Virtual Observatory Alliance (IVOA) which guarantees a high level of data accessibility and handling.

Reduced data are of fundamental importance for astronomical archives as they enhance their use by the community and provide a higher visibility of the project results. In parallel to the procedure implemented by the CAB Data Center for the community to return GTC reduced data to the archives (more than 50 observing programmes have already submitted more than 7000 reduced science files to the archive), we are using an upgraded version of Alambic[1] to process OSIRIS raw images. Source extraction is performed using SExtractor[2] and PSFEx[3] providing aperture, PSF and model photometry. OSIRIS images are astrometrically calibrated using 2MASS as a reference catalogue. Two approaches for the photometric calibration (search for counterparts in large surveys using the same photometric filters or photometric standards observed the same night) are currently being assessed. Morphometric parameters (ellipticity, FWHM, $\hat{\epsilon}$) for the extracted sources will also be provided in the catalogue.

The first release of the OSIRIS catalogue is expected to be delivered in Autumn 2015 and will contain sources from more than 16000 images observed in the period 2009-2013.

[1] 2002SPIE.4847..123V, [2] 1996A&AS..117..393B, [3] 2013ascl.soft01001B13.



P102 Mauricio Solar

Technical University Federico Santa Maria

The ChiVO Library: Advanced Computational Methods for Astronomy.

The main objective of the Advanced Computational Astronomy Library (ACALib) is to ensemble a coherent software package with the research on computational methods for astronomy performed by the first phase of the Chilean Virtual Observatory between years 2013 and 2015. During this period, researchers and students developed functional prototypes, implementing state of the art computational methods and proposing new algorithms and techniques. This research was mainly focused on spectroscopic data cubes, as they strongly require computational methods to reduce, visualize and infer astrophysical quantities from them, and because most of the techniques are directly applicable either to images or to spectra.

The ACALib philosophy is to use a persistent workspace abstraction where spectroscopic data cubes can be loaded from files, created from other cubes or artificially generated from astrophysical parameters. Then, computational methods can be applied to them, resulting in new data cube instances or new data tables in the workspace. The idea is to provide not only API bindings for the workspace and the cubes, but also web-services to use the library in cloud-based frameworks and in the Virtual Observatory.

In a nutshell, ACALib is integrating and testing several cube manipulation routines, stacking procedures, structure detection algorithms, spectral line association techniques and a synthetic data cube generation module. The library is developed in python, strongly rooted in astropy modules and using efficient numerical libraries such as numpy and scipy, and machine learning libraries like scikitlearn and astroML.

In the near future, we plan to propose ACALib as an astropy affiliated package, and to create a CASA add-on to ease the usage of our methods. Also, we are exploring bleeding-edge computational methods to include to ACALib, such as deep learning networks, and developing new prototypes for other types of astronomical data, such as light curves in the time-domain.

P103 Yihan Song

National Astronomical Observatories, Chinese Academy of Sciences

Sky Subtraction on Bright Night for LAMOST

LAMOST is a large sky area multi-object fiber spectroscopic telescope. It surveys on the both dark nights and bright nights. Sky background subtraction is important and difficult on the bright nights. In this poster, we talk about some problems which are encountered during our work.



P104 Travis Stenborg
Macquarie University

Efficient Sky Subtraction from Emission Nebula Slit Spectra in C OpenMP

A new, efficient sky subtraction technique for emission nebula slit spectra reduction is presented. The technique yields reduced spectra more accurately baselined to zero flux than obtained with typical contemporary community practice. This accurate baseline gives reliable emission line ratios for use in determining nebula properties, such as extinction from the Balmer decrement. A parallel programming implementation is described applied to planetary nebula observations, but applies equally well to any astronomical objects exhibiting sparse emission line spectra such as supernova remnants, HII regions, etc., saved into a pixel array.

P105 Christian Surace
Laboratoire d'Astrophysique de Marseille

Gazpar: Galaxy Photometric Redshifts (Z) and Physical Parameters

GAZPAR is a WEB service offered to the community to help in measuring galaxy photometric redshifts and physical parameters. This service relies on two public codes developed maintain and support at LAM: Le Phare and CIGALE. These codes are based on SED (Spectral Energy Distribution) fitting techniques applied to multi-color imaging data ranging from the UV to IR.

The user can upload its own catalogue and get back not only the photometric redshifts and/or the physical parameters but also diagnostics and a scientific feedback.

P106 Giuliano Taffoni
INAF-OATs

Workflows and Science Gateways for Astronomical Experiments

Workflow and science gateway technologies have been adopted by scientific communities as a valuable tool to carry out complex experiments. They offer the possibility to perform computations for data analysis and simulations, whereas hiding details of the complex infrastructures underneath. In this talk we describe our experiences in creating workflows oriented science gateways based on gUSE/WS-PGRADE technology. This technology allows astronomers to develop workflows and science gateways using any workflows management system (e.g. Kepler, Taverna, Mouter) and even combining them together. Astronomers can use their preferred workflows system and recycle workflows. Although the major obstacle of workflow recycling is that workflow systems are not normally compatible, our adoption of our framework allows to overcome this limitation. We notice that this approach improves efficiency and reliability by reusing tested methodologies, it increases the lifetime of workflows and it reduces development time for new workflows and consequently science gateways. We provide a few examples of real production environments developed using



gUSE/WS-PGRADE technology and some perspectives for future implementation (e.g. the workflow system for EUCLID space mission, or E-EET data reduction science gateways).

P107 Mark Taylor
University of Bristol

TOPCAT's TAP Client

TAP, the Table Access Protocol, is a Virtual Observatory (VO) protocol for executing queries in remote relational databases using ADQL, an SQL-like query language. It is one of the most powerful VO-based tools, but also one of the most complex to use, with an extensive stack of associated standards. This complexity, along with patchy service implementation, has meant that science use of TAP has been fairly limited to date, even though the TAP standard has been endorsed in its current form since 2010.

We present here significant improvements to the client and GUI for interacting with TAP services from the TOPCAT table analysis tool, introduced in the recent version 4.3, but also available for standalone or embedded use. As well as the management of query submission and response decoding required for basic TAP operations, the GUI attempts to provide the user with as much help as possible in locating services, understanding service metadata and capabilities, and constructing correct and useful ADQL queries. The implementation and design are, unlike previous versions, both usable and performant even for very large TAP services, for instance TAPVizieR which hosts around 30,000 tables.

It is hoped that TOPCAT's enhanced TAP user interface, alongside parallel developments in other available TAP clients, evolution of associated standards, and continuing improvements in service implementations, will lead to more widespread use of TAP in making optimal use of the vast and increasing amount of astronomical data which is exposed using this protocol.

P108 Jeff Valenti
STScI

Science Concept for an Improved JWST Mosaic Planning Tool

The James Webb Space Telescope (JWST) will spend a significant fraction of its mission lifetime obtaining mosaic data. In the Science Operations Design Reference Mission, which is a plausible pool of JWST science programs, approximately one third of the science time is used for mosaics, including 50% of the NIRCcam imaging time and 70% of the MIRI imaging time. The baseline mosaic capability for JWST produces a rectangular grid of pointings with overlap, skew, and rotation parameters. JWST uses a sunshield for passive cooling, so the instantaneous range of allowed roll angles is narrow and depends on when an observation is scheduled. This means baseline mosaics are often valid only in relatively narrow scheduling windows, especially when using rectangular NIRCcam tiles. Our new

flexible mosaic concept allows users to specify an irregular sky region of interest, possibly containing interior subregions that need not be mapped. A simple heuristic algorithm automatically determines a tile pattern that fully covers the sky region at one roll angle. A first pass determines the minimum number of tiles needed to completely cover the sky region as a function of roll angle. A second pass determines the fraction of the sky region covered by a fixed number of tiles as a function of roll angle. The proposer requests and the TAC approves a fixed number of tiles, simplifying subsequent accounting. More importantly, the mosaic can be rescheduled without contacting the observer. The tool also provides the user with feedback on guide star availability for all tiles as a function of roll angle. This simplified mosaic approach makes it easier for proposers to submit nearly complete programs, shortening the time between proposal submission and program execution.

P109 Juan C. Vallejo

GMV for ESA

BepiColombo MPO Instrument Pipelines Integration in the Science Operations Control System Framework

BepiColombo is an interdisciplinary ESA-JAXA mission to explore the planet Mercury consisting of two separate orbiters: ESA's Mercury Planetary Orbiter (MPO) and JAXA's Mercury Magnetospheric Orbiter (MMO). The ESA orbiter payload comprises 11 instruments covering different scientific disciplines developed by several teams. The Science Ground Segment (SGS), located at the European Space Astronomy Centre (ESAC), will be in charge of preparing the science operations for MPO including data processing and distribution to the instrument teams and the scientific data archiving in a central archive accessible to the science community.

Although traditionally, Instrument Teams are responsible for processing, analyzing and preparing their science data for the long-term archive, in BepiColombo, the SGS will play a key role in these activities; having the full responsibility for the production of the first level of MPO science data (un-calibrated). In some cases the teams will develop totally their pipeline SW and in others, the SGS will co-develop it with team's collaboration. For all the instruments, the SGS will integrate and execute, the software used for production of un-calibrated science data and for the rest of data levels will require a primary-redundant pipeline configuration where some instrument pipelines will be operated from the instrument team's data center's, having a replica that can be run from the SGS, while others will be executed as primary pipelines from the SGS, with the SGS adopting, in all cases, the pipeline orchestration role. This role requires the implementation of a framework that copes with distributed data processing pipelines, and developed in different coding languages.

We describe the methods and techniques used by the SGS to integrate and execute the pipeline SW developed in different, but limited, set of coding languages and operating systems in the SGS pipeline execution environment, which includes monitoring, and control functionalities services. This pipeline execution environment is developed in JAVA using standard practices and resources within a continuous integration environment. A service oriented modular architecture approach and the possibility to monitor and control the execution of SW developed in different platforms and coding language provide an excellent opportunity to be reused for other type of pipelines.



P110 Ger van Diepen
ASTRON

LOFAR Fast Transient Imaging

Interferometric synthesis is one of the major observation modes of LOFAR. It uses a few thousand baselines and many thousands of frequency channels resulting in a data rate of a few GB/sec. Apart from storing these data in MeasurementSets on disk for continuum or spectral line imaging, the data can also be used for transient detection by forming images for each time sample. This paper describes the flagging, calibration and imaging steps to form such images that can be fed into another pipeline (TraP) for transient detection. A posteriori processing is also possible using the stored MeasurementSet as the data source.

P111 Dany Vohl
Swinburne University of Technology

An interactive, comparative and quantitative 3D visualisation system for large-scale spectral-cube surveys using CAVE2

As the quantity and resolution of spectral-cubes from optical/infrared and radio surveys increase, desktop-based visualisation and analysis solutions must adapt and evolve. Novel immersive 3D environments such as the CAVE2 at Monash University can overcome personal computer's visualisation limitations. CAVE2 is part advanced 2D/3D visualisation space (80 stereo-capable screens providing a total of 84 million pixels) and part supercomputer (80 TFLOP/s of integrated GPU-based processing power). We present a novel visualisation system enabling simultaneous 3D comparative visualisation of $O(100)$ spectral-cubes. With CAVE2 operating as a large 3D tiled-display augmented by our newly developed web-based controller interface, astronomers can easily organise spectral-cubes on the different panels, apply real-time transforms to one or many spectral cubes, and access quantitative information about the displayed data. We also discuss how such a solution can help accelerate the discovery rate in varied research scenarios.

P112 Yijing Wang
Beijing Normal University

Sparse Expression - A Case of Radio Astronomy Image Compression

Yijing WANG¹, Guanyin GAO¹, Dan WU², Lei QIAN, Xianchuan YU¹, Wenwu TIAN²

¹ Beijing Normal University, Beijing 100875, China; ² National Astronomical Observatories, Chinese Academy of sciences, Beijing 100012, China yuxianchuan@163.com chuan.yu@ieee.org wangyijing@mail.bnu.edu.cn twu@bao.ac.cn wudan@bao.ac.cn gaoguanyin0099@126.com lqian@nao.cas.cn

With radio astronomy data exploding, lossless compression, one existing astronomical image data compression algorithms, can guarantee to restore the image without distortion, but the compression ratio is too low, and it has brought great challenge to the radio astronomy image



storage and transmission. In view of this situation, this paper proposes a sparse expression astronomical image compression algorithm. The algorithm uses the K-SVD algorithm to get KSVD dictionary through a complete DCT atoms library updates adaptively, and representing the astronomical data sparsely using the dictionary, then compressing sparse coefficient obtained by improved run-length algorithm coded and stored as a binary stream. Experiments show that for FITS format radio astronomy data processing, when the compression ratio is 5: 1, the difference between raw data and the decompressed data after a lossy compression is minimal, the mean square error is from 0.0026 to 0.0337, it will not affect interpreting of the very weak information of the data.

P113 Craig Warner

University of Florida

Highly Configurable, Near Real-Time Data Pipelines

As we proceed into the era of very large telescopes, there are many challenges thrust upon the development of next-generation data pipelines: 1) the desire to process data in near real-time, providing data quality feedback concurrent with ongoing observations and thus maximizing observing efficiency and precious telescope time; 2) the need to handle ever increasing volumes of data as array sizes continue to increase; and 3) the flexibility to be able to run on different platforms and hardware, ranging from laptops and single board computers to high-end workstations with powerful Graphics Processing Units (GPUs). In addition, reduction algorithms may differ vastly for each specific instrument. Thus, either separate data pipelines must be developed for each instrument, continually re-inventing the wheel, or a common framework is needed that is both highly configurable and easily extensible.

We present superFATBOY, a next-generation data pipeline that is currently used to reduce imaging data for the Canarias InfraRed Camera Experiment (CIRCE) at the Gran Telescopio Canarias (GTC) telescope. superFATBOY can be configured and extended to reduce data from virtually any near-IR or optical instrument. It was designed to be able to harness the power of massively parallel algorithms developed using Nvidia's Compute Unified Device Architecture (CUDA) platform to enable near real-time data processing while retaining the ability to seamlessly run on machines without CUDA-enabled GPUs simply by changing one parameter in the input XML configuration file.

superFATBOY is highly configurable, due to the design choice of using XML configuration files as input. The XML configuration files are comprised of three sections: <queries> describes the data, <processes> describes the processes (in order) that will be applied to the data, and <parameters> describes global settings. It is also easily extensible support for different data formats and new processing algorithms can be added by simply extending the base classes of fatboyDataUnit or fatboyProcess, respectively, and overriding a few key methods. Access to the common framework (for example, the ability to find a particular set of calibration frames, recursively process them, and combine them into a master calibration frame to be applied to a data frame) is provided through a well-defined API. This allows superFATBOY to be easily tailored virtually any near-IR or optical instrument. Furthermore, since superFATBOY is a Python package, it can be either run from the command line or imported from within another Python program, such as a quick-look display tool.



Finally, we note that superFATBOY will be the basis for the upcoming MIRADAS data reduction pipeline. MIRADAS is a near-IR, multi-object, high resolution, cross-dispersed spectrograph that is scheduled to be installed on the GTC in 2019.

P115 Matthew Whiting
CSIRO

The Development and Testing of the Early ASKAP Processing Pipelines

The Australian Square Kilometre Array Pathfinder (ASKAP) requires high-performance processing pipelines to be able to calibrate and image observations spanning its large field-of-view. This talk will present some of the work we have undertaken to develop pipelines that produce images and catalogues that meet the scientific expectations. This work has spanned complex science simulations, data challenges and, now that we have the initial 6-antenna commissioning telescope (BETA), comparisons between the ASKAP processing software ASKAPsoft and other, more mature, reduction packages. We show the results of such comparisons, and detail the future of the ASKAPsoft pipelines as ASKAP moves towards its Early Science program in 2016.

P116 Andrew Williams
Curtin University

Monitor and Control software for the Murchison Widefield Array.

The Murchison Widefield Array (MWA) is a low frequency (80-300Mhz) radio telescope operating in Western Australia at the Murchison Radio-astronomy Observatory. It has 2048 dual-polarisation dipoles laid out in groups of sixteen on 128 'tiles', each with an analogue beam-former. Sixteen digital receiver boxes in the desert each digitise and process eight of these dual-polarisation tiles. The receivers transmit the data (24 coarse channels, each 1.28MHz wide) via fibre to a hybrid correlator in the CSIRO control building composed of four custom FPGA cards for digital filtering, and twenty four servers with nVidia cards to do the cross-multiply using CUDA on GPU's.

In normal operation, the raw data rate of 168 GBytes/sec at the receivers is reduced to around 0.8 GBytes/sec of correlated visibility data sent via fibre to Perth for archival. There is also the capacity to capture short periods of uncorrelated data at ~7.8 GBytes/sec for high time resolution observations, using a 50Tb array of 10,000 rpm disks.

This poster describes the structure of the MWA software system and the computing infrastructure it runs on.



P117 Duncan Wright

University of New South Wales

Implemented 2D PSF Extraction in High Resolution Spectrographs

The poster will outline the implementation of 2D point-spread function extraction as applied to the HARPS and HERMES (Australia) high-resolution fibre spectrographs. It will demonstrate the benefits of such an extraction as compared to a typical 1D extraction using real data.

P118 Chao Wu

National Astronomical Observatories, Chinese Academy of Sciences

Ground Wide Angle Cameras of SVOM

The Ground Wide Angle Cameras (GWAC) is dedicated ground instrument on GRB optical emission observation and optical transient search in SVOM project. Thanks to its wide field of view (5000 Sq. deg) and high cadence (15 sec), GWAC is an excellent time-domain astronomical observation instrument. We will review the instruments performance and its data processing challenges on large volume data management and real time processing. Our progress and preliminary solution are also presented in the poster.

P119 Jian Xiao

Tianjin University

Calculating Non-equilibrium Ionization in Heterogeneous CPU-GPU Systems

Non-equilibrium ionization (NEI) is an important phenomenon related to many astrophysical processes, and typical NEI simulations require solving a large number of ordinary differential equation (ODE) at each global time step. However solving the associated ODEs even with modern methods is still time-consuming. This paper presented a GPU-optimized approach to accelerate NEI computation. We also proposed a dynamic load balance strategy on hybrid multiple CPUs and GPUs architecture to maximize performance. Comparing with the 24 CPU cores (2.5GHz) parallel scheme, our approach on 4 Tesla C2075 GPUs achieves a speedup of up to 15.



P120 Masafumi Yagi

National Astronomical Observatory of Japan (NAOJ)

Ghost in the Subaru Prime Focus Camera

In a data reduction of narrow-band images for a study on extended objects, stray lights of bright objects (so called ghosts) are often troublesome. For a better detection and measurement of astronomical objects, we investigated the nature of the ghosts in the Subaru Prime Focus Camera (Suprime-Cam) and tried to purify. We first investigated the observed data, detected the ghosts, and classified them into several types according to their shape and the relative position to the bright source object -- circular, cometary, horseshoe, arched, etc --. They are compared with the mechanical design of the focal optical unit and the instrument, the filter information, and with a ray-tracing calculation, and we identified the reflecting surfaces of some types of the ghosts. It was realized that the behavior of the ghost was sometimes different from the calculation from the mechanical and optical design. Ghosts let us know a possible tilt of surfaces, a change of the position of components, and an extra vignetting which are not expected from the design. Then, we derived several empirical relations about the shape and the position of the ghosts, and implemented a software to fit and subtract the ghosts in the Suprime-Cam data semi-automatically. The performance of the software applied to the real data will be presented.

P121 Michael Young

Indiana University

Big Data Challenges in the Blanco DECam Bulge Survey (BDBS)

As part of a multi-year effort to survey 200 square degrees of the Southern Milky Way Bulge in SDSS ugrizY utilizing the Dark Energy Camera (DECam) on the Blanco 4m telescope, the Blanco DECam Bulge Survey (BDBS) has taken >350k source frames. Utilizing a distributed architecture executing dophot, we have extracted ~15 billion source detections from these frames. With one of the primary goals of the survey being the creation of a catalog as a community resource, we have explored a number of ways to facilitate the querying and dissemination of this dataset. Given the wide and deep nature of the data, a traditional RDBMS is unsuitable for all but the simplest of queries. Here we will present our efforts to leverage the open-source Apache Hadoop/HDFS/Hbase stack, a widely recognized industry-standard approach to Big Data problems. Running on relatively inexpensive hardware, we will demonstrate how solutions designed for the commercial web have already addressed many of the concerns facing scientists in the Big Data future.



P122 Yanxia Zhang

National Astronomical Observatories, Chinese Academy of Sciences

K-nearest neighbour approach for photometric redshift estimation of quasars

We apply k-Nearest Neighbour (KNN) approach to predict photometric redshifts of quasars with SDSS, WISE and UKIDSS photometric data and compare the effect of various distances of KNN on the performance of estimating photometric redshifts. The results show that the Mahalanobis distance of KNN is better than other distances when predicting photometric redshifts. It is necessary to reduce the dimension of data to improve the performance of approaches with high dimensional data. We use LASSO on our sample for feature selection and then employ KNN to estimate photometric redshifts. The experimental result indicates that LASSO+KNN is an efficient method to increase the performance and efficiency of regressors when dealing with high dimensional data.