

Astroinformatics I (Semester 1 2024)

Introduction & Course Logistics

Nina Hernitschek

Centro de Astronomía CITEVA
Universidad de Antofagasta

April 22, 2024

What you will learn in this class

Astronomers spend about 100 % of their time working with a computer.

This course will **prepare you** for:

- How to use specialized computer programs to do research
- How to use a computer efficiently to do research
- How to write programs for typical research tasks

Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

Tools

Survey

Outlook

What you will learn in this class

Astronomers spend about 100 % of their time working with a computer.

This course will **prepare you** for:

- How to use specialized computer programs to do research
- How to use a computer efficiently to do research
- How to write programs for typical research tasks

What this course does **not** cover:

- advanced programming skills
- machine learning
- usage of \LaTeX

(You might find some of that in other courses offered.)

Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

Tools

Survey

Outlook

Course Logistics

course content:

- lecture: Monday 11 - 13h
- tutorial: Wednesday 12 - 14h
- graded project practice
- project presentation

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Course Logistics

Motivation
Course Logistics
What is Astroinformatics?
Astronomical Data
Tools
Survey
Outlook

course content:

- lecture: Monday 11 - 13h
- tutorial: Wednesday 12 - 14h
- graded project practice
- project presentation

grading:

- 4 graded project practices of 2 hours each (60% of total grade)
content: application of what has been learned in class, report in English to produce in one week
evaluation: capacity to solve a problem using astronomical software tools, written expression
- participation (10 % of total grade)
- project presentation (30 % of total grade)

Course Logistics

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

course content:

- lecture: Monday 11 - 13h
- tutorial: Wednesday 12 - 14h
- graded project practice
- project presentation

grading:

- 4 graded project practices of 2 hours each (60% of total grade)
content: application of what has been learned in class, report in English to produce in one week
evaluation: capacity to solve a problem using astronomical software tools, written expression
- participation (10 % of total grade)
- project presentation (30 % of total grade)

contact and course material:

- e-mail: nina.hernitschek@quantof.cl
- github: https://github.com/ninahernitschek/astroinformatica_I_2024_1

Course Logistics

Motivation
Course
Logistics
What is
Astroinformatics?
Astronomical
Data
Tools
Survey
Outlook

April 22 Lecture 1: Introduction, **April 24** Tutorial 1

May 7 (!) Lecture 2: Linux, **May 8** Tutorial 2

May 13 Lecture 3: Astronomical Survey Data, **May 15** Tutorial 3

May 20 Graded Project Practice 1

May 27 Lecture 4: Programming Introduction, **May 29** Tutorial 4

June 3 Lecture 5: Python Introduction, **June 5** Tutorial 5

June 10 Graded Project Practice 2

June 17 Lecture 6: Data Exploration & Visualization, **June 19** Tutorial 6

June 24 Lecture 7: Astronomical Packages **June 26** Tutorial 7

July 1 Graded Project Practice 3

July 8 Lecture 8: Software Projects, **July 10** Tutorial 8

July 15 Graded Project Practice 4

July 22 Project Presentation

Rules for Coding, Presentations, Report

coding:

- use meaningful variable names
- use comments and documentation
- If you have a question when something doesn't work, summarize what you tried - often this will even lead to the solution.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Rules for Coding, Presentations, Report

Motivation
Course Logistics
What is Astroinformatics?
Astronomical Data
Tools
Survey
Outlook

coding:

- use meaningful variable names
- use comments and documentation
- If you have a question when something doesn't work, summarize what you tried - often this will even lead to the solution.

project practice reports and project presentation:

- use \LaTeX
- figures: all own figures should be in vectorized pdf format
- for all data: data description (incl. citation)
- own work: properly cite what is not your own work; discuss how the previous work is similar to or different from your own work
- implementation: medium-level implementation description with libraries/ software frameworks (incl. citation)
- discussion: reflect your approach (strengths, weaknesses, limitations), lessons learned
- bibliography

Textbooks

I can recommend the following textbooks and online resources:

Python cookbook; Beazley, D. & Jones, B.; O'Reilly Media.

A Beginner's Guide to Working with Astronomical Data. M. Pössel, The Open Journal of Astrophysics, vol. 3, issue 1, id. 2, 2020.

Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

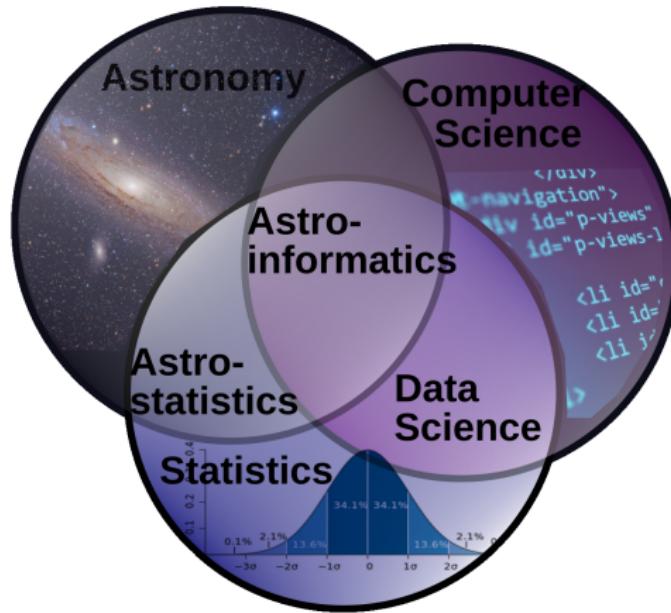
Tools

Survey

Outlook

What is Astroinformatics?

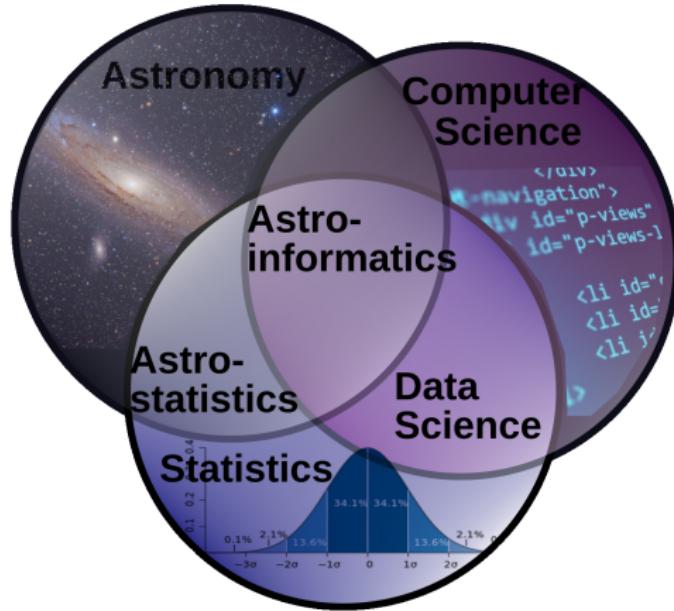
Astroinformatics is an **interdisciplinary** field:



- Motivation
- Course Logistics
- What is Astroinformatics?
- Astronomical Data
- Tools
- Survey
- Outlook

What is Astroinformatics?

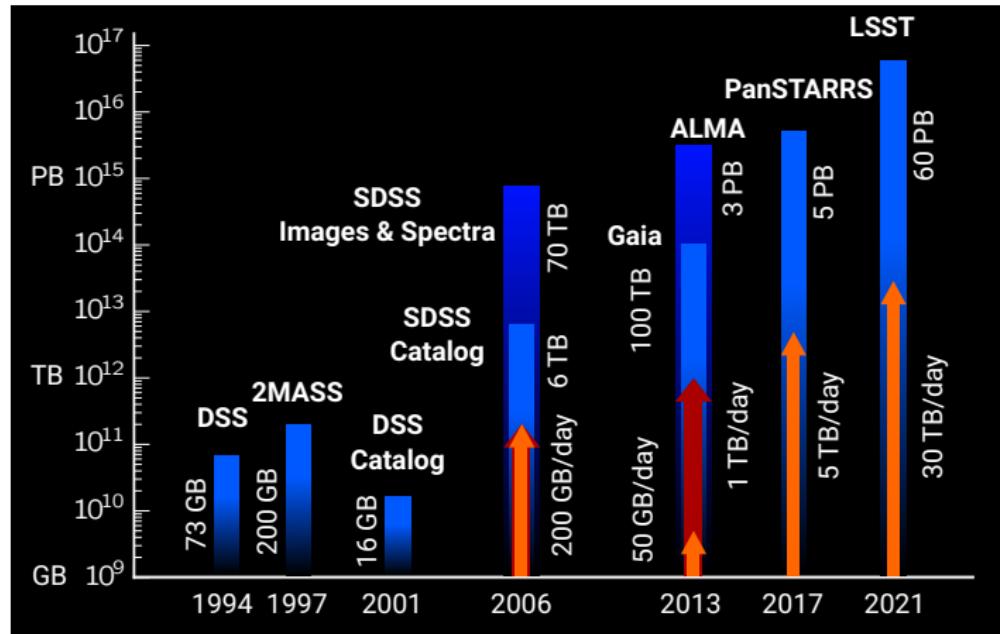
Astroinformatics is an **interdisciplinary** field:



Astroinformatics is primarily focused on developing the tools, methods, and applications for research in data-heavy astronomy.

Why we need Astroinformatics?

Modern All-Sky Surveys come with an increasing data volume:



Why we need Astroinformatics?

upcoming large all-sky surveys like (but not only) LSST

Challenge:

- enormous data volume, < 60 PB total, 15 - 30 TB / night
- follow-up opportunities should be identified immediately

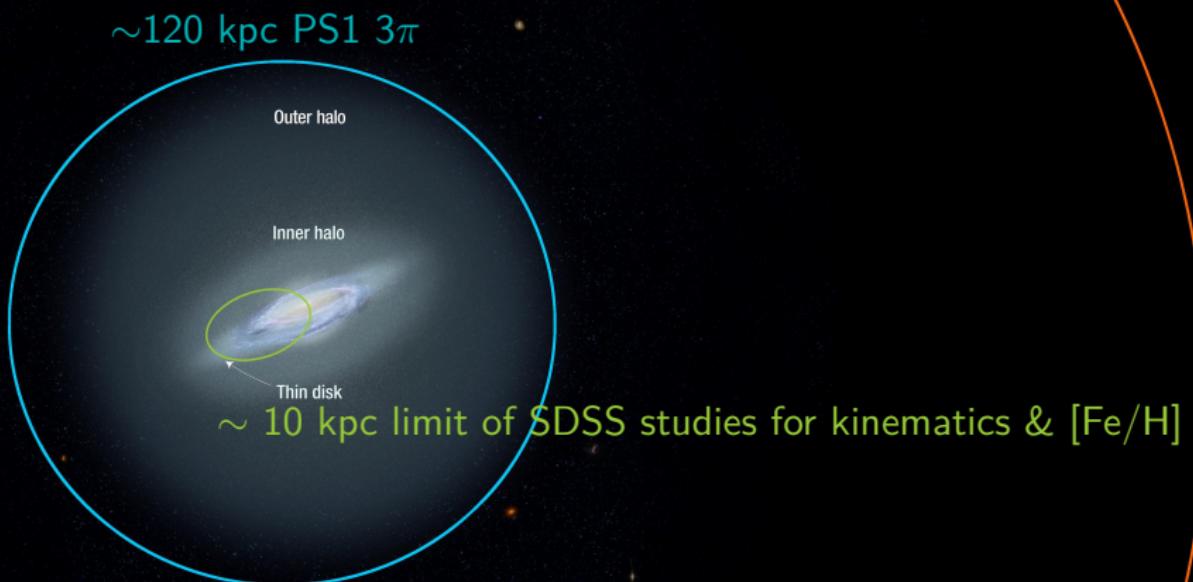
Chance:

large data volume
enables for

population studies

finding rare ‘one-in-a-million’,
‘one-in-a-billion’ events, often
called *anomalies*





Why we need Astroinformatics?

Astronomers need to solve equations, sometimes symbolic, sometimes numerically. Those are for example:

- Equations of movement of celestial bodies. (Compare to: The discovery of Neptune was made by solving equations by hand.)
- Computation of the stellar interior.
- Hydrodynamical simulations of which some now contain billions of particles.
- Machine learning for data from large all-sky surveys, for such as classification.
- Survey strategy planning, e.g. LSST survey strategy uses machine learning.
- Astroengineering problems, e.g. adaptive optics.

Motivation

Course

Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Example Use Cases

how I use computers/ computer programs in research

- planning observations

Motivation

Course
Logistics

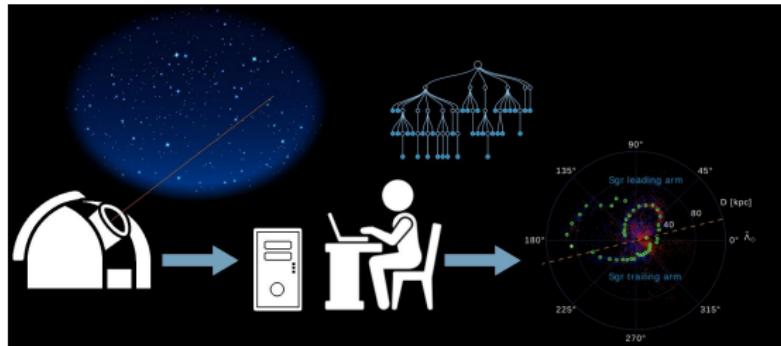
What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

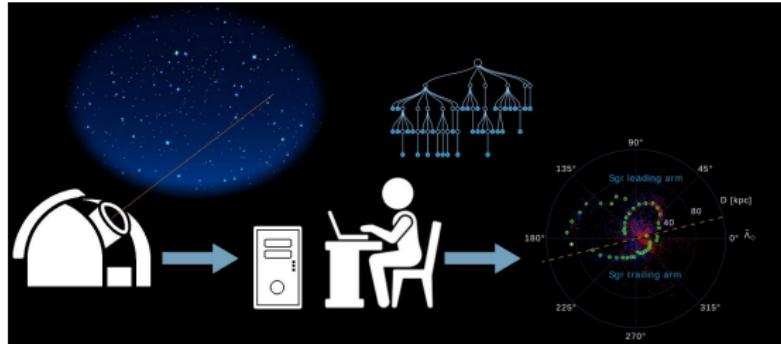
Outlook



Example Use Cases

how I use computers/ computer programs in research

- planning observations
- downloading, plotting, preprocessing, merging data



Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

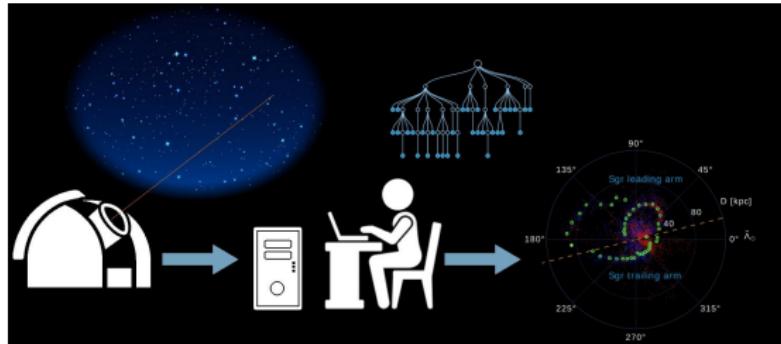
Survey

Outlook

Example Use Cases

how I use computers/ computer programs in research

- planning observations
- downloading, plotting, preprocessing, merging data
- classifying data (machine learning)



Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

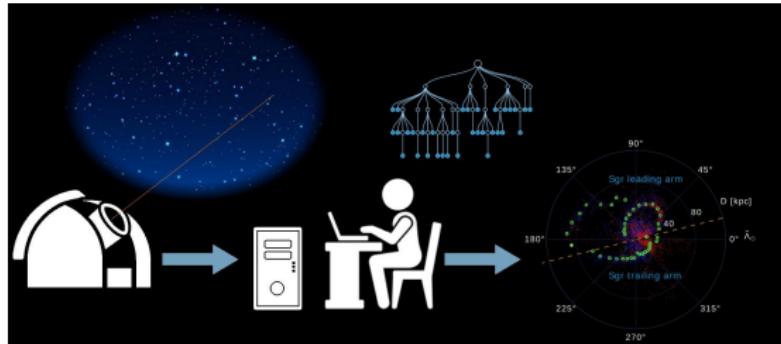
Survey

Outlook

Example Use Cases

how I use computers/ computer programs in research

- planning observations
- downloading, plotting, preprocessing, merging data
- classifying data (machine learning)
- creating complex visualizations



Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

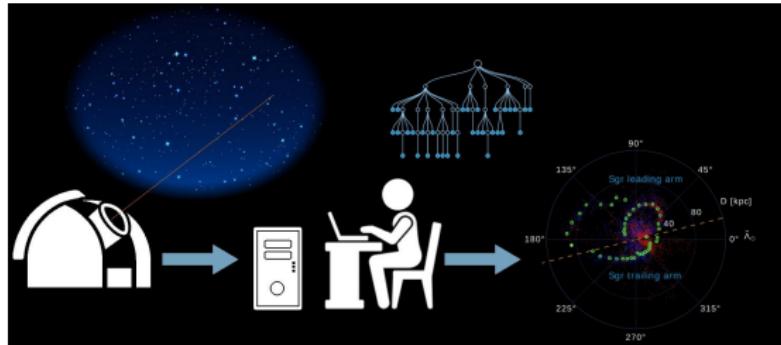
Survey

Outlook

Example Use Cases

how I use computers/ computer programs in research

- planning observations
- downloading, plotting, preprocessing, merging data
- classifying data (machine learning)
- creating complex visualizations
- developing a LSST broker software



Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Simulations

first astronomical simulation by Holmberg (1941):

1941ApJ...94..385H

THE ASTROPHYSICAL JOURNAL

AN INTERNATIONAL REVIEW OF SPECTROSCOPY AND
ASTRONOMICAL PHYSICS

VOLUME 94

NOVEMBER 1941

NUMBER 3

ON THE CLUSTERING TENDENCIES AMONG THE NEBULAE

II. A STUDY OF ENCOUNTERS BETWEEN LABORATORY MODELS OF
STELLAR SYSTEMS BY A NEW INTEGRATION PROCEDURE

ERIK HOLMBERG

ABSTRACT

In a previous paper¹ the writer discussed the possibility of explaining the observed clustering effects among extragalactic nebulae as a result of captures. The present investigation deals with the important problem of whether the loss of energy resulting from the tidal disturbances at a close encounter between two nebulae is large enough to effect a capture. The tidal deformations of two models of stellar systems, passing each other at a small distance, are studied by reconstructing, piece by piece, the orbits described by the individual mass elements. The difficulty of integrating the total gravitational force acting upon a certain element at a certain point of time is solved by replacing gravitation by light. The mass elements are represented by light-bulbs, the candle power being proportional to mass, and the total light is measured by a photocell (Fig. 1). The nebulae are assumed to have a flattened shape, and each is represented by 37 light-bulbs. It is found that the tidal deformations cause an increase in the attraction between the two objects, the increase reaching its maximum value when the nebulae are separating, i.e., after the passage. The resulting loss of energy (Fig. 6) is comparatively large and may, in favorable cases, effect a capture. The spiral arms developing during the encounter (Figs. 4) represent an interesting by-product of the investigation. The direction of the arms depends on the direction of rotation of the nebulae with respect to the direction of their space motions.

Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

Tools

Survey

Outlook

Simulations

first astronomical simulation by Holmberg (1941):

The paper is a study of the tidal disturbances appearing in stellar systems which pass one another at small distances.

Every mass element is represented by a small light-bulb, the light being proportional to the mass, and the total light along the x and y axes is measured by a combination of a photocell and a galvanometer. The measured values represent the components of the gravitational force. The components are obtained by adding up the attractions due to individual mass elements, each multiplied by the cosine of the corresponding projection angle.

Motivation

Course

Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

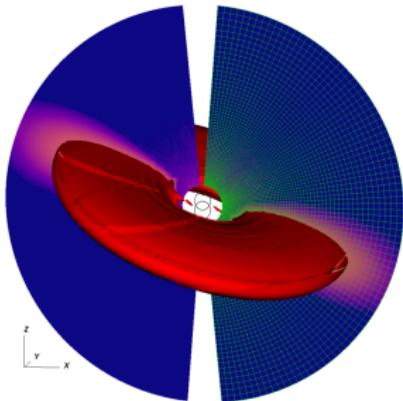
Simulations

modern astronomical simulations run on the computer:

Grid-based simulations:

Here, space is divided into cells, for instance a space-filling set of small cubes. For each cell, properties such as density and temperature are tracked while they change as the simulation runs.

In more complex simulations, the grid itself can also change in ways that are adapted to making the simulation more efficient (*adaptive grid*).



The evolution of an inclined disk around an eccentric binary was simulated using the grid-based code ATHENA++. Colors show a slice of density along the xz plane, with red contour showing a 3D isosurface. The orbits of the binary stars are visible in the center.

Credit: Rabago et al. (2022).

Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

Tools

Survey

Outlook

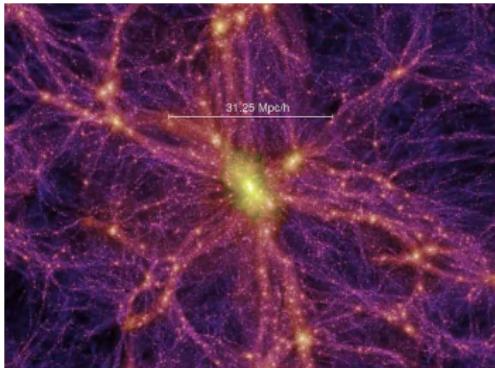
Simulations

modern astronomical simulations run on the computer:

N-body simulations:

A N-body simulation is a simulation of massive particles under the influence of physical forces, usually gravity and sometimes other forces. Depending on the simulation, a point particle could represent a lump of gas, or a star, or a group of a few 10^4 or 10^5 stars in a galaxy, or a lump of dark matter.

Every particle is characterized by a position in space (and potential other properties) which change under the physical forces as the simulation runs.



The Millennium Simulation simulates the universe until the present state, where structures are abundant, manifesting themselves as stars, galaxies and clusters. The dark matter density field is colour-coded by density and local dark matter velocity dispersion. Credit: Springel et al. (2005).

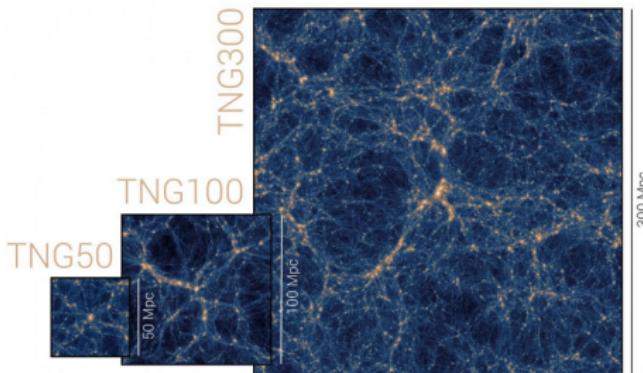
Simulations

Cosmological simulations:

There are many N-body simulations in cosmology. One is IllustrisTNG, a series of magnetohydrodynamical simulations of galaxy formation.

The smallest but most detailed of the TNG runs, TNG50, follows the fate of a cube that, in the present universe, has a side-length of 50 Mpc.

Within this volume, dark matter is represented by 10 billion point particles. 10 billion point particles are representing gas.

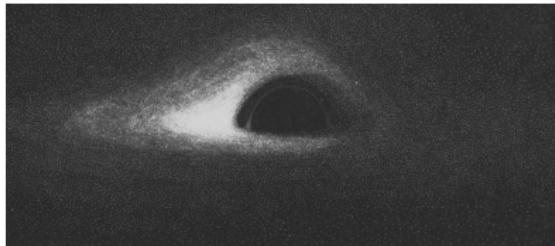


All the results and data from the three IllustrisTNG simulation volumes – TNG50, TNG100 and TNG300 – are publicly available (<http://www.tng-project.org/>). The data volume sums up to **more than one Petabyte**.

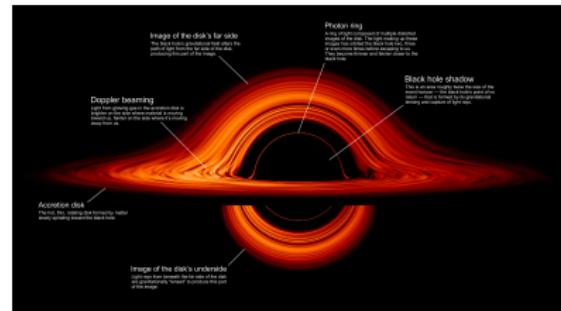
Simulations

relativistic raytracing now and then

Raytracing is the simulation of the path of light through spacetime. This rendering technique can simulate physically accurate reflections, refractions, shadows, and indirect lighting. Also different spacetime geometries can be taken into account to for example generate realistic images of the neighborhood of a black hole, such as the accretion disk.



Visualization of a black hole, calculations done by using a computer, data points drawn by hand, J.-P. Luminet, (1979)



Visualization of a black hole by using a raytracing algorithm, NASA (2019)

Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

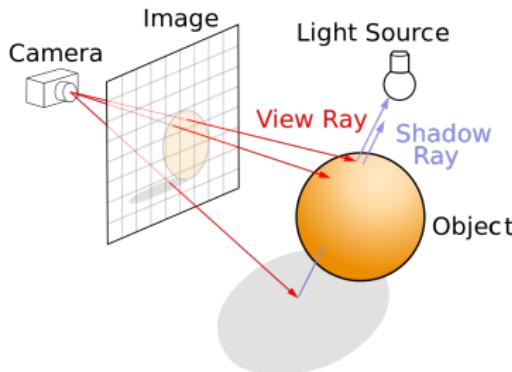
Tools

Survey

Outlook

Simulations

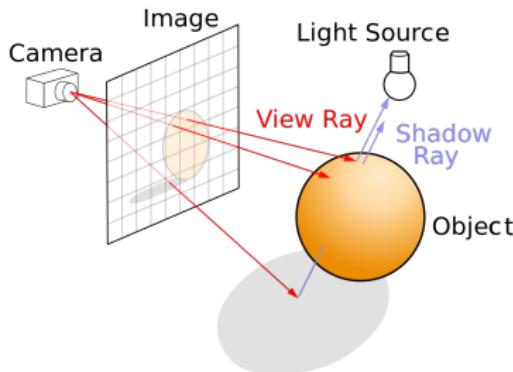
The Raytracing Algorithm



An image is constructed by extending rays through the image plane into a scene, bouncing them off objects and towards light sources to approximate the color value of pixels. Scenes in raytracing are described mathematically. Light rays might be bent by non-Euclidean spacetime geometry.

Simulations

The Raytracing Algorithm



An image is constructed by extending rays through the image plane into a scene, bouncing them off objects and towards light sources to approximate the color value of pixels. Scenes in raytracing are described mathematically. Light rays might be bent by non-Euclidean spacetime geometry.

Sending rays in the other direction than light travels in reality might at first be counterintuitive, but doing so is many orders of magnitude more efficient. As most light rays from a given light source do not make it into the viewer's eye, a "forward" simulation could potentially waste a tremendous amount of computation on light paths that are never recorded.

Survey Strategy

Observation planning (efficiently scheduling astronomical observations) is a very important topic in observational astronomy:

- visibility duration (location, season)
- moon influence
- pointing
- multiple objects per field
- weather
- target (object) characteristics
- exposure time necessary

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Survey Strategy

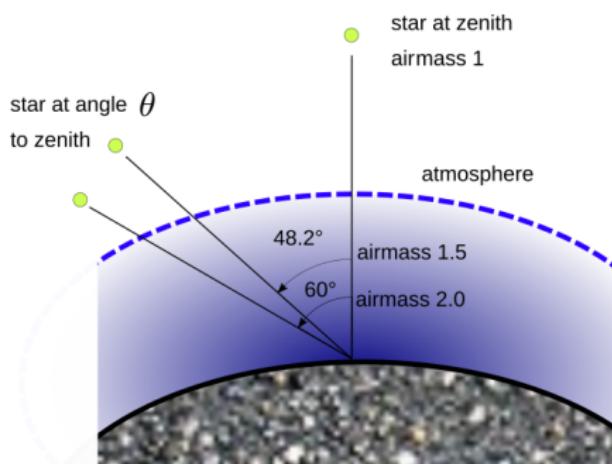
Motivation
Course Logistics

What is Astroinformatics?

Astronomical Data Tools Survey Outlook

In astronomy, **airmass** is a measure of the amount of air along the line of sight when observing through Earth's atmosphere.

The closer your target is to the horizon, the more air you have to look through, and the more degraded your view gets. The amount of air directly overhead is called **one airmass**.



If θ is the angle between zenith and the observed star (or other object), then a simple model of airmass X is $X = 1/\cos(\theta)$.

Survey Strategy

Motivation
Course Logistics
What is Astroinformatics?
Astronomical Data
Tools
Survey
Outlook

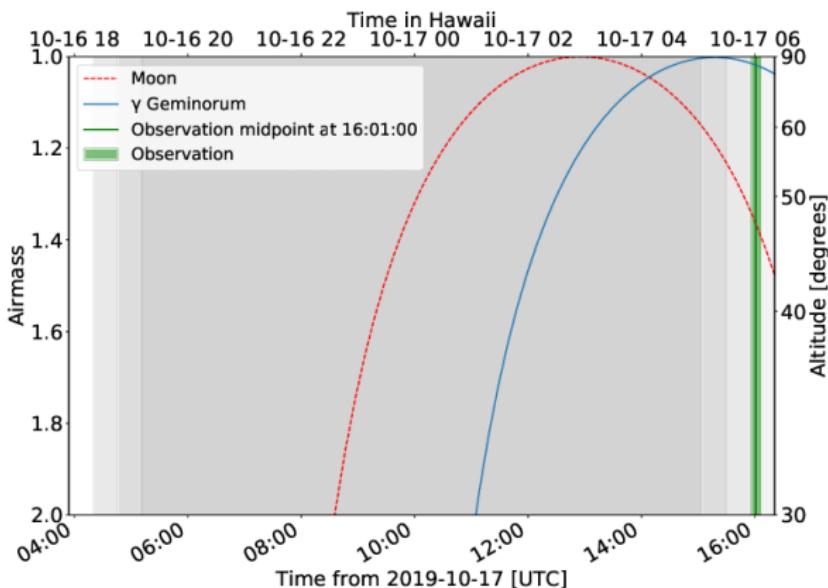


Illustration of time vs. airmass/altitude of the observation at 6:01 UTC -10 on the 17th of Oct. 2019 from Mauna Kea Observatory (green vertical region), for γ Geminorum (solid blue line) and the moon (red dashed). Exposure started at the very end of observing time, close to sunrise.

Survey Strategy

What is the case for individual observations or small surveys is also the case for large all-sky surveys:

Survey strategies needs to be **optimized**.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

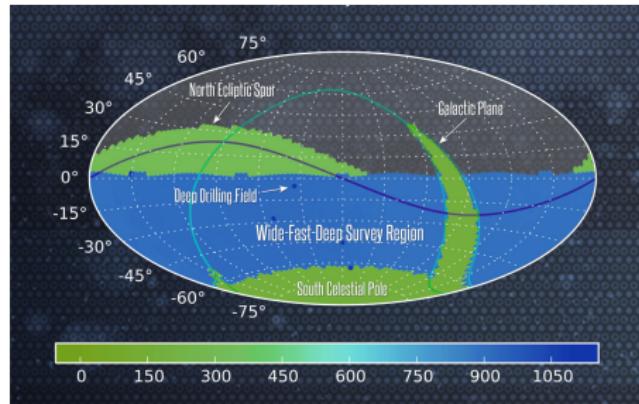
Survey

Outlook

Survey Strategy

The LSST Survey

- 10-year photometric *ugrizy* survey (near-UV, optical, near-IR)
- depth of $r \sim 27.5$ mag
- 1000 images/night = 15 TB/night, 10 million transients/night
- start of operations: 2024



LSST survey strategy, number of visits incl. sub-surveys. (credit: www.lsst.org)

Survey Strategy

The LSST Survey

Synthesizing the requirements to accomplish the four primary science objectives of Rubin Observatory,

- Probing dark energy and dark matter
- Taking an inventory of the Solar System
- Exploring the transient optical sky
- Mapping the Milky Way

results in the following **constraints** for the LSST Survey Strategy:

Motivation

Course

Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Survey Strategy

The LSST Survey

Synthesizing the requirements to accomplish the four primary science objectives of Rubin Observatory,

- Probing dark energy and dark matter
- Taking an inventory of the Solar System
- Exploring the transient optical sky
- Mapping the Milky Way

results in the following **constraints** for the LSST Survey Strategy:

Cosmological parameter estimation by many techniques requires uniform coverage of 18,000 square degrees of sky. Obtaining accurate photometric redshifts in every field requires a specified number of visits in each filter.

Motivation

Course

Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Survey Strategy

The LSST Survey

Synthesizing the requirements to accomplish the four primary science objectives of Rubin Observatory,

- Probing dark energy and dark matter
- Taking an inventory of the Solar System
- Exploring the transient optical sky
- Mapping the Milky Way

results in the following **constraints** for the LSST Survey Strategy:

Weak lensing shear measurements benefit from allocating times of best seeing to observations in the *r* and *i* bands. Maximizing signal-to-noise ratios requires choosing the next filter based upon the current sky background.

Motivation

Course

Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Survey Strategy

Motivation
Course Logistics
What is Astroinformatics?
Astronomical Data
Tools
Survey
Outlook

The LSST Survey

Synthesizing the requirements to accomplish the four primary science objectives of Rubin Observatory,

- Probing dark energy and dark matter
- Taking an inventory of the Solar System
- Exploring the transient optical sky
- Mapping the Milky Way

results in the following **constraints** for the LSST Survey Strategy:

Supernova cosmology requires frequent, deep photometry in all bands, with z and y observations even during dark time.

Survey Strategy

The LSST Survey

Synthesizing the requirements to accomplish the four primary science objectives of Rubin Observatory,

- Probing dark energy and dark matter
- Taking an inventory of the Solar System
- Exploring the transient optical sky
- Mapping the Milky Way

results in the following **constraints** for the LSST Survey Strategy:

Detecting the motion of solar system objects and transients, characterizing variability on various timescales, and acquiring the best proper motions and parallaxes place further demands upon the distribution of revisit intervals and observation geometries to each point on the sky.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Survey Strategy

Motivation
Course Logistics
What is Astroinformatics?
Astronomical Data
Tools
Survey
Outlook

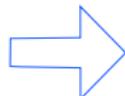
The LSST Survey

Synthesizing the requirements to accomplish the four primary science objectives of Rubin Observatory,

- Probing dark energy and dark matter
- Taking an inventory of the Solar System
- Exploring the transient optical sky
- Mapping the Milky Way

results in the following **constraints** for the LSST Survey Strategy:

Detecting the motion of solar system objects and transients, characterizing variability on various timescales, and acquiring the best proper motions and parallaxes place further demands upon the distribution of revisit intervals and observation geometries to each point on the sky.



complex requirements where a strategy is chosen and then observations are scheduled by **optimization and machine learning**

Data Access

In order to do science, we need (in most cases) access to astronomical data.

For this, Astroinformatics entails



development of services
and tools to enable users
to access the data

using those services to
access the data

Motivation

Course
Logistics

What is
Astroinforma-
tics?

Astronomical
Data

Tools

Survey

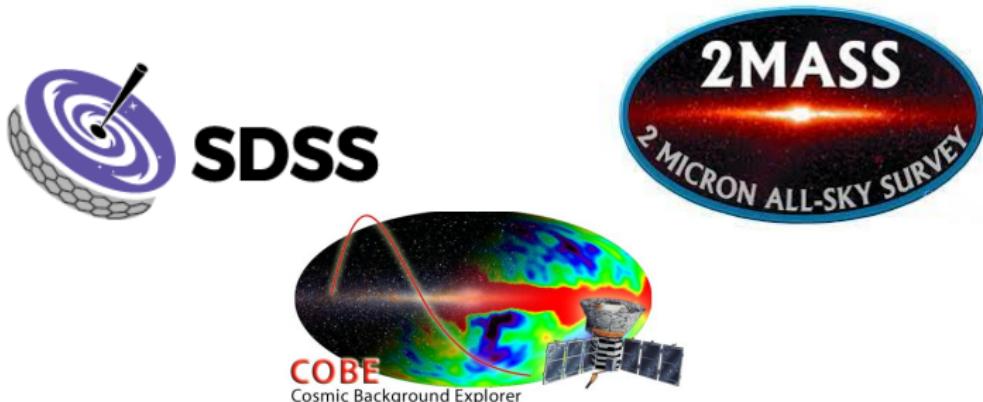
Outlook

Data Access

Motivation
Course Logistics
What is Astroinformatics?
Astronomical Data
Tools
Survey
Outlook

Some astronomical survey data are shared via the observatory that acquired them, for example, the all-sky data acquired with Planck, Wilkinson Microwave Anisotropy Probe, and COBE.

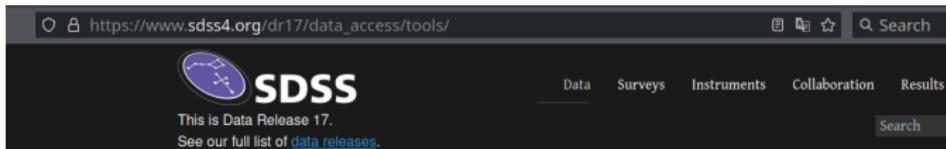
Other surveys serve their own data. Examples include the SDSS, 2MASS, UKIDSS.



Surveys as Astronomical Data Sources

Example: SDSS

<https://www.sdss4.org>



Available Tools for Data Access

Table of Contents

Available Tools for Data Access

Catalog Archive Server (CAS)

Science Archive Server (SAS)

Direct file downloads

This page describes several tools for accessing the SDSS data: the Catalog Archive Server (CAS), the Science Archive Server (SAS), and direct data file access via rsync, wget, or http.

Catalog Archive Server (CAS)

Catalog data are available through [SkyServer](#) (also known as the Catalog Archive Server, or CAS). SkyServer also offers interactive tools to browse through SDSS images with links to associated spectra and catalog data about the objects on the images.

[DR17 SkyServer](#) Interactive query tools for CAS

[SkyServer Imaging Query Form](#) Simple image query form interface

[SkyServer Spectroscopic Query Form](#) Simple spectroscopic query form interface

[SkyServer Cross-ID](#) Match list of RA/Dec or IDs to imaging or spectroscopic catalog

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Surveys as Astronomical Data Sources

Example: SDSS

<https://skyserver.sdss.org>

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

The screenshot shows a web browser window for the SDSS SQL Search tool at <https://skyserver.sdss.org/dr18/SearchTools/sql>. The page has a sidebar on the left with links like Home, Visual Tools, Search Tools, etc., and a main content area with a code editor and a sidebar of sample queries.

SQL Search

```
1 -- This query does a table JOIN between the Imaging (PhotoObj) and spectra
2 -- (SpecObj) tables and includes the necessary columns in the SELECT to upload
3 -- them to the SAS(Science Archive Server) for FITS file retrieval.
4 SELECT TOP 10
5 p.objid,p.ra,p.dec,p.u,p.g,p.r,p.i,p.z,
6 p.run, p.rerun, p.cancel, p.field,
7 t.z,t.zerr,s.z,s.zerr, s.z as redshift,
8 s.plate, s.mjd, s.filterid
9 FROM PhotoObj AS p
10 JOIN SpecObj AS s ON s.bestobjid = p.objid
11 WHERE
12 p.u BETWEEN 0 AND 19.6
13 AND g BETWEEN 0 AND 20
14
```

Output Format

Sample Queries

- Basic SQL
- SQL Jujitsu
- Black Hole Mapper
- Milky Way Mapper
- Miscellaneous
- Variability Queries
- General Astronomy
- Galaxies
- Stars
- Quasars
- BOSS
- APOGEE

Surveys as Astronomical Data Sources

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Example: Gaia

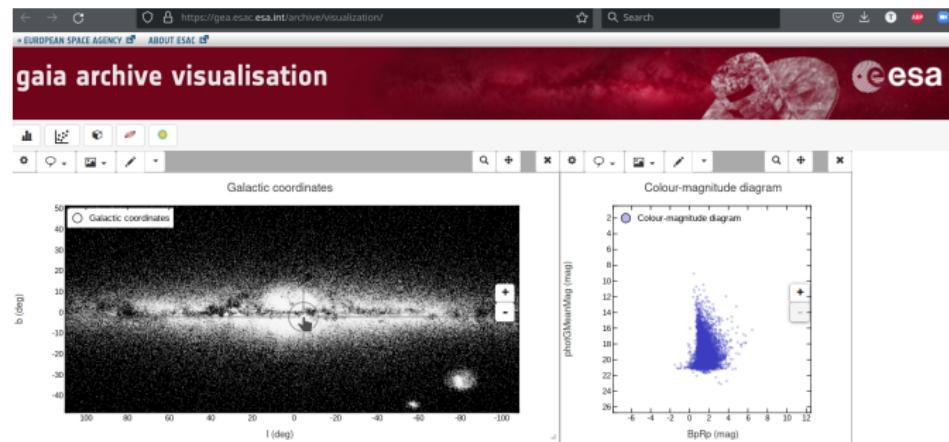
<https://gea.esac.esa.int/archive/>

The screenshot shows the homepage of the Gaia ESA Archive. At the top, there is a navigation bar with links for HOME, SEARCH, SINGLE OBJECT, VISUALISATION, and HELP. Below the navigation bar is a banner featuring the Gaia logo and the text "Welcome to the Gaia ESA Archive". The main content area is titled "Top Features" and includes several sections: "Gaia Mission" (with a link to "Mission Alerts, Information, and Resources for the Gaia community"), "Gaia DR3" (with a link to "Direct access to Gaia DR3 papers, license issues, tools, including data, etc."), "Gaia FPR" (with a link to "Direct access to all Information of the Finalised Product Release"), "Download" (with a link to "Direct bulk download of Data files in BCDV format"), "Software Tools" (with a link to "Software tools for reprocessing of spectra, calibration of data, etc."), "Auxiliary Data" (with a link to "Small data sets related to Gaia DR3, such as parallax, proper motion, absolute magnitude, astrometry, etc."), "Citation" (with a link to "How to cite and acknowledge the use of Gaia data and where to find DOIs"), and "Partners" (with a link to "Partner data centres also serving Gaia data").

Surveys as Astronomical Data Sources

Example: Gaia

<https://gea.esac.esa.int/archive/visualization>



Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Services

With a lot of different surveys existing, a **common scripted interface** to tie all these services together is a good way to make the data more accessible.

Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

Tools

Survey

Outlook

It also provides **reproducibility** of analysis done for publications. A centrally maintained library also safeguards against lost links to data archives, moving some of the responsibility for maintaining long-term reproducibility from each individual researcher to the broader community.

Services

The SIMBAD astronomical database provides basic data, cross-identifications, bibliography and measurements for astronomical objects outside the solar system. SIMBAD can be queried by object name, coordinates and various criteria. Lists of objects and scripts can be submitted. Links to some other on-line services are also provided.

The **SIMBAD** database currently contains information for about 13,000,000 astronomical objects (stars, galaxies, planetary nebulae, clusters, novae and supernovae, etc.).

The only astronomical objects specifically excluded from SIMBAD are Solar System bodies (planets, satellites, asteroids, comets)

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Services

<https://simbad.unistra.fr/simbad/>

The screenshot shows the SIMBAD Script execution interface at <https://simbad.cds.unistra.fr/simbad/sim-fscript>. The top navigation bar includes links for Portal, Simbad, Vizier, Aladin, X-Match, Other, and Help. Below the navigation is a toolbar with tabs for other query modes (Identifier, Coordinate, Criteria, Reference, Basic, Script submission, TAP, Output options), and the Help tab is currently selected.

The main area has a blue header "Enter the name of an ASCII file containing a script:" followed by a file input field with "Browse..." and "submit file" buttons, and checkboxes for "file output" and "compressed".

The bottom half contains a large text area for "Type your script in:" with "submit script" and "clear" buttons, and checkboxes for "file output" and "compressed". To the right is a detailed help section titled "Short reminder of the script commands. A more complete description can be found behind the help anchor". It lists various commands with their descriptions:

control:	
# comment	a comment line must have a '#' as its first char
output outfile=action...	output definition. 'outfile' can be 'console', 'error' or 'script'
result type	'action' can be 'off' or 'merge([line header])' (merge into the data output)
utable [name] {fieldlist}	define a votable output with a fieldlist
votable open [name]	use the defined votable output for the next queries
votable close	close the current votable
format {obj ref} [name] "..."	define the output formats for objects or references
format name	set the named output format as current
format {obj ref} default	reset the output formats for obj or ref to the default
format display [name]	display all format or the one defined by its name
set limit #rows	limits the number of returned rows (0 removes any limitation)
set radius val[d m s]	radius for coordinates and around queries (Ex: 1.5d, 10m, 3s)
set frame name	input frame for coord. queries (ICRS,FK5,FK4,GAL,SGAL,ECL)
set equinox val	Equinox value for coord. queries (Ex: 2006.5)
set epoch val	EPOCH value for coord. queries (Ex: B1950, J2000)
echo [-n] text	default for coord. queries: radius=10m, frame=ICRS, equil=2000.0, epc without a value, a set xxx command resets the default
echodata [-n] text	simple text to display on the console. option -n: no newline at the end
quit	'echodata' displays on the data output. option -n: no newline at the end quit. Allows to quit before the end of a script

Motivation
Course
Logistics

What is
Astroinformatics?

Astronomical
Data

Tools

Survey

Outlook

Services

The **VizieR** service provides mostly catalogs associated with papers or authors.

<https://vizier.cds.unistra.fr/viz-bin/VizieR>

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Screenshot of the VizieR service interface:

The page title is "VizieR". The URL is "https://vizier.cds.unistra.fr/viz-bin/VizieR".

Header menu: Portal, Simbad, VizieR, Aladin, X-Match, Other, Help.

Search Criteria section:

- Preferences: max: 50, HTML Table, All columns checked, Compute checked, Mirrors: CDS, France.
- Find catalogs among 24534 available: Catalog, author's name, words(s) from title, description, etc. e.g.: AGN, Veron, I/239, or bibcodes..
- Search for catalog by column descriptions (UCD):
- Search for catalogs containing additional data

Find by Position section:

- Target Name (resolved by Sesame) or Position: Clear, J2000, Target dimension: 2 arcmin, Go!
- NB: The epoch used for the query is the original epoch of the table(s): Radius (radio button selected), Box size.
- More about VizieR link.

Tools related to VizieR section:

- Catalogue collection: Search VizieR catalogues available via various services (FTP, VizieR, TAP, ...)
- CDS Portal: Access CDS data through VizieR, Simbad and Aladin using the CDS portal
- Spectra, Images in VizieR: Search Spectra, images in VizieR
- Photometry viewer: Plot photometry (sed) including all VizieR
- TAP_VizieR: query VizieR using ADQL (a SQL extension dedicated for astronomy)
- CDS cross-match service: fast cross-identification between any 2 tables, including VizieR catalogues, SIMBAD

Community brokers

The newest generation of large survey telescopes such as the Vera C. Rubin Observatory (which will conduct the 10-year Legacy Survey of Space and Time (LSST)) are designed to produce large streams of transient alerts from difference imaging. They will be processed, annotated and classified by a new generation of astronomical software for time-domain surveys, called **alert brokers**.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Community brokers

The newest generation of large survey telescopes such as the Vera C. Rubin Observatory (which will conduct the 10-year Legacy Survey of Space and Time (LSST)) are designed to produce large streams of transient alerts from difference imaging. They will be processed, annotated and classified by a new generation of astronomical software for time-domain surveys, called **alert brokers**.

Brokers are accessible over the internet, providing real-time information crucial for:

- follow-up
- observations depending on accurate & up-to-date light curve information

Community brokers

The newest generation of large survey telescopes such as the Vera C. Rubin Observatory (which will conduct the 10-year Legacy Survey of Space and Time (LSST)) are designed to produce large streams of transient alerts from difference imaging. They will be processed, annotated and classified by a new generation of astronomical software for time-domain surveys, called **alert brokers**.

Brokers are accessible over the internet, providing real-time information crucial for:

- follow-up
- observations depending on accurate & up-to-date light curve information

Brokers are currently developed tested with the Zwicky Transient Facility (ZTF). Usually, brokers can be accessed from a website as well as a Python interface.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

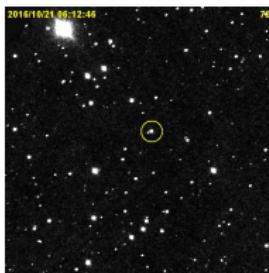
Survey

Outlook

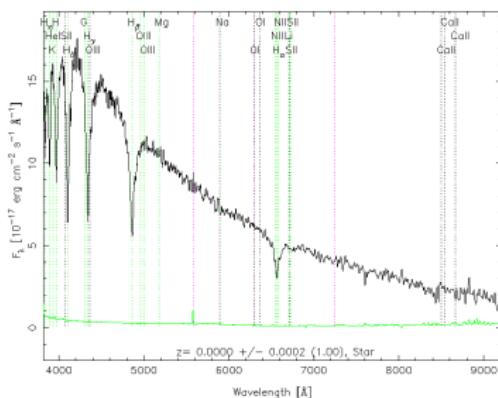
Astronomical Data

astronomical data typically comes in one of three forms:

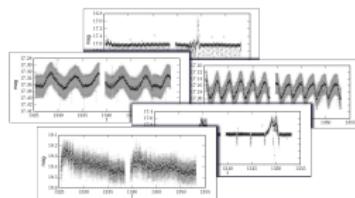
images



spectra



light curves



Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

Tools

Survey

Outlook

Image Data

An image can be seen as 2D array of pixels, where each pixel value denotes a brightness value. In an ordinary color image, each pixel will have three brightness values for the contribution from red, green, and blue (RGB).

As astronomers use many specialist filters, astronomical images can have even more color values per pixel. They are typically displayed as one grey-scale image per channel.

A commonly used file format is FITS (Flexible Image Transport System). Such files can be opened with specialized software or can be read in with your own code.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

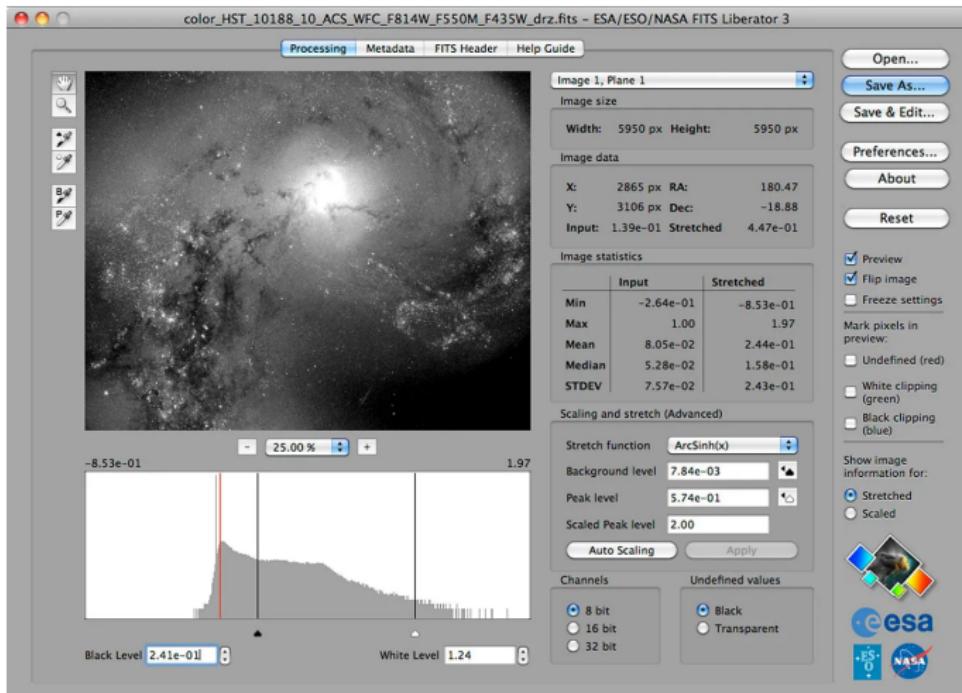
Tools

Survey

Outlook

Image Data

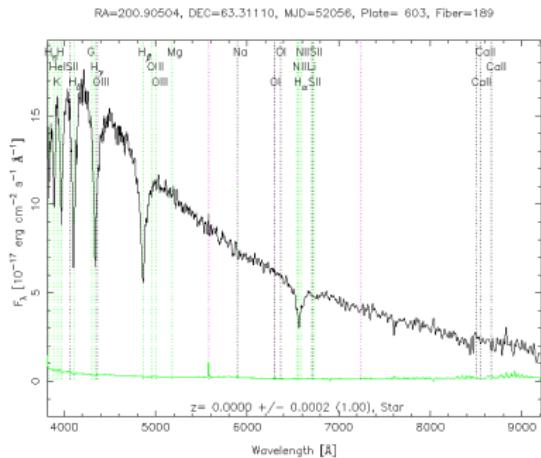
Motivation
Course
Logistics
What is
Astroinformatics?
Astronomical
Data
Tools
Survey
Outlook



Spectra

Motivation
Course
Logistics
What is
Astroinformatics?

Astronomical
Data
Tools
Survey
Outlook

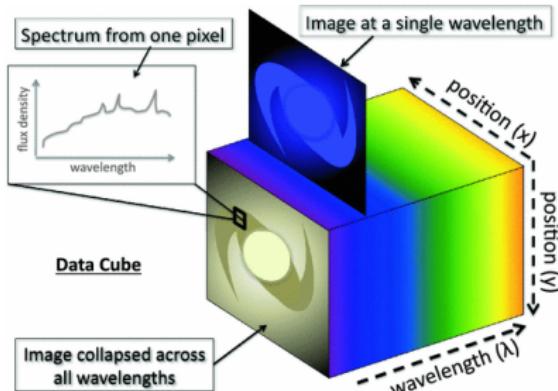


A SDSS spectrum. The plot includes additional information, such as the object, time of observation, the plate and fiber the spectrum was taken. In green, the positions of specific spectral lines are highlighted.

Data Cubes

A data cube can be thought of as enhanced version of an image where each pixel contains not a single brightness value (or three for RGB), but a whole spectrum received from the region of the sky within that pixel.

An example for a data cube is data from **Integral Field Spectroscopy (IFS)**. IFS combines spectrographic and imaging capabilities: take a single exposure of a 2D region and obtain spatially resolved spectra.



Integral Field Spectroscopy, schematic. (Credit: Ch. M. Harrison (2016))

Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

Tools

Survey

Outlook

Catalog Data

Motivation

Course
Logistics

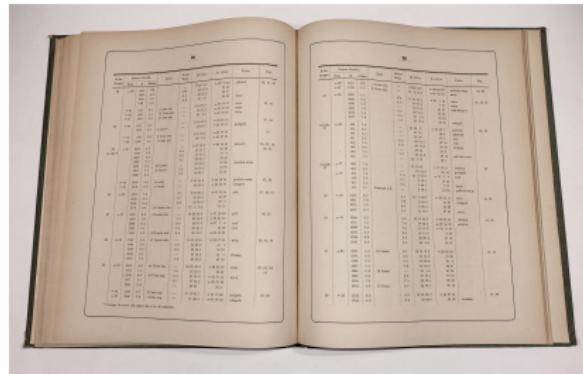
What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook



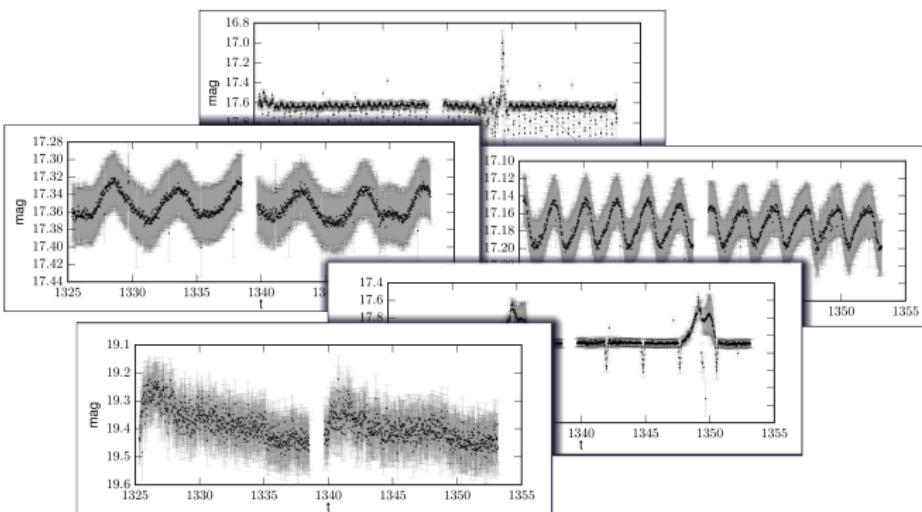
Catalogs group astronomical data together in the form of tables containing data of multiple objects, or of the same object at multiple times, or both. The first astronomical catalogs were books containing coordinates, names and magnitudes of objects.

The catalog from the "Bonner Durchmusterung" (survey by Bonn observatory) from 1862

Catalog Data

Catalogs group astronomical data together in the form of tables containing data of multiple objects, or of the same object at multiple times, or both.

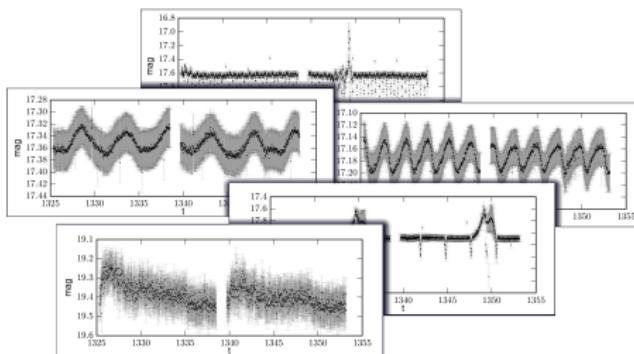
Photometry of an object taken at multiple times is called a **light curve**.



Catalog Data

Catalogs group astronomical data together in the form of tables containing data of multiple objects, or of the same object at multiple times, or both.

Photometry of an object taken at multiple times is called a **light curve**.



Light curves are given in tabular format, typically containing the columns time, magnitude, magnitude error and in the case of multiple photometric bands also indicating the band.

Light curves can be saved as individual tables each containing data from a single object, or as part of a larger table containing the light curves for multiple objects, where as an additional column an object ID is given.

Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

Tools

Survey

Outlook

Other Astronomical Data

The list so far is not complete.

For instance, in **interferometric imaging**, for trying to reconstruct an image by combining coherently the measurements of different telescopes (aperture synthesis), the raw data will be time-stamped data from the single telescopes, and the between those telescopes.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Other Astronomical Data

The list so far is not complete.

For instance, in **interferometric imaging**, for trying to reconstruct an image by combining coherently the measurements of different telescopes (aperture synthesis), the raw data will be time-stamped data from the single telescopes, and the between those telescopes.

What we have seen in general is that **providing and accessing**, as well as **processing** nowadays astronomical data requires some amount of informatics background.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Tools

Typically, we need some software tools to work with astronomical data. Here we can distinguish three types:

Application software is software written for a specific set of tasks. We might, for example, have software to load data tables (acting similar to spreadsheet programs), to load astronomical images, and so on.

A **programming language** is a tool for writing your own custom applications.

A **data access language** allows for selecting data from (usually remote) databases.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

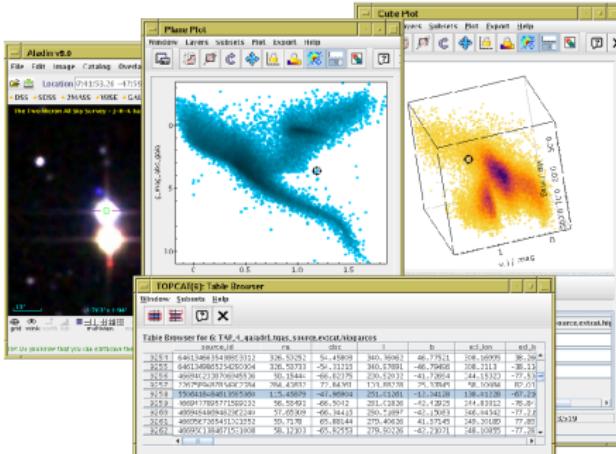
Outlook

Application Software

commonly used software include the following:

TOPCAT is a standard tool for viewing, manipulating and plotting tables, including cross-matching catalogs. Its main advantage is that it allows for fast access to large datasets (millions of rows/hundreds of columns) - tables which cannot efficiently explored with non-astronomical standard tools.

TOPCAT is available at <https://www.star.bristol.ac.uk/mbt/topcat/>



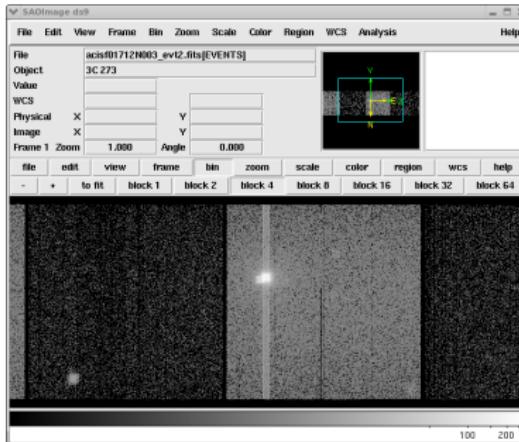
Application Software

commonly used software include the following:

SAOImage DS9 is an astronomical imaging and data visualization application. DS9 supports FITS images and tables, region manipulation, and many scale algorithms and colormaps. The highly configurable and extensible software can communicate with external analysis tools.

DS9 is available at

<https://sites.google.com/cfa.harvard.edu/saoimageds9>



Data Access and Programming Languages

Commonly used **programming languages** in astronomy are:

- Python
- C/C++
- IDL
- FORTRAN

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Data Access and Programming Languages

Commonly used **programming languages** in astronomy are:

- Python
- C/C++
- IDL
- FORTRAN

Python has become the language of choice for astronomers and astrophysicists working with data analysis and visualization.

Python can be combined with C++ for efficiency. Also code in FORTRAN (usually older code) can be integrated.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Why use Python?

The **Python ecosystem** provides a single environment that is sufficient for the vast majority of astronomical analysis. It does so on several levels:



Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Why use Python?

The **Python ecosystem** provides a single environment that is sufficient for the vast majority of astronomical analysis. It does so on several levels:



- open source, which means it is free (as opposed to proprietary languages (like IDL) which require you to buy a license)
- powerful language, many scientific libraries available.
- strong set of 3rd party analysis tools that are professionally and actively developed.
- robust methods for binding with C, C++ and FORTRAN libraries (speed, legacy)
- standard library support for web, GUI, databases, process management, etc.
- very active user and developer community

Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

Tools

Survey

Outlook

Data Access and Programming Languages

In contrast to programming languages which enables us to write software to analyze data, data access languages (sometimes also known as *database languages*) enable us to select data from large data sources, usually databases.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Data Access and Programming Languages

Example: Gaia

In addition to the bulk download of data, one can write Astronomical Data Query Language (ADQL) queries which are similar to SQL queries.

example use case: Cone search sorted by angular separation

retrieve a sample of filtered sources (brighter than $G=20.5$ mag and with a parallax measurement) in a circular region centred on the LMC [$(\text{R.A.}, \text{Dec}) = (81.28, -69.78) \text{ deg}$] with a search radius of 5 arcmin. The output should be ordered by the angular separation, from small to large.

Target table: `gaiadr3.gaia_source`

Query:

```
SELECT *, DISTANCE(81.28, -69.78, ra, dec) AS ang_sep
FROM gaiadr3.gaia_source
WHERE DISTANCE(81.28, -69.78, ra, dec) < 5./60.
AND phot_g_mean_mag < 20.5
AND parallax IS NOT NULL
ORDER BY ang_sep ASC
```

Motivation

Course
Logistics

What is
Astroinformatics?

Astronomical
Data

Tools

Survey

Outlook

Survey

A rapid survey to better understand your preknowledge:

**Linux/Unix
Experience?**



- Motivation
- Course Logistics
- What is Astroinformatics?
- Astronomical Data
- Tools
- Survey**
- Outlook

Survey

A rapid survey to better understand your preknowledge:

**Linux/Unix
Experience?**

**Programming
Experience?
Languages?**



Motivation
Course
Logistics
What is
Astroinformat-
ics?
Astronomical
Data
Tools
Survey
Outlook

Survey

A rapid survey to better understand your preknowledge:

Motivation

Course

Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Linux/Unix
Experience?



Programming
Experience?
Languages?

Expectations
Projects?
Plan attending other courses?

Survey

A rapid survey to better understand your preknowledge:

Motivation

Course

Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook

Linux/Unix
Experience?



Programming
Experience?
Languages?

Expectations
Projects?
Plan attending other courses?

Questions?

An Outlook: Using Linux

We have seen which topics fall under the relatively broad umbrella of *astroinformatics*.

In the next lecture we will see how to use the **Linux command line** for manipulating files.

Motivation

Course
Logistics

What is
Astroinformat-
ics?

Astronomical
Data

Tools

Survey

Outlook