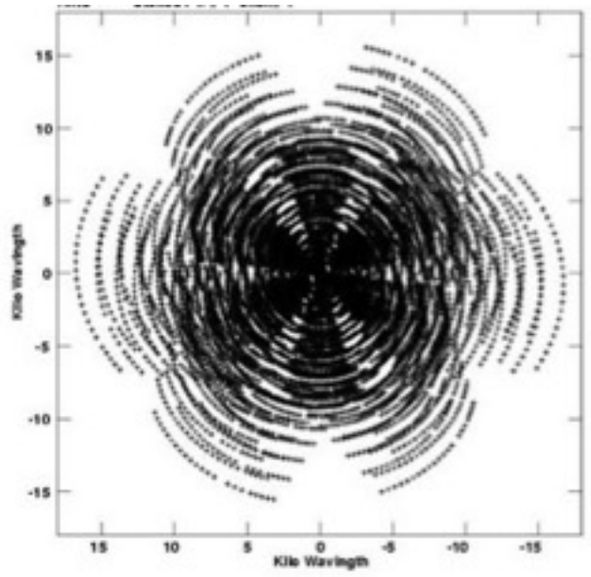
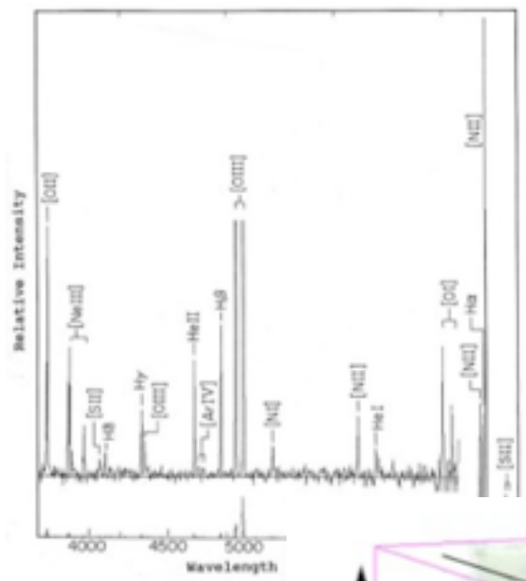
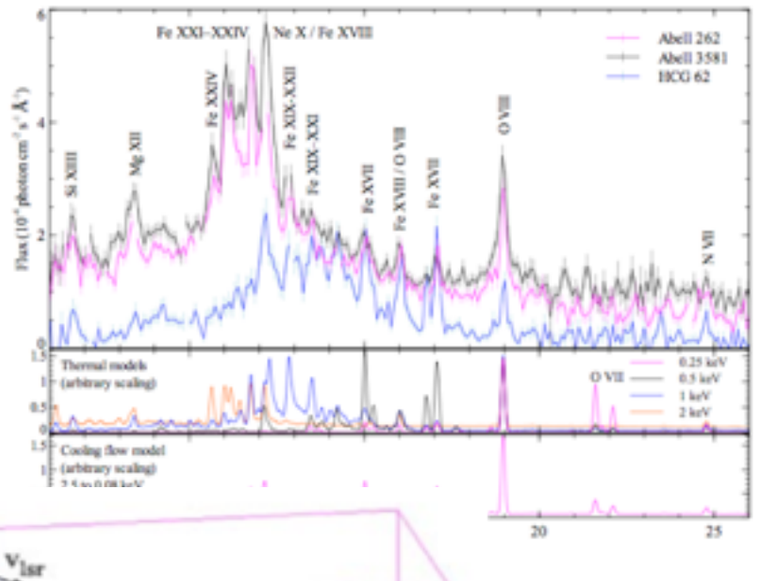
Virtual Observatory:
an Astronomy's answer to Big Data

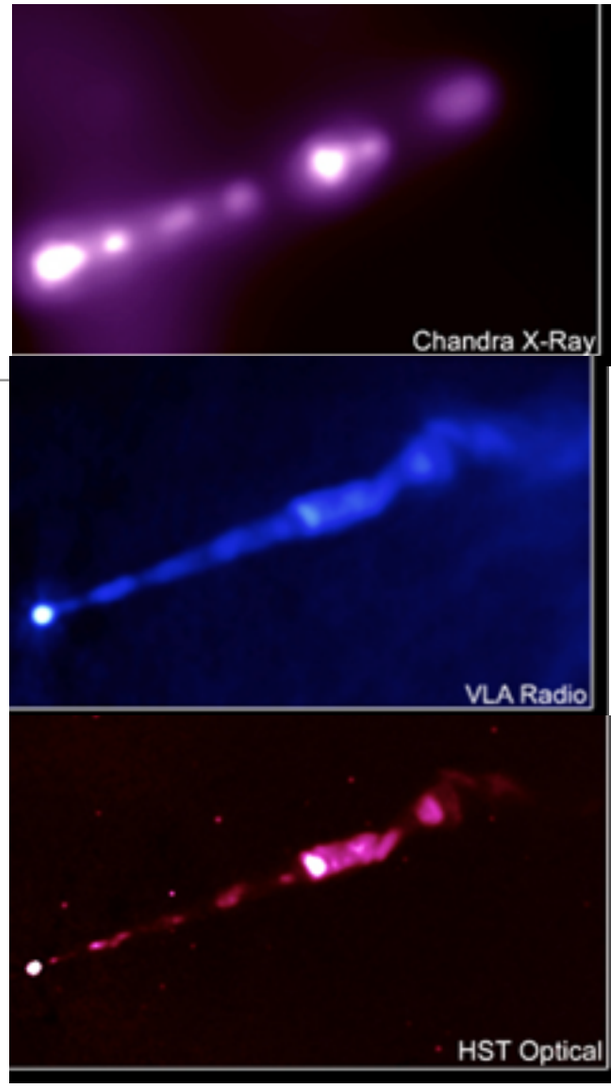
by Paula Coelho, IAG/USP
on behalf of BRAVO

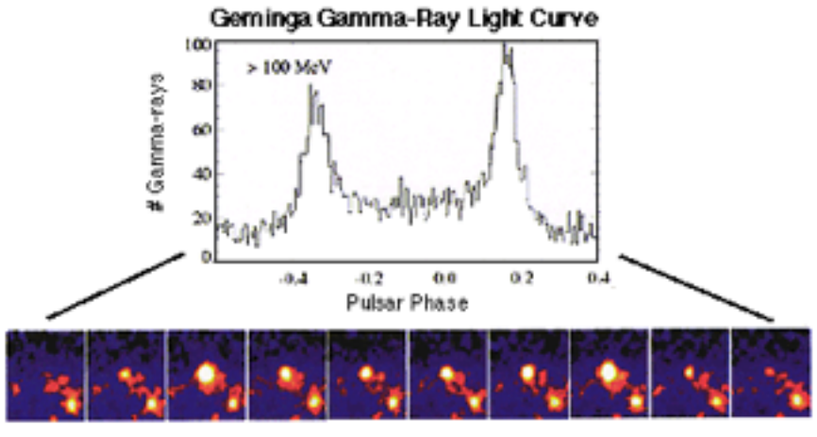
Data in Astronomy

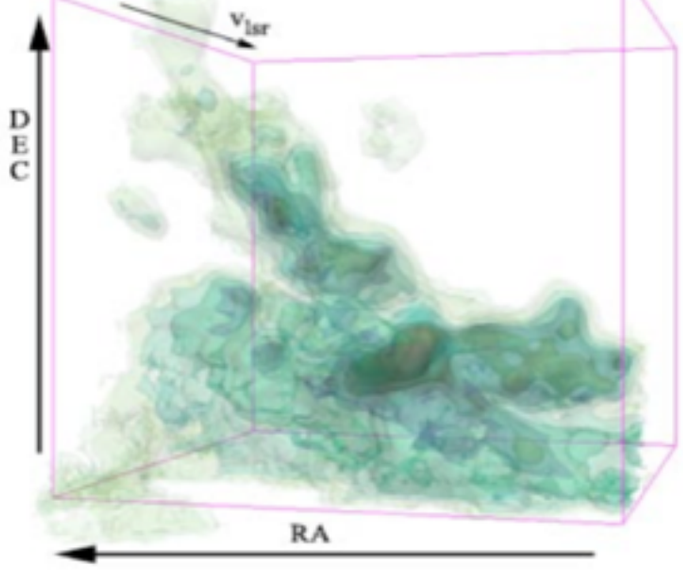












jdate	designation	ra	dec	sup_ra	sup_dec	glon	glat	density	r_k20fe
2451305.6569	12552517+2134339	12 55 25.2	21 34 33.9	193.854874	21.576124	312.417426	84.374259	2.51	5.4
2451700.6751	12554924+2123581	12 55 49.2	21 23 58.2	193.955109	21.399385	313.052422	84.184820	2.66	5.0
2451261.6020	12571719+2120180	12 57 17.2	21 20 18.1	194.321625	21.338383	316.206268	84.058667	2.24	5.9
2451261.6020	12572936+2132520	12 57 29.4	21 32 52.1	194.372269	21.547800	317.124576	84.251945	2.24	5.0
2451261.6020	12572893+2137370	12 57 28.9	21 37 37.1	194.370453	21.627054	317.296215	84.329266	2.35	5.0
2451261.7924	12562741+2131175	12 56 27.4	21 31 17.6	194.114243	21.521484	314.721401	84.277947	2.54	8.1
2451261.6020	12573991+2146420	12 57 39.9	21 46 42.1	194.416367	21.778374	318.097097	84.466693	2.36	5.0
2451261.7972	12564252+2148223	12 56 42.5	21 48 22.4	194.177277	21.806303	315.909776	84.544554	2.51	5.0
null	12564369+2140575	12 56 43.7	21 40 57.6	194.182068	21.682659	315.683000	84.423000	null	214.8
2451261.7924	12561052+2148274	12 56 10.5	21 48 27.5	194.043808	21.807701	314.635368	84.571294	2.78	7.4
2451261.7972	12571196+2146234	12 57 12.0	21 46 23.5	194.299911	21.773294	316.993658	84.486836	2.51	5.5
2451261.6020	12572147+2140450	12 57 21.5	21 40 45.1	194.339417	21.679213	317.136181	84.386833	2.35	5.3
2451700.6751	12554548+2153222	12 55 45.5	21 53 22.2	193.939529	21.889559	313.784434	84.669874	2.66	9.0

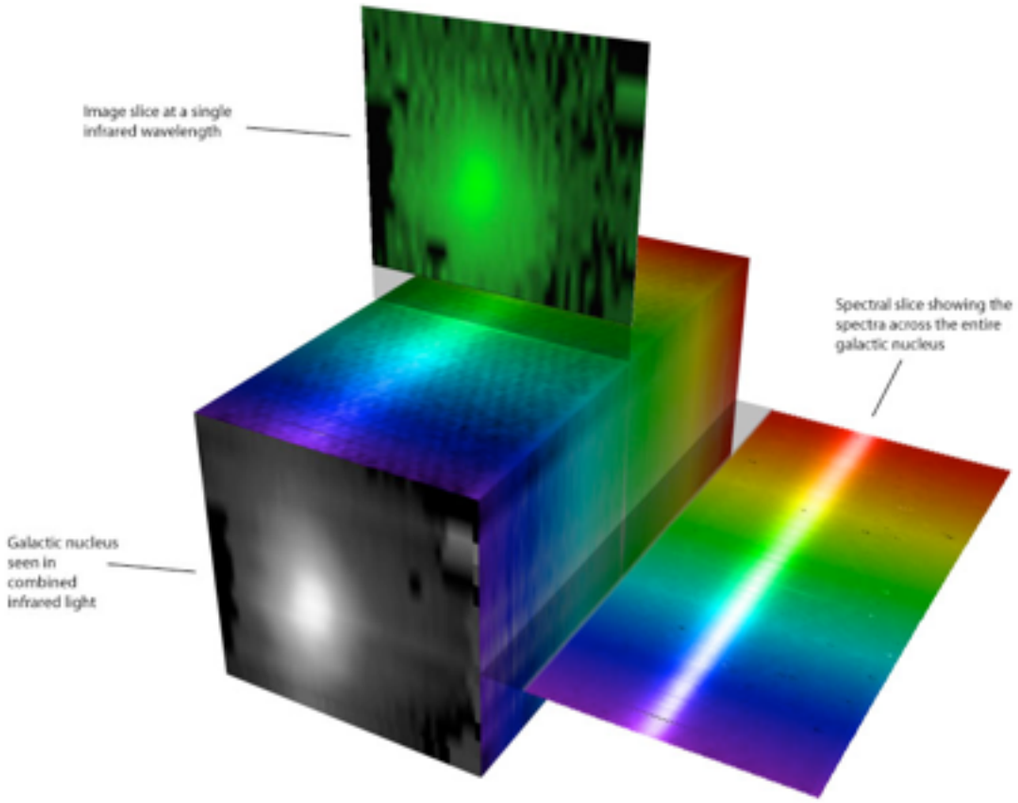
1D, 2D, 3D; intensity/polarization vs. energy, time, position, velocity. tables, DBs, catalogs, x-ray event lists, radio visibility measurements… various data processing levels, from raw to “science-ready”

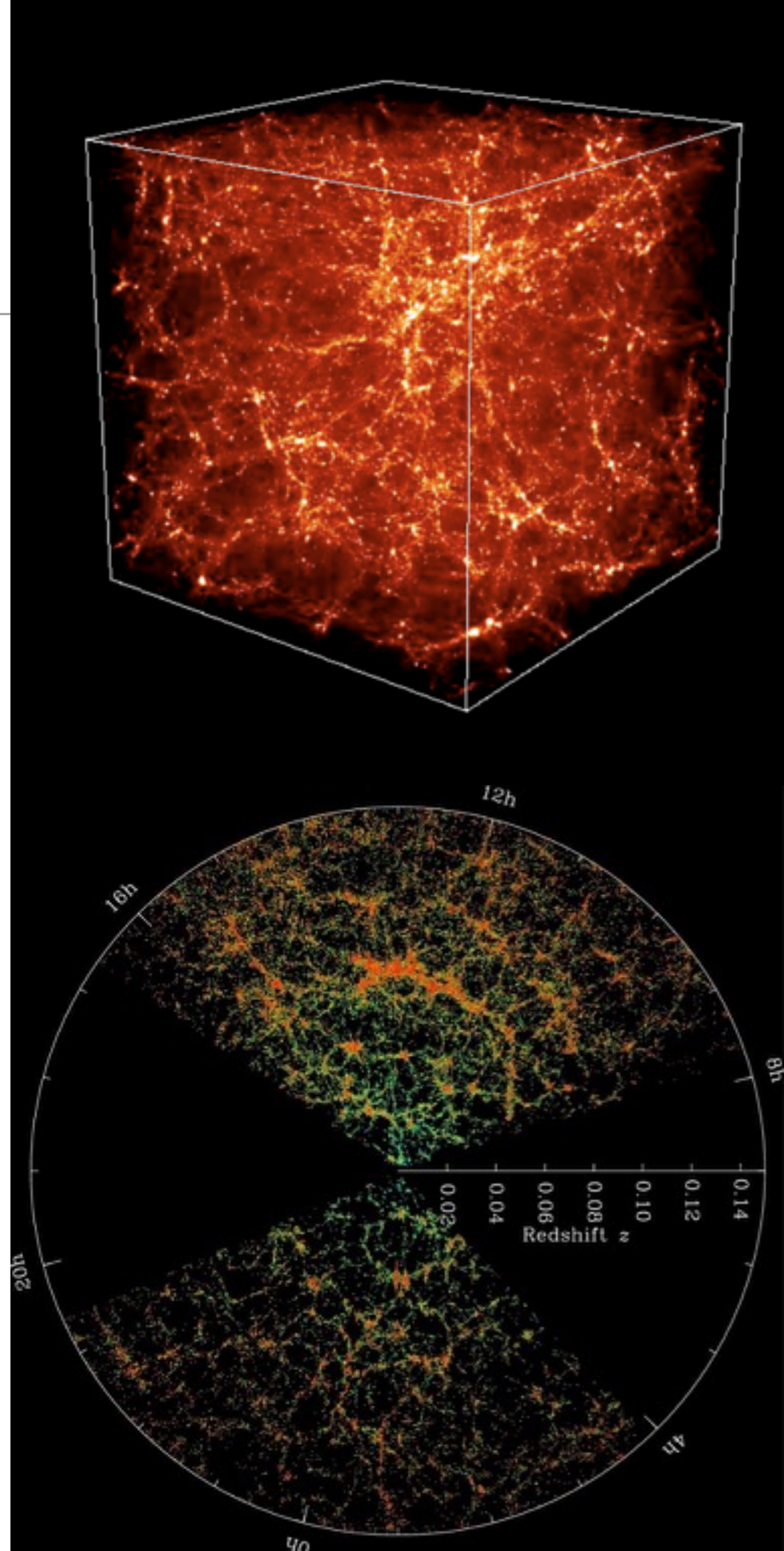
A new era in astronomy

- Past: observations of small, carefully selected samples (often with a priori prejudices) of objects in one or a few wavelength bands

A new era in Astronomy

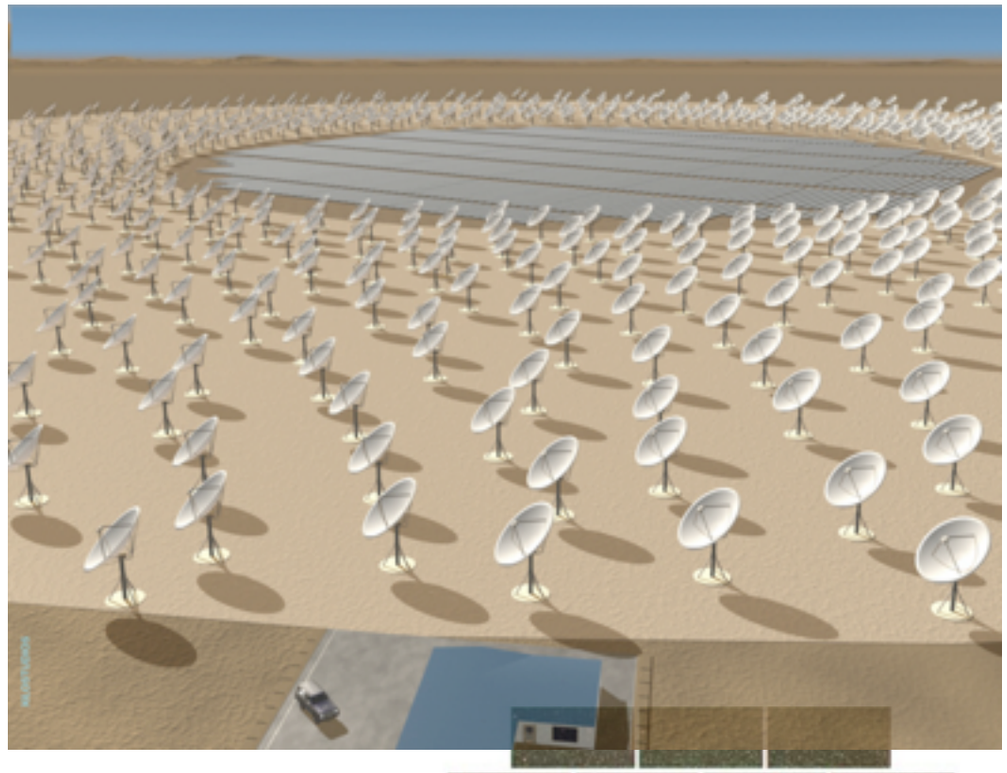
- Now: multi-wavelength data for millions of objects, allowing us to:
  - **discover** significant patterns from the analysis of statistically rich and unbiased databases
  - **understand** complex astrophysical systems via confrontation between data and sophisticated numerical simulation



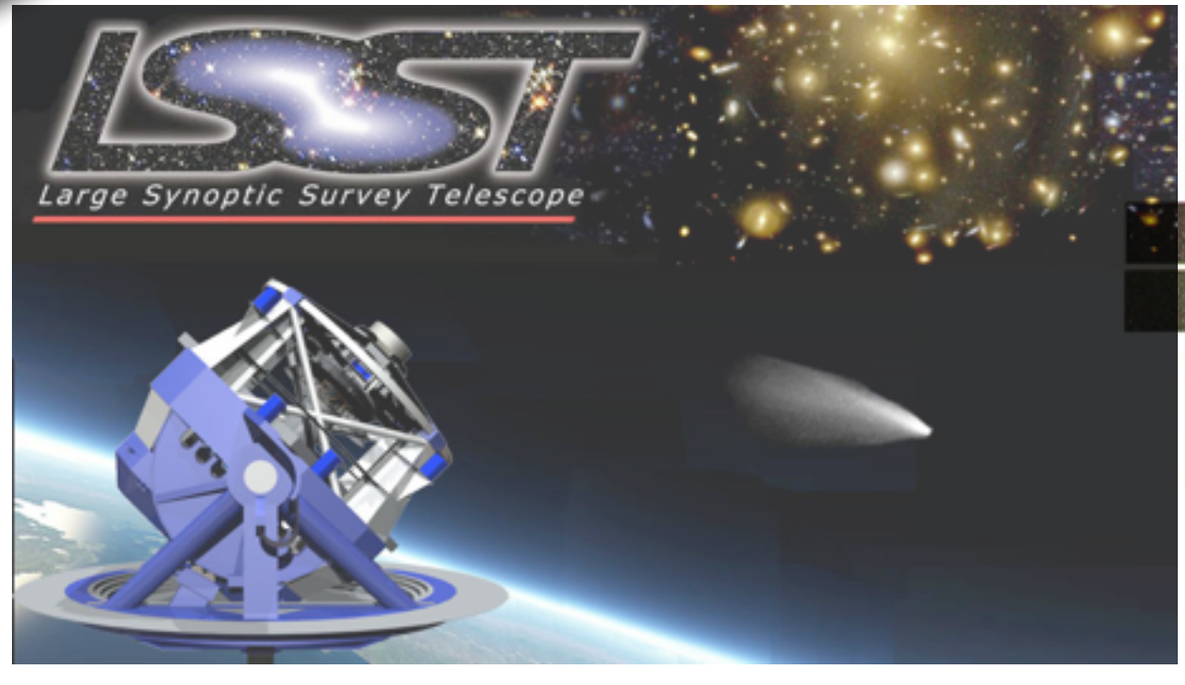


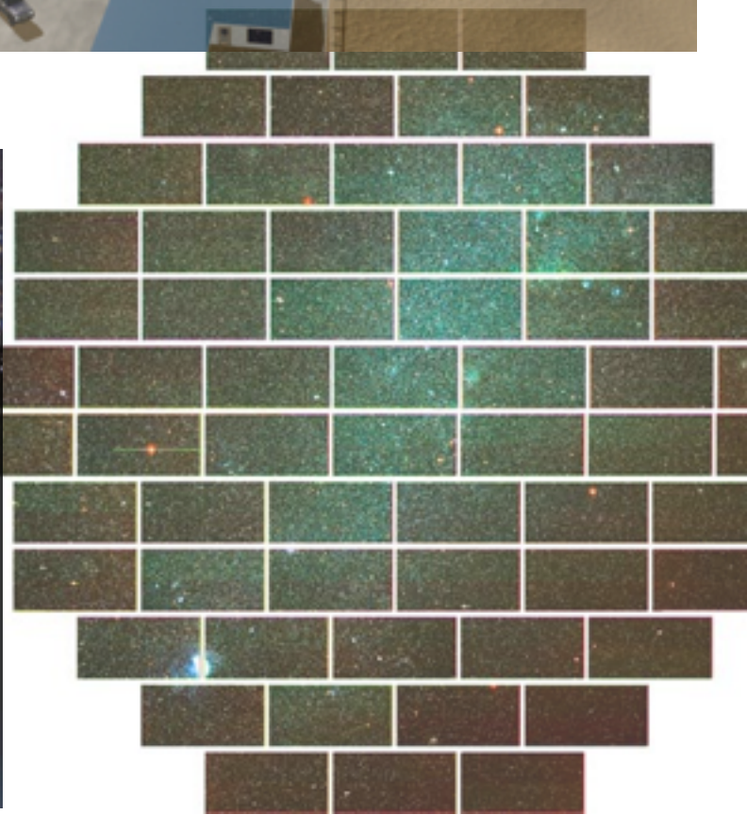
Survey Science: Big Data in Astronomy

SDSS, DES, J-PAS,
LSST, Gaia, SKA…



gaia





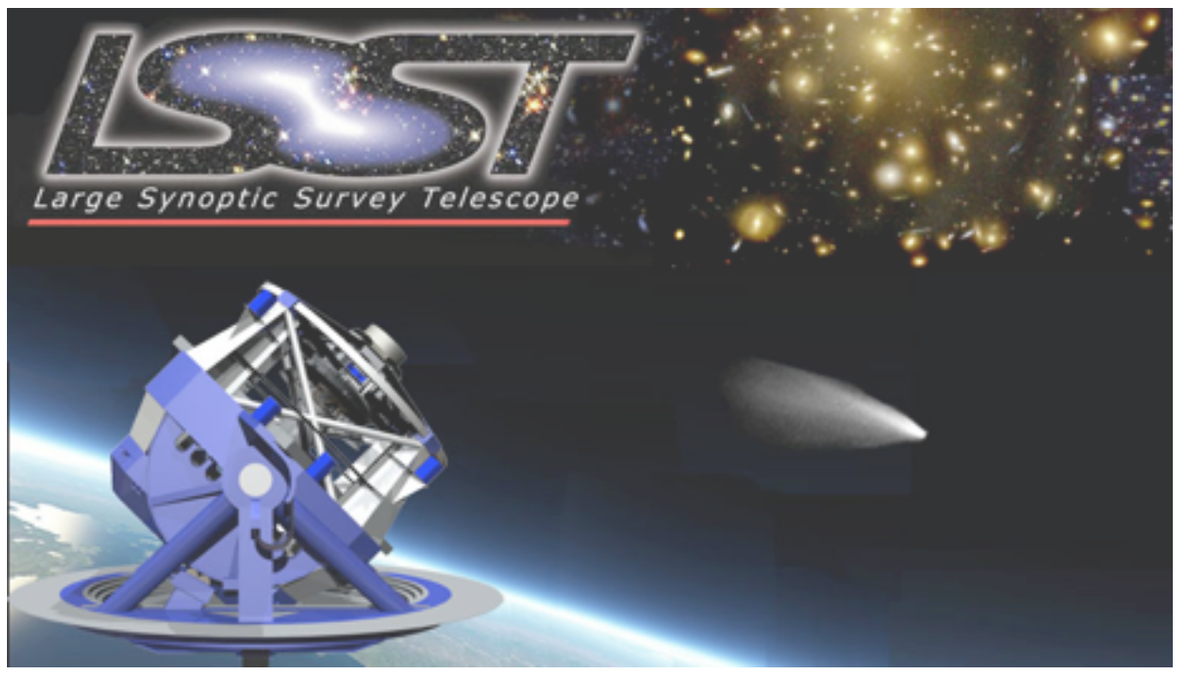
Exploding Data Rates

Survey science: Big Data in Astronomy

The Large Synoptic Survey Telescope (LSST) is a planned wide-field survey that will observe the entire available sky every few nights.

15 TB/night, 7 PB/year, 200 PB total, ~1M transient events/night

(in comparison, SDSS 170 GB/night)



And the scary future… SKA

The Square Kilometre Array is a project to build a radio telescope tens of times more sensitive and hundreds of times faster at mapping the sky than today's best radio astronomy facilities.

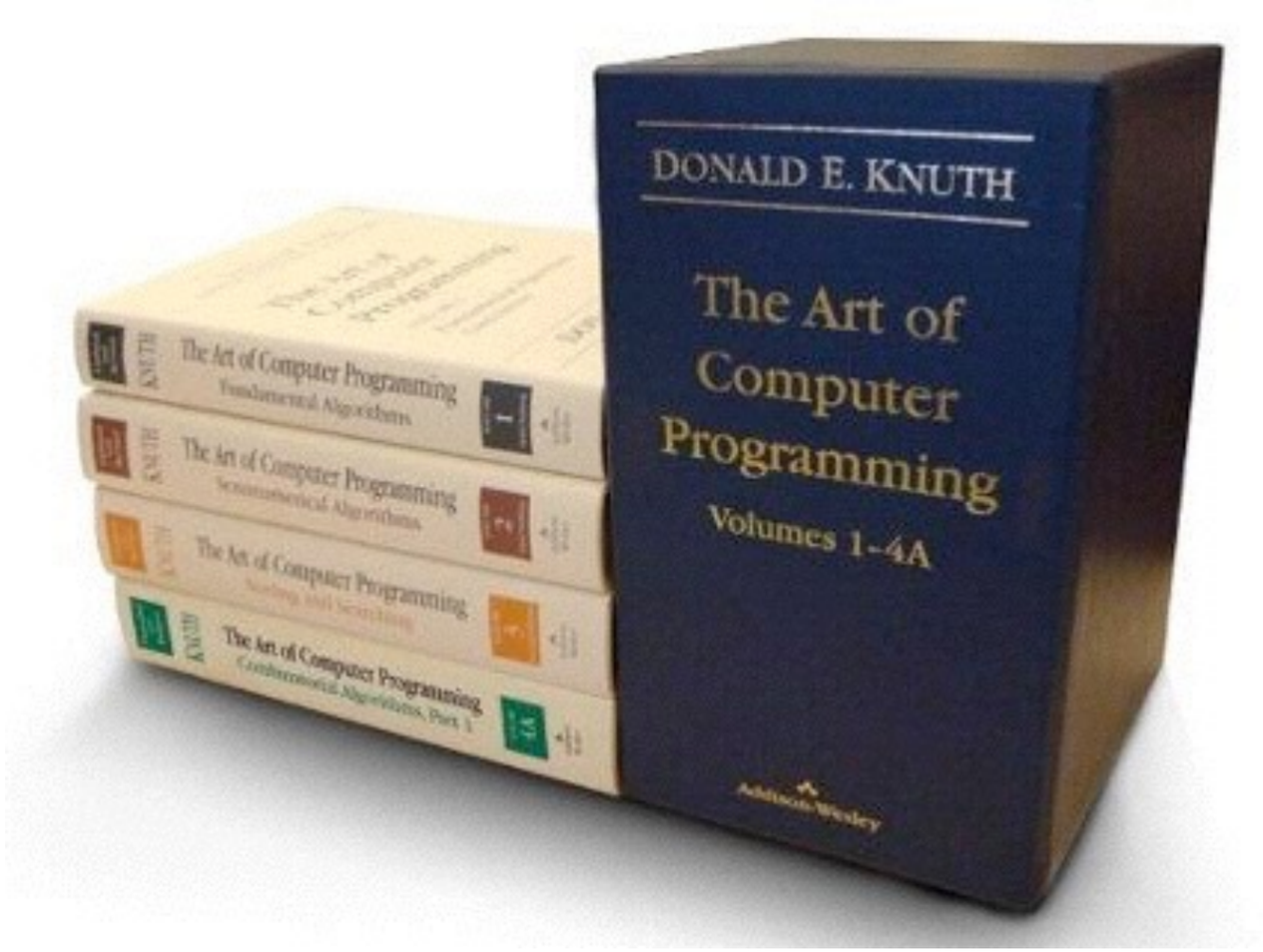
1 exabyte of raw data in a single day; more than the entire daily Internet traffic!!

What do these projects have in common?

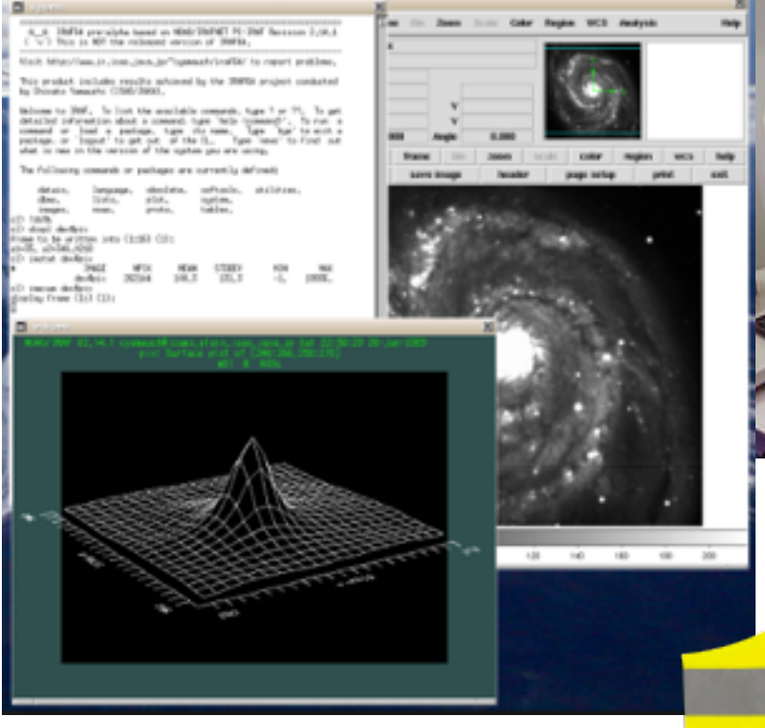
Data, data, data!

- Volume
- Complexity
- Interdependency

But with data...

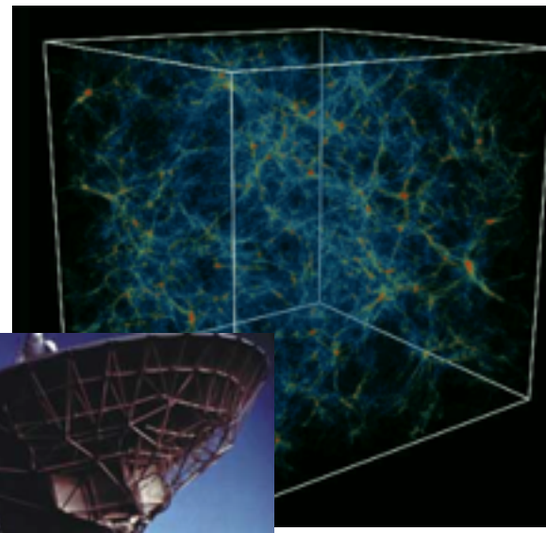


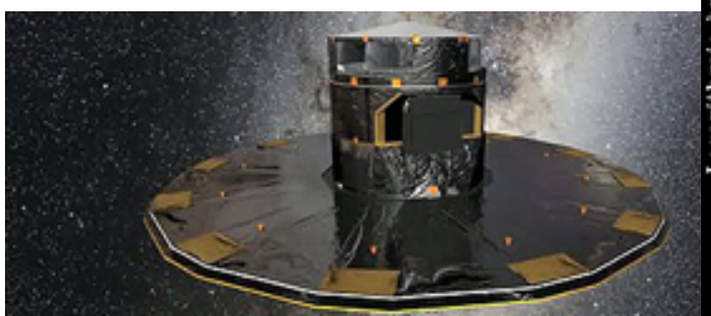
should come the software!

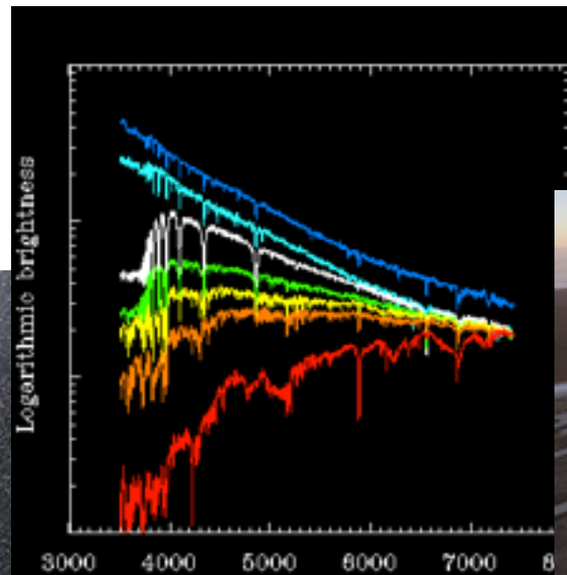


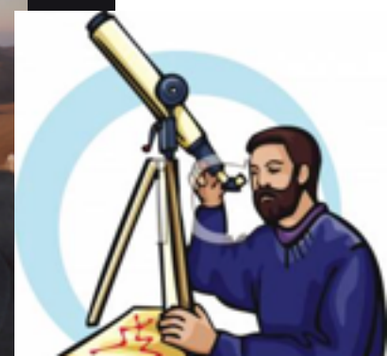
ASTRONOMER AT WORK

???









Data – Software

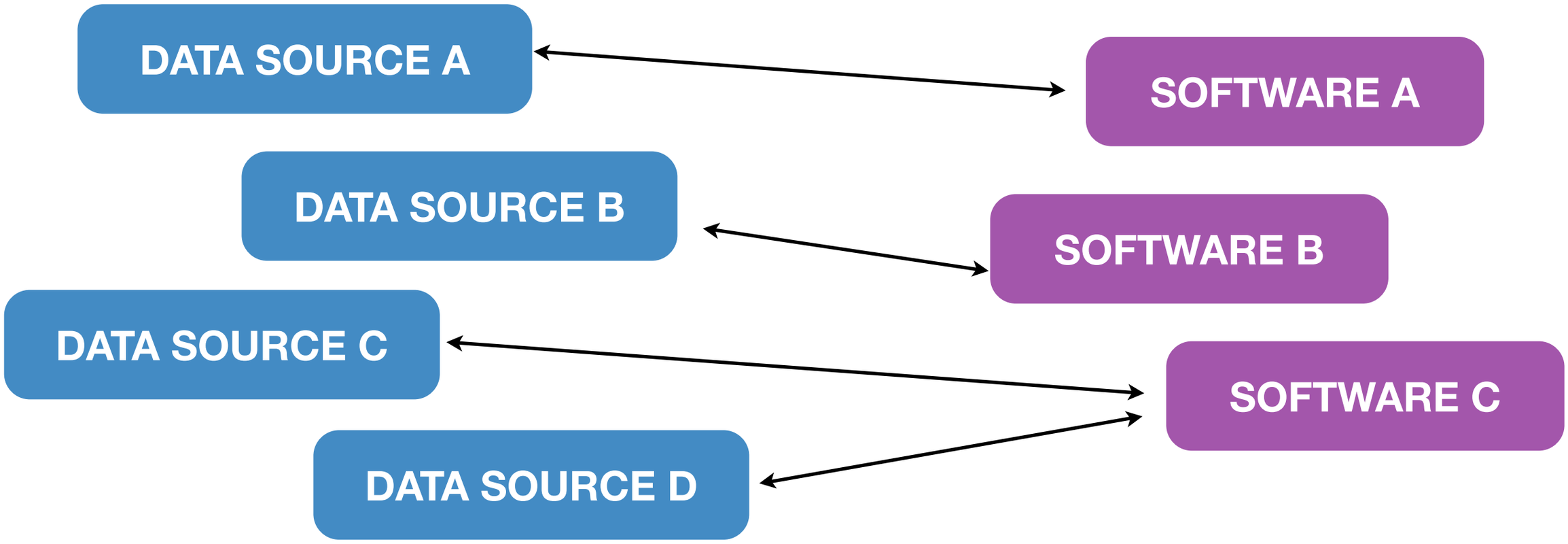
DATA
The language in which we register Nature (even if simulated!)

SOFTWARE
The language in which we interpret Nature (even if simulated!)

Data – Software

DATA
The language in which we register Nature (even if simulated!)

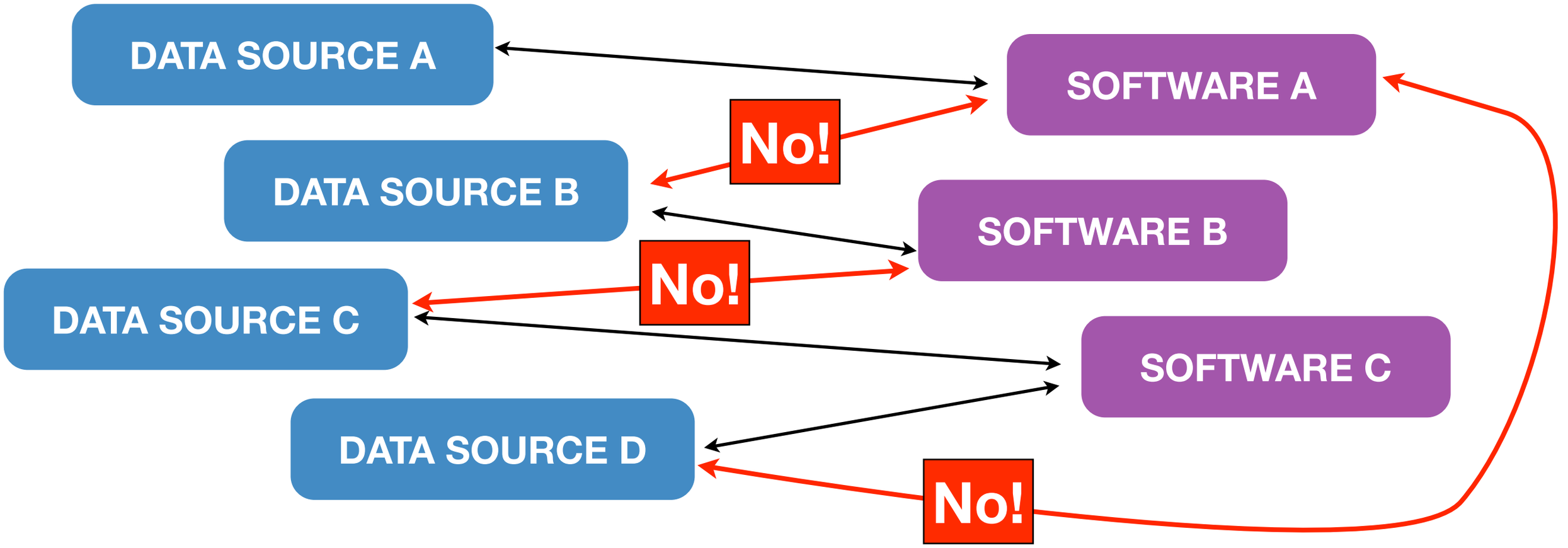
SOFTWARE
The language in which we interpret Nature (even if simulated!)



Data – Software

DATA
The language in which we register Nature (even if simulated!)

SOFTWARE
The language in which we interpret Nature (even if simulated!)

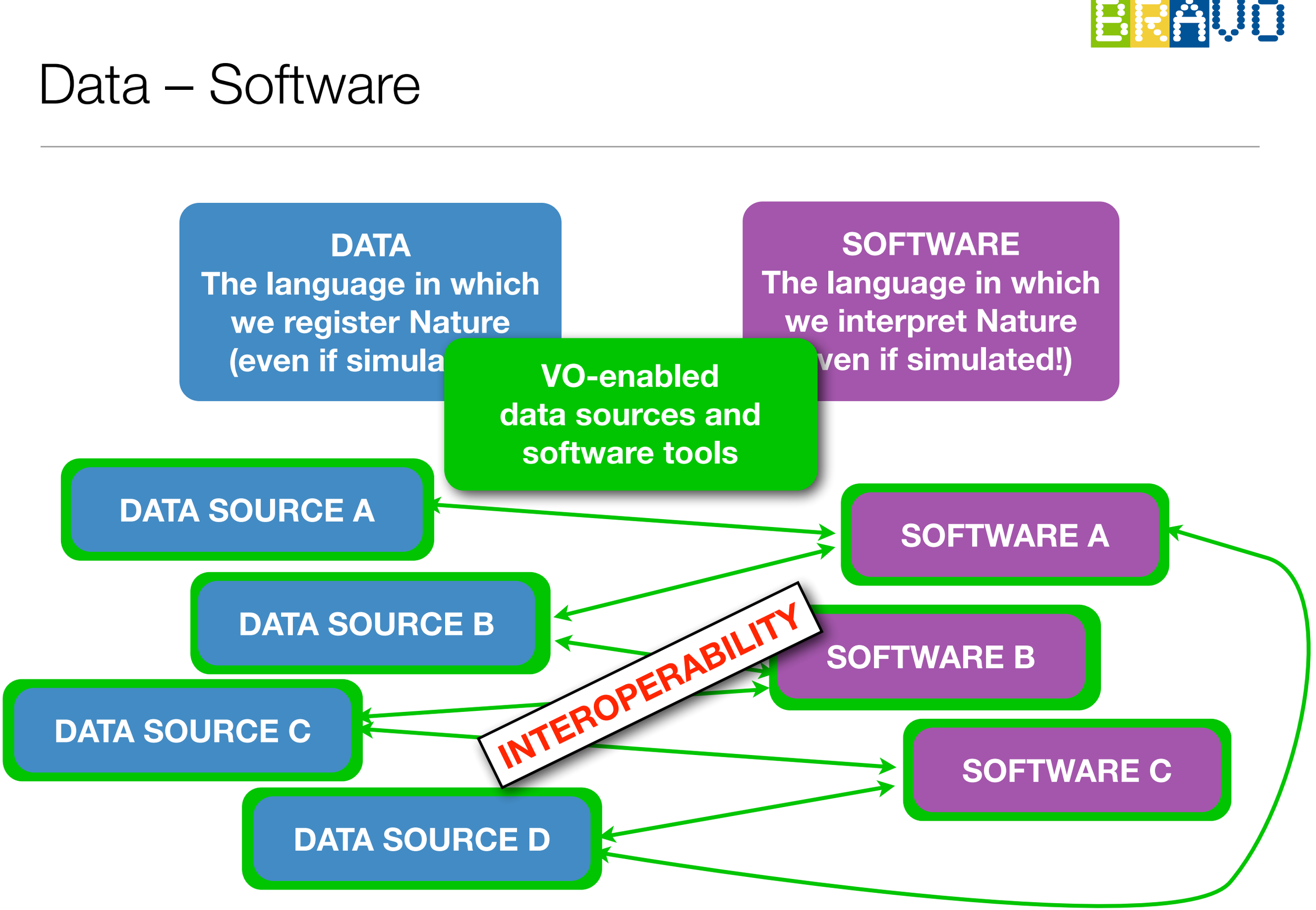


Data – Software

DATA
The language in which we register Nature (even if simula

SOFTWARE
The language in which we interpret Nature ven if simulated!)

VO-enabled data sources and software tools



The Virtual Observatory (VO) is …

- the vision that astronomical datasets and other resources should work as a seamless whole (IVOA)

- a collection of interoperating data archives and software tools which utilize the internet to form an environment in which astronomical research projects can be conducted (Wiki)

- a data discovery, access, and integration facility (B. Hanisch)

The Virtual Observatory (VO)

It provides a common language (standards, protocols, data models etc.) for communicating and exchanging data within Astronomy, and indicate where the data is stored!

Goal: To enable science! Transparent and distributed access to data available worldwide, allowing scientists to discover, access, analyze, and combine nature and lab data from heterogeneous data collections in a user-friendly manner.

An ambitious goal and no pre-existing model to follow…

IVOA

Global interoperability means world-wide agreement

Many projects and data centres worldwide are working towards this goal since 2002.

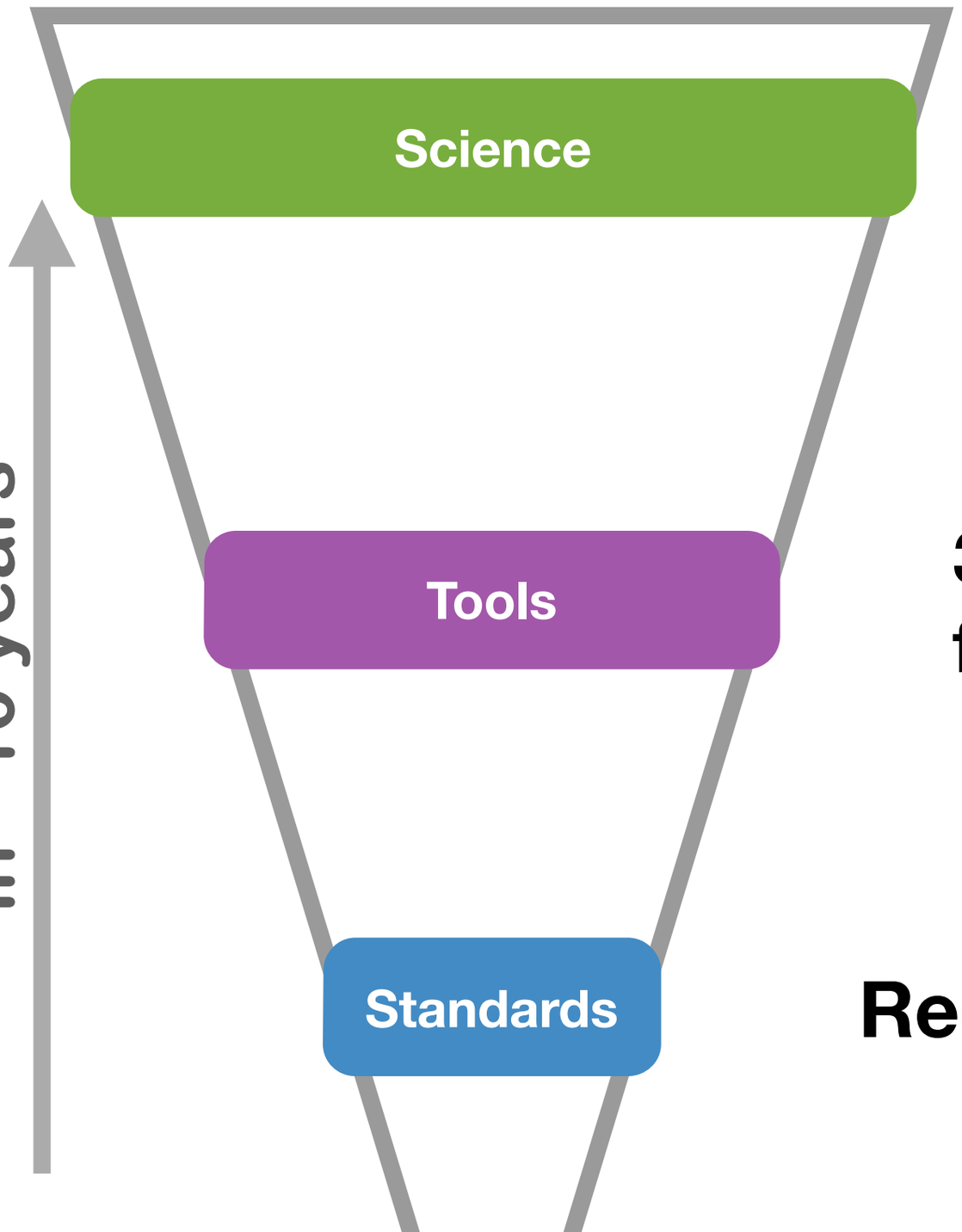
The IVOA is a standards body created by the VO projects to develop and agree the vital interoperability standards upon which the VO implementations are constructed.

International Virtual Observatory Alliance

http://www.ivoa.net

17 country members + 2 institutions

VO: bottom-up approach

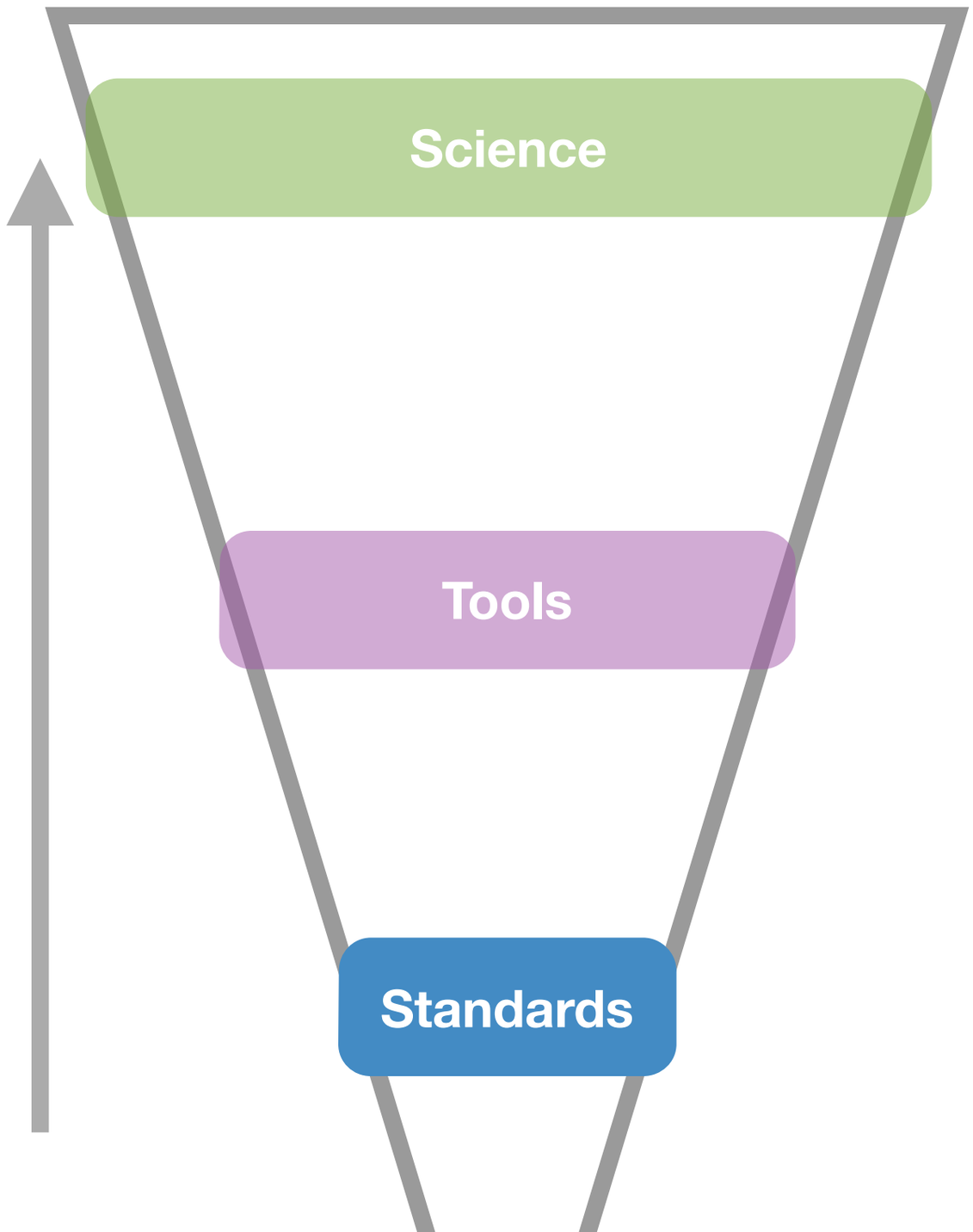


About 300 refereed papers

30+ VO-enabled applications, from 10+ VO projects

Registry, Protocols, Data Models

VO: bottom up approach

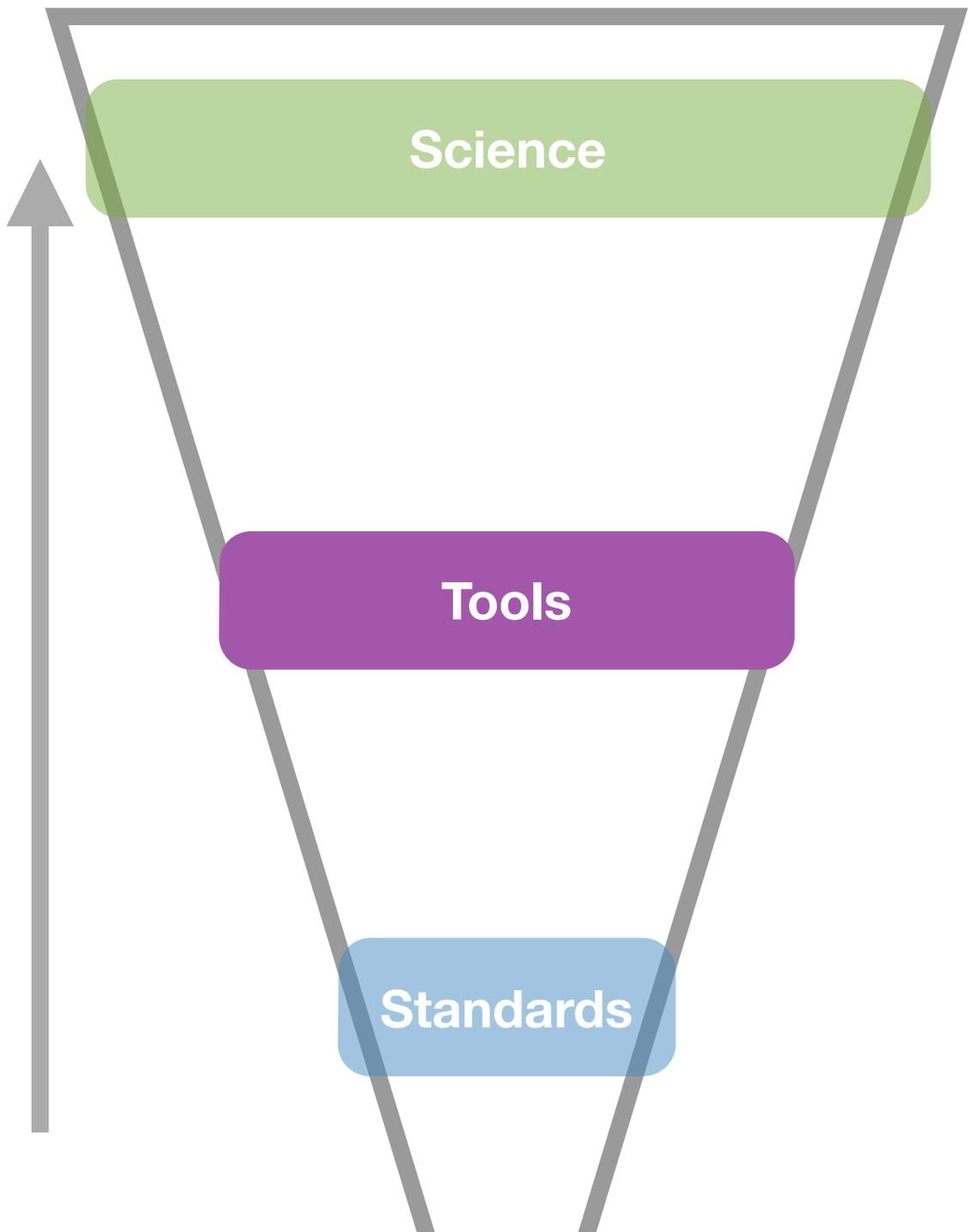


Registry, Protocols and Data Models

Core standards established

Priorities now on multi-dimensional data and time domain astronomy

VO: bottom up approach



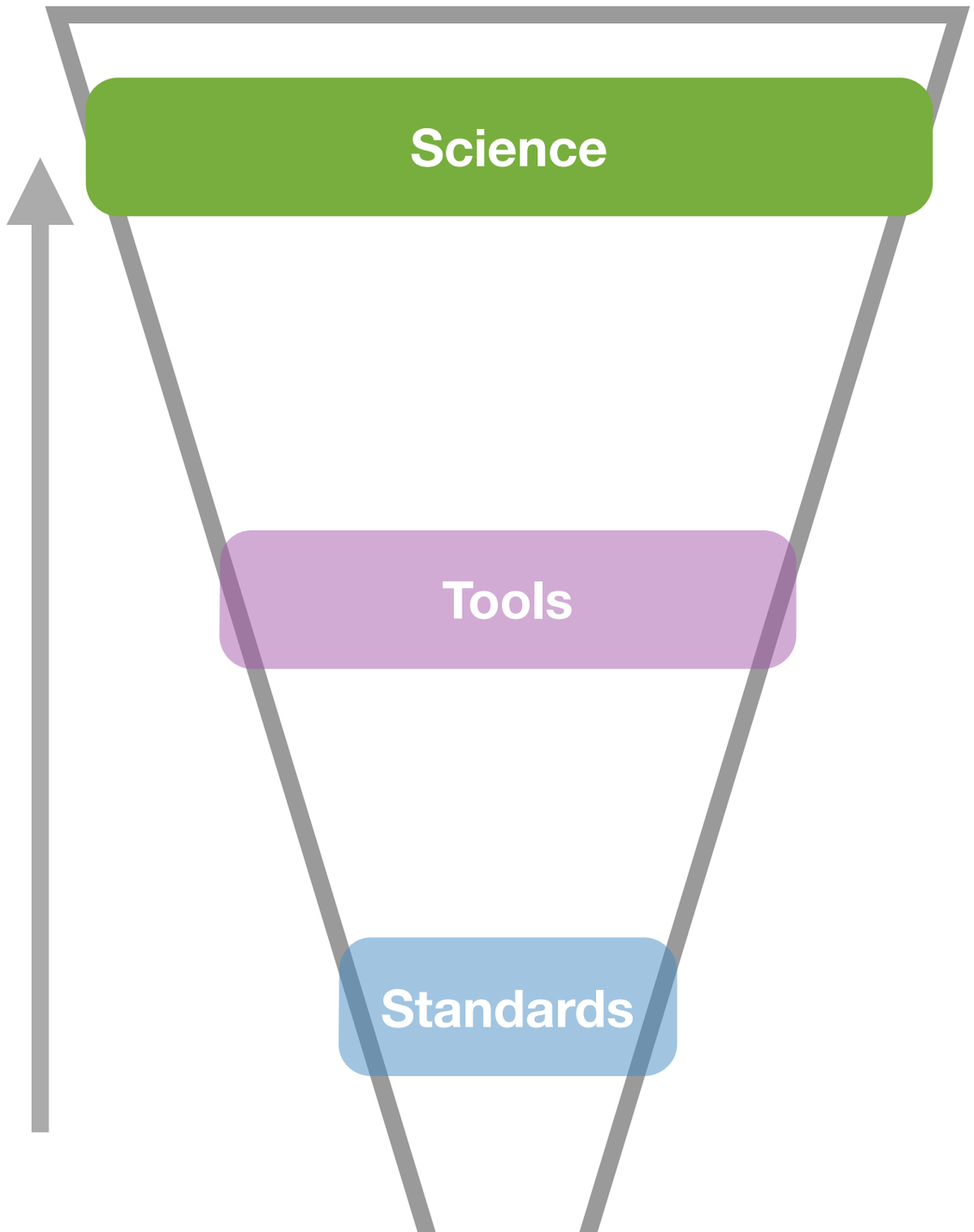
30+ VO-enabled applications, from 10+ VO projects

Many with hundreds of downloads

Web apps used frequently

Many users are unaware that data requests are being handled by VO services

VO: bottom up approach



About 300 refereed papers, several of them with innovative ways

Part of Astronomer's everyday tool kit

'VO' not well cited, but tools are!

Really is just the beginning...

sciops.esa.int

Research & Science Home | ESA Public Web Site | Sci-Tech Portal

esa Research Science Portal

Astrophysics Missions | Planetary Exploration Missions | Solar Terrestrial Science Missions | Fundamental Physics Missions | Science Faculty

Science Archives Team

ESA Science Archives

People

Science Archives at ESAC

ean Space Astronomy Centre (ESAC) hosts most of ESA astronomy and planetary missions' archives. This currently

uster Science Archive	→ Cluster Mission
A Hubble Science Archive	→ HST Mission
OSAT Science Archive	→ EXOSAT Mission
erschel Science Archive	→ Herschel Mission
O Data Archive	→ ISO Mission
anck Legacy Archive	→ Planck Mission
A's Planetary Science Archive	→ (regrouping data from Rosetta, Mars Express, Venus Express, Huygens, Smart-1 and Giotto for the time being)

Latest

- 6th M NAVCAM 2014) a new Arc Full deta
- 22nd Legacy

odsweb.u-strasbg.fr

Portal | Simbad | VizieR | Aladin | X-Match | Other | Help | Search CDS Web pages

Home

About CDS

People

Support
Help and Tutorials
Developer's corner
Publication support

myCDS

Virtual Observatory projects
IVOA
Euro VO
CoSADIE
Past projects:
Euro-VO ICE
- AIDA - DCA -
VOTech
OV France

Other projects
RDA Europe
ARCHES
ASTRODEEP

Authorities
Strasbourg

Centre de Données astronomiques de Strasbourg

Strasbourg astronomical Data Center

Entry point to all services | Object database | Catalogue database | Interactive sky atlas

Object/position | Obj/position/bibcode | Keywords, target, ... | Object/position

Other services

X-match | Dictionary | Sesame | SimPlay

Late

- Cat 201
- Cat 201
- Cat 201
- Cfa
- Sor
- HIP

www3.cadc-ccda.hia-iha.nrc-cnrc.gc.ca

Government of Canada | Gouvernement du Canada

Canada.gc.ca | Services | Departments | Français

Canadian Astronomy Data Centre

Canada

Telescope Data Products | Advanced Data Products | Services | Advanced Search | Login

CADC Home

Search for data by target | Search

Advanced Search

Telescope Data Products

Gemini | CFHT | JCMT

HST | BLAST | MOST

Advanced Data Products

MegaPipe | HLA

IRIS | CGPS

Services

Meetings | Community

SSOIS | CANFAR

And future surveys, such as J-PAS and Gaia

The VO concept elsewhere

- Space Science
  - Virtual Heliophysics Observatory (HELIO)
  - Virtual Radiation Belt Observatory (ViRBO)
  - Virtual Space Physics Observatory (VSPO)
  - Virtual Magnetospheric Observatory (VMO)
  - Virtual Ionosphere Thermosphere Mesosphere Observatory (VITMO)
  - Virtual Solar-Terrestrial Observatory (VSTO)
  - Virtual Sun/Earth Observatory (VSEO)
- Virtual Solar Observatory
- Planetary Science Virtual Observatory
- Deep Carbon Virtual Observatory

And currently this model is being exported to the National Institute of Standards and Technology (NIST) and to Neurosciences

BRAVO

- BRazilian Astrophysical Virtual Observatory

- 2006: BRAVO is born as a collaboration between **Divisão de Astrofísica** and **Laboratório de Computação @ INPE**

- 2009: BRAVO becomes IVOA partner and associate with INCT-A

- 2015: team 6 integrants and 5 collaborators http://bravo.iag.usp.br

Would you like to join us?
Let me know!

Mission

- To **stimulate** and to **encourage** the **projects** of the different local groups, facilitating the necessary coordination and collaboration for the **development and deployment** of the tools, systems, and organizational structures;

- To **organize workshops and schools** aiming at the dissemination of the VO concepts and the qualification of people capable to use and to work on the development of new VO services and tools;

- To **act as a partner of the IVOA** and as an intermediate between the IVOA and the various Brazilian groups working with VO.

To organize workshops and schools aiming at the dissemination of the VO concepts

2011

2014

- NAT/LAC/BRAVO/IAG initiative

To organize workshops and schools aiming at the dissemination of the VO concepts

2012

Workshop de eScience na Astronomia Brasileira

- One day meeting aiming at
  - To **stimulate** the **approximation** between **Astronomy and Computing** (CS and Eng.) research **communities**.
  - Open a communication channel between these societies.

- Result: The **creation** of the **Astronomy Track** at the **Brazilian Computer Society** annual meeting!

To organize workshops and schools aiming at the dissemination of the VO concepts

2012

To organize workshops and schools aiming at the dissemination of the VO concepts

- To **develop a minimal SIAP** service from scratch and to **demonstrate** that this works **integrated** with Microsoft's **World Wide Telescope**;

- **21 teams**

- **To the winners:** Internship stage at Microsoft Research and/or Caltech and Two AMD notebooks;

- **winner team were 2 undergrad students at UFRJ!**

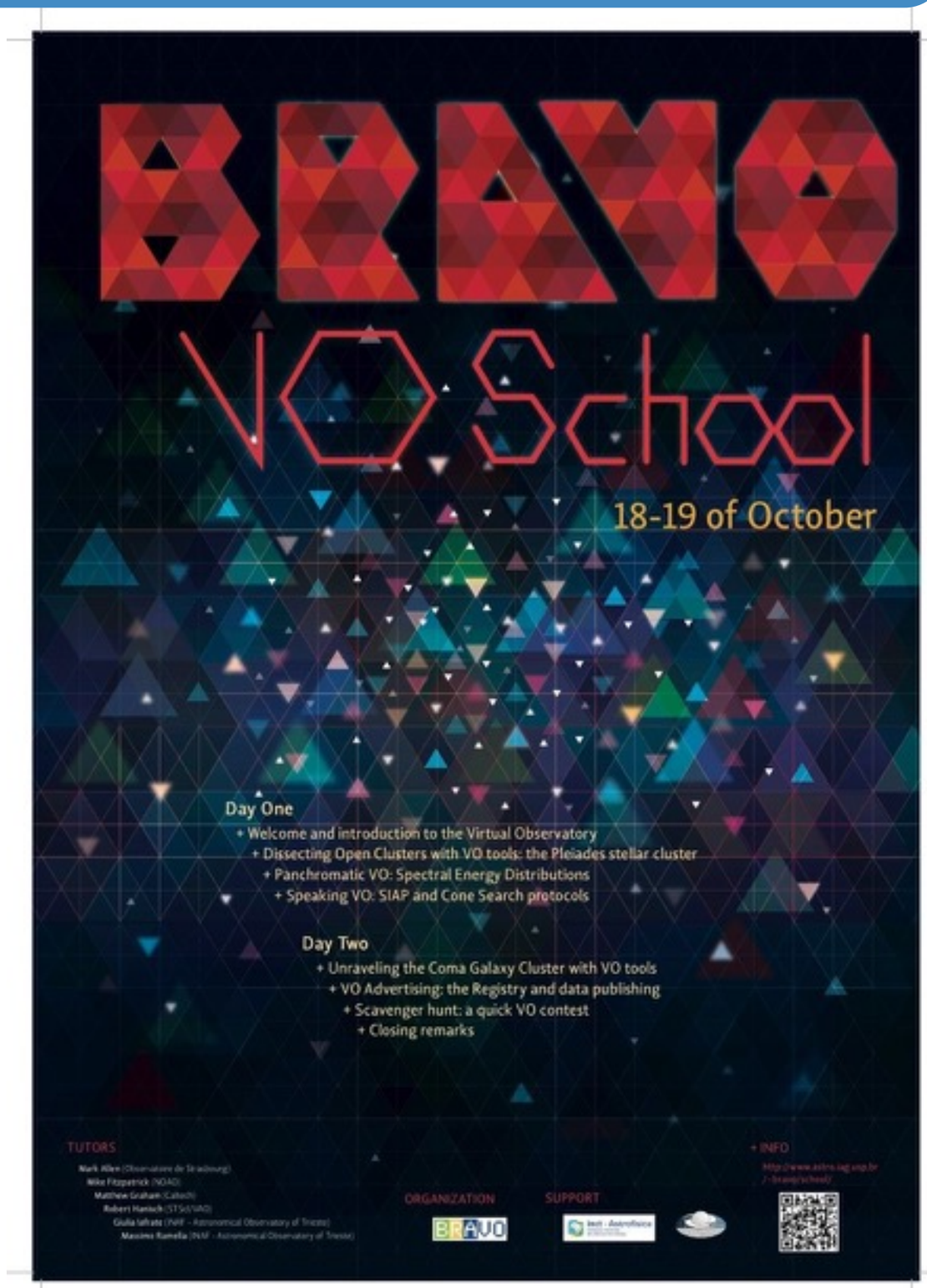
To organize workshops and schools aiming at the dissemination of the VO concepts

- **Aim**: To begin a gradual empowering of the community

- Very good answer from the community. **Over-subscribed**!

- ~35 participants

- **Tutors**: Mark Allen, Mike Fitzgerald, Mathew Graham, Bob Hanisch, Giulia Iafrate, Massimo Ramela



To act as a partner of the IVOA and as an intermediate between the IVOA and the Brazilian groups working with VO

To act as a partner of the IVOA and as an intermediate between the IVOA and the Brazilian groups working with VO

VO-DAY

- *The International Virtual Observatory Alliance*
  Mark Allen
- *The Research Tools of the Virtual Astronomical Observatory*
  Robert Hanisch
- *What the VO can do for you?*
  Massimo Ramela
- *An introduction to VO-IRAF*
  Mark Fitzpatrick
- *The transient sky and the Virtual Observatory*
  Mathew Graham
- *The VO in the classroom*
  Giulia Iafrate

To act as a partner of the IVOA and as an intermediate between the IVOA and the Brazilian groups working with VO

- IVOA centered discussions
- VO-driven presentations.
- Brazilian talks about:
  - J-PAS and S-MAPS
  - DES Brazil
  - South-Pol
  - ... and BRAVO.

To stimulate and to encourage the projects, facilitating the development and deployment of the tools, systems, and organizational structures

2012

We negotiated community access to the IAG/NAT's Alphacrucis cluster, together with the emergent groups

2011

GINA

O uso de GPUs (Graphical Processing Units) tem aberto novas possibilidades para computação astronômica, provendo paralelização em grande escala.

GINA (GPUs para o Instituto Nacional de Ciência e Tecnlogia de Astrofísica) é um cluster experimental, com o objetivo de ser um ambiente de desenvolvimento e testes de aplicativos que fazem uso de GPUs na área de Astronomia.

Está aberto para o uso de toda a comunidadade astronômica brasileira

GPUs para o INCT-A

BRAVO BRAZILIAN VIRTUAL OBSERVATORY

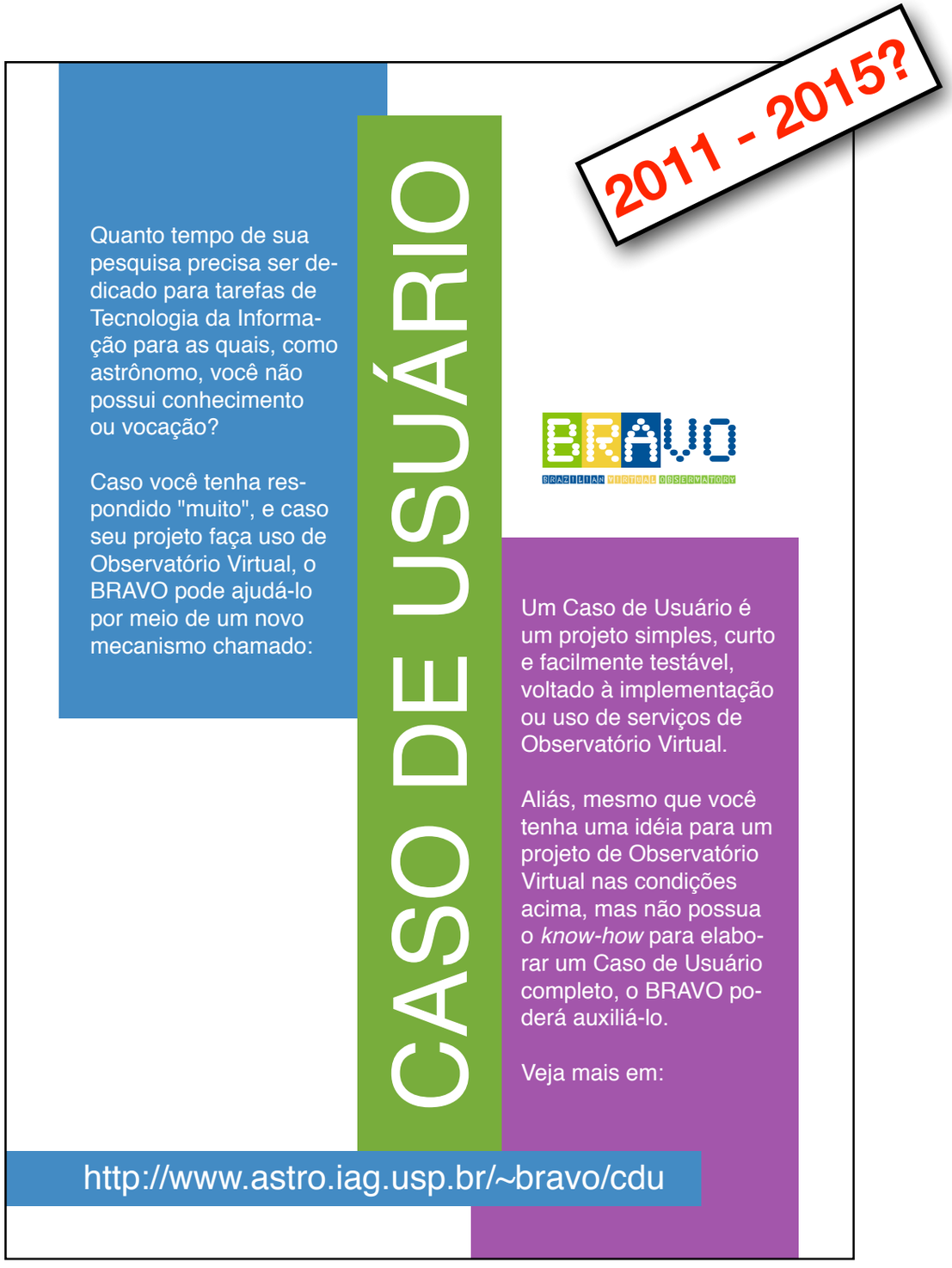
O computador GINA possui atualmente 2 nós, cada um com:
· 2 GPUs NVIDIA Tesla C2050 (448 núcleos cada)
· 2 CPUs Intel Xeon X5650 (6 núcleos cada)
· 48 GB de RAM

Os softwares já disponíveis incluem:
· CUDA C
· CUDA Fortran
· IDL (com GPULib)
· R (com R+GPU)
· Python (com PyCUDA)
· MPI, OpenMP, TORQUE

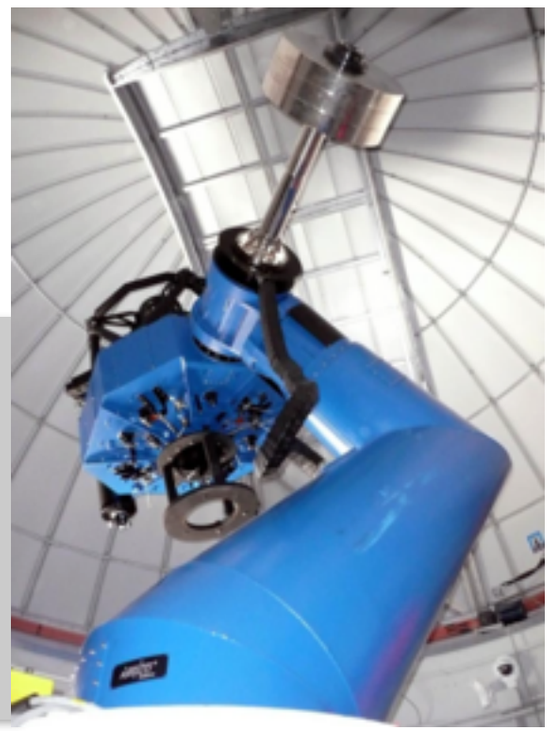
Veja mais em:

http://www.astro.iag.usp.br/~gina

To stimulate and to encourage the projects, facilitating the development and deployment of the tools, systems, and organizational structures



- A Database for the SOUTH POL survey
- Photometric Redshifts Portal
- Grid of Synthetic Spectra (3x)
- A-Plus (S-Plus + J-Plus), ~40TB/year



To stimulate and to encourage the projects, facilitating the development and deployment of the tools, systems, and organizational structures

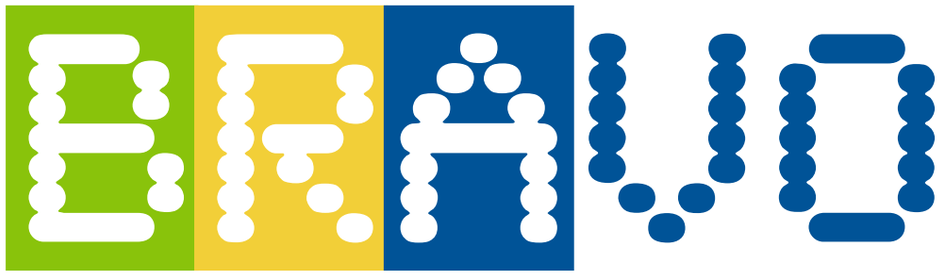
- First BRAVO's **dedicated servers** !
- To host several VO compliant archives and services (catalogues, images and theoretical spectra) **Currently being deployed**
- Software architecture being designed, **advisors:** SVO (Spain), CADC (Canada), NOAO (US), NOVA (Argentine)

BRAVO nowadays

- **Synergy with J-PAS and A-Plus surveys**
  - 3 telescopes, 2 surveys, **17000 square degrees**. A consortium of Spanish and Brazilian institutions, funding agencies and universities. The team of scientists and engineers includes more than 100 people from Brazil, Spain, the U.S.A. and other countries.

- **Archiving and publishing of T80S data** (S-Plus), and negotiating mirror of T80N data (J-Plus)

- Strong support from VO communities abroad, with special thanks to Spanish teams at CEFCA and Spanish VO



We are an alliance of people in Brazil who believes in worldwide, interoperable and distributed access to astronomical data and software.

Would you like to join us? Let us know!

Obrigada!

bravo@iag.usp.br