Data Intensive Astronomy

Astronomy Data Center,
National Astronomical Observatory of Japan

Masatoshi OHISHI
masatoshi.ohishi@nao.ac.jp
Data Intensive Astronomy

- **Data deluge**
  - Huge Image data
  - Wide spectral range
  - Transient data
  - time-domain

- **New paradigm in astronomical research by introducing data management and advanced data analysis**
Planned Future Projects

- ALMA
- JWST
- LSST
- LOFAR
- SKA
- TMT
- Pan-STARRs
- E-ELT

45 PB/yr x 10 yr ~ 450 PB

~ 200 TB/yr

~ a few TB/night, only object params stored

Credit: LSST Corporation

Courtesy TMT Observatory Corporation
x2 increase/1.5 year

A graph showing the increase in data volume (TB) over time (Year). The data points are labeled as follows:

- 2MASS
- SDSS (DR8)
- ALMA (1 year)
- Pan-STARRS (2013)
- LSST (1st year)
- LSST (5 year)

The graph indicates an exponential increase in data volume, with a significant rise predicted for LSST in the 5-year horizon.
IAU GA
Beijing
August, 2012

- Scientific Impact of Past and On-Going Large-Scale Observations and Surveys to Astronomy
- Current Status and Challenges of Future Large-Scale Observations and Surveys (1) Near- and Mid-future projects, (2) Far-future projects
- Data Management and Data Access: Past, Present and Future

- Advanced Data Analysis in the Data Intensive Astronomy Era
- Synergy of Data Intensive Astronomy with other field
- Expectation on Scientific Insights in the Data Intensive Astronomy
- Education, Public outreach related with Data Intensive Astronomy

Its program and presentation files are available at http://www.adc.nao.ac.jp/SpS15/program.html
Virtual Observatory

- Infrastructure for **efficient** research environment
- **International standards** for data publication & access
- Sharing data worldwide, **Maximize** scientific return

![Diagram of Virtual Observatory](http://jvo.nao.ac.jp/)

**Data Services**
- SDSS
- 2MASS
- HST
- Subaru
- ...

**Analysis Services**

**Web portal**

**Applications**
Astronomical Virtual Observatories

More than 10,000 resources, including Subaru SupCAM and HDS, are accessible

Images, spectra, and catalog data can be retrieved
Science output by JVO

• Concentration of galaxies around SMBH
  – 10,000+ SMBH & $7 \times 10^7$ galaxies collected by JVO
Scientific Results from SMOKA (Subaru-Mitaka-Okayama-Kiso Archives)
HSC (Hyper Suprime-Cam)

• A new prime-focus camera for Subaru
  – Cosmology (weak lensing)
  – FL in October, 2012

**Suprime-Cam**
- FoV: 34′ × 27′ (10 × 2k4k CCDs)
- Data: 185MB/shot (≈30GB/night)
- Survey Area: 1～10 sq. deg.

**HSC**
- FoV: 1.5 d × 1.5 d (104 × 2k4k CCDs)
- Data: 2GB/shot (≈300GB/night)
- Survey Area: ～2000 sq. deg.
http://jvo.nao.ac.jp/portal/alma.do

JVO Portal

Search Result

ALMA VO Web Viewer

Web Service

User’s PC Application

ALMA VO Desktop Viewer

User PC Environment

FITS Data Retrieval

JVO Service

VO Search

Select subset of the data cube and download for local application

Examine downloaded data from several aspects. Comparison of the data taken from other VO services is available.

ALMA VO Service

ALMA EA-ARC
~12.6 TB of memory and ~6 PB of storage, in total
Available Data @ ADC, NAOJ

• Raw data from Subaru and others (SMOKA)
  http://smoka.nao.ac.jp/
• Japanese Virtual Observatory – Reduced data from all around the world
  http://jvo.nao.ac.jp/portal/
  – ALMA Archival Data
    http://jvo.nao.ac.jp/portal/alma.do
• Hinode Science Archive
  http://hinode.nao.ac.jp/SDAS/index_e.shtml
• Nobeyama Data Archive
  http://nrodb.nro.nao.ac.jp/
• VERA, RISE, Nobeyama radio heliograph and others
Contact me if you are interested in the “Data-Intensive Astronomy”

masatoshi.ohishi@nao.ac.jp