

# LOFAR Archive Information System

*Kor Begeman<sup>a</sup>, Andrey Belikov<sup>a</sup>, Danny Boxhoorn<sup>a</sup>, **Fokke Dijkstra<sup>b</sup>**  
Dennis van Dok<sup>c</sup>, Hanno Holties<sup>d</sup>, Zheng Meyer-Zhao<sup>b</sup>, Adriaan Renting<sup>d</sup>,  
Floris Sluiter<sup>e</sup>, Edwin Valentijn<sup>a</sup>, Willem-Jan Vriend<sup>a</sup>*

*a OmegaCEN, Kapteyn Astronomical Institute, University of Groningen, NL*

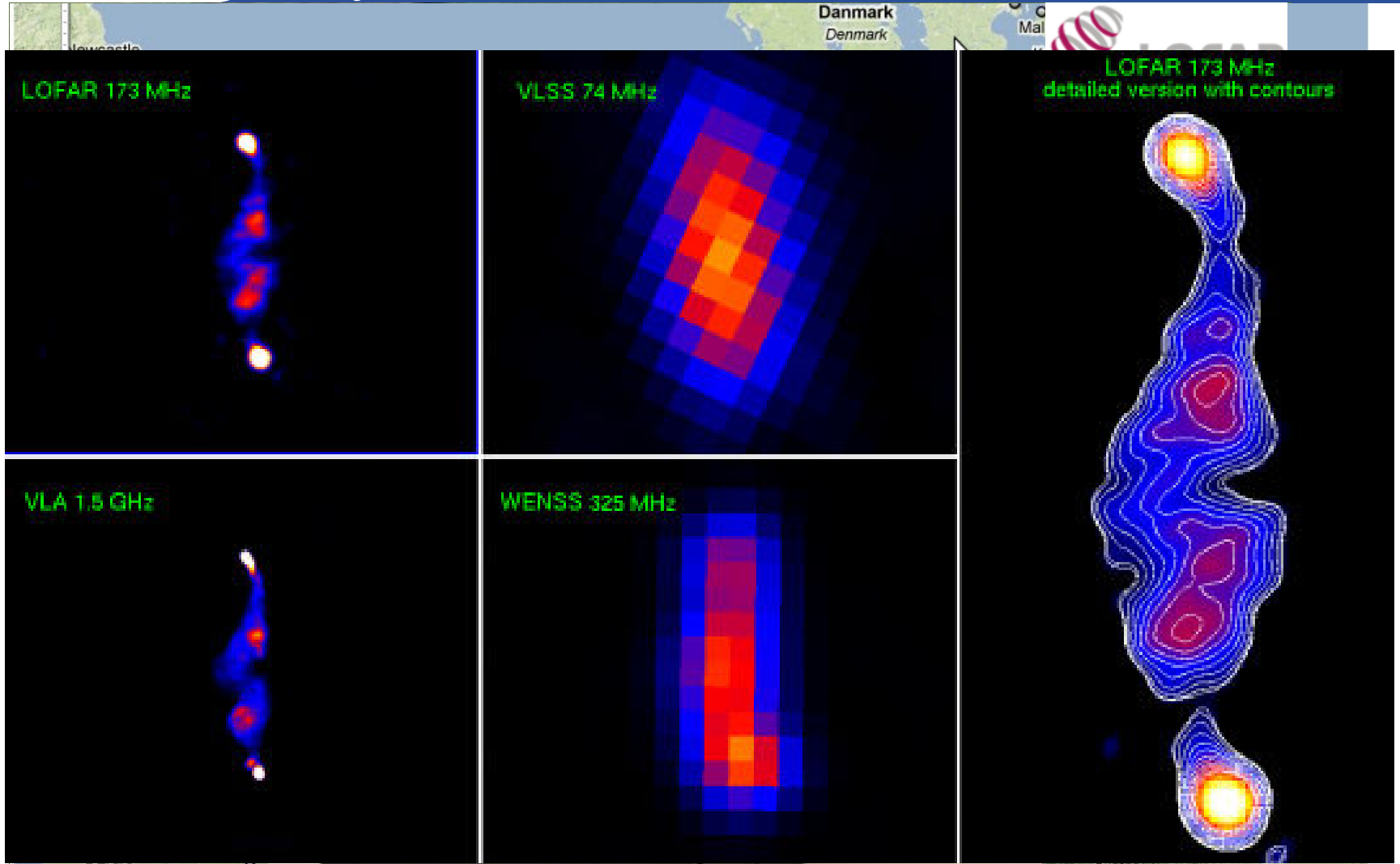
*b Donald Smits Centre for Information Technology, University of Groningen, NL*

*c Nikhef, Amsterdam, NL*

*d Astron, Dwingeloo, NL*

*e SARA, Amsterdam, NL*

*EGEE User Forum, Uppsala, 14 April 2010*



60 hour observation of 3C61.1 using 20 HBA stations, compared with observations of other instruments

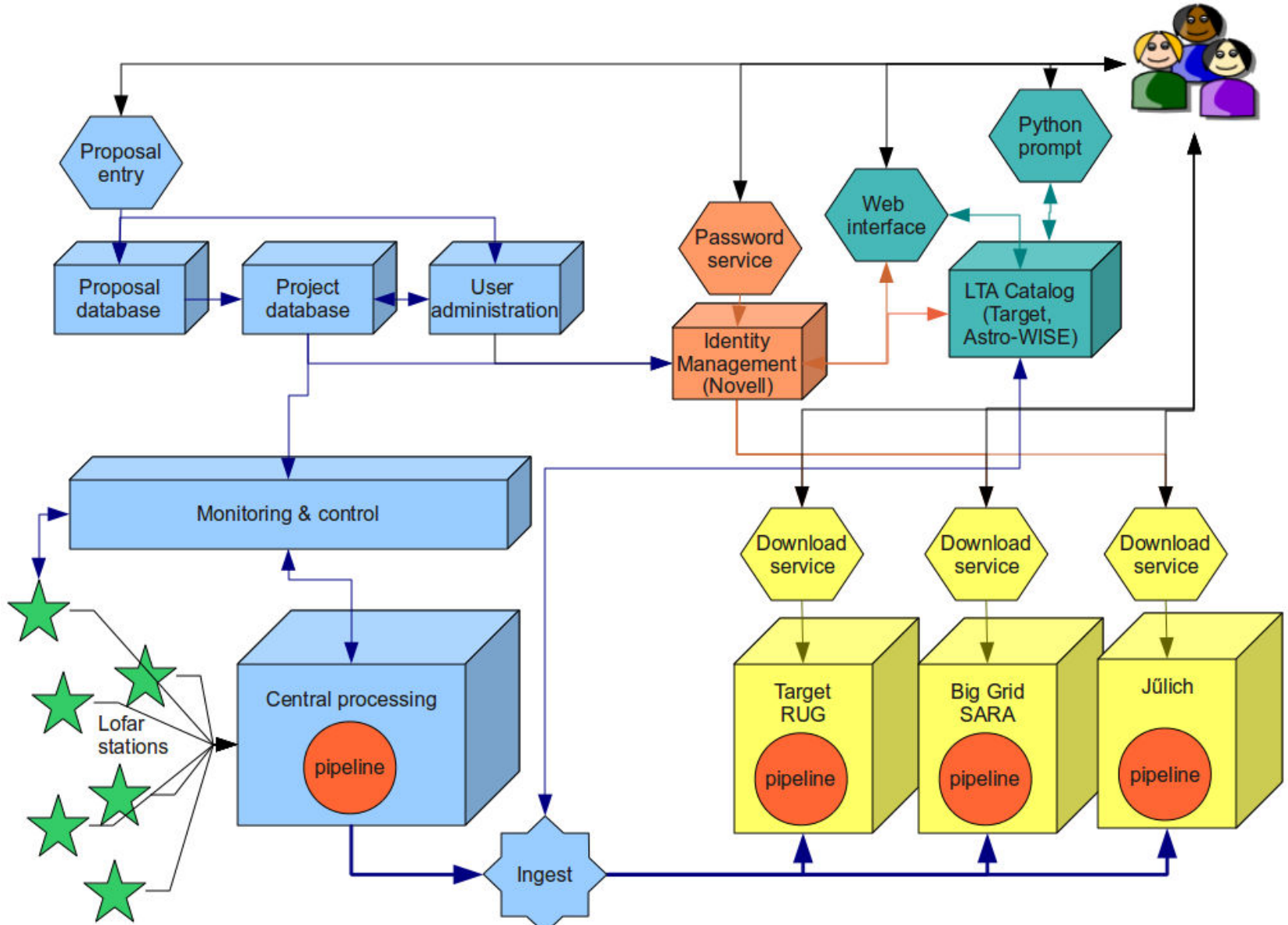
- **Several petabytes per year need to be archived**
  - Storage & Bandwidth requirements
  - Processing within the archive
- **Integrated environment**
  - Processing & Storage
  - Data mining
- **Distributed archive**
  - Groningen, NL – Target
  - Amsterdam, NL – BiG Grid @ SARA
  - Jülich, DE – Grid (planned)
  - ?

- **LOFAR archive will make use of different platforms at different locations**
  - Grid (BiG Grid NL NGI) (<http://www.biggrid.nl>)
    - Processing
    - Storage (SRM protocol)
  - Target (Data storage and data mining project, NL) (<http://www.rug.nl/target>)
    - Metadata catalog + processing engine (Astro-WISE)
    - Processing
    - IBM GPFS storage platform
  - Other partners in the near future
    - Forschungszentrum Jülich, DE, Grid technology



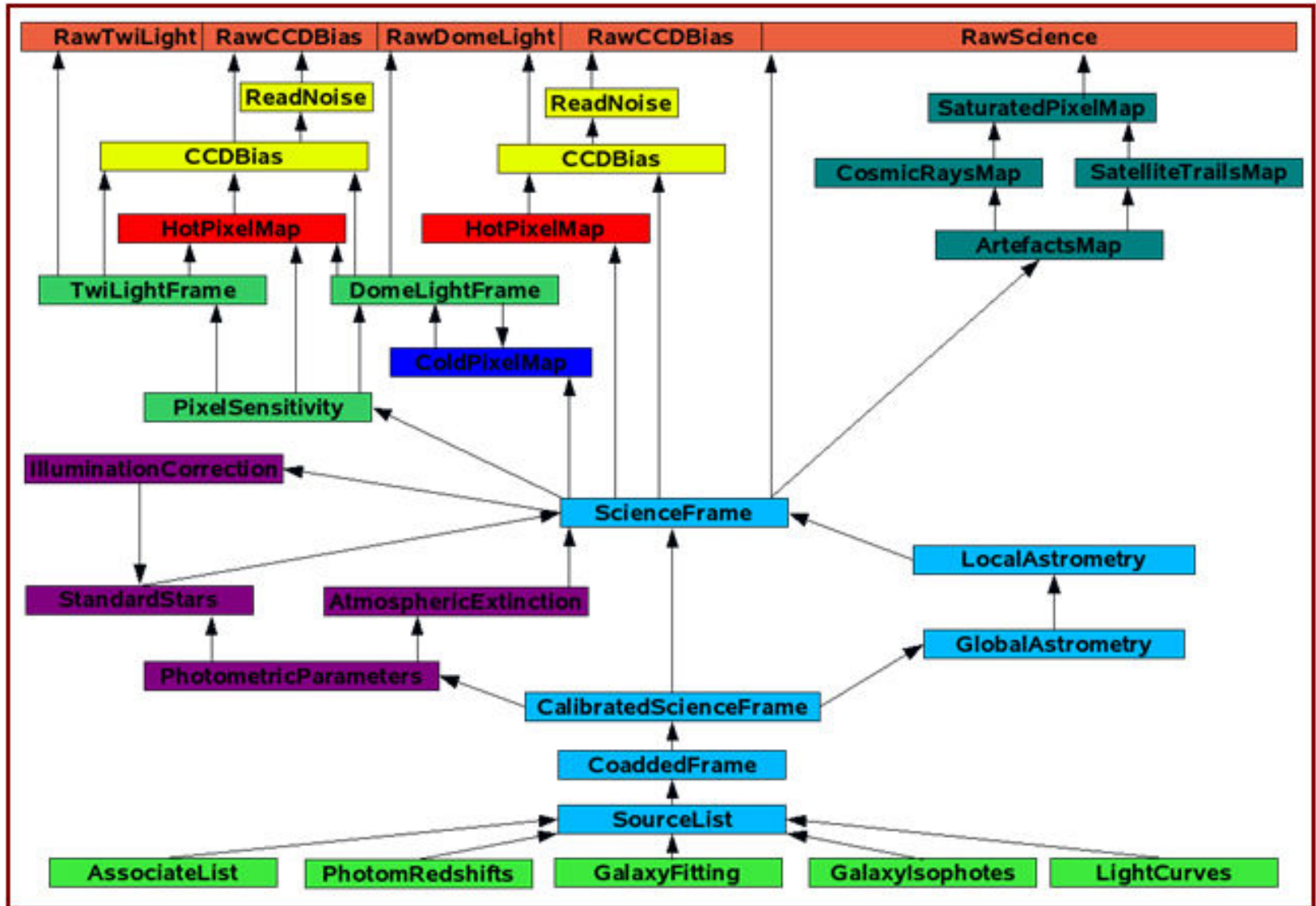
**BiG Grid**  
the dutch e-science grid





- **Complete environment**
  - Metadata catalog
  - Processing environment
  - Data mining
- **Key component of the Target project**
- **Target processing**
  - On the fly processing to generate requested & up-to-date results
  - Full data lineage stored in db!
- **Extended to handle Grid storage and processing**





- **LOFAR long term archive uses Grid resources**
  - Computing & storage
- **Astro-WISE adapted to handle these**
  - Grid credentials handled using external tools
    - Certificate Management Wizard (previously called MyProxy uploader) (<http://www.ngs.ac.uk/tools/certwizard>)
    - jLite (<http://code.google.com/p/jlite/>)
    - Proxy server to store user credentials for Astro-WISE DPU
  - Job submission described at Istanbul conference
    - Distributed Processing Unit (DPU) on gLite UI using proxy from Proxy server

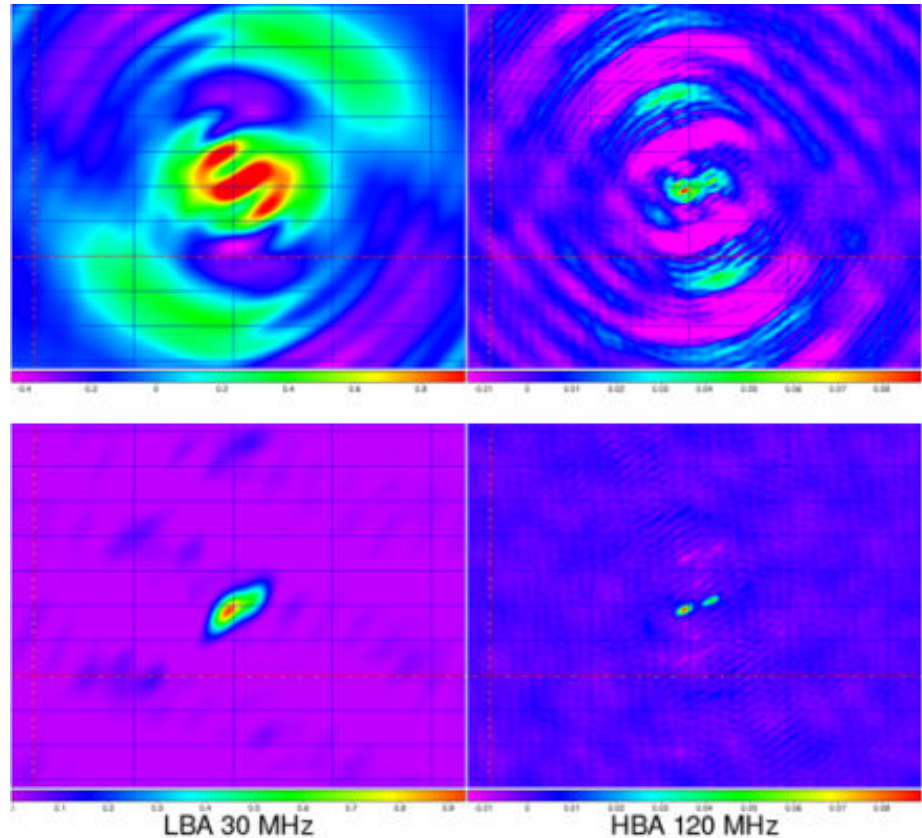


- **Requirements**

- Support multiple protocols
  - Astro-WISE data servers
  - TarGet GPFS storage
  - SRM Grid storage
  - ?
- Store entries to files in Astro-WISE db
- Allow for and keep track of replicas
- Access rights on metadata may differ from those on file location information
- Integrate with Python environment
  - Using lightweight Grid clients, e.g. dCache SRM client
- Described in more detail at EGEE '09 Barcelona

- **LOFAR instrument will become operational in the coming months**
- **Archive must also be operational**
  - Currently Catalog operational
  - Grid storage in use
  - Target storage in development
  - Processing under construction
- **Port main pipelines to Grid / Target**
  - Pipelines currently deployed on LOFAR Central Processing
  - Integrate with Target processing framework
  - Porting effort maybe substantial

- LOFAR Archive almost operational
- Successful integration of LOFAR Archive with EGEE
- Storage and processing of Astro-WISE catalog integrated with EGEE
- Power of Astro-WISE available for LOFAR user community
- Future work on porting and integration of LOFAR processing pipelines



First LOFAR observation of Cygnus A with three stations  
(Copyright: ASTRON)