

Space is big. Really big. So is the data.

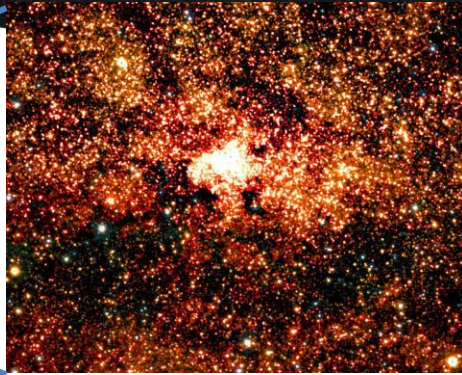
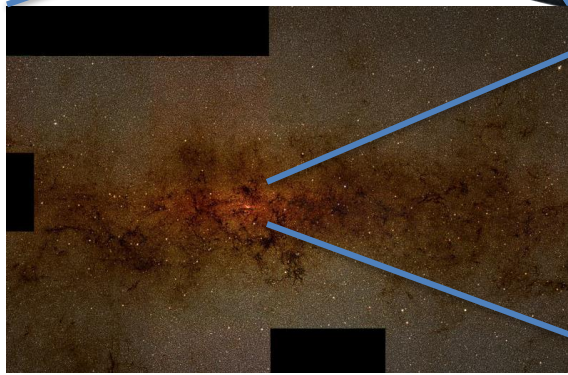
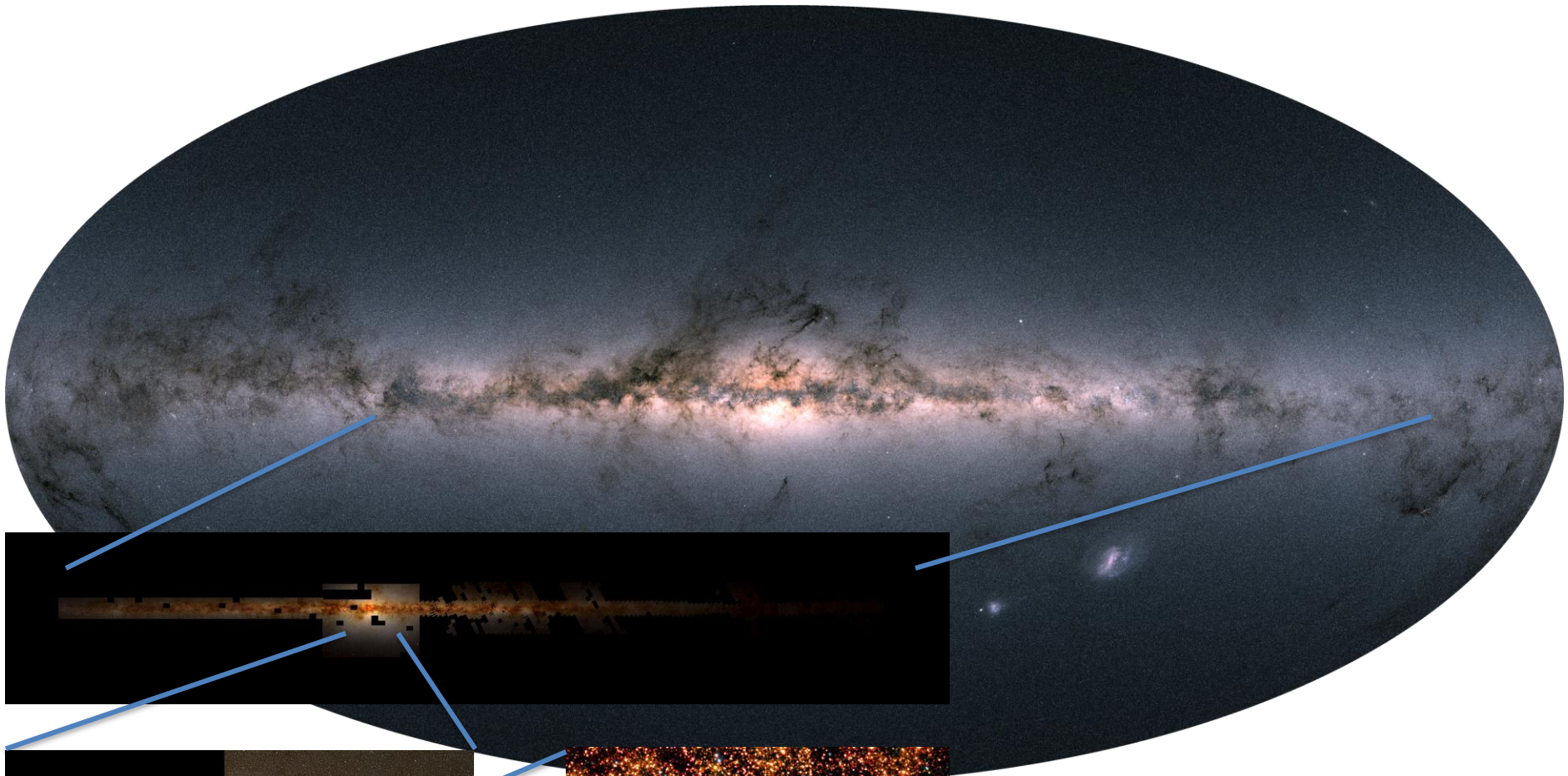
(... you just won't believe how vastly,
hugely, mind-bogglingly big ...)

Nigel Hambly

Wide Field Astronomy Unit, Institute for Astronomy
School of Physics and Astronomy
University of Edinburgh



Big Surveys = Big Data



<http://djer.roe.ac.uk/vsa/vv/v/mosaic/lb.html>

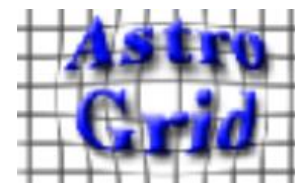
WFAU at IfA

- Formed in 1999, built on the long tradition of astronomical imaging science and research at the ROE on Blackford Hill
- Production and digitisation of legacy photographic surveys
- Progressed to managing data from new generation optical / near infrared surveys
- Latterly involved in data processing SW engineering for large space-based survey missions (European Space Agency)
- Now preparing for the next generation of ground and space-based surveys
- Recently moved into new offices in the Higgs Centre for Innovation, co-located with the UK Astronomy Technology Centre
- Current head count: 14 Research Software Engineers



WFAU projects: past

- SuperCOSMOS digitisation/dissemination programme (pre-2008)
 - <http://ssa.roe.ac.uk>
 - 20 TB imaging data
 - 4 TB, 10^{10} row catalogues (all sky)
- United Kingdom Infrared Sky Surveys science archive
 - <http://wsa.roe.ac.uk>
 - 230 TB imaging data
 - 17 TB, 10^{10} row catalogues ($\frac{1}{4}$ sky)
- Astrogrid: SW infrastructure for the global Virtual Observatory
 - <http://www.astrogrid.org>



WFAU projects: present

- United Kingdom Infra-Red Telescope Hemisphere Survey
- European Southern Observatory's visible and infrared survey telescopes public surveys

- <http://osa.roe.ac.uk>

- <http://vsa.roe.ac.uk>

- 310 TB imaging data

- 136 TB, 10^{11} row catalogues ($\frac{1}{2}$ sky)



- Gaia-ESO spectroscopic surveys

- <http://ges.roe.ac.uk>



- Gaia (European Space Agency)

- <http://sci.esa.int/gaia/>

- ≈ 100 TB data

- 10^{12} row catalogues (all sky)



- Further SW infrastructure work for the Euro-VO and International Virtual Observatory Alliance

- Tools, infrastructure and standards for inter-operability



WFAU projects: in prep. / future

- ESA-Euclid



- <http://sci.esa.int/euclid/>

- ≥ 1000 TB imaging data ($\geq \frac{1}{2}$ sky)

- systems engineering, data processing SW engineering, Data Processing Centre preparation

- Large Synoptic Survey Telescope



- <https://www.lsst.ac.uk>

- $\approx 60,000$ TB imaging data ($\geq \frac{1}{2}$ sky)

- UK Data Access Centre preparations

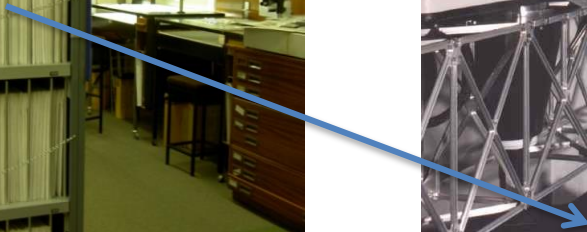
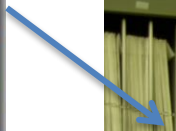
The IT Future for WFAU ...

- Larger collaborations
 - International survey projects
 - Professional Research Software Engineering
- Data management
 - No longer purely in-house
 - Distributed systems
- Computing model is evolving
 - IT base no longer purely relational
 - Scale-out usage scenarios need HPC / HTC
 - Cloud and Grid-based technologies

The importance of RSEs

- Data processing
 - Research background = domain knowledge
 - IT training = professional programming standards
 - Collaborative development
- Data serving
 - Familiarity with state-of-the-art technologies
 - Abilities in system administration and network infrastructure
 - Maintenance of high availability services
- Data preservation
 - Long term curation, e.g. migrate data onto new media
 - Ensures survey legacy value is maintained

Data preservation (I)



Data preservation (II)

<http://ssa.roe.ac.uk>

SuperCOSMOS Science Archive

Home | Overview | Browser | Access | Cookbook | Links | Credits

SSA - SuperCOSMOS Science Archive

LATEST NEWS: A new database of ~200 plates in the single survey field no. 287 has been included in the SSA - for more details click on [F287 here](#) or in the navigation bar below left.

The SuperCOSMOS Science archive holds the object catalogue data extracted from scans of photographic Schmidt survey plates.

At around 4 terabytes in size, the database contains nearly 6.4 billion individual object detections which are merged into just under 1.9 billion multi-colour, multi-epoch sources and covers the whole sky in three wavebands (BRI), with one colour (R) represented at two epochs.

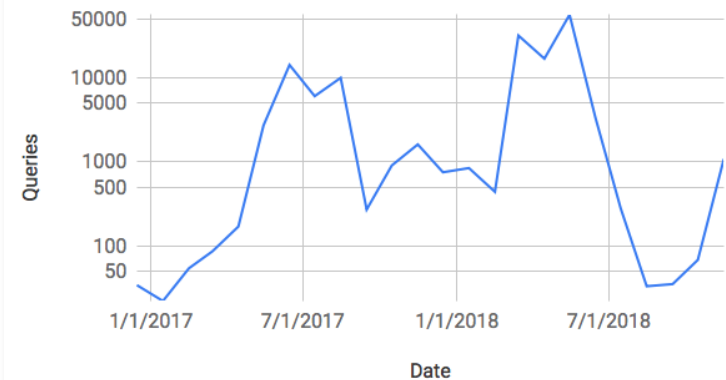
Access to the data has previously been made available through the [SuperCOSMOS Sky Survey](#) pages. The SSA is based on the same underlying data but it is housed in a relational database (Microsoft SQL Server 2005). This platform allows users more power and control over how they can access the data.

A short description of the SSA database structure and content is given in the [Data Overview](#), for full details see the [Schema Browser](#).

Users wishing to access the data should first read the general introductory notes under [Data Access](#).

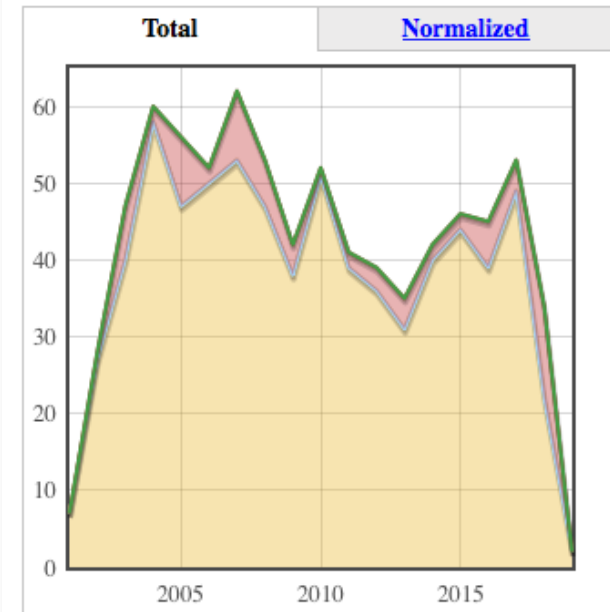
Home | Overview | Browser | Access | Cookbook | Links | Credits
Radial | MenuQuery | FreeSQL | CrossID

Queries vs. Date




(... plus 1000s of image requests per week ...)

Citations per year



- Ref. citations to ref. papers
- Ref. citations to non ref. papers
- Non ref. citations to ref. papers

Data serving (I)



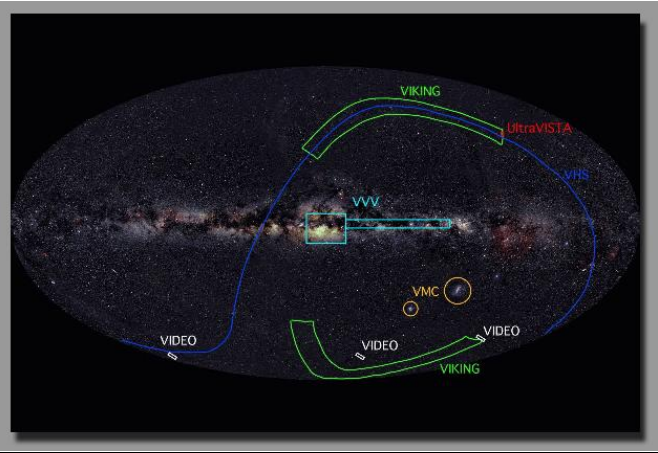
VISTA Science Archive

[Home](#) | [Overview](#) | [Browser](#) | [Access](#) | [Login](#) | [Cookbook](#) 

VSA - VISTA Science Archive

The VISTA Science Archive (VSA) holds the image and catalogue data products generated by VIRCAM on the Visible and Infrared Survey Telescope for Astronomy (VISTA). The primary contents of the archive originate from the VISTA Public Surveys. Survey science-ready catalogue data will be released in phases, while standard flat-file data products (both images and derived single passband catalogues) become available continually after routine observation and processing operations. Information on the various archive releases can be found on the [surveys page](#).

The history of archive releases, updates and bug fixes is recorded under the [release history](#) page. Users wishing to receive email announcements of such entries should subscribe to the VSA_Annoucelist (contact vsa-support@roe.ac.uk).



Picture: Sky coverage of VISTA surveys, overlaid on a 2MASS image of the whole sky.
Credit: VISTA

[Home](#) | [Overview](#) | [Browser](#) | [Access](#) | [Login](#) | [Cookbook](#)
[Listing](#) | [FreeSQL](#)
[Links](#) | [Credits](#)

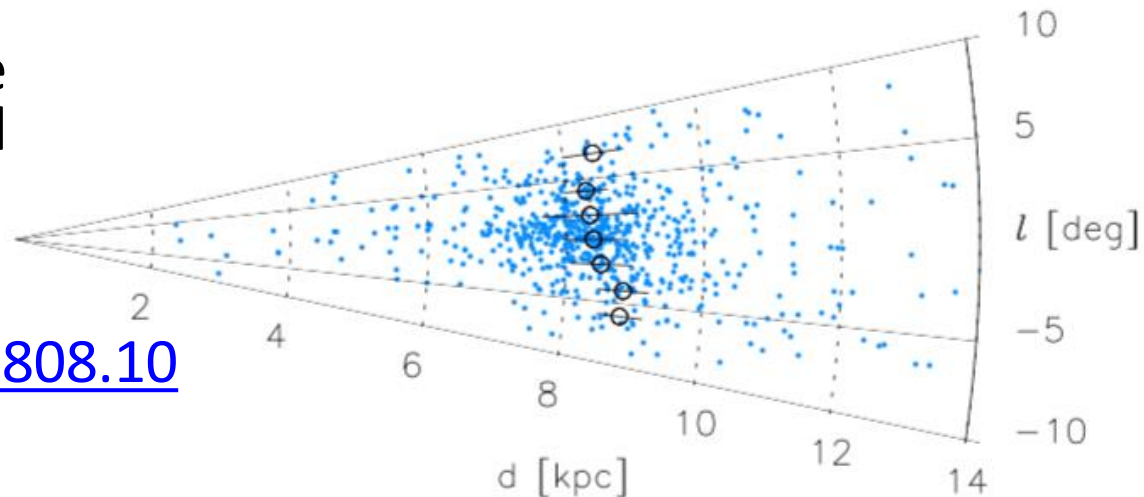
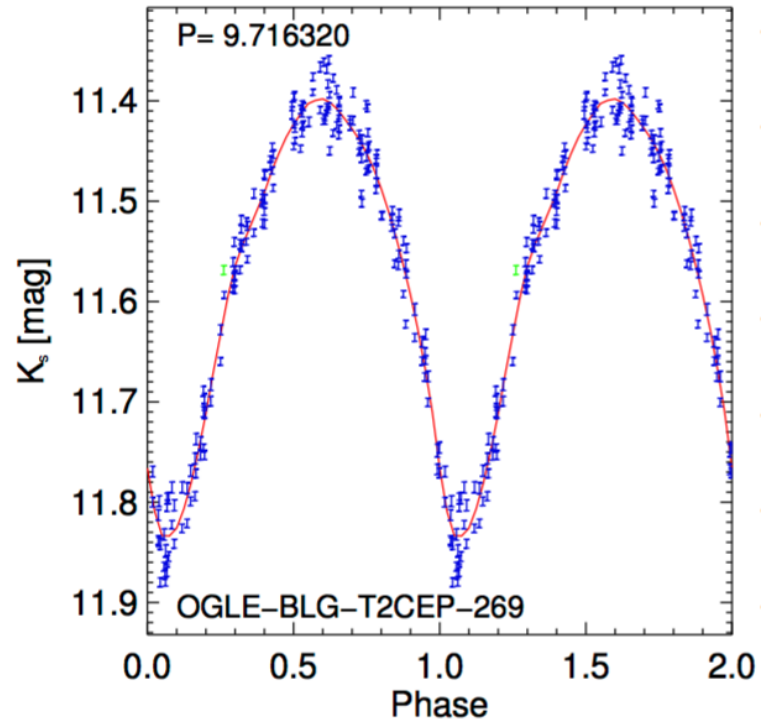
WFAU, Institute for Astronomy,
Royal Observatory, Blackford Hill
Edinburgh, EH9 3HJ, UK

vsa-support@roe.ac.uk
23/1/2015

- e.g. VISTA Variables in the Via Lactea (VVV)
 - Time domain infrared imaging survey
 - Yields light curves, stellar motions, ...
- $\approx 10^9$ unique sources with $\approx 10^{11}$ epoch measures (5th Data Release)
 - 60 billion row match table
 - Final DR “VVVX” 2x bigger
- <http://surveys.roe.ac.uk/vsa>

Data serving (II)

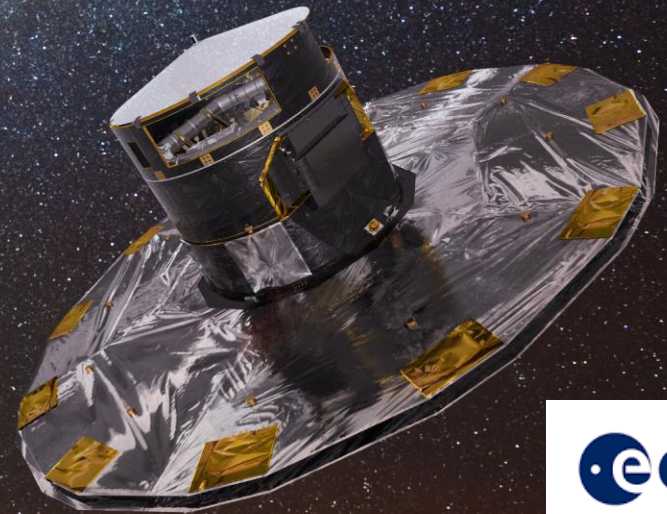
- Recent science highlight:
Braga et al. (2018)
 - Infrared imaging penetrates the interstellar gas and dust (which obscures visible light)
 - Time-resolved measurements yield light curves
 - Identify “Cepheid” variables which have a Period-Luminosity correlation
 - Spatial structure of the Galactic Bulge mapped



- <https://arxiv.org/pdf/1808.10838.pdf>

Data processing (I)

- European Space Agency's Gaia mission
 - Launched December 2013
 - WFAU staff: 3 (out of 100s spread around Europe)
 - Provide core processing SW for the “ground segment”
 - Data rate: up to 8 Mbit/s
 - Key goal: 3-d view of the Galaxy via trigonometric parallax measurements for 10^9 stars



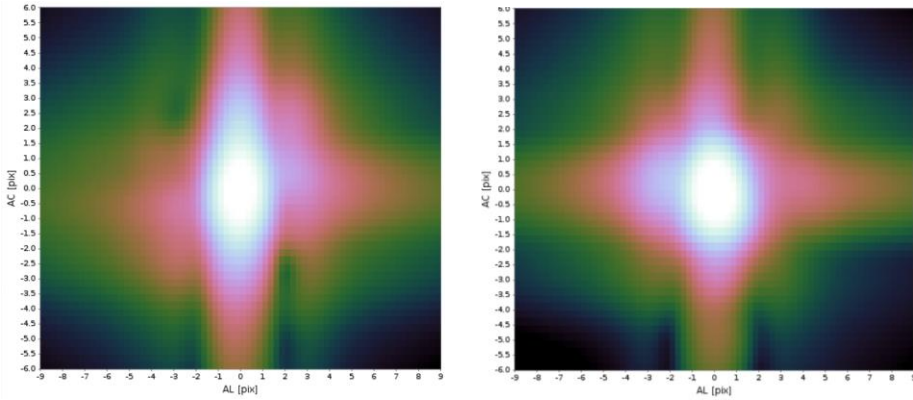
GAIA MISSION STATUS

1592 days in routine phase

60,065 GB of science data gathered

115,062,237,531 transits observed

Data processing (II)



WFAU staff provide key instrumental calibration software for Gaia:

- Optical
- Electronic

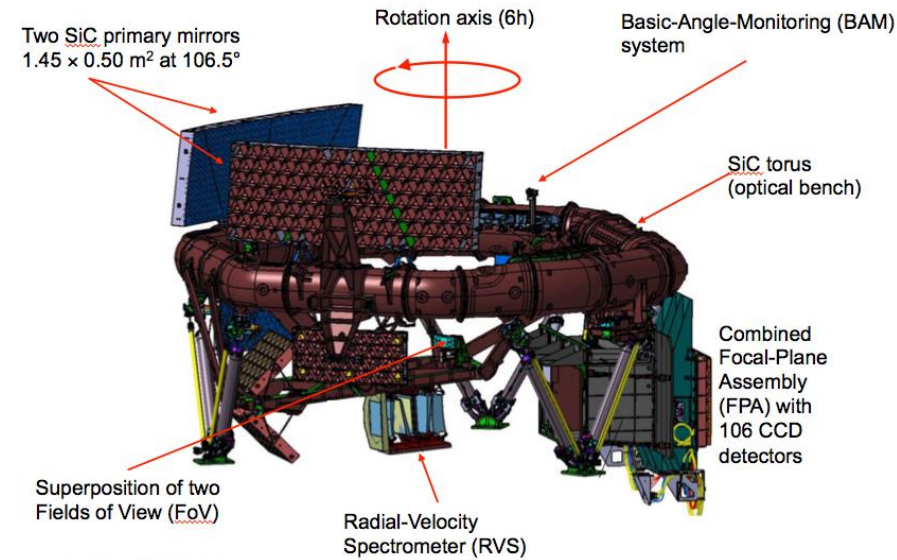
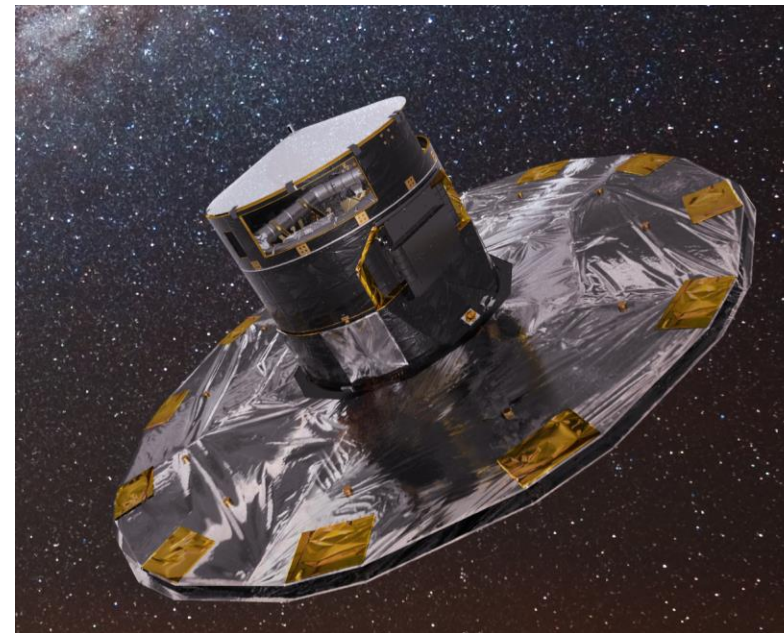
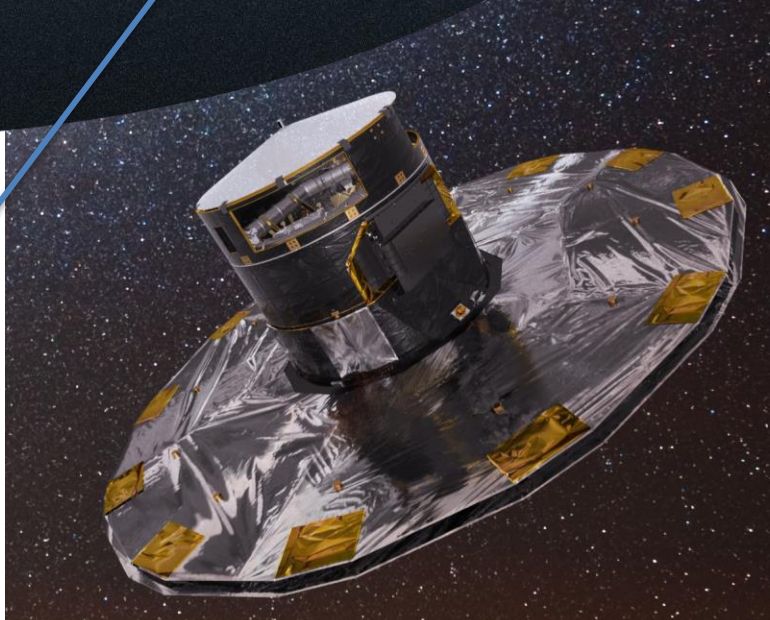
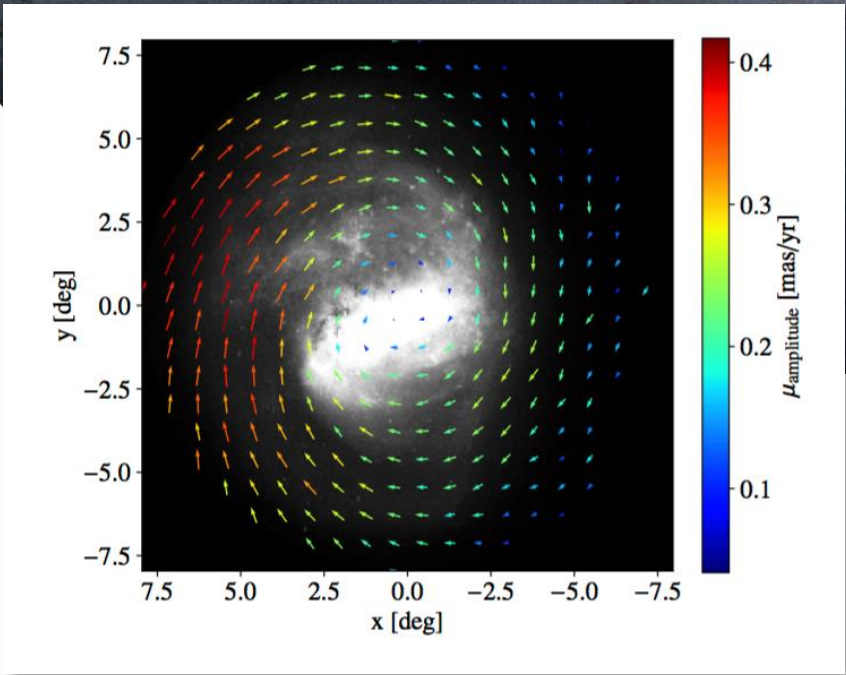
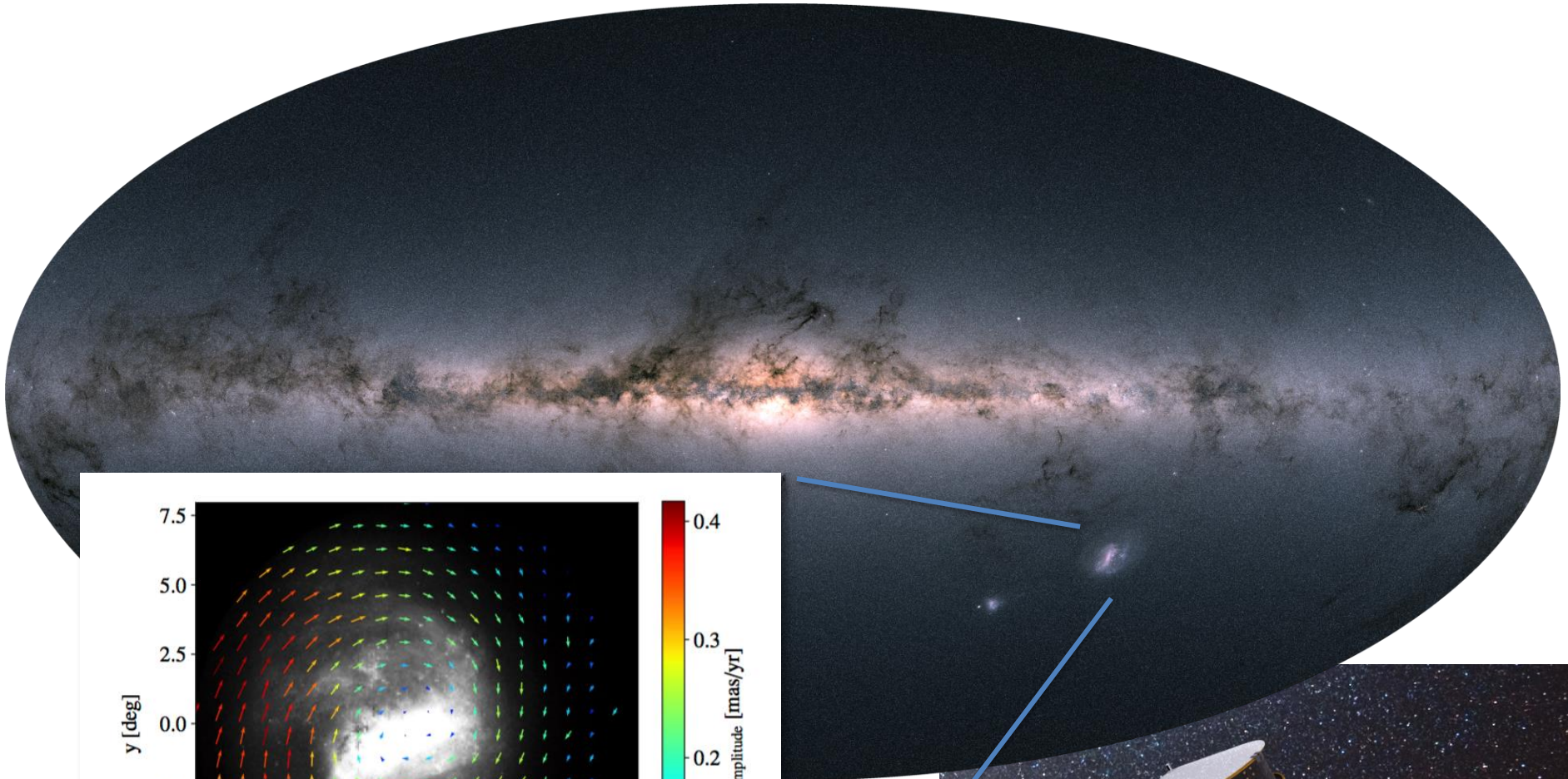


Figure courtesy Airbus DS / ESA



WFAU SW runs at processing centers around Europe

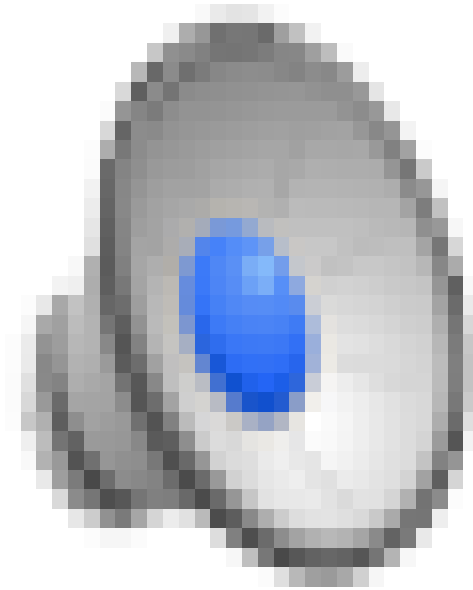




Helmi et al. (2018)

<https://arxiv.org/pdf/1804.09381.pdf>

Our locality in three dimensions ...



Acknowledgement: Gaia Data Processing and Analysis Consortium (DPAC); Gaia Sky;
S. Jordan / T. Sagristà, Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Germany

Merry
Christmas

