GENIUS

Legacy of the project

gaia

Gaia
DPAC
Data Processing & Analysis Consortium

WP2 User community

- Consolidated user requirements on the archive
- A good start on the requirements on long term archiving of the data and software in view of potential future reprocessing.
- Requirements for the seamless interoperation between Gaia and data at other wavelengths. These will be used in the future to design the multi-wavelength interface.

WP3 System design

We take seriously our responsibility for dissemination of results and SW coming out of the R&D programme using well-established Digital Curation mechanisms (and avoiding ad-hoc, ephemeral web resources):

- Software and associated documentation in GitHub
- Mainstream electronic publishing (e.g. article in Astronomy & Computing)
- Participation and integration in the international Virtual Observatory (developments of which adhere to the first two points)
- Re-use of knowledge in other areas and projects (e.g. taking forward FP7 GENIUS experience into H2020 ASTERICS)

Examples from WP3

- ✓ GitHub resources (whole or in part):
  - https://github.com/stvoutsin/pyrothorn
  - https://github.com/stvoutsin/tap-autocomplete
  - https://github.com/stvoutsin/taplib
- ✓ IVOA contributions:
  - http://wiki.ivoa.net/internal/IVOA/InterOpMay2016-DAL/adql-20160509.pdf
- ✓ A&C paper (submitted end Jan 2017):
  - *"Use of Docker for deployment and testing of astronomy software"*
- ✓ Data Centres using GENIUS-developed software:
  - ESAC Science Data Centre
  - Wide Field Astronomy Unit (University of Edinburgh, UK)
  - In fact anywhere that uses IVOA infrastructure, e.g. ADQL parser

WP4: VO tools

TopCat will be maintained by M. Taylor at Bristol Univ. Continued collaboration with DPAC ensured.

VOSA and Clusterix will be maintained in the framework of the Spanish Virtual Observatory and the CAB.

- VOSA will provide access to future Gaia releases. In particular the information on distances will be fundamental for the determination of masses and ages of very different types of astronomical sources.

- Clusterix will also include the future Gaia releases. In this case, the information on radial velocities, proper motions and distances will be key to assess the membership probability in known cluster as well asto discover new clusters and associations.

WP4: Data mining

The DM work developed during GENIUS will continue in many ways:

The current approach when working with large datasets is no longer valid due to the size of the data to be analyzed. In GENIUS we have explored a new way to work with huge data volumes based on state-of-the-art technologies for Big Data, mainly Spark and its surrounding services.

In a near future, Gaia archive will implement a similar (if not the same) platform such as GDAF. The platform is highly scalable so it could be expanded to different mission archives beyond Gaia, such as Euclid.

Other working packages will use similar technologies for their current computational processes such as data validation.

Similar platforms will be used not only for astronomy but for many scientific data processing developments, such a particle physics, where the data size and computing needs will require new methodologies to work with large datasets.

WP4: Visualization

The visualisation portal for the Gaia archive has been developed from scratch with support from GENIUS. It is currently a unique service providing and environment for interactive visual exploration of a data-set with over one billion objects with many dimensions. We know of no other service currently providing this in a scientific context. The service is scalable, having been able to handle over 7500 users in the afternoon of the day of DR1. This was made possible by GENIUS.

There is also a legacy here: The current service is expansible. It will be the platform that will host increased functionalities for future data releases (e.g. 3D, data overlays and more). The system has captured the interest of other collaborations and may be adopted by other missions and surveys.

This Gaia source density map has become the iconic image of Gaia Data Release 1 and was made possible by the GENIUS support. It was a worldwide success featured in all important media. It also has a legacy: It has inspired new ideas for visualising the next data releases.

Finally, the GENIUS funding was crucial for consolidating the Lisbon team as a group recognised by the quality of their work in scientific visualisation.

WP5: Validation

The development of a full validation code
- 130 000 source lines of code
- Plus the infrastructure for the code execution
- Being reused for DR2 and following data releases

Tools trained on validation then scientific exploitation
- E.g. Kullback-Leibler divergence→ 2017ApJ...836..234S

A published paper (A&A 599, A50):
- Full validation of DR1, about 35p.

Training of new developers
- in an ESA project environment

WP6: support

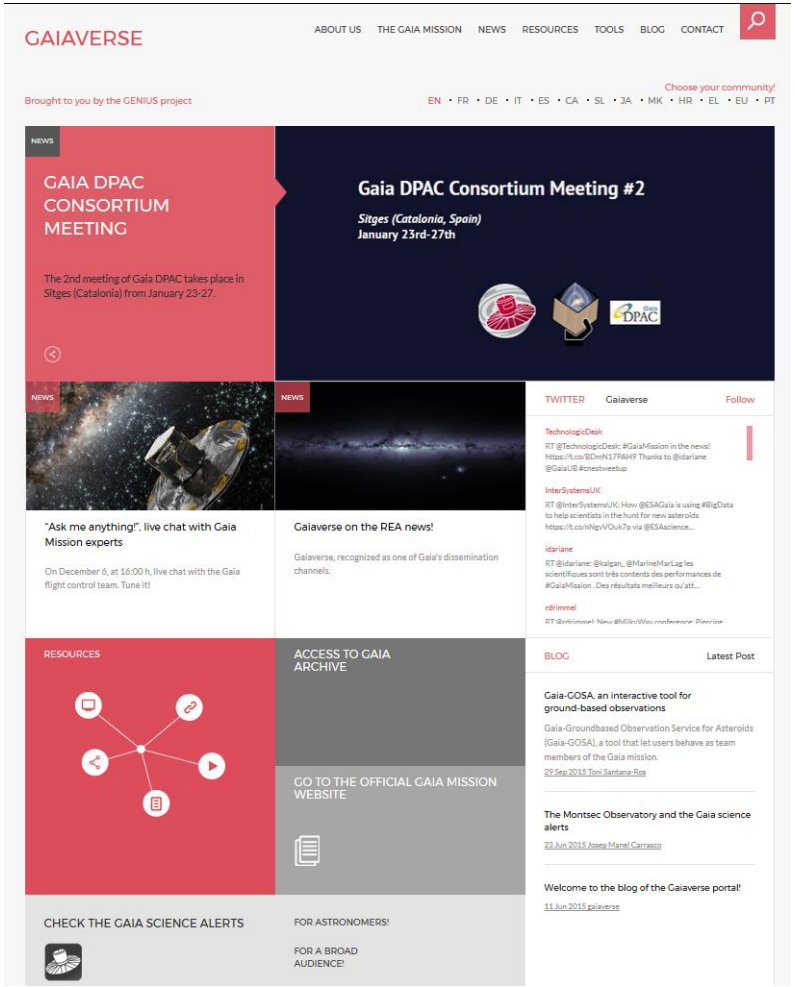
Several GOG simulations run. Simulations stored and available for validation and scientific usage.

Simulation environment installed at DPCB and available for future simulations.

Science Alerts system up and running. Will continue in operations in the coming years.

WP7: dissemination

- For the society:
  - Access to relevant news, resources and tools about the Gaia mission in many languages and in a single website.
- For the Gaia community:
  - Easy to use and maintain website based on open source Content Management System (CMS), WordPress.
  - Self-Management, no dependency on a third party
  - Flexibility and autonomy for each language



Conclusions

GENIUS has contributed in many ways to the design and implementation of the Gaia archive.

These contributions are now an integral part of the archive and will both contribute to the exploitation of the Gaia data and pave the way for future archive enhancements.