GENIUS
Final Report
for WP2

gaia

Gaia
DPAC
Data Processing & Analysis Consortium

Anthony G.A. Brown
Leiden Observatory, Leiden University

WP2: Tailoring to the end user community

- Gaia will provide an unprecedented stereoscopic map of our Milky Way and the nearby universe
  - >1 billion stars, 350000 solar system objects, 500000 quasars, 1–10 million galaxies, tens of thousands of exoplanets
  - For nominal 5-year mission catalogue and archive ‘finished’ in ~2022
- Will be key astronomical archive for decades to come
  - Tremendous discovery potential when combined with other archives

WP2: Tailoring to the end user community

- Research and invest effort in:
  - Taking into account user requirements (T2.2)
  - Confronting complex models with a complex catalogue (T2.3)
    - explore the ‘bring processing to the data’ concept
  - Seamless inter-operation with other data archives, in particular across wavelength domains (T2.4)
  - Can we facilitate future reprocessing? (T2.5)
    - Requirements on preservation of raw data, calibration data, and processing software
  - Explore concept of ‘living archive’ (T2.6)

GENIUS WP2 contributors

- Leiden: T2.1 (Brown), T2.2/2.5/2.6 (Costigan), T2.3 (Hypki), T2.2 (Massari)
- l UCAM: T2.2 (Lead: Walton)
- l KU: T2.2 (Lead: Yamada)
- l FFCUL: T2.2 (Lead: Moitinho)
- l INAF: T2.4 (Leads: Smart, Marrese)

- Technical coordination was done through the regular coordination mechanisms in CU9
- Costigan/Hypki interfaced to astronomical community by participation in conferences/workshops
  - Presentation of GENIUS and its goals
  - Ask Gaia catalogue users about their requirements

T2.2 Requirements Gathering

- UL/UCAM contribution
  - Organized the beta-testing of the archive for Gaia DR1
  - Science alerts requirements analysis
  - Requirement for outreach facilities in archive system
  - Supported GAVIP developments
  - Input received through http://great.ast.cam.ac.uk/Greatwiki/GaiaDataAccess used in assessing usage scenarios, ranking these and checking Gaia archive performance against the requirements
- KU contribution
  - Academic outreach for Gaia in Japan
  - Support for the development of the Gaia archive mirror in Japan (http://jvo.nao.ac.jp/portal/gaia.do)
- FFCUL contribution
  - Requirements for Gaia DR1 archive visualization services

T2.2 Requirements Gathering: Conclusions

- Requirements analysis and updates uncovered no major missing user requirements
  - Will continue to keep an eye on development of user requirements for future releases, e.g., through interaction with users at the various Gaia data workshops
  - Beta-testing got going a bit late for Gaia DR1, but infrastructure now in place for timely testing of the GaiaDR2 archive and its facilities

Legacy

- Consolidated list of user requirements for Gaia archive
- Beta-testing infrastructure in place for future Gaia DRs

T2.3 Confronting complex models with complex catalogues

- UL contribution
  - Proposal for API to interact with Gaia data in sophisticated ways
  - Upload simulations/models to user space
  - Upload code to carry out data analysis or model-catalogue comparison
  - Query archive from within code
  - Save and download results
  - Share data and code with other users
- Implementation proposed to be through thick-server/thin-client approach following REST approach
- Documented in GAIA-C9-TN-LEI-HYP-001/002

T2.3 Confronting complex models with complex catalogues: Conclusions

- This task did not lead to satisfactory results
  - Lack of interaction with other GENIUS WPs, in particular WP3/4
  - Led to vague general requirements from which no real progress can be made
  - Development of concrete prototype would have been a more fruitful approach

Lesson learned

- More effort in WP4 (work on concrete model-catalogue comparison case), followed by derivation of specific requirements on archive would have been better

T2.4 Seamless data retrieval across archives and wavelength domains

- INAF contribution
  - Requirements for multi-wavelength cross-match facility
    - Census of catalogues to cross-match against Gaia: radio, sub-mm, infrared, X-ray, γ-ray
    - Developed algorithms for multi-wavelength cross-match
    - Prototyped web interface
  - Implementation to be in form of
    - Pre-computed matches to large archives
    - X-match algorithms for x-matching of smaller data sets (millions of objects) from user-provided samples
    - Available through web portal from Gaia-DR3 onward
  - Documented in series of tech-notes by Marrese and Fabrizio

T2.4 Seamless data retrieval across archives and wavelength domains: Conclusions

- List of archives to cross-match against Gaia data available
- Clear proposal for the x-match across wavelength domains
- Path to making the service available identified

Legacy

- Pre-computed cross-matches with large astronomical catalogues
  - already well-received by community for Gaia DR1
- State-of-the-art cross-match algorithms

T2.5 The living archive

- UL Contribution
  - Investigated ideas and requirements for implementing living archive concept
    - Including survey of existing ‘living’ archives
  - Documented in GAIA-C9-TN-LEI-GCO-003

T2.5 The living archive: Conclusions

- Critical areas to keep in mind if this idea would be pursued for Gaia archive
  - Restricting scope (no duplication of efforts already existing in other archives)
    - Focus specifically on improving over Gaia-only results
  - Quality control of user-provided data
  - Ease of process to add data

Legacy

- Drop this proposal from consideration for Gaia archive functionalities
  - Creating and sharing of user tables is possible
  - VO-space interface also available

T2.6 Re-processing of archived (raw) data

- UL contribution
- All DPAC CU leaders were invited to provide input
  - which data should be preserved as a starting point for re-processing
  - which software (processing and simulation) should be preserved
  - which calibration data should be preserved
- Preservation of documentation also considered

T2.6 Re-processing of archived (raw) data: Conclusions

- Document (GAIA-CU9-TN-LEI-GCO-006) available that summarizes all inputs received
- Will form basis for further work on this topic by DPACE chair and Gaia Mission Manager

Legacy

- Good starting point for discussions and implementation of long term Gaia data and software preservation