WP-200, tailoring to the end user community

Anthony Brown

Sterrewacht Leiden brown@strw.leidenuniv.nl

- Motivation for WP-200
- Sub-WP descriptions



Future proofing the Gaia archive and catalogue

- Gaia will provide an unprecedented stereoscopic map of our Milky Way and the nearby universe
 - ▶ 1 billion stars, 300 000 solar system objects, millions of galaxies, 500 000 quasars, 10 000 exo-planets, ...
 - catalogue 'finished' in ~ 2022
- It will be *the* astronomical data archive for decades to come
 - tremendous discovery potential when combined with other archives
- Research and invest effort in the following:
 - take into account user requirements (WP-220)
 - confronting complex models with a complex catalogue (WP-230)
 - seamless inter-operation with other data archives and catalogues (WP-240)
 - facilitate reprocessing (WP-250)
 - keep raw data, calibration data, and processing software available
 - ▶ make the archive 'live' (WP-260)
 - bring the processing to the data

WP-220: Analysis and working out of requirements

Gaia data base and archive

- General public oriented with advanced features for professionals
- Design driven by user requirements
- Allow arbitrary queries
 - with option to account for star-to-star covariances
- Data mining
- Visualisation (data space is > 10-dimensional!)
- Transparent integration with other catalogues or sky surveys
- Access with variety of 'devices'



WP-220: Analysis and working out of requirements

Requirements gathering process

- Ideas have been collected through GREAT wiki pages
 - http://great.ast.cam.ac.uk/Greatwiki/GaiaDataAccess
- Many ideas collected already
- Reviewed and ranked according to
 - Urgency, generality, science impact, scale, expected frequency
 - ranking will be used to prioritize CU9 efforts
- All of this is summarized in a publicly available document (Gaia web site → Publications → Public DPAC documents)
 - http://www.rssd.esa.int/doc_fetch.php?id=3125400
- New ideas and use cases are still welcome

WP-220: Analysis and working out of requirements

From use cases to requirements and specifications

UL 9SM Analyze data access scenarios and turn into precise requirements on and specification for the Gaia archive

- Input to WPs 300/400/500
- include analysis of outreach needs

FFCUL 2SM Analysis of use cases related to visualization

- KU Analysis of use requirements specific to Nano-JASMINE data archive
 - in particular look at combined Nano-JASMINE/Hipparcos catalogue
 - exchange lessons learned
- UCAM Update requirements (use cases) gathering process and analyse requirements specific to 'science alerts'

WP-230: confronting complex models with complex catalogues

• Test Milky Way models against entire Gaia catalogue

- Gaia catalogue highly constraining for models
- model must explain data across all stellar populations over a large fraction of the Galaxy's volume
- This is not easy
 - large amount of data
 - large range in observational errors
 - correlations between errors on different quantities and between sources
 - complex (non-linear) relations between measured quantities and natural model parameters
- Rely on 'forward modelling'
 - good model provides correct 'predicted catalogue'

WP-230: confronting complex models with complex catalogues

Research into following issues (UL 12 SM)

- Can tools be provided to facilitate the projection of models into the Gaia data space?
 - Galactic distribution functions \longrightarrow Gaia observables
 - take into account the knowledge of the Gaia instrument
 - build on CU2 experience
- Can tools be provided to facilitate the model-catalogue comparison?
 - example: likelihood generators that know about the Gaia error properties, including correlations
 - optimization tools
- Contact the scientific community on their requirements
- Work out practical implications

WP-240: Seamless data retrieval across archives



WP-240: Seamless data retrieval across archives

Research and development of: (INAF 16 SM)

- Advanced inter-operation of data archives
 - beyond simple 'cross-matching'
 - across wavelength domains
- Provide user with seamless data retrieval options
 - as if working in a single archive
 - work out requirements and implementation options
- Build on existing VO tools?
- Results feed into archive design, data exploitation, and data validation

WP-250: The living archive

Gaia catalogue and archive released in ~ 2022 will not be 'final'



Photo Courtesy of the Isaac Newton Group of Telescopes, La Palma

- Updates should be allowed so as to incorporate:
 - updated classification or parametrization of stars
 - better distance estimates for faint stars
 - ground-based follow-up observations
 - independent information on, e.g., double stars
- Implications for maintenance, quality, security, keeping mirrors in sync





WP-250: The living archive

Research and development of: (UL 12 SM)

- Incorporation of new information in controlled manner
 - quality control
- Tracing the catalogue/archive history of a given source
 - track updates in source parameters, classification, multiplicity, etc
- Publication and dissemination of updates
 - keeping archives in sync

WP-260: Re-processing of archived data

Hipparcos reprocessing (van Leeuwen & Fantino 2005) 2 0.5 Precision 0.2 Brightness 0.1 10 5

Re-processed Hipparcos catalogue following better understanding and modelling of S/C attitude

Science goals

- Raw data reprocessing based on better algorithms, better calibration models etc
- Alternative processing of specific stars, groups of stars, or even entire catalogue
- Reprocessing data based on new and independent information

WP-260: Re-processing of archived data

Research and development of: (UL 12 SM)

Data curation

- All raw data
- Calibration data and models
- Intermediate data products
- Processing software curation
 - including documentation of algorithms and code
- Facilitating the use of raw or intermediate data
- Tracking versions of processing results