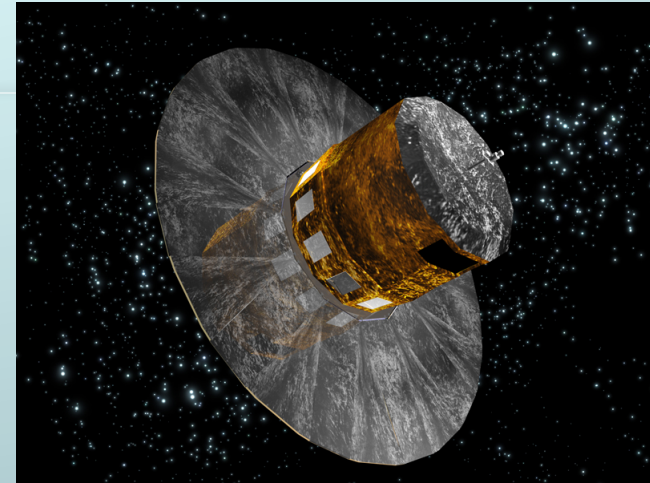


ASDC: l'accesso ai dati



ASDC Gaia Team: G. Giuffrida, S. Marinoni, P.M. Marrese
R. Buonanno (Senior Scientist)

Outline:

- *What is meant by data access*
- *Gaia Archive Data releases*
- *GAP/CU9: DPAC environment*
- *GAP/CU9: Coordinated italian participation*
- *Coordinated italian participation to ESA AO Response*
- *An example: Multi-wavelength cross-match with external archives*
- *Auxiliary data in CU9*
- *Gaia Auxiliary data in Italy*

=

Data distribution and support to scientific analysis

Gaia data will be immediately available to the scientific community →

no proprietary data rights for Gaia

Hereafter: Gaia data =
Gaia catalogue data, not
including raw data

Gaia data: very **large, rich and complex** in content

(astrometry, photometry, spectroscopy, non single stars and solar system objects, variable stars, and higher level astrophysical parameters of Gaia objects, epoch data)

Complex data will answer complex astrophysical questions

The implementation and operation of the Gaia archive is a technological challenge and opportunity, but ...

archive realization must be driven by science considerations →

science requirements have been gathered from the astronomical community and we will continue to request new science cases during the mission

=

Data distribution and support to scientific analysis

Exposing Big Data to the Community → Access + explore + analyse

The *science services* provided to the astronomical community form a layer on top of the catalogue and archive providing :

- quick and easy **basic access** to the catalogue data
- **advanced science enabling tools**
- efficient inclusion of **complementary ground- or space-based surveys**

Advanced science enabling tools:

- Visualization
- Data analysis
- Statistics
- Data mining



Astroinformatics

Launch date **October 2013** (possible opportunity in September)

Baseline operational period: **5 years** (possible 1–2 year operational extension)

Final Gaia post mission analysis will complete within 2–3 years after Gaia ceases to operate (**2021-2023**)

~ 6 Months for cruise to L2, commissioning, DPAC systems initialization

First full sky coverage after 6 months of nominal scanning

Disentangling parallaxes and proper motions requires at least 18 Months of data collection

Processing, calibration, validation take time

Each data release requires time to go from DPAC internal database to public archive (3 Months)

Full data releases from late 2015 with a number of major releases foreseen at regular points thereafter

Following schedule assumes smooth operations

All values prior to final release may be truncated at some confidence level

Each release updates the previous and contains significant new additions

L=Launch (9-10/2013)

L+22M Positions + G magnitude (~ all sky, single source)

- Includes more often scanned Ecliptic pole regions
- Hundred Thousand Proper Motions (Hipparcos-Gaia, ~ 50 μ as/yr)

L+28M more tentative some radial velocities for bright stars + BP/RP integrated

L+40M full astrometry (α , δ , π , μ_α , μ_δ), orbital solutions, ($G_{BP} - G_{RP}$), some BP/RP Spectrophotometry and astrophysical parameters, RVS spectra

L+65M Updates on previous release — including more sources, source classifications, multiple astrophysical parameters, variable star solutions and epoch photometry for them, solar system results

End+3yr Everything

Gaia data processing

Gaia Data Processing :

- **ESA** Announcement of Opportunity for the Gaia Data Processing
06/11/2006
- **DPAC** (Data Processing and Analysis Consortium) Proposal for the Gaia
Data Processing 04/2007

- **DPAC:**

DPAC is a consortium of over **400 scientists and software engineers** who are organizing themselves to process the data that will arrive from Gaia satellite.

The Consortium as a whole is coordinated by the DPAC Executive (**DPACE**)

DPAC structure and activities CUs & WGs:

- CU1:** System Architecture
- CU2:** Data Simulations
- CU3:** Core Processing
- CU4:** Object Processing
- CU5:** Photometric Processing
- CU6:** Spectroscopic Processing
- CU7:** Variability Processing
- CU8:** Astrophysical Parameters
- CU9: Catalogue Access (to be activated)**

Responsible for preparation and access to the Gaia Catalogue (including early releases)

- WGs:** Calibration Task Force
- GBOG
- Offset Instability Task Force
- Radiation Task Force
- GAP**
- Ops Steering Group

Scientific & technological activities

Coordinated italian participation: scientists, archive scientists, IT engineers

WHAT: Italian Gaia scientific community =>

Past: preparation of answer to ESA call for GAP participation.

Response Letter => group of italian scientists coordinated by **ASDC**

Present: preparation of answer to ESA AO for CU9 participation

AO Response submitted 10 january 2013

Future: participation to CU9 activities

WHY: Final aim =>

allow the italian scientific community to fully exploit the Gaia data

HOW:

Gaia with no proprietary period => participating to the realization of the data reduction, access and analysis tools is a clear advantage and a strategic activity

STRATEGY:

GAP => CU9 => italian mirror + advanced applications => scientific exploitation

Coordinated italian participation:

Why ASDC?

- Institutional role of support to the scientific community – specifically for space mission data
- Expertise
 - Data center since 1996 (since 2000 as ASDC)
 - Experience in Archive Management
 - Multi-wavelength cross-match (radio to X-ray)
 - WEB based scientific analysis tool
 - On site IT engineers support (Telespazio)

Coordinated italian participation:

Andrea Baruffolo	INAF-OAPd
Roberto Buonanno	Univ.of Rome Tor Vergata, ASDC-INAF
Alberto Cellino	INAF-OATo
Paolo Giommi	ASDC-ASI
Giuliano Giuffrida	ASDC-INAF
Giuseppe Longo	Univ.of Naples, Federico II
Silvia Marinoni	ASDC-INAF
Paola Maria Marrese	ASDC-INAF
Luciano Nicastro	INAF-IASFBo
Riccardo Smareglia	INAF-OATs
Richard Smart	INAF-OATo
Alessandro Spagna	INAF-OATo
Alessandro Sozzetti	INAF-OATo
Antonella Vallenari	INAF-OAPd

GBOG contribution:

Elena Pancino	INAF-OABo
----------------------	-----------

Gaia Archive Preparation => CU9:

Start CU9 specific activities so that DPAC view on Catalogue access could be studied and that costs could be evaluated before CU9 activation

Gaia Archive Preparation

- Informal starting of activities (as a DPAC wg)
- **ESA Call for Letters of Interest to participate in the preparatory activities for the Gaia archive access 01/03/2011**
- Several response Letter, accepted by ESA 04/2011
- **GAP structure and activities:**
 - WP1: Management and Coordination - the Plan and WBS**
 - WP2: Requirements Gathering - Consolidate requirements**
 - WP3: Architecture and Technical Development - Software and Hardware for the Archive**
 - WP4: Data Validation and Verification. Global catalogue validation**
 - WP5: Data Documentation. Documentation of the Archive and contents**
 - WP6: Operations and Support (how will the archive run).**
 - WP7: Outreach and academics.**

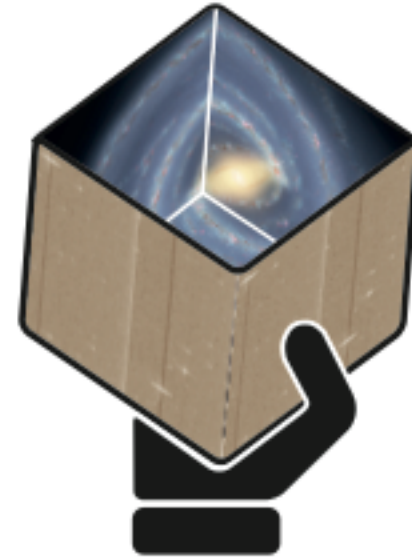
ESA AO for CU9 participation
(issued 19 November 2012)

AO Response
(submitted 10 January 2013)

CU9 kickoff meeting
(april/may 2013)

Num	Title
WP910	Management
WP920	Documentation
WP930	Architecture, Design and Development
WP940	Validation
WP950	Operations and Services
WP960	Education and Outreach
WP970	Science Enabling Applications
WP980	Visualisation

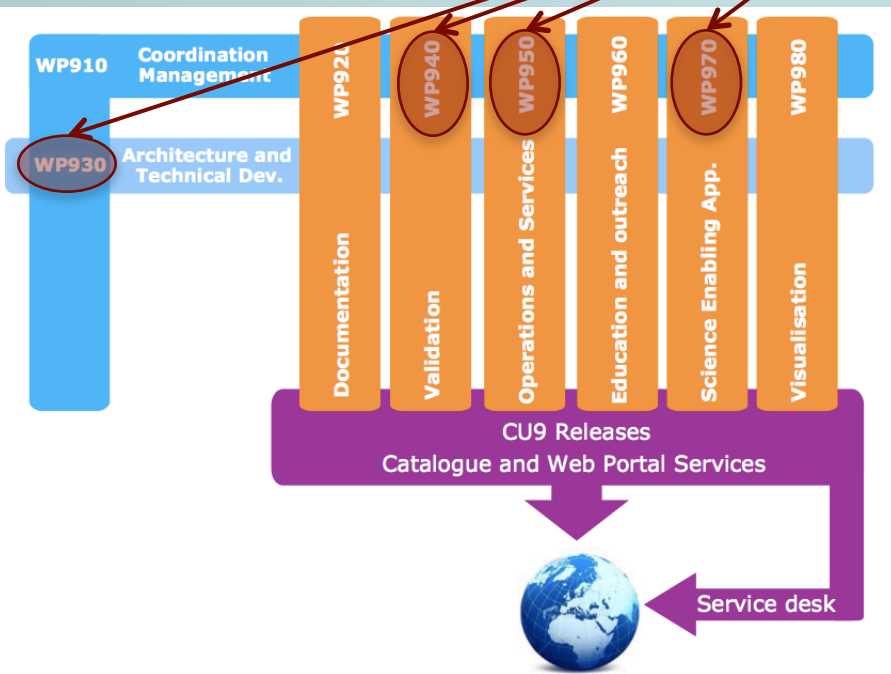
Roma, 14/02/2013



Delivering the promise of Gaia

**Response to ESA's Announcement of
Opportunity**

Proposal for the Gaia Archive



Institute	Effort (SY)
Univ. Barcelona	50.1
ESAC	45.2
IT(ASDC INAF Univ)	43.1
ARI	28.1
ObsPM	26.2
MPIA	17.0
Lisbon Univ.	14.1
IoA Cambridge	11.1
Univ. Coruna	10.5
Vienna	10.0
Geneva	9.7
OCA	8.0
INTA	7.5
AIP	7.0
CDS	6.8
Groningen	6.4
Innsbruck	5.0
ARI/HdA	4.9
IfA Edinburgh	4.0
Besançon	4.0
KU Leuven	3.0
Lund	3.0
UNINOVA	3.0
PO	2.2
Royal Obs Belgium	2.1
ESTEC	2.0
Groningen/Leiden	2.0

WP 930 Architecture, Design and Development

Gaia DPAC WP: GWP-T-934-0000

Title: Archive architecture: database collaboration

Provider(s): AIP, Univ. Edinburgh, ASDC-INAF

Gaia DPAC WP: GWP-T-935-0000

Title: Archive architecture: interface design (TAP+Server side)

Provider(s): AIP, ARI/ZAH, Univ. Edinburgh, ASDC-INAF

WP 940 Validation

Extrasolar planets, Open clusters, Solar system objects, all sky statistics to validate Gaia source list

Science from quality assured data

Gaia DPAC WP: GWP-T-943-0000

Title: **Validation: models**

Provider(s): Observatoire de Besancon, Observatoire de Paris, University of Barcelona, **ASDC-INAF**, Observatoire de la Cote d'Azur

Gaia DPAC WP: GWP-T-944-0000

Title: **Validation: external archives**

Provider(s): Observatoire de Paris, **ASDC-INAF**, Observatoire de la Cote d'Azur, INTA-Istituto Nacional de Tecnica Aeroespacial, CDS-Centre de donnees astronomiques de Strasbourg

Gaia DPAC WP: GWP-T-945-0000

Title: **Validation: statistical tools**

Provider(s): Observatoire de Paris, University of Barcelona, **ASDC-INAF**, University of Groningen, University of Corunna

WP 950 Operations and Services

Gaia DPAC WP: GWP-T-952-2000

Title: Operations: services operations

Provider(s): ARI/ZAH, **ASDC INAF Univ**, CDS,
Groningen, CIESCA

Gaia DPAC WP: GWP-T-954-0000

Title: Operations: service monitoring and feedback

Provider(s): **ASDC INAF Univ**, CDS

Gaia DPAC WP: GWP-T-957-0000

Title: Operations: auxiliary data

Provider(s): **INAF-OABO**

WP 970 Science Enabling Applications

Gaia DPAC WP: GWP-T-973-0000

Title: Science enabling applications: data mining

Provider(s): Univ. Barcelona, **ASDC-INAF**, AIP, Univ. Coruna, KU Leuven, INTA

Gaia DPAC WP: GWP-T-974-0000

Title: Science enabling applications: auxiliary and external data

Provider(s): **ASDC-INAF, INAF-OABo**, Laboratoire d'Astrophysique de Bordeaux, MPIA, INTA, Univ. Coruna, CDS

Multi-wavelength cross-match with external archives

- Science case driven first ranking and classification is provided in AB-026
- Minimum requirement nearest-neighbour+list of neighbours, but also with figure of merit (not only position, but density, magnitude, ...)
- Pre-computed and integrated in Gaia query system
- Multi-wavelength (different problems ---> different tools)
- Tools made available to final user

Why pre-computed and integrated

GDAS-GA-20

I would like to make selections of Gaia objects based not only on Gaia data, but also based on other major catalogues available at that time, such as Pan-STARRS and SDSS

GDAS-ST-15

One wish would be to have more than only cross-matches to external catalogs but also the data values, so I assume there is a link to e.g. WISE or VISTA Hemisphere survey, say a K-band magnitude.

Query: return the epoch radial velocities of all stars with $ABS(a \log P + b - K\text{-mag}) < \sigma$, where a, b, σ are user-supplied, K-mag would come from the cross-matched external catalog and P is the Gaia-derived period.

Why tools available to final user

GDAS-BR-3

I have a list of targets of interest from SDSS/UKIDSS/VISTA/PanSTARRS/whatever. For each one, give me the astrometric parameters (position at my specified equinox and epoch, proper motion, parallax and errors/quality statistics indicating the accuracy and significance of the data)

- GBOG (Ground-Based Observations for Gaia) is a DPAC WG chaired by C.Soubiran (deputy E.Pancino)
- Mission: to coordinate and facilitate the effort of gathering and preserving ground-based data in support of Gaia (GBOG Terms of Reference CS-004)
- **GBOG in CU9: preserve Gaia-DPAC auxiliary data and to publish them, when necessary and appropriate, together with Gaia data releases**
- Auxiliary data within Gaia DPAC can be:
 - New ground-based data obtained to calibrate Gaia or to train its algorithms
 - New critical compilations of literature data (or hybrids with new data)
 - New models or synthetic data-sets (or hybrids with the above)
- A first census has just been completed (EP-009): ``A census of potential GBOG auxiliary data products for GAP/CU9'' by E. Pancino et al., 2013, GAIA-CD-TN-OABO-EP-009.

ASDC Support activity to Gaia DPAC

Archive, management and access

- Spectro-photometric data gathered by OABo for the photometric external calibration (CU5)
- Synthetic spectral library calculated by OAPd for the astrophysical characterization of Gaia targets (CU8)

ASDC-SPSS archive&database (final size 500-600 GB):

Homogeneization of data, collection and reconstruction of metadata, web interface

- **Raw data**: calibration, science, tests for instruments characterization, ~ 10^5 frames in 410 nights
- **Reduced data** (various intermediate data reduction levels for both photometry and spectroscopy, ~ 1.5×10^5 frames expected)
- **Data products** (photometric catalogues, short and long term light curves, absolute photometry, night quality parameters, synthetic magnitudes, flux tables)
- **Auxiliary data** (filters and grisms characterization, reference data)

Foreseen Releases:

- **DPAC only** : dynamic, CU5-DU13 working space
- **Public** : static, one for each Gaia Releases