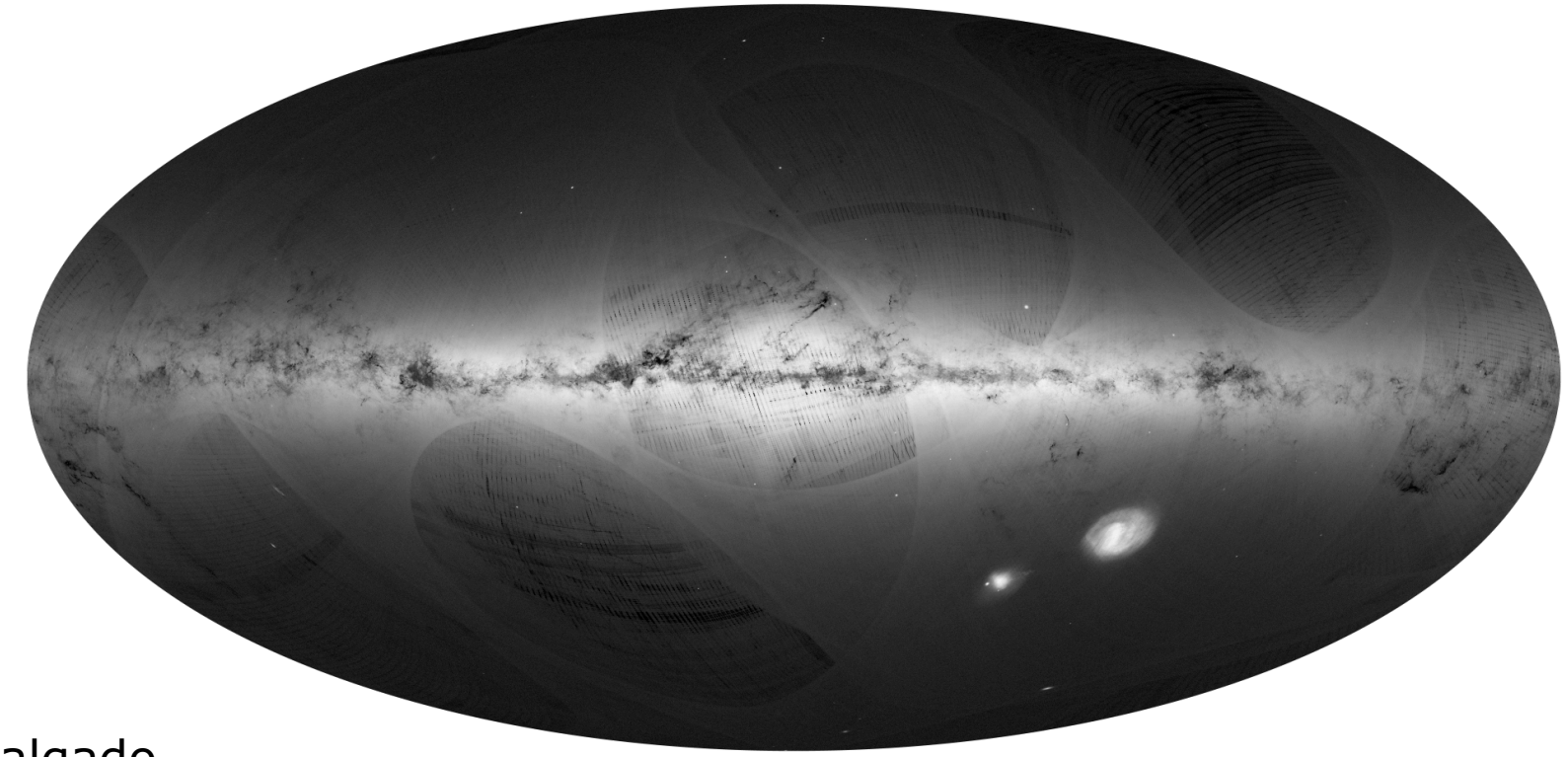
Gaia Archive for Data Release 2: Access to Time Series



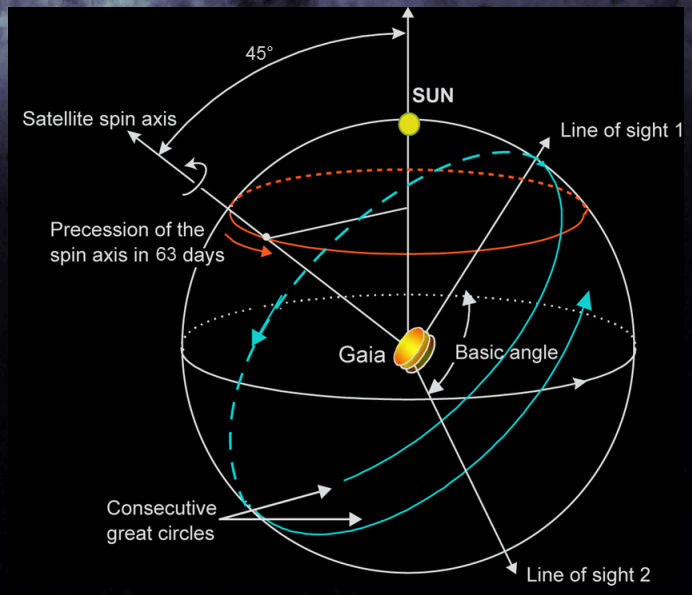
J. Salgado

J. González-Núñez, R. Gutiérrez-Sánchez, J.C. Segovia, J. Durán, A.Mora, B. Merín, C. Arviset

24/10/2017

Gaia operations

- Gaia in routine operations since July 2014
- Scanning operations with observing strategy of continuous measuring
- Dead-time: orbit maintenance, micrometeoroids, decontaminations, ground station weather
- Nominal 5-year mission ends mid-2019
- Estimated end of mission due to cold gas exhaustion end-2023 (±1year)



Gaia DR2 release

April 2018

- Five-parameter astrometric solutions for all sources with acceptable formal standard errors (>10^9 anticipated), and positions (α, δ) for sources for which parallaxes and proper motions cannot be derived.
- G and integrated GBP and GRP photometric fluxes and magnitudes for all sources
- Median radial velocities for sources brighter than GRVS=12 mag
- For stars brighter than G=17 mag estimates of the effective temperature and, where possible, line-of-sight extinction will be provided, based on the above photometric data
- Photometric data for a sample of variable stars
- Epoch astrometry for a pre-selected list of >10,000 asteroids

Release Plan

- **DR2 (~2018):**

5 parameter astrometric solutions ~10^9 sources

- **DR3 (mid to late 2020):**

BP/RP/RVS spectra

- **DR5 (~2022):**

All epoch and transit data for all sources

Data Volumes

(Mid 2017)	Volume
Size	~200 TB
Size (estimation 5-year ext.)	1.4 PB
Number of objects (sources, transits…)	1000 billion
Number of objects (est. 5-year ext.)	6000 billion

Archive content @ DR1

Gaia

• Gaia DR1 catalogue	1.1×10^9	rows
• TGAS	2.0×10^6	rows

External Catalogues

• Hipparcos & Hipparcos new red.	1.2×10^6	rows
• IGSL (Initial Gaia Source List)	1.2×10^9	rows
• 2MASS	4.7×10^8	rows
• Tycho2	2.5×10^6	rows
• UCAC4	1.1×10^8	rows
• Hubble Source Catalogue v1.0	2.9×10^7	rows

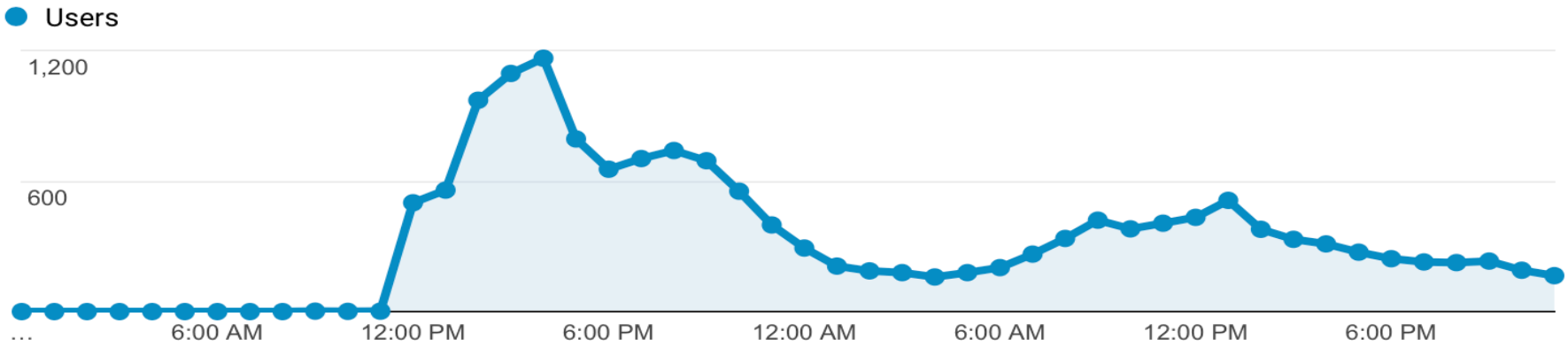
Crossmatches

- Crossmatch tables between Hipparcos, 2MASS, Tycho2… and Gaia expressed as neighbourhood and best neighbour, e.g:

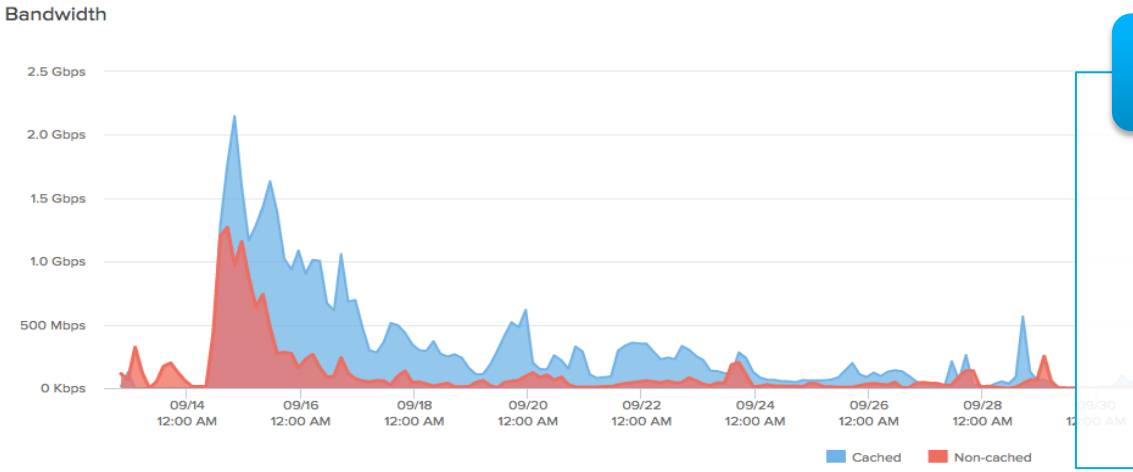
• AllWise-Gaia neighbourhood	3.1×10^8	rows

Ops for Gaia DR1

Archive UI usage: First 24 hours



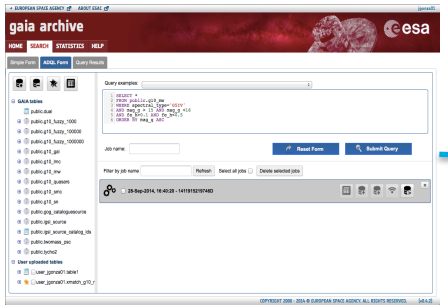
File CDN download: First 15 days



File download

- Total volume downloaded: **73 TB**
- # download requests: **9 millions**

DataLink/SSAP int. with TAP+



TAP+

DataLink

* DR2

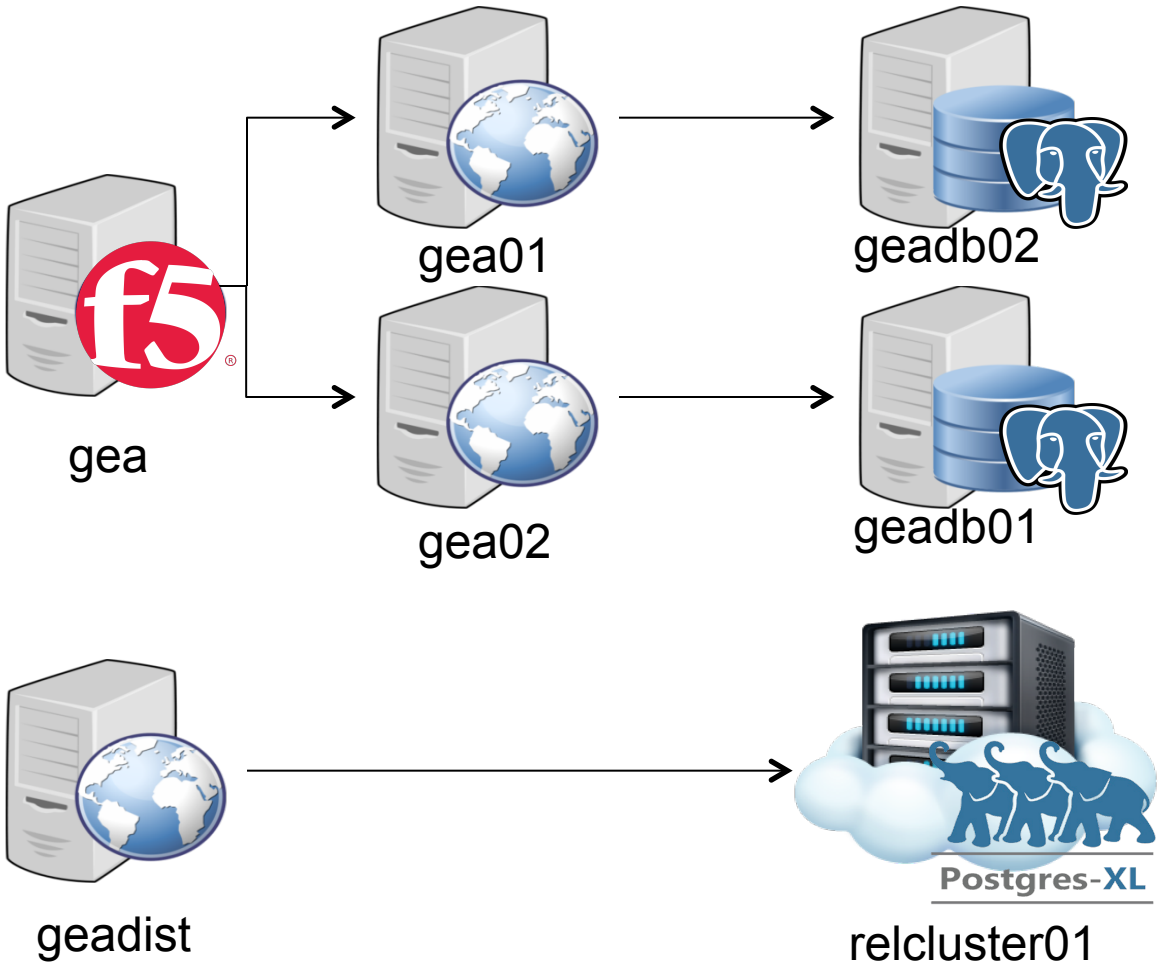
TAP+
Data Mining
SSAP?

* DR3

Catalogues

Time Series
Spectra

DR2 Architecture



DataLink in UI

1508331006595D

›n_strength_k1	scan_direction_strength_k2	scan_direction_strength_k3	scan_direction_strength_k4	solution_id	source_id	access_url	epoch_photometry_url
	0.6618591	0.9311092	0.5040036	1635655487712133120	2745170132975985280	Datalink	Open link
	1	1	1	1635655487712133120	3631112484733690880	Datalink	Open link
	1	1	1	1635655487712133120	5958736269568260992	Datalink	Open link
	0.46685013	0.48544493	0.90699345	1635655487712133120	4102394882451939328	Datalink	Open link
	1	1	1	1635655487712133120	4265328383883268224	Datalink	Open link
	0.17035252	0.49307513	0.91825473	1635655487712133120	4081668950814014848	Datalink	Open link
	0.67557305	0.9733705	0.36128184	1635655487712133120	5872417975329082880	Datalink	Open link
	0.54799855	0.56486565	0.763219	1635655487712133120	3085246365030241152	Datalink	Open link
	0.7591276	0.48959264	0.18620124	1635655487712133120	4586542493902615936	Datalink	Open link
	1	1	1	1635655487712133120	168240023674627968	Datalink	Open link
	0.36942342	0.5841497	0.88677555	1635655487712133120	4102436079781587584	Datalink	Open link
	0.51372313	0.023877695	0.47217703	1635655487712133120	5877527199762893568	Datalink	Open link
	0.76711327	0.45930234	0.6195379	1635655487712133120	260224891577522560	Datalink	Open link
	0.20396036	0.3596186	0.43628216	1635655487712133120	5875352949859083776	Datalink	Open link
	0.29051554	0.2945069	0.94581664	1635655487712133120	3425066720108469888	Datalink	Open link
	0.18453127	0.15257531	0.22539945	1635655487712133120	5206075716572505344	Datalink	Open link
	0.41418886	0.46120533	0.7954249	1635655487712133120	3776911604932087296	Datalink	Open link
	1	1	1	1635655487712133120	3983882887765462784	Datalink	Open link
	0.09667761	0.26300278	0.81443006	1635655487712133120	971228415004375808	Datalink	Open link
	0.25292978	0.5054329	0.90796155	1635655487712133120	4052834498901393792	Datalink	Open link

1-20 of 100 Gaia Data Model Show query in ADQL form

Time Series Serialization

- No format decided for Time Series Serialization
- Format proposed from ESDC to Gaia Consortium in line with ongoing note **Time Series Cube DM** (Jiri Nadvornik et al)
- Problems:
  - No agreement in how to make reference to DMs
    - **VO/DML ?**
  - **Quite preliminary note**
  - When characterization metadata is added it could be very verbose
  - Supported by applications?

Time Series Serialization (II)

```
<?xml version="1.0"?>
<VOTABLE version="1.1"
    xmlns="http://www.ivoa.net/xml/VOTable/v1.1"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
    <DESCRIPTION>This is the calibrated FoV transit photometry from CU5.  All instrumental effects should have been calibrated out and all transits should be on the same photometric
        system. This table is complemented by the FinalCalPhotFovTransit table which extends this, including additional CCD-level epoch photometry for all FoV transits whose sources
        show a high noise level. Transits that are selected to appear in the FinalCalPhotFovTransit table will not appear in the CalPhotFovTransit table. </DESCRIPTION>
    <PARAM arraysize="*" datatype="long" name="sourceId" ucd="meta.id;meta.main" value="2745170132975985280"/>
    <PARAM arraysize="*" datatype="long" name="solutionId" ucd="meta.version" value=""/>
    <PARAM arraysize="*" datatype="short" name="nTransits" ucd="meta.number" value="0"/>
    <RESOURCE>
        <TABLE>
            <DESCRIPTION>Transit Time Serie</DESCRIPTION>
            <GROUP name="ndcube:TimeSeriesCube">
                <GROUP name="independent_axes">
                    <GROUP name="dateTimeAxis">
                        <FIELDref ref="transitTime"/>
                    </GROUP>
                    <GROUP name="spectralAxis">
                        <FIELDref ref="bandName"/>
                    </GROUP>
                </GROUP>
                <GROUP name="dependent_axes">
                    <GROUP name="fluxAxis">
                        <FIELDref ref="transitFlux"/>
                    </GROUP>
                    <GROUP name="magnitudeAxis">
                        <FIELDref ref="transitMag"/>
                    </GROUP>
                </GROUP>
            </GROUP>
            <FIELD ID="bandName" datatype="char" name="bandName" ucd="instr.bandpass" utype="ssa:DataID.Bandpass"/>
            <FIELD ID="waveLength" datatype="double" name="Wavelength" ucd="em.wl.effective" unit="Angstrom" utype="photdm:PhotometryFilter.SpectralAxis.Coverage.Location.value"/>
            <FIELD ID="transCurveWavelengthMin" datatype="double" name="transCurveWavelengthMin" ucd="em.wl;stat.min" unit="Angstrom" utype="photdm:
                PhotometryFilter.SpectralAxis.Coverage.Bounds.start"/>
            <FIELD ID="transCurveWavelengthMax" datatype="double" name="transCurveWavelengthMax" ucd="em.wl;stat.max" unit="Angstrom" utype="photdm:
                PhotometryFilter.SpectralAxis.Coverage.Bounds.stop"/>
            <FIELD ID="transCurve" arraysize="*" datatype="char" name="transCurve" ucd="DATA_LINK" utype="photDM:PhotometryFilter.transmissionCurve.access.reference"/>
            <FIELD ID="transitId" datatype="long" name="transitId" ucd="meta.id" unit="" utype="">
                <DESCRIPTION>Transit Identifier</DESCRIPTION>
            </FIELD>
            <FIELD ID="transitTime" datatype="double" name="transitTime" ucd="VOX:Image_MJDateObs" unit="d" utype="spec:Spectrum.Data.TimeAxis.Value">
                <DESCRIPTION>Field-of-view transit averaged observation time in units of Barycentric JD (in TCB) in days - 2 455 197.5, computed as follows. First the observation time
                    is converted from On-board Mission Time (OBMT) into Julian date in TCB (Temps Coordonnee Barycentrique). Next a correction is applied for the light-travel time to
                    the Solar system barycentre, resulting in Barycentric Julian Date (BJD). Finally, an offset of 2 455 197.5 days is applied (corresponding to a reference time T0 at
                    2010-01-01T00:00:00) to have a conveniently small numerical value.  Although the centroiding time accuracy of the individual CCD observations is (much) below 1 ms,
                    this per-FoV observation time is averaged over typically 9 CCD observations taken in a time range of about 44 sec.</DESCRIPTION>
            </FIELD>
            <FIELD ID="transitFlux" datatype="float" name="transitFlux" ucd="em.opt;phot.flux;stat.mean" unit="e-/s" utype="spec:Spectrum.Data.SpectralAxis.Value">
                <DESCRIPTION>The average G flux value for the FoV transit. The calculation only uses accepted transits. This could include SM and AF fluxes.</DESCRIPTION>
            </FIELD>
            <FIELD ID="transitFluxError" datatype="float" name="transitFluxError" ucd="em.opt.B;phot.flux;stat.error" unit="e-/s">
                <DESCRIPTION>The error on the mean G Flux.</DESCRIPTION>
            </FIELD>
            <FIELD ID="transitMag" datatype="float" name="transitMag" ucd="phot.mag" unit="mag">
                <DESCRIPTION>G-band mean magnitude for the field-of-view transit, computed from the fluxGFov field using magnitude zero-point defined in ExtPhotZeroPoint.</DESCRIPTION>
            </FIELD>
            <FIELD ID="photometryFlagNoiseData" datatype="boolean" name="photometryFlagNoiseData" ucd="meta.code.status" unit="">
                <DESCRIPTION>G band flux scatter larger than expected by photometry processing (all CCDs considered).</DESCRIPTION>
            </FIELD>
```

Time Series Summary

1. Independent Axes
    a. Time (several times depending on magnitude but aggregated in one column)
    b. Spectral Coordinate (depending on Consortium decision)
    c. Possibly Source identifier

2. Dependent Axes
    a. Flux
    b. Magnitude
    c. Errors

3. Other possible metadata
    a. Band characterization

Connection to Data Models

- No general data model for IVOA Time Series
- Fluxes are in e-/s. It would be better to have measurements in physical units that could be compared easily with other catalogues.
- Magnitude zero points: Zero Points provided by the consortium are corrections from the instrumental response
  - Proposed to connect to Filter Profile Service, Transmission curves and Vega zero points calculated by SVO
  - ***GAIA BP 3478.8 Jy***
  - ***GAIA G 2861.3 Jy***
  - ***GAIA0 RP 2461.2 Jy***
  - SVO has been requested to provide zero points in erg/cm^2/s/A
- Also, connections to SpectralDM
- Source DM?

But….. VO/DML? Utypes mapping?

Gaia DM Array based

Adaptation of Gaia TAP+ to support arrays:

- SELECT get_double_array_element('{{1.0,2.6},{0.8,0.1}}','[1][2]') from public.dual
  - Different function names for different data types
  - "Not nice" second argument

Possible evolution:

- SELECT get_array_element('{{1.0,2.6},{0.8,0.1}}',1,2) from public.dual
  - Type casting can be done at function level (possible loose of control of the casting)
  - Currently, only function overload can be limited to only 2 new arguments but it does not look very scalable
- Also,… Gaia DM is quite based on arrays for epoch data (one array for times and other for magnitude/fluxes)
  - Correspondence is done by array element order
  - Probably, not very optimal for data mining

Conclusions

- Epoch data is arriving for quite important missions
- Not clear VO format for, e.g., Time Series
- Some important aspects on VO compatibility should be addressed fast:
  - VO time series serialization
  - Approaches to consume these data through VO protocols (e.g. Time Series Access and/or ADQL/TAP evolution)
  - Agreed mapping of complex data models (VO/DML mapping?)
  - Relevant data models

- Promote notes or endorse formats/protocols could be a solution but normal speed of standardization needs a boost in this case

Questions?

http://archives.esac.esa.int/gaia/

THANKS!

Gaia DPAC Data Processing & Analysis Consortium

QUASAR SCIENCE RESOURCES, S.L.

gaia European Space Agency