Source list evolution
in the Gaia data releases

Jordi Portell

on behalf of the Gaia CU3/DPCB teams

Gaia "source identifiers"

- NOT unique/maintained across different Data Releases!
- ESA/Gaia announcement prior to DR2:

Users of Gaia data should be aware that the source list for Gaia DR2 should be treated as independent from Gaia DR1. Although the majority of sources in Gaia DR1 can be identified with the same source in Gaia DR2 through the Gaia source identifier, the improved source list will lead to the following changes in linking the observations to the source identifiers for a substantial fraction of sources in this list:

- The merging of groups of observations previously linked to more than one source will lead to a new source associated to the merged observations (with a new source identifier) and the disappearance of the original sources (along with their source identifiers).
- The splitting of groups of observations previously linked to one source will lead to new sources associated to the split groups of observations (with new source identifiers) and the disappearance of the original source (along with its source identifier).
- The list of observations linked to a source may change (and hence the source characteristics may change), while the source identifier remains the same.

A means to trace sources across the Gaia DR2 and Gaia DR1 source lists will be provided.

More on Gaia Source Identifiers

- From the official Gaia DR2 documentation:

GAIA SOURCE IDENTIFIERS

Sources in the Gaia Catalogue are all identified through the Gaia Source Identifier, i.e. the `source_id` field in the various tables in the Gaia Archive. The construction of the source identifiers is explained in the archive documentation (for Gaia DR1, see the data model section). In particular the `source_id` number contains rough information about the source position on the sky.

As explained in the announcement about the Gaia DR2 schedule, there are various reasons why the identifier of a specific source may change or disappear when going from the Gaia DR1 to the Gaia DR2 source list. Users of Gaia data should thus be aware that the source list for Gaia DR2 should be treated as independent from Gaia DR1.

For the same reason, Gaia source names have the data release number incorporated. Gaia source names are therefore constructed as follows:

```
Gaia DRx yyy....yy
```

Resulting in an example source ID: Gaia DR1 5425628298649940608 for some source in Gaia DR1.

A table to trace sources from Gaia DR1 to Gaia DR2 is provided given that source identifiers used in Gaia DR1 may have changed or disappeared in Gaia DR2. In future releases, the source list is expected to become progressively more stable.

Sources can appear, disappear and change!

Source merging:

Initial situation   Clustering and source linkage   Final resolution

"Gaia DR1 XXX", "Gaia DR1 YYY"

→ New "Gaia DR2 ZZZ"

→ Source IDs XXX and YYY do not exist anymore
(there is no "Gaia DR2 XXX" or "Gaia DR2 YYY")

Sources can appear, disappear and change!
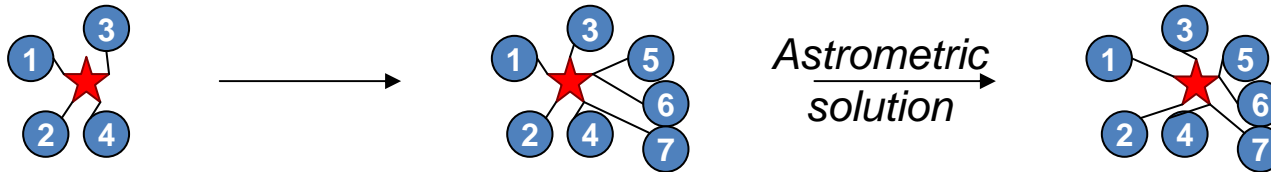
Source splitting:

Initial situation → Clustering and source linkage → Final resolution

"Gaia DR1 XXX"

→ New "Gaia DR2 YYY" + "Gaia DR2 ZZZ"

→ "Gaia DR1 XXX" does not exist anymore
(there is no "Gaia DR2 XXX")

Sources can appear, disappear and change!

"Gaia DR1 XXX"

→ "Gaia DR2 XXX"

→ Same "XXX" number, but the actual source features may be (very) different!

→ Again: do NOT rely on the source identifier.

Sources can appear, disappear and change!
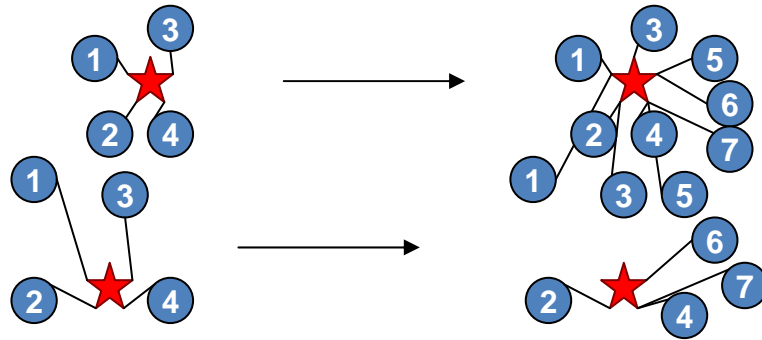
"Gaia DR1 XXX"

→ "Gaia DR2 XXX"

→ Same "XXX" number, but the actual source features may be (<u>very</u>) different!

→ Again: do NOT rely on the source identifier.

Some numbers on DR1→DR2 changes

- ~70-80% of sources for G < 16mag have changed from DR1 to DR2!
- Overall, ~25% of Gaia DR1 sources have been deleted
- But, again:
  do NOT rely on the ~75% sources for which the SourceID has not changed

- So... What can you do?

Gaia Archive cross-match tools

Gaia Archive cross-match tools: example

- Look for our star (e.g. Beta Eridani) in the Gaia DR1 GaiaSource table
    - → SourceID = 3211922645852214144

- Go to the Archive and run this:

```
SELECT TOP 100 * FROM gaiadr2.dr1_neighbourhood WHERE gaiadr2.dr1_neighbourhood.dr1_source_id =
3211922645852214144
```

- Result: another Source ID

dr1_source_id	dr2_source_id	angular_distance arcsec	magnitude_difference mag	rank
3211922645852214144	3211922645853255552	0.010352397	0.2551555	1.6474571

- With this you can now access all its DR2 features

Final remark: DR3!

- Source list changes between DR2 and DR3 are expected to be much less
  - Just ~2-3% of DR2 source IDs may be deleted in DR3
- But again: do not rely on getting the same SourceID!

Thank you