Identification/characterization of objects of unknown nature (OUTLIERS) observed by Gaia.

Minia Manteiga
Universidade da Coruña

and GGG team

Reunión RIA "La contribución de las ICTS al proyecto Gaia de la ESA".
Madrid 21-23 marzo 2012

gaia

1.- Gaia sample

- ~ 10^9 stars (1% of the Galaxy)
- ~ 10^5 WDs
- ~ 10^6 unresolved galaxies
- ~ 500.000 quasars
- Many solar system objects

→DPAC-CU8 "Astrophysical parameters" main objective: sort objects by astronomical classes.

→It will manage a standardized set of labels to classify the observed sources.

2.- Available information

- position, magnitude
- BP-RP low resolution espectra
- astrometry (π,pm)
- variability
- Hi-res spectrum RVS (V:6-17)

3.- CU8 Classifiers

i) Supervised Classification

- **DSC** --> stars, binaries, galaxies, qsos, object pairs, outliers
- Dedicated : UGC, QSOC, MSC, ESP

"Templates" (semi-empirical or synthetic)
+
pattern recognition algorithms
supervised training
metrics and minimum distances
→ labelling (classification) and parameterization

Processing algorithms are mainly based on BP/RP spectrophotometry

3.- CU8 Classifiers

gaia

ii) Unsupervised classification

Observed "natural" classes

- OCA --> class 1, class 2, class 3,.. HMAC
- OA--> oclass1, oclass2, oclass3,.. HSC

4.- Outliers

- Objects with membership probabilities to any Gaia class below a threshold (67%)

- Objects outlying from the clusters centroids obtained by the classification algorithms

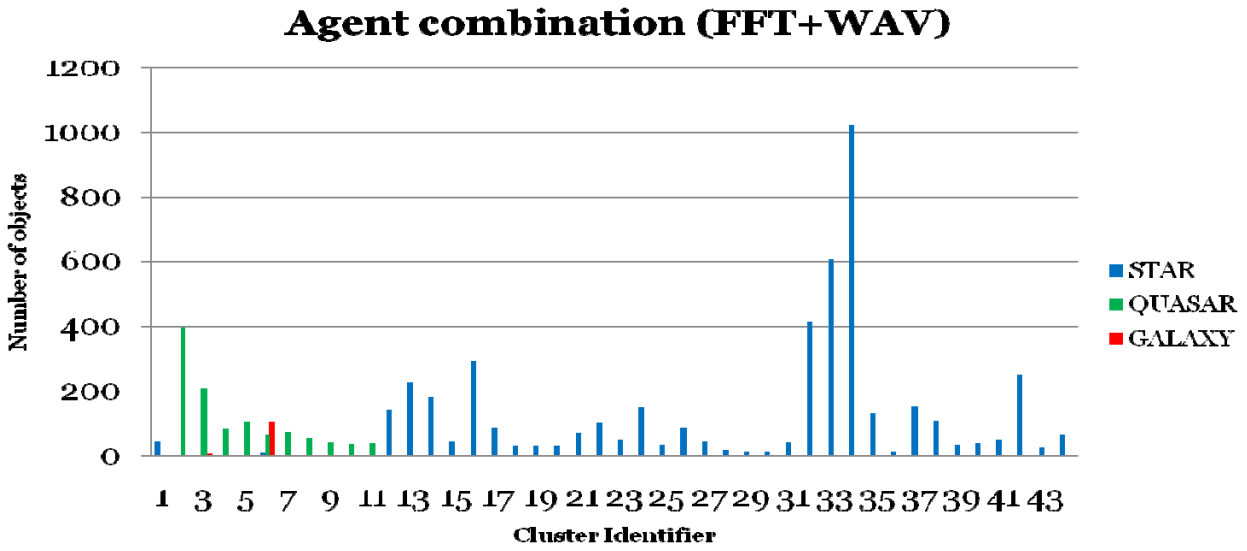
DPAC estimation: ~5% of sources

- new classes of variable stars
- rare stars (brief phases of stellar evolution)
- abnormal abundance patterns
- atypical or very faint extragalactic objects
- multiple systems
- ¿?

5.- Outlier Analysis

1- Segmentation:

- Gather objects in homogeneous clusters, in order to better identify their physical nature and study them in detail
- Clustering techniques based on unsupervised ANNs



5.- Outlier Analysis

2- Labelling

- Distribution of probabilities assigned by DSC inside outliers´ homogeneous clusters

- Statistical description of clusters by "mean" observed properties.

- Cross-correlation with external databases
  UKIDSS SDSS 2MASS

5.- Outlier Analysis

3- Analysis

Clusters with identified physical nature :
→ Feeded back to DSC (and improve it)

Remaining clusters:
→ Study "by hand"
→ Complementary observations

6.- Conclusions

- Forseen number of outliers: ~5 million from which ~30% could be identified and feeded back to DSC

- We expect to group most of the remaining and reduce the number of cases of study

- Telescope time will be needed to confirm and characterize their physical nature
  → high-medium resolution espectroscopy