

The GaiaPortal and the cross-match algorithms and software developed at SSDC: unique solutions to complex scientific and technological challenges. (Paola Maria Marrese - Silvia Marinoni)

Tuesday 18 June 2019 09:25 (25 minutes)

Although each astronomical catalogue on its own can be a very powerful scientific tool, it is the combination of archives with each other that truly opens up amazing possibilities for modern astronomical research and more closely meets its requirements. The advanced interoperation of archives should leave the user with the feeling of working with one single data archive. The first step for seamless data access is the computation of the cross-match between surveys. The complexity and scientific issues related to cross-matching has become very popular now that the combined use of large data sets from different surveys and/or wavelength domains is more and more common. The cross-matching of astronomical catalogues is a complex and challenging problem both scientifically and technologically, especially when matching large surveys which include several millions or billions of sources. There are different approaches to the combination of astronomical catalogues, and cross-match algorithms can be very different. It is important to correctly define both the scientific problem one is faced with and the objectives of the cross-match. The Gaia catalogue, with its high accuracy astrometry and high angular resolution, is used in the cross-match algorithms developed in SSDC as the link between other publicly available surveys obtained either from ground or from space, large or small, old, recent and future.

Gaia data releases, with almost two billions of sources, a largely heterogeneous dataset included in the catalogue and thus complex metadata, represent an excellent and challenging example of how to implement the techniques necessary for the management, the access and the scientific exploitation of big data. SSDC is one of the four ESA partner data centres for the distribution of Gaia data. Through the GaiaPortal at SSDC, users are allowed to access a huge distributed archive of complex high-level data including Gaia, large optical/NIR public surveys and the results of their cross-match.

The GaiaPortal web interface is almost unique, especially if compared to the solutions adopted by ESA and other partner data centres. GaiaPortal is built to be multi-wavelength and its distinctive characteristic and strength lie in allowing users to interrogate highly composite data without worrying neither to have in depth knowledge on the structure and the organization of the data in the database, nor to correctly write intricate SQL based queries.

Session Classification: Session 2a: From data to science and back: current experiences and future perspectives - Chair: A. Antonelli