

Distributed Repository per Applicazioni Mediche

L. Corradi, I. Porro, A. Schenone, M. Fato

Università degli Studi di Genova, DIST – Viale Causa 13, Genova, Italy

e-mail: ivan.porro@unige.it

Current trends in modern e-Science require the adoption of advanced tools in order to collect, organize, annotate and access data. This is evident in life science and especially in neuroinformatics. Nowadays within this context it is more and more frequent to face common issues like the storage and the sharing of large amounts of produced data, and the availability of dedicated computational resources.

Another key aspect is multimodal and multiscale data integration: this means that there is need for a platform able to manage data from different image modalities (MRI, CT, functional, but also physiological time series signals) on a variety of scales (organ/body image and signals, tissue/sub-organ images, molecular and genomic data). These modalities need to be accessible through a unified Web interface where studies (collection of data about a patient) from different specialties (structural and functional neuroradiology, oncology, psychology, and psychiatry) for a given patient are available.

According to that, there is need for a collaborative environment to store and share data under appropriate security policies and to follow patients diagnostic path by using hierarchical structures defining all different events (visits, diagnostic examinations, screenings and all data analysis steps). This can provide both a detailed archive and an improved repeatability of experiments. For this reasons a Grid approach has been considered. In fact the added value provided is the possibility of exploiting both computational (also taking advantage of dedicated HPC nodes) and data Grids infrastructures to analyze large datasets of distributed data. Furthermore it is worth noting that an added relevance related to the use of Grid computing for the presented application consists in the opportunity of sharing data and resources.

Furthermore, an intensive study has been performed to improve flexibility in the database schema supporting the platform architecture. In particular, as regards data and metadata entities and relationships, the XCEDE (XML-based Clinical and Experimental Data Exchange) model has been taken as a reference, that describes a complete metadata hierarchy for clinical experiments. Data and metadata hierarchies used by the BIRN (Biomedical Informatics Research Network) initiative have also been taken into account.