Project context and objectives

Earth observation data has increased considerably over the last decades as satellite sensors collect and transmit back to Earth many gigabytes of data per day. The aim of project TELEIOS is to increase the usability of the terabytes of satellite images lying dormant in archives by automating the relevant data management, integration and knowledge discovery tasks.

The main innovation of project TELEIOS is the development of a Virtual Observatory infrastructure that goes beyond the current state of the art Earth Observation portals and Image Information Mining systems.

This will be achieved by combining advanced image mining, database, geospatial and semantic web technologies, and pursuing the following technical innovations:

- Extensions of SQL based on the concept of a multidimensional array. These extensions are appropriate for querying scientific databases in general and satellite image databases in particular.
- Knowledge discovery techniques for satellite images and other relevant data sets, especially geospatial ones.
- Geospatial extensions to RDF and SPARQL for specifying and querying satellite image annotations. These annotations can capture image metadata and knowledge discovered from the images, and be linked to relevant data sources available on the web as linked data.
- Query processing and optimization techniques for multidimensional array queries, and queries for the extensions of RDF and SPARQL.
- Scalable implementations of the techniques developed in TELEIOS on top of MonetDB, an open-source database system for high-performance applications. Scalability will be shown using very large data sets, e.g., terabytes of image data, billions of triples of image annotations.

The technologies developed in TELEIOS will be demonstrated in the following use cases:

- A Virtual Observatory for TerraSAR-X data.
- Real-time fire monitoring based on continuous acquisitions of satellite images and geospatial data.

Work performed and results achieved

From September 2010 to August 2011 we initiated the research programme of TELEIOS. In these twelve months we focused on capturing the user requirements, developing a preliminary version of the TELEIOS architecture and infrastructure, and studying analytically and experimentally the individual models, languages, algorithms, and software components to be developed in the project.

Our technical achievements in this period are:

- Identification and analysis of the user requirements of the two TELEIOS use cases, i.e., "A Virtual Observatory for TerraSAR-X data" and "Real-time fire monitoring based on continuous acquisitions of satellite images and geospatial data".
- Review of the state-of-the-art technologies related to the TELEIOS scientific challenges.
- Detailed analysis and experimentation with the software and hardware components available by project partners at the beginning of the project.
- Detailed specification of the first version of the TELEIOS architecture based on the requirements of the two TELEIOS use cases, the systems available by partners, and the software components developed in the project.
- Preliminary version of the TELEIOS infrastructure with an emphasis on the NOA use case.
- The development of SciQL, an extension of SQL with the concept of a multidimensional array that will serve as the query language for satellite images.
- The development of the ability to run continuous SciQL queries over EO data represented as arrays by adopting and extending DataCell, the MonetDB stream engine.
- The data models stRDF, stRDF^{i} and the query language stSPARQL. These are geospatial extensions of RDF and SPARQL for specifying and querying satellite image annotations capturing image metadata and knowledge discovered from the images, and being linked to relevant data sources available on the web as linked data. stRDF^{i} allows the representation of incomplete, indefinite or qualitative geospatial information.
- A detailed study of the computational complexity of evaluating stSPARQL queries over stRDF databases.
- A detailed study of the semantics of the model stRDF^{i} and the computational complexity of evaluating queries expressed in a subset of stSPARQL over stRDF^{i} databases.
- Porting the system Strabon to MonetDB. Strabon is an implementation of stRDF/stSPARQL developed in the project SemsorGrid4Env by NKUA.
- Preliminary query processing and optimization techniques for stSPARQL implemented in the MonetDB version of Strabon.
- Preliminary query processing and optimization techniques for SciQL implemented in MonetDB.
- The design of the TELEIOS framework for knowledge discovery from satellite images, their metadata, and relevant GIS data.
- The application of the TELEIOS knowledge discovery framework to high resolution TerraSAR-X images.
- The development of two prototype image search engines exploiting the principles of the TELEIOS knowledge discovery framework. The first search engine is based on support vector machines (SVN), and the other on content-based image retrieval using the fast compression distance similarity measure.

We also undertook various dissemination activities. The most important dissemination activity is the first TELEIOS User Community workshop at ESA/ESRIN in Frascati, Italy. In this workshop we presented the TELEIOS vision to the wider EO community, and captured the requirements of invited EO scientists that are not participants of TELEIOS but are also working on problems similar to the ones we study in our use cases.

Other dissemination activities include 2 published scientific papers related to the data model of TELEIOS, one in the "Proceedings of the EDBT/ICDT 2011 Workshop on Array Databases" and one in the "Proceedings of the IJCAI 2011 Workshop on Benchmarks and Applications of Spatial Reasoning", a poster presentation that illustrates the goals and challenges of TELEIOS for the "Seventh Conference on Image Information Mining: Geospatial Intelligence from Earth Observation (ESA-JRC-EUSC)" conference, as well as a number articles about TELEIOS in traditional and electronic press, 3 radio interviews, and a television interview.

Finally, we prepared the first version of the exploitation plan for the expected results of TELEIOS.

Expected final results

The expected scientific and technological results of TELEIOS are the following:

- The data models stRDF, stRDF^i and the query language stSPARQL. These are geospatial extensions of RDF and SPARQL appropriate for specifying and querying satellite image annotations capturing image metadata and knowledge discovered from the images, and being linked to relevant data sources available on the web as linked data.
- Theoretical results for the data models stRDF, stRDF^i and the query language stSPARQL.
- Query processing and optimization techniques for the data models stRDF, stRDF^i and the query language stSPARQL, and efficient implementation in MonetDB.
- The query language SciQL, an extension of SQL with the concept of a multidimensional array.
- Query processing and optimization techniques for SciQL, and efficient implementation of them on top of MonetDB.
- Query processing and optimization techniques for continuous SciQL queries using DataCell, the MonetDB stream engine.
- The TELEIOS framework for knowledge discovery techniques from satellite images, their metadata, and relevant geospatial data sets.
- Algorithms for knowledge discovery and data mining in the above framework.
- The TELEIOS Earth Observatory architecture.
- The TELEIOS Earth Observatory infrastructure.
- A deployment of the NOA use case and the DLR use case in the TELEIOS Earth Observatory infrastructure.

Impact of final results

TELEIOS will enable the wider EO community to deal with the challenges posed by the huge scale and complexity of EO data.

The TELEIOS infrastructure will enable better leveraging of the skills of EO scientists, improved quality and quantity of output, and reduced time and cost for their research. Using the TELEIOS infrastructure, EO scientists will be able to search for satellite images using semantic criteria, integrate satellite images and other relevant data sets, and use them in applications cutting across domains and organizational boundaries.

TELEIOS will allow organizations with large sets of EO data (ESA, our partners DLR and NOA, other national space agencies etc.) to mine the content of their archives and extract meaningful knowledge that can be used to develop new applications.

TELEIOS will develop a new generation of EO portals that will function as Virtual Observatories for archived satellite products (Earth Observatories). In addition, TELEIOS will develop more precise real-time fire monitoring techniques based on continuous acquisitions of satellite images and geospatial data.

Through a focused set of research, development, dissemination and exploitation activities, TELEIOS will contribute to the objectives of international efforts in the area of EO such as GEOSS and GMES. The research and technological contributions of TELEIOS will advance the state of the art in the areas of Semantic Web and Linked Data, Scientific Data Management and Image Information Mining.

Project public website

The public website of TELEIOS can be found at http://www.earthobservatory.eu/. The website reflects the public image of the project through a clean functional design, and provides up-to-date information about the progress of the project.