TELEIOS Annual Report

http://www.earthobservatory.eu/

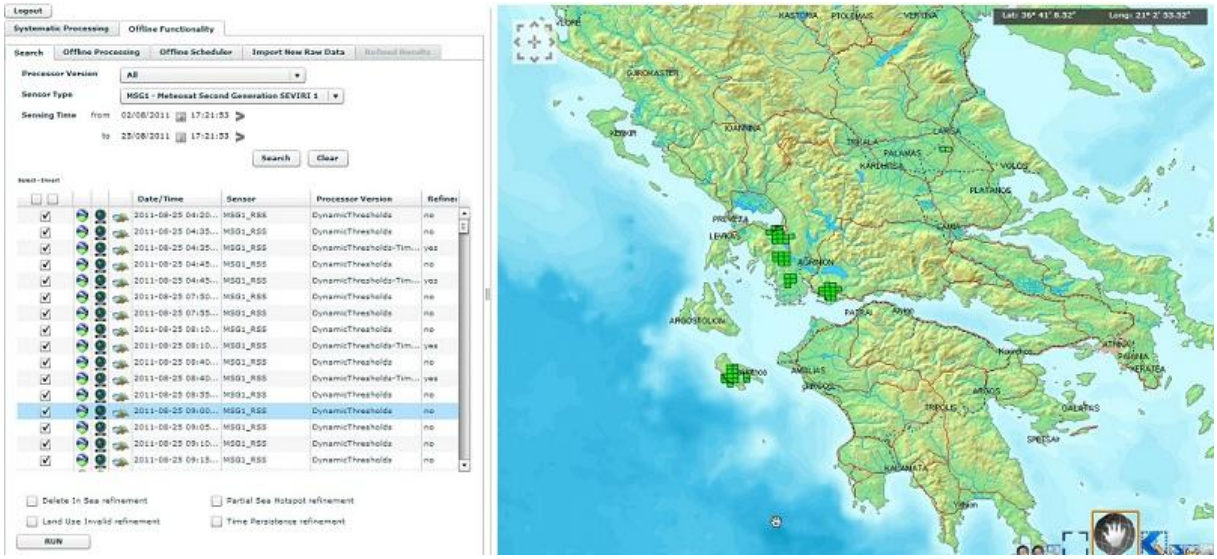
Earth Observation data has increased considerably over the last decades as satellite sensors collect and transmit back to Earth many gigabytes of data per day. The aim of project TELEIOS is to increase the usability of the terabytes of satellite images lying dormant in archives by automating the relevant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

The main innovation of project TELEIOS is the development of new user access, data management, image mining and semantic annotation technologies that go beyond systems currently deployed in Earth Observation data centers. The advances of TELEIOS build on new breakthroughs of TELEIOS partners in scientific databases, Semantic Web and satellite image information mining.

Summary of Activities

The period between January 2012 to December 2012 was very important for TELEIOS since various system components (SciQL, data vault, Strabon, user interface) were completed and integrated into the first version of the TELEIOS infrastructure. The infrastructure was presented in the 2nd user community workshop that took place in Darmstadt and it received very encouraging feedback by the TELEIOS user community. The progress that has been made is reflected in the large number of scientific publications and in the opening up of the Fire Monitoring application which is being showcased by the present version of the infrastructure which is now available to the TELEIOS users. The individual system components and the integrated infrastructure will be constantly improved in terms of functionality and performance for the rest of the project.

The first version of the TELEIOS Infrastructure

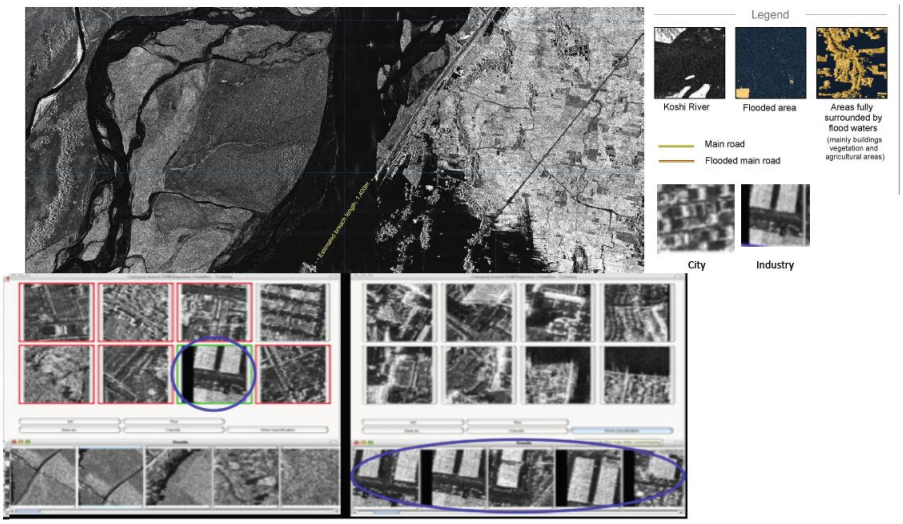


TELEIOS is developing an infrastructure on which newer and more effective methods for EO data management are implemented and experimented with. In the last year we designed and implemented the first version of the TELEIOS infrastructure in the context of the use case "Real-time fire monitoring based on continuous acquisitions of Earth Observation images and geospatial data". This version exposes the fire monitoring service offered by the National Observatory of Athens (NOA) via a web interface and utilizes the implementation of SciQL on MonetDB[1] and the semantic geospatial RDF store Strabon[2]. The query language SciQL is used to encode all the talks of the processing chain for hotspot identification in an MSG/SEVIRI image. stSPARQL queries executed in the system Strabon are then used to refine the generated hotspots using linked geospatial data. For example, linked geospatial data representing the coastline of Greece is utilized to refine hotspot products by discarding the part of their geometries that lie in the sea. Such inaccuracies are due to the low spatial resolution of the MSG/SEVIRI infrared imager that is used for monitoring the hotspots. Hotspot refinement using TELEIOS technologies together with the high temporal resolution of the imager (one image every 5 minutes) allow NOA to develop maps that show the advancement of a fire front on the ground very quickly. The fire monitoring application is now available online[3] to the users and an evaluation questionnaire has been prepared so that feedback can be collected and taken into consideration for the final version of the TELEIOS infrastructure. The final version of the TELEIOS infrastructure will be used by NOA during the fire season of 2013.

[1] http://www.monetdb.org
[2] http://www.strabon.di.uoa.gr
[3] http://teleios.space.noa.gr/

Knowledge Discovery from EO images

In this period we completed the first implementation of the TELEIOS framework for knowledge discovery from satellite images, their metadata, and relevant GIS data. The core technological contribution of the framework is developing advanced methods to structure the heterogeneous feature space of satellite images, their metadata descriptors and related geospatial data sources. In the first step of this process, we enhance traditional raw data processing by utilizing content extraction methods to produce image descriptors, such as texture features. In the next step, we have been developing methods to map those into semantic and symbolic representations with the use of two KDD concepts that have been developed and demonstrated; an active learning search engine implemented on a Support Vector Machine (SVM), and a Compression Based Image Retrieval system being implemented on MonetDB. Both these KDD modules consult ontologies to guide the annotation process. The annotations produced are expressed in RDF and are made available as linked data so that they can be easily combined with other publicly available linked data sources (e.g., GeoNames, LinkedGeoData, DBpedia) to allow for the expression of rich user queries. In this period 39 TerraSAR-X scenes have been analysed with our techniques, 80% of the respective patches (45000 in total) have been annotated and 320 semantic categories have been identified. Since satellite images (as opposed to multimedia images) are much bigger in size and much more difficult to analyze, the present analysis already constitutes a significant achievement. By the end of the project we expect to have processed more images and, in this way, we will set a new international standard for image mining of TerraSAR-X images.



Scalable storage and query processing for Linked Geospatial Data

In the past 12 months we studied query processing and optimization algorithms for the model stRDF and the query language stSPARQL. stRDF is an extension of RDF that allows the representation of geospatial data that changes over time, and is accompanied by stSPARQL, an extension of the query language SPARQL 1.1 for querying stRDF data. Both stRDF and stSPARQL use OGC standards (Well-Known Text and Geography Markup Language) for the representation of temporal and geospatial data. The developed techniques have been implemented in the system Strabon, available as open source at http://www.strabon.di.uoa.gr/. In the current version of Strabon, PostgreSQL and MonetDB are supported as the underlying spatial DBMS. Strabon has been compared experimentally with systems offering similar functionalities and has been established as one of the most rich, scalable and efficient semantic geospatial DBMS.

In this period the temporal features of the initial stRDF proposal have been revised regarding the representation of the valid time of triples. Similarly, the query language stSPARQL has been extended with a new set of functions for querying triples annotated with their valid time. These temporal features of stRDF and stSPARQL have been implemented by extending Strabon. A preliminary performance evaluation that has been carried out shows that the temporal component of Strabon performs efficiently. More experiments are in order with significantly larger datasets.

We have also fully developed the data model RDF^i which extends RDF with the ability to represent incomplete information about property values that are unknown or partially known using constraints. Following ideas from the incomplete information literature, we developed a semantics for this extension of RDF, called RDF^i, and studied SPARQL query evaluation in this framework. In addition, we defined the fundamental concept of certain answer to SPARQL queries over RDF^i databases and presented an algorithm for its computation. We also presented preliminary complexity results for computing certain answers by considering equality, temporal, and spatial constraint languages. Last, we demonstrated the usefulness of RDF^i in geospatial Semantic Web applications by studying many examples and comparing the modeling capabilities of RDF^i with the ones of related Semantic Web systems.

Image processing and analysis using SciQL and data vaults

Traditional approaches in databases store image data either in separate file repositories outside the database, or as “black-box” BLOBs inside the database. An inconvenience of these approaches is that they require the use of external programs or user-defined functions to allow access to the image content. By keeping this content outside the control of the DBMS, the optimization and execution engines of the DBMS are not utilized. With the introduction of SciQL by TELEIOS, a user-friendly high-level declarative language that provides efficient array manipulation primitives, the optimization and execution engines gain transparent access to the image content. Therefore, by using SciQL, we can express low level image processing

operations such as cropping, re-scaling, geo-referencing etc. in the context of the database. In addition, the algorithms followed to execute these processing operations can be optimized by the extended query optimizer of the DBMS. Finally, by integrating relational tables and arrays, we can simultaneously utilize both image content and image metadata during query processing and knowledge discovery.

We have also developed the data vault, a mechanism that provides a true symbiosis between a DBMS and existing (remote) file-based repositories such as the ones used in Earth Observation applications. The data vault keeps the data in its original format and place, while at the same time enables transparent data and metadata access and analysis using the SciQL query language.

In this period, the first implementation of SciQL and the data vault has been completed in MonetDB and has been demonstrated in the context of the Fire Monitoring application of NOA.

User Involvement, Promotion and Awareness

In the 2nd year of TELEIOS, we disseminated actively the ideas and technologies that have been developed in the project. Twenty two scientific publications were presented in various forums, eight of which were joint publications, authored by more than one partners of the consortium.

The scientific contributions of the TELEIOS technologies in the field of the Semantic Web were presented by the NKUA team in the 11th International Semantic Web Conference (ISWC 2012), at Boston, USA on November 2012. A research paper on Strabon, a demo on the fire monitoring application of TELEIOS, and a workshop paper on the RDF[i] framework were presented. Furthermore, a demo of the fire monitoring application based on technologies developed in the projects TELEIOS and the related Greek project SWeFS won 3rd place in the Semantic Web Challenge 2012.[4]

The TELEIOS co-ordinator Manolis Koubarakis co-organized the "Terra Cognita 2012" workshop that took place in conjunction with the 11th International Semantic Web Conference (ISWC 2012), at Boston, USA on November 2012. The workshop brought together Semantic Web and geospatial data researchers and one of its goals was to encourage the use of semantics in geospatial applications. More information about the workshop can be found in its proceedings available at http://www.strabon.di.uoa.gr/terracognita/.

TELEIOS partner DLR involved internal users in TELEIOS and coordinated the involvement of relevant stakeholders with ESA. DLR also organized the 8th ESA-EUSC-JRC Image Information Mining conference at their site in Oberpfaffenhofen, Germany. The TELEIOS consortium had a strong participation to the conference with two papers. The KDD concept was also the topic of invited lectures and tutorials presented at IGARSS 2012 and several

[4] http://challenge.semanticweb.org/2012/

European Summer Schools.

We also organized the 2nd TELEIOS User Community Workshop at the premises of our partner Fraunhofer in Darmstadt, Germany on 10/11 May, 2012. In this workshop, the TELEIOS infrastructure was presented to the users and received encouraging feedback that has been taken into consideration regarding the final version of the TELEIOS system.

Furthermore, the TELEIOS partner NKUA created the Greek Linked Open Data portal[5], where some of the datasets that have been created and used within the project are published and can be browsed and queried using Strabon. This site is now actively extended with more linked open datasets of interest to Greece.

Future Work

The remaining period from January to August 2013 will complete the work of TELEIOS. We will concentrate on developing the final version of our technologies and apply them to our use cases. Our main goals in this period will be:

- The continuous improvement of the current infrastructure in accordance to feedback received from the users, leading to the final version of the TELEIOS infrastructure.
- An extended implementation of stSPARQL in MonetDB which demonstrates the advantages of column stores for geospatial data processing.
- Benchmarking and evaluation of stSPARQL and SciQL
- The final version of the TerraSAR-X Earth Observatory for our partner DLR and its evaluation by DLR users.
- The final version of the NOA fire monitoring service, its evaluation and its operational use by NOA in the fire season of 2013 in Greece.

Further Information

- TELEIOS website
- MonetDB website
- Strabon website
- Greek Linked Open Data
- http://teleios.space.noa.gr/

[5] http://linkedopendata.gr/