GET-IT, a software suite for easy, interoperable sharing of ecological data in the Long Term Ecological Research Network

Lanucara Simone[1], Martina Zilioli[2], Oggioni Alessandro[3] and Paola Carrara[4]

1. Introduction

The European Network for Long Term Ecological Research (LTER-Europe - http://www.lter-europe.net) is a network of more than 400 ecological interest sites [1] on which researchers carry out decadal-scale ecological investigations. Leading research institutions, universities and governmental agencies manage the sites and produce long term ecological data, which are planned with a long-term view, carefully considering data preservation along time, storage of specimens, locations of measurement. Sharing these data and related metadata can: (I) improve the comparability of ecological data collected around the world; (II) foster their (re)use outside the research context; (III) augment existing data collections and verify research results [2].

As already observed by different authors [3], [4], data and metadata sharing in LTER can be enabled by developing information technology (IT) infrastructures that allow to search, manage and share data on environmental change and ecosystems.

IT infrastructures can be implemented with different architectures (by example they can be distributed, centralized, hybrid), software tools (proprietary or open source), and varying degrees of interoperability, from systemic to semantic [5]. Since 2012, IT facilities have been developed to enable LTER-Europe site managers to register, manage and share ecological information on sites, people (researchers involved) and data. The main facility developed has been Drupal Ecological Information System (DEIMS - https://data.lter-europe.net/deims) for managing and discovering data, sites and people by their metadata [6]; sharing of the data itself is offered via a centralized repository of excel sheets established during the LIFE+ Project EnvEurope [7]. This IT solution lacks in basic systemic interoperability for ecological data sharing, a gap that obstacles information exchange both inside the LTER-Europe Network and outside, towards other users and research networks. In fact, it does not conform to the rules established by the Infrastructure for Spatial Information in the European Community

[1] National Research Council of Italy, Institute for Electromagnetic Sensing of Environment, lanucara.s@irea.cnr.it

[2] National Research Council of Italy, Institute for Electromagnetic Sensing of Environment, zilioli.m@irea.cnr.it

[3] National Research Council of Italy, Institute for Electromagnetic Sensing of Environment, oggioni.a@irea.cnr.it

[4] National Research Council of Italy, Institute for Electromagnetic Sensing of Environment, carrara.p@irea.cnr.it

Directive (INSPIRE; [8]) for the sharing of environmental data and metadata through interoperable Open Geospatial Consortium (OGC) standard services for data deployment and discovery.

Besides sharing ecological data and metadata through OGC international standards and in accordance with the INSPIRE Directive, it should be also advisable to propose solutions targeted to ecological researchers, e.g. not requiring skilled IT personnel for (meta)data entry and management. A possible solution to face the aforementioned challenges, tested and adopted in the Italian Long Term Ecological Research Network (LTER-Italy) and in the Project European Long Term Ecosystem and Socio-Ecological Research Infrastructure (eLTER H2020 - http://www.lter-europe.net/lter-europe/projects/eLTER) is based on *Geoinformation Enabling Toolkit StarterKit* ® (GET-IT; [9]).

In this short paper, we first briefly describe GET-IT, a software suite to share ecological data and metadata through OGC international standards and in accordance with the INSPIRE Directive. Then, in Section 3, we report a test of GET-IT to share ecological data and sensor metadata in LTER-Italy and we present the work in progress for the research infrastructure of eLTER by exploiting GET-IT. Section 4 describe the use of GET-IT in other disciplinary domains, present the results, the current issues, future works and it concludes the paper.

2. Geoinformation Enabling Toolkit StarterKit ®

GET-IT is a free and open source software suite developed by our research group to enable non IT-skilled researchers from various scientific domains to create their own OGC standard services (Table 1) for distributing geospatial data, observations and metadata of sensors according to the INSPIRE Directive.

Web service	**Version**
OCG Web Map Service (WMS)	1.1.1, 1.3.0
OGC Web Feature Service (WFS)	1.0.0, 1.1.0, 2.0
OGC Web Coverage Service (WCS)	1.0, 1.1.0, 1.1, 1.1.1
OGC Styled Layer Descriptor (SLD)	1.0
OGC Tile Map Service (TMS)	1.0.0
OGC Web Map Tiling Service(WMTS)	1.0.0
OGC Catalogue Services for the Web (CSW)	2.0.2
OGC Sensor Observation Service (SOS)	1.0.0, 2.0.0
OGC Sensor Model Language (SensorML)	1.0.1, 2.0.0
OGC Observation and Measurement (O&M)	1.0.0, 2.0.0

Tab. 1. Main interoperable and OGC standard services supported by GET-IT

GET-IT fosters the creation of autonomous nodes of a distributed Spatial Data Infrastructure (SDI) for an observational network. In fact, it allows users to straightforwardly share on the web their ecological observations (but also maps, if provided) and metadata of the sensors used. This last characteristic improves the quality of the data collected and distributed, as it allow to examine and compare the ways in which data are collected. The software suite has been developed by improving the facilities of GeoNode (http://geonode.org), a widely known geospatial Content Management System that is capable of dealing with maps, geospatial data and also with geotagged documents. Metadata creation is enabled by EDI [10], the GET-IT metadata editor which allows easy and friendly instrument registration through graphical user interfaces (GUI) and auto completion facilities linked to domain vocabularies. In fact, EDI facilitates the harmonization of metadata exploiting thesauri based on Simple Knowledge Organization System (SKOS), which are basic elements for representing knowledge in the context of the Semantic Web. Furthermore a GUI (Figure 1) enables uploading and sharing observed data, in compliance with the Observations and Measurements (O&M) OGC standard. Sensor metadata management and standard OGC deploy of observations are completely new features of GET-IT, not included till now in GeoNode.

In order to facilitate GET-IT installation, it is distributed as a virtual machine, based on the Ubuntu operating system.

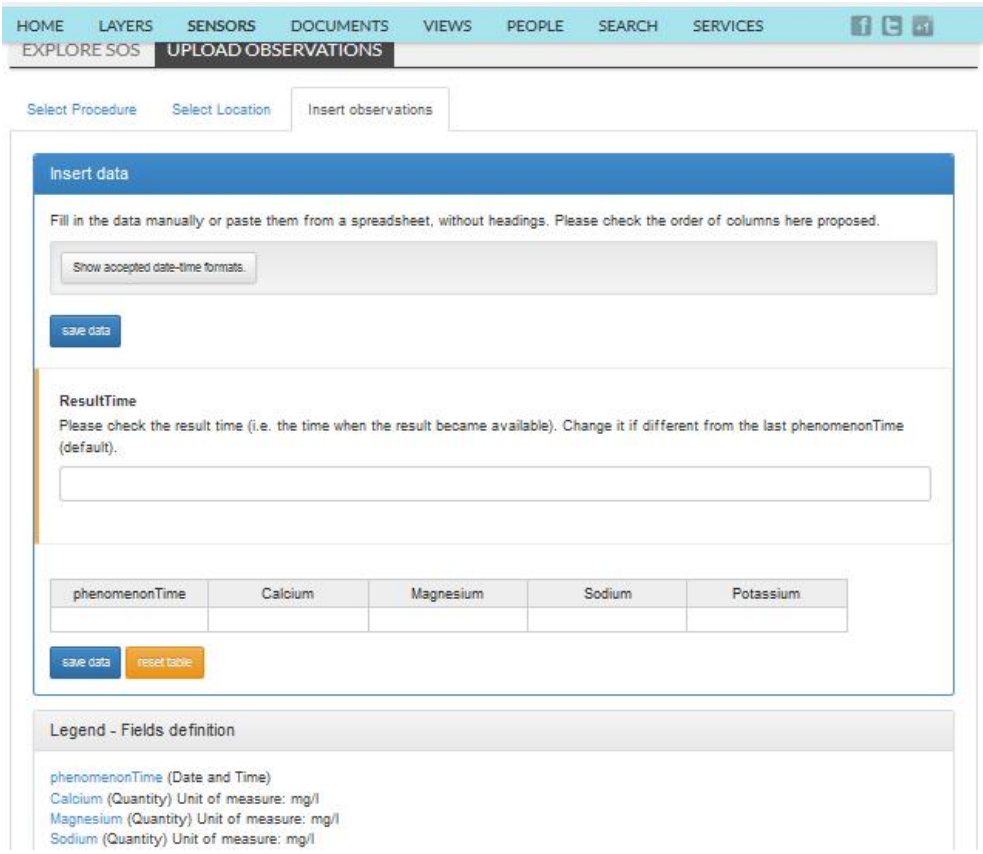

Fig. 1: GET-IT GUI for uploading observed data.

3 Test and work in progress

In the following sections we first (3.1) describe a test to share ecological data and sensor metadata in LTER-Italy using GET-IT, then (3.2) we present the work in progress for sharing ecological data in LTER-Europe.

3.1 Sharing ecological data in LTER-Italy by GET-IT: a test

We tested GET-IT in the context of LTER-Italy during the national Project NextData funded by the Italian Ministry of University and Research (MIUR). This project aims to share ecological data and sensors metadata regarding mountain LTER-Italy sites, by deploying an interoperable, distributed SDI. The challenges of the test were: (i) to develop a distributed IT infrastructure, where dataset could be hosted by autonomous repositories of the LTER sites; (ii) to allow standardized, interoperable data distribution; (iii) to harmonize dataset metadata by ecological thesauri, (iv) to propose friendly IT solutions. During the test, the mountain ecological research sites of LTER-Italy (both lacustrine and terrestrial) have been equipped with interoperable repositories for their data and metadata. In fact, two

independent nodes of a distributed SDI were deployed by GET-IT, in particular one for mountain lacustrine sites (http://sk.ise.cnr.it/) and one for mountain terrestrial sites (http://nextdata.get-it.it). By exploiting the suite, researchers themselves easily uploaded and shared 70 long term ecological observations covering a time range variable between fifty and thirty years with a time resolution between monthly and quarterly. The ecological sensors and related observation metadata were also semantically enriched by EDI using the EnvThes environmental vocabulary [11], in order to harmonize with terms used by all LTER-Europe Network sites in the DEIMS facility. Also, the metadata community practices adopted in the LTER-Europe community (DEIMS) are still in use in order to harvest and allow the discover of distributed GET-IT.

3.2 Sharing ecological data in LTER-Europe: work in progress.

With respect to the previous Project EnvEurope, which conceived the first IT facilities for LTER-Europe, the Project eLTER H2020 works towards a further implementation of services for the LTER-Europe community, enhancing data distribution and accessibility. The project aims to develop the core for a common information management infrastructure by sharing data from different distributed repositories. The management of data for most of the LTER-Europe sites lies within the responsibility of the respective site managers, thus maintaining a great autonomy in the data management process. Therefore, being ecological data in LTER-Europe extremely heterogeneous, data integration, harmonization and accessibility are challenging.

To tackle these challenges and following the test previously described, we have proposed to implement an OGC service-based architecture compliant with the INSPIRE Directive by exploiting GET-IT. As in LTER-Italy, GET-IT is tested to develop a LTER Europe distributed SDI in 22 European countries. LTER-Europe community has implemented tools for metadata creation, but some challenges about complete interoperability have yet to be gathered: metadata authoring of sites and instruments, data discovery and data visualizations through OGC standards. Purpose of GET-IT is to meet the needs of LTER researchers concerning: 1. sharing of data (observations and maps) in a standard way; 2. creating a distributed network where the users could share data using a friendly interface; 3. giving the information of data (metadata) using a common vocabulary and unique interface; 4. giving a user friendly tool and software for share data.

4 Results and conclusions

GET-IT is a free and open source software suite that facilitates the creation of nodes of a distributed SDI for an observational network. It allows users to straightforward share on the web their data, sensors and metadata which result semantically enriched. Using GET-IT, in the framework of the project NextData “Data-LTER-Mountain”, LTER-Italy researchers shared seventy new long term observations and relative metadata on mountain lakes, covering a time range variable between fifty and thirty years with a time resolution between monthly and quarterly. After the NextData project, GET-IT has been accepted by the LTER-Italy network and adopted for the interoperable sharing of ecological observations in eLTER. In fact, GET-IT is the tool adopted to develop the SDI nodes of a distributed research

infrastructure for 162 research sites in 22 European countries. GET-IT is also used to share marine, agricultural and geological data, sensors and metadata. In fact:

- in the Italian project RITMARE (http://www.ritmare.it/), GET-IT was created and is used to share marine data and relative metadata on the Mediterranean sea belonging to different domains (physical oceanography, chemical oceanography, geophysics, geology, ecosystem, molecular life sciences, fishing and agriculture, coastal system) [9], [13];
- in the project Space4Agri (http://space4agri.irea.cnr.it), that aims to use Earth Observation data (space, air and in situ), GET-IT is used to support the planning and management of the agricultural sector in Lombardy (a region of Italy) by sharing agricultural data collected by unmanned aerial sensors, and/or data produced by researchers as a result of applying scientific analysis on high quality remote sensing data [14];
- in the geological domain, GET-IT is used to semantically enrich and share metadata of geological sensors, their collected data and geological maps of a coastal zone of Southern Italy [5].

Therefore, researchers of different disciplinary domains can easily create, manage, edit and share geospatial data and metadata (of data and sensors), using the GET-IT OGC service-based architecture compliant with the INSPIRE Directive; they can distribute their own dataset avoiding unnecessary and harmful duplication of data.

GET-IT, has not still been used to exchange biodiversity data and metadata, though different researches are testing how to share biodiversity data exploiting O&M standard [15] [16] and we will investigate how to enable GET-IT to share them.

Beside the LTER network, various ecological communities share ecological data with different practices, software and standards, i.e. National Ecological Observatory Network (NEON - http://www.neonscience.org/), Global Lake Ecological Observatory Network (GLEON - www.neonscience.org/), Global Biodiversity Information Facility (GBIF - http://www.gbif.org/). Enabling them to OGC service-based architecture compliant with the INSPIRE Directive, by GET-IT, can improve the comparability of ecological data collected around the world, foster their reuse outside the research context and avoiding duplication of data. Further researches could investigate the sharing practices of other ecological communities and try to enable them to the exchange of ecological data globally by OGC standards.

References

[1] M. Mirtl and K. Krauze, “Developing a new strategy for environmental research, monitoring and management: The European Long-Term Ecological Research Network’s (LTER-Europe) role and perspectives,” *Nat. Conserv. Manag. From Idea to Pract. Results*, vol. PWZN "Prin, no. November, pp. 36–52, 2007.

[2] W. K. Michener, “Ecological data sharing,” *Ecol. Inform.*, vol. 29, pp. 33–44, 2015.

[3] W. K. Michener, “Meta-information concepts for ecological data management,” *Ecol. Inform.*, vol. 1, no. 1, pp. 3–7, 2006.

[4] R. D. Sutter, S. B. Wainscott, J. R. Boetsch, C. J. Palmer, and D. J. Rugg, “Practical guidance for integrating data management into long-term ecological monitoring projects,” *Wildl. Soc. Bull.*, p. n/a-n/a, 2015.

[5] S. Lanucara, A. Oggioni, G. Modica, and P. Carrara, “Interoperable Sharing and Visualization of Geological Data and Instruments: A Proof of Concept,” in *Computational Science and Its Applications – ICCSA 2017. ICCSA 2017. Lecture Notes in Computer Science*, vol. 1, 2017, pp. 584–599.

[6] K. L. Vanderbilt, C.-C. Lin, S.-S. Lu, A. R. Kassim, H. He, X. Guo, I. S. Gil, D. Blankman, and J. H. Porter, “Fostering ecological data sharing: collaborations in the International Long Term Ecological Research Network,” *Ecosphere*, vol. 6, no. 10, p. art204, 2015.

[7] T. Kliment, J. Peterseil, a Oggioni, A. Pugnetti, and D. Blankman, “Life+ EnvEurope DEIMS-improving access to long-term ecosystem monitoring data in Europe,” *EGU Gen. Assem. Conf. Abstr. 15*, vol. 15, p. 4920, 2013.

[8] Eu, “Directive 2007/2/EC of the European Parliament and of the council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE),” *Off. J. Eur. Union*, vol. 50, no. January 2006, pp. 1–14, 2007.

[9] C. Fugazza, A. Basoni, S. Menegon, A. Oggioni, F. Pavesi, M. Pepe, A. Sarretta, and P. Carrara, “RITMARE: Semantics- Aware harmonisation of data in Italian marine research,” *Procedia Comput. Sci.*, vol. 33, pp. 261–265, 2014.

[10] F. Pavesi, A. Basoni, C. Fugazza, S. Menegon, A. Oggioni, M. Pepe, P. Tagliolato, and P. Carrara, “SOFTWARE METAPAPER EDI – A Template-Driven Metadata Editor for Research Data,” *J. Open Res. Softw.*, vol. 4, no. 1, pp. 1–10, Oct. 2016.

[11] H. Schentz, J. Peterseil, N. Bertrand, M. Bastiano, D. Blankman, M. Frenzel, and U. Grandin, “EnvThes - interlinked thesaurus for long term ecological research , monitoring , and experiments,” 2013.

[12] S. Lanucara, P. Carrara, A. Oggioni, M. Rogora, and L. Kamburska, “Exploiting Observations & Measurement data standard for distributed LTER-Italy freshwater sites . Water quality issues .,” in *PeerJ Preprints 4:e2233v2*.

[13] A. Oggioni, M. Bastianini, P. Carrara, T. Minuzzo, and F. Pavesi, “Sensing Real-time Observatories in Marine Sites A Proof-of-Concept,” *SensorNets 2014, 3rd Int. Conf. Sens. Networks*, pp. 111–118, 2014.

[14] T. Kliment, G. Bordogna, L. Frigerio, D. Stroppiana, A. Crema, M. Boschetti, and S. Sterlacchini, “Supporting a Regional Agricultural Sector with Geo & Mainstream ICT – the Case Study of Space4Agri Project Key words,” *Agris on-line Pap. Econ. Informatics*, vol. VI, no. 4, p. 2020, 2015.

[15] A. Oggioni, P. Tagliolato, K. Schleidt, P. Carrara, S. Grellet, and A. Sarretta, “Biodiversity Data

Interoperability Issues: on the Opportunity of Exploiting O{\amp}M for Biotic Data Management," in *American Geophysical Union, Ocean Sciences Meeting 2016, abstract #OD34A-2487*, 2016.

[16] D. De Pooter, W. Appeltans, N. Bailly, S. Bristol, K. Deneudt, M. Eliezer, E. Fujioka, A. Giorgetti, P. Goldstein, M. Lewis, M. Lipizer, K. Mackay, M. Marin, G. Moncoiffé, S. Nikolopoulou, P. Provoost, S. Rauch, A. Roubicek, C. Torres, A. van de Putte, L. Vandepitte, B. Vanhoorne, M. Vinci, N. Wambiji, D. Watts, E. Klein Salas, and F. Hernandez, "Toward a new data standard for combined marine biological and environmental datasets - expanding OBIS beyond species occurrences," *Biodivers. Data J.*, vol. 5, p. e10989, Jan. 2017.

The activities described in this paper have been partially funded by the H2020 eLTER and the Italian Projects RITMARE and NextData.