# Exploiting Observations & Measurement data standard for distributed LTER-Italy freshwater sites. Water quality issues.

**Lanucara Simone[1], Carrara Paola [1], Oggioni Alessandro[1], Rogora Michela[2], Kamburska Lyudmila[2] and Rossetti Giampaolo[3]**

**[1]Istituto per lo Studio Elettromagnetico dell'Ambiente, Consiglio Nazionale delle Ricerche (CNR-IREA). Via Bassini, 15 - 20133 Milano (MI)**
**[2]Istituto per lo Studio degli Ecosistemi, Consiglio Nazionale delle Ricerche (CNR-ISE). Largo Tonolli 50 - 28922 - Verbania Pallanza (VB)**
**[3]Università di Parma. Parco Area delle Scienze, 11/a - 43124 Parma (PR)**

Corresponding author:
Lanucara Simone[1]

Email address: lanucara.s@irea.cnr.it

## ABSTRACT

Water quality is a multi-source, multi-purpose problem that needs exploiting observations, often taken by a number of heterogeneous bodies. This problem has been tackled within the Italian Long Term Ecological research network (LTER-Italy) in an experiment aimed at testing how ecological observations of mountain lakes water can be shared by OGC (Open Geospatial Consortium) standard services of the Sensor Web Enablement (SWE) initiative. A friendly and easy implementation of these services is fostered by the usage of the open source software *Geoinformation Enabling Toolkit StarterKit* ® (GET-IT[1]). It has been used in the experiment to create SOS services, upload observations and create SensorML metadata of the involved sensors. This contribution describes the experiment and presents its results.

Keywords: Free and Open Source Software for Geospatial, Spatial Data Infrastructure, Sensor Web Enablement, Ecological research.

## INTRODUCTION

Water monitoring needs frequent in situ measurements of different parameters that can vary depending on many factors, including its purpose. Many institutions generally perform in situ measurement activities; measures are stored and deployed following heterogeneous practices. This usually prevents sharing that calls for homogeneous practices. Since the emergence of

[1] http://get-it.it

Spatial Data Infrastructures (SDI), distribution and heterogeneity of geodata sources is no more an issue, provided actors agree in adopting recommendations and techniques enabling interoperability. Many good practices show how a change of perspective can be made (Finney, 2007; Budhathoki & Nedovic-Budic, 2008; Box & Rajabifard, 2009; Shamsi et al., 2011).

OGC dictates regulations fundamental in this respect and SWE is its proposal for interoperability of observations from fixed and mobile sensors. Creating SWE basic components, i.e. metadata of sensors and data, plus services for deploying, is not at all, in this moment, an easy task for limnologists, expert in lake research. This contribution presents the methodology adopted to this purpose in the LTER-Italy community for researches on mountain lakes. The methodology is based on open and free software and proved successful in creating SWE components and enabling interoperability, without requiring skilled IT profiles. The next section describes the approach and its tools; then results are shortly presented and discussed.

## METHOD

In the context of the NextData "Data–LTER-Mountain" project[2] we tested a methodology and workflow to share sensors, observations, datasets and metadata of Italian LTER sites, developing a distributed SDI. The architecture of SDI includes two interoperable and independent nodes enabled to OGC standard web services. The first node, installed at CNR ISE in Verbania Pallanza[3], is related to repository and web services on the mountain lake LTER sites Lakes Paione Inferiore and Superiore, located in the Alps (Marchetto et al., 2004) and Lake Scuro and Santo Parmense, in the Appennines. A second node is installed at CNR IREA Milan[4], for repository and web services on terrestrial mountain LTER sites.

For developing an SDI there are several free and open source software solutions (Steiniger et al. 2013): PostgreSQL-PostGIS for geospatial database, Geoserver or MapServer for sharing geospatial data, GeoNetwork Opensource or pyCSW for sharing geospatial data and services catalogues, 52° North SOS or istSOS for sharing sensors and observations. Anyway, these single software components have to be installed, configured and integrated to implement a SDI, which is not an easy task for limnologists, expert in lake research.

A software solution, Geonode (Benthall et al., 2010; Winslow, 2010), integrates several software before mentioned (PostgreSQL-PostGIS, Geoserver, GeoNetwork Opensource); however, while enabling a straightforward implementation of SDI, it misses semantic enablement and SWE.

To address these issues the free and open source software suite GET-IT was developed, exploited in this experiment for the implementation of SDI nodes, that allows a friendly and easy implementation of a semantic and SWE enabled SDI.

Developed by a joint research group of CNR IREA – CNR ISMAR, under the flagship project RITMARE[5]; it facilitates the creation of nodes of a federated SDI for an observational network. The suite allows users to straightforward share on the web (by OGC standards) their observations

[2] http://www.nextdataproject.it/sites/default/files/docs/PP1-LTER-Mountain.pdf

[3] http//sk.ise.cnr.it

[4] http//nextdata.get-it.it

[5] http://www.ritmare.it

and metadata on them and on sensors used. It consists of a virtual machine, based on the Ubuntu operating system; the basic software used is GeoNode[6], a widely known geographic content management system (Benthall et al., 2010; Winslow, 2010). GeoNode has been edited and have been added new facilities, both client and server side, for the creation, semantically enabled, and the management of observations and metadata of sensors (Fugazza et al., 2014).

The distribution on the web of sensor metadata and observations collected by sensors is performed in GET-IT by means of standard OGC Sensor Observation Services (SOS), i.e. a component of Sensor Web Enablement (SWE). This initiative, thanks to its high level of abstraction and associated use of schemes such as SensorML (Sensor Metadata Language; Boots et al., 2007) and O&M (Observations & Measurements; Cox, 2013) allows to create, store and share sensor metadata and observations gathered by the sensors.

Information on the sensors, registering mountain lakes ecological observations (i.e. DX-500 Ion Chromatograph, ICS-3000 Ion Chromatograph, ION450 Ion Analyser, Titrator TIM 900, double-beam UV-Visible spectrophotometer, etc.), originally on spreadsheets or printed on paper, have been modeled according to the scheme SensorML and in eXtensible Markup Language (XML). This modeling phase was carried out by the definition, for each instrument, of: identification code, sensor type, manufacturer, operator, classification, inputs, outputs, parameters, and characteristics. Metadata creation has been performed by GET-IT metadata editor, called EDI (Tagliolato et al., 2016)[7], which allows ease and friendly instrument registration (SensorML editing) through graphical user interfaces (GUI) and auto completion facilities linked to vocabularies. In particular, parameters definition have been borrowed from the terms present in the EnvThes[8] environmental vocabulary, in order to harmonize and semantically enrich the metadata with respect to the LTER community.

SOS service interfaces provided by GET-IT enable the interoperable sharing of sensor metadata; SOS operation InsertSensor registers the instruments in the repository, and simultaneously shares sensor information in the form of interoperable SOS services, so that it is possible to retrieve it in XML format through the SOS operation DescribeSensor.

The following figure (1) shows an example of SOS request in hypertext transfer protocol (HTTP), i.e. SOS DescribeSensor request for metadata of a DX-500 Ion Chromatograph, and the relative response in XML.

But GET-IT lets users search and view information on the instrument also in the form of a human readable html page. The following figure (2) shows the same description of DX-500 Ion Chromatograph of figure 1 in the user friendly way provided by GET-IT: information is the same but the presentation is deeply changed.

[6] http://geonode.org

[7] http://edidemo.get-it.it

[8] http://www.enveurope.eu/news/envthes-environmental-thesaurus

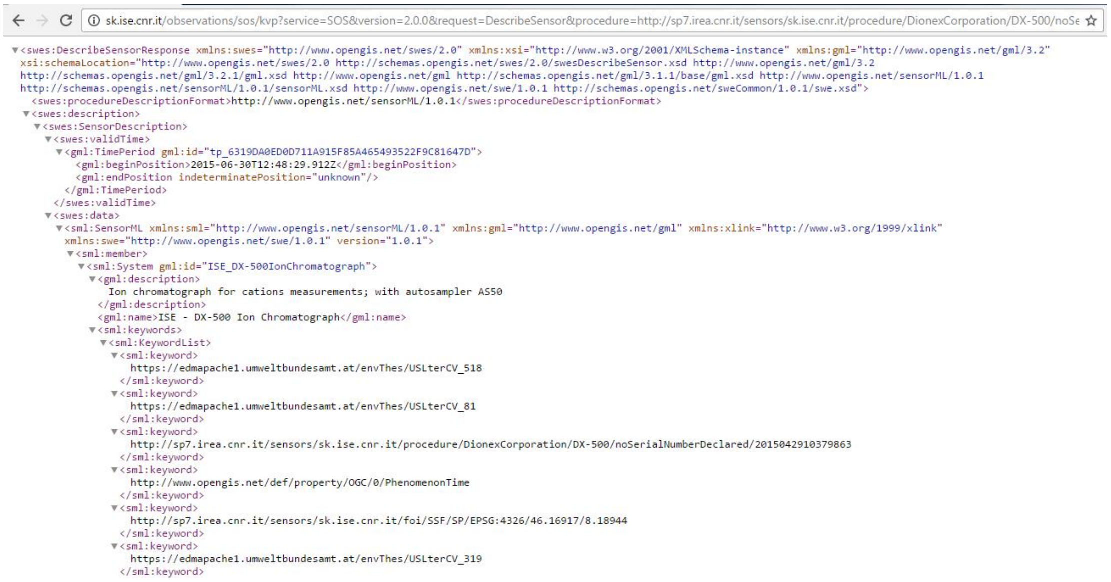


*Figure 1: SOS DescribeSensor request in HTTP for metadata of a DX-500 Ion Chromatograph and relative response in XML.*

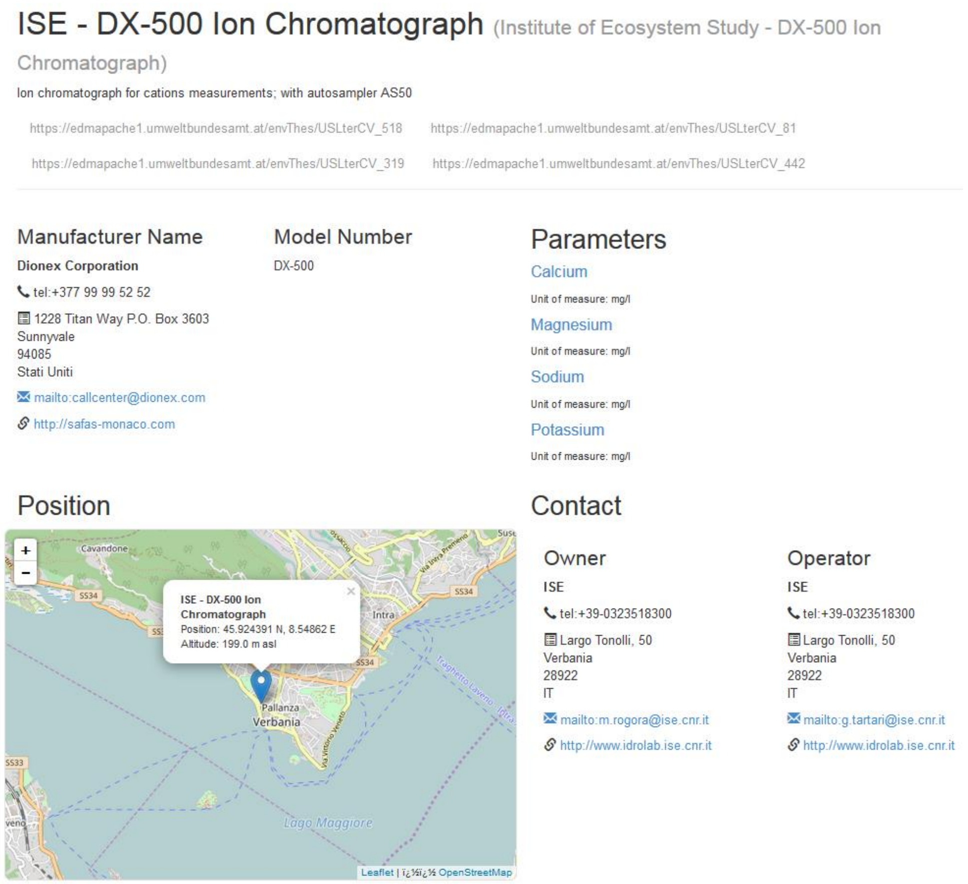


*Figure 2: HTML page, provided by GET-IT, for metadata of a DX-500 Ion Chromatograph.*

After sensor metadata registration and sharing, the next step is the inclusion of related collected data, originally on spreadsheets or paper. First, we modeled data according to O&M, after which each measure can be composed by the following elements: feature of interest, phenomenon time, result time, procedure, observed property, result. In this work, parameters measured (i.e. observed properties in O&M) have been borrowed from the terms present in the EnvThes vocabulary, in order to harmonize and semantically enrich the metadata.

Once modeled data following O&M, SOS services interfaces made available by GET-IT GUI enable their upload and interoperable distribution; SOS InsertObservation registers data to distribute them in the form of interoperable SOS services, GetObservation retrieves them in XML format. GET-IT also provides friend GUI for InsertObservation (fig. 3) and for GetObservation in an interactive web map (fig. 4).

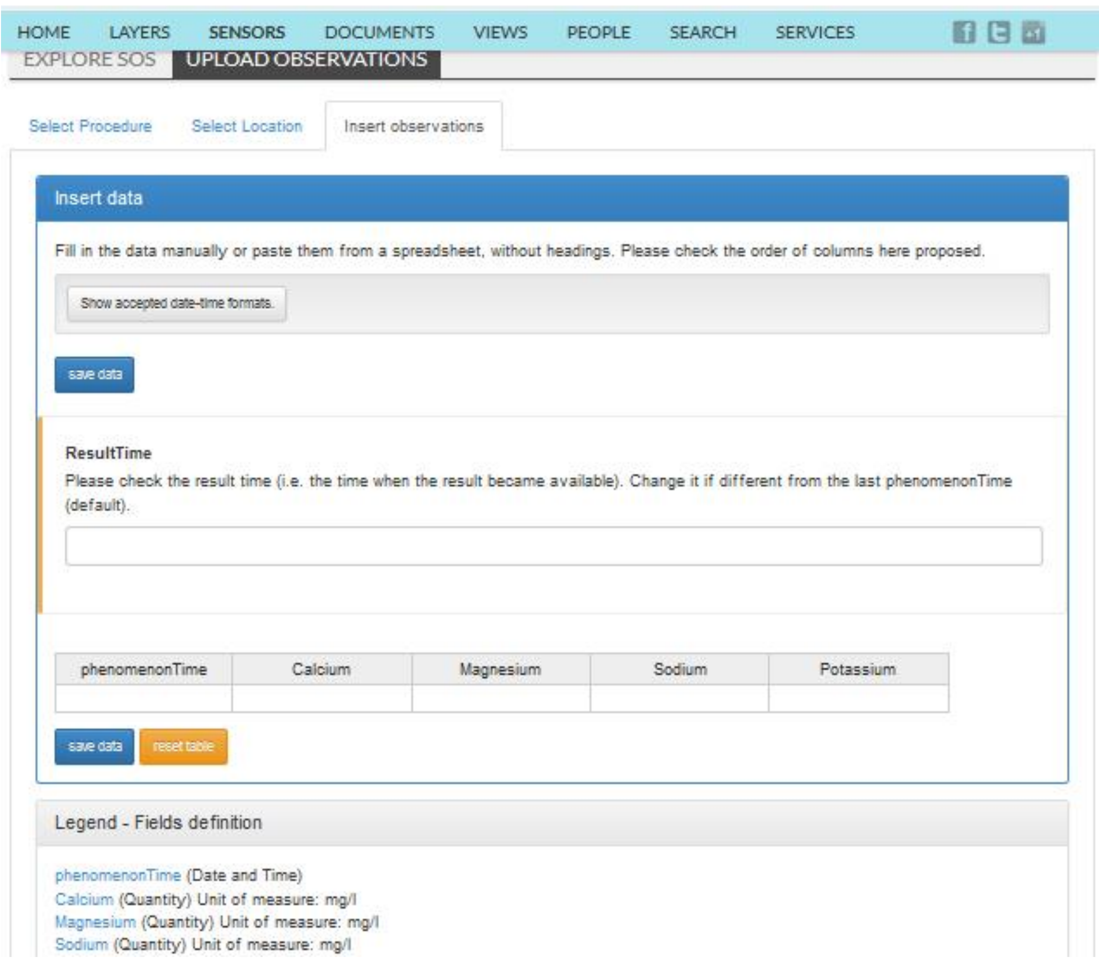


*Figure 3: GET-IT form for InsertObservation.*

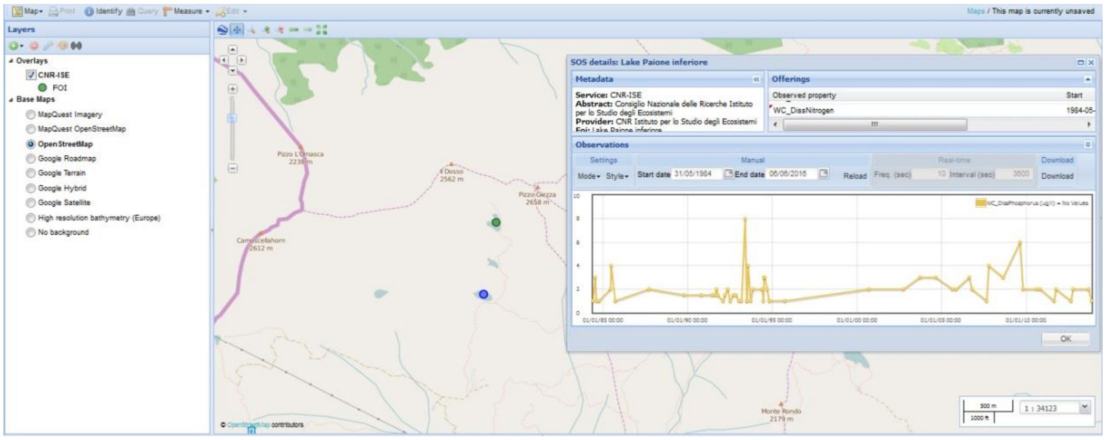


*Figure 4: GET-IT web map and graphical dialogue window to show results of GetObservation.*

## RESULTS AND CONCLUSIONS

GET-IT is a free and open source software suite, based on a collaborative development process, available on github (https://github.com/SP7-Ritmare/starterkit). It facilitates the creation of nodes of a federated SDI for an observational network and allows users to straightforward share on the web their observations, sensors and metadata, semantically enabled.

Using GET-IT and its SOS services, in the framework of the project NextData "Data-LTER-Mountain", LTER-Italy researchers making observations on mountain lakes shared seventy new observations and relative metadata, available at the CNR ISE SDI node. Observations of the sites Lake Paione Inferiore and Lake Paione Superiore cover a time range of about thirty years with a biyearly resolution; observations of the Lakes Scuro and Santo Parmense cover a time range variable between fifty and thirty years with a time resolution between monthly and quarterly. The result is considerable bearing in mind that, compared to the amount of observations and metadata at the start of the project, there was an increase of 488% in the number of observations and relative metadata distributed for a total number of records in the repository of approximately 60,000.

By this test we proved that, exploiting the existing GET-IT nodes, researchers can easily create, manage, edit and share sensors metadata and ecological observations of mountain lakes water and terrestrial mountain based on OGC SWE initiative; LTER-Italy researchers can distribute their own dataset in other projects, external to the network, in an interoperable way, avoiding unnecessary and harmful duplication, by means of web portals compatible with OGC standards. They also can use controlled vocabulary resources of EnvThes in SensorML and O&M, with semantic harmonization of output parameters for sensors and measured parameters for observations. In particular users can edit metadata of the sensor (SensorML), within GET-IT and througt EDI editor, where the sensor type, parameters and, unit of measure are derived from controlled vocabulary terms (more details in Tagliolato et al., 2016). EDI metadata editor allows the users to create XML documents, semantically enriched, througt simply graphical user interface.

In addition, LTER-Italy researchers can easily enable other nodes of GET-IT, consulting on-line software documentation[9], for manage, upload and share ecological observations of all network research sites.

The experiment carried out has helped LTER-Italy researchers to realize the benefit of metadating, which is often considered a useless and time-consuming transaction, though IT developers must still do more to facilitate the semi-automatic and simplified insertion of metadata as much as possible.

The future work focuses on improving the capabilities of GET-IT implementing semantic aware graphical user interfaces that enable to add, discover and display not only two-dimensional time-series but also other types of observations which are included in O&M schema, i.e. observations in three-dimensional water columns, wind directions and plots. In addition, GET-IT, in the context of eLTER H2020 project[10], will help advancing the development of European LTER infrastructures, by developing the SDI nodes of a distributed research infrastructure for 162 sites in 22 European countries that will provide data on long-term trends in environmental change.

## REFERENCES

Benthall B., Gill S. 2010. SDI Best Practices with GeoNode. In: Proceedings of Free and Open Source Software for Geospatial Conference (FOSS4G 2010).

Boots M., Robin A. 2007. OGC SensorML: Model and Xml Encoding Standard. Open Geospatial Consortium.

Box P., Rajabifard, A. 2009. SDI governance: Bridging the Gap Between People and Geospatial Resources.

Budhathoki N. R., Nedovic-Budic Z. 2008. Reconceptualizing the role of the user of spatial data infrastructure. Geojournal 72, 149–160.

Cox S. 2013. Geographic information – Observations and measurements - OGC and ISO 19156.

Finney K. T. 2007. A ‘bottom up’ governance framework for developing Australia’s marine Spatial Data Infrastructure (SDI). Data Science Journal 6, 64–90.

Fugazza C., Basoni A., Menegon S., Oggioni A., Pavesi F., Pepe M., Sarretta A., Carrara, P. 2014. RITMARE: Semantics- Aware harmonisation of data in Italian marine research. In: Procedia computer science (Vol. 33, pp. 261–265). Amsterdam: Elsevier. doi:10.1016/j.procs.2014.06.041.

Marchetto A., Mosello R., Rogora M., Manca M., Boggero A., Morabito G., Musazzi S,. Tartari G.A., Nocentini A.M., Pugnetti A., Bettinetti R., Panzani P., Armiraglio M., Cammarano P., Lami A. 2004. The chemical and biological response of two remote mountain lakes in the Southern Central Alps (Italy) to twenty years of changing physical and chemical climate. J. Limnol. 63: 77-89.

Shamsi S. A. A., Ahmad A., Desa G. 2011. Development of critical successful factors model for spatial data infrastructure implementation. In: International Colloquium on Signal Processing & Its Applications. doi:10.1109/CSPA.2011.5759875.

[9] http://getit.readthedocs.io/en/latest/tutorials/index.html

[10] http://www.lter-europe.net/elter/about

Steiniger S., Hunter A. J. S. 2013. The 2012 free and open source GIS software map – A guide to facilitate research, development, and adoption. Computers, Environment and Urban Systems, 39, 136–150. http://doi.org/10.1016/j.compenvurbsys.2012.10.003.

Tagliolato P., Oggioni A., Fugazza C., Pepe M., Carrara P. 2016. Sensor metadata blueprints and computer-aided editing for disciplined SensorML. IOP Conference Series: Earth and Environmental Science, 34(1), 12036. doi:10.1088/1755-1315/34/1/012036.

Winslow D. 2010. GeoNode Architecture: wrangling $100 million worth of open source software to make SDI building a walk in the park. In: Proceedings of Free and Open Source Software for Geospatial Conference (FOSS4G 2010).