

## Research, stakeholders and results transfer in the Pacific islands: first results of a multi-organisation project

**A. Couturier**<sup>a</sup>, **S. Pouillet**<sup>a</sup>, **I. Rouet**<sup>a,c</sup>, **Touraivane**<sup>a</sup>, **M. Mangeas**<sup>b</sup>, **M. Allenbach**<sup>a</sup>  
and **M. Menezo**<sup>b</sup>

<sup>a</sup> *University of New Caledonia - Noumea, New Caledonia*

<sup>b</sup> *Institute of Research for Development - Noumea, New Caledonia*

<sup>c</sup> *Rouetis R&D – New Caledonia*

*Email: isabelle.rouet@univ-nc.nc*

**Abstract:** In many Pacific islands, acquiring and accessing knowledge is a strategic issue. Local authorities especially have to manage development to guarantee resources supply, environmental and sanitary protection, and to face up to natural hazards, such as volcanoes, cyclones and recently sea level rise due to global warming. Many research programs are conducted in the Pacific islands, producing a wealth of scientific findings, but those are often scattered on different mediums (hard drives, CDs, intranets...), leaving local authorities with incomplete, hard to exploit results. This paper introduces a solution to help decision-makers access scientific data from a centralized, integrated source, so that data is easy to find, to identify and thus to use. This solution is being developed by two French organizations: the University of New Caledonia (UNC) and the Research Institute for Development (IRD). It uses a web-based, interoperable architecture. The major issues to overcome to provide such an architecture are: availability, accessibility, security, and metadata management (copyrights, authors, protocols ...), with all of these in a multilingual context. After several years of development, it is now time to assess the methods and tools used, what targets have been reached, what problems have been encountered along with their solutions, and what still needs to be done.

**Keywords:** *New Caledonia, interoperability, geographic data, storage, sharing*

## INTRODUCTION

Scientific data collected in the course of research programs is valuable to a wide audience, from political decision-makers to PhD students. Unfortunately a large portion of this data remains unexploited, dormant on a few mediums accessible to only few individuals, and is therefore vulnerable to damage and loss. It is sometimes faster to reconduct parts of a previous work rather than to find its reported results. In order to avoid such a waste of time and money, a centralized information system was developed, accessible through the internet, and relying on open standards. The goal of our project is long-term data storage along with easy information sharing among the major actors of development in New Caledonia and neighbor countries.

After the presentation of our project, the Spatial Data Infrastructure (SDI) that was built will be described and the first feedback discussed.

## 1. THE COLLABORATIVE PROJECT FOR SOUTH PACIFIC

University of New Caledonia (UNC, <http://www.univ-nc.nc/>) and Institute of Research for Development (IRD, <http://www.ird.nc/>) are the leading Research organisations in New Caledonia. Their scientists lead numerous research programs in the South Pacific, such as those for the National Center for Technological Research (CNRT, <http://www.cnrt.nc/>) on “Nickel and its environment” or for the South Pacific integrated Observatory for environment and terrestrial and marine biodiversity (GOPS, <http://www.observatoire-gops.org/>).

A wealth of results like GIS maps and publications are produced, and they need to be easily transferable to not only the first people concerned (usually those who financed the project), but also to all the people who would be interested in these results (and who have permission).

Major data portals exist but they are either commercial or subject to tight restrictions concerning data storage and use. The research programs’ conditions may not allow their resulting data to be deposited in such portals because of access restrictions. Combining their data with those coming from other sources is cumbersome because their search engines are restricted to the data they have.

This problem is not specific to the South Pacific, and as such the solution must be compatible to a large extent with existing solutions. Our choice is to develop a regional platform based on interoperability and open source principles. International and European technical specifications have to be respected. Different kinds of users need to work with the proposed services, and they have different expectations and hold various positions in the community. In this case, scientists, decision-makers and students are the main end-users considered.

## 2. INFORMATION SYSTEM ELEMENTS

In building our information system, the tools we chose had to be interoperable and easily accessible over the Internet through the HTTP protocol.

Interoperability basically means different systems can work together. It is possible only if all the systems involved follow the same set of rules, standards and protocols. The promise of interoperability is to have easier information sharing.

### 2.1. Geographic data sharing

We store our geographic vector data in a PostGIS database (<http://postgis.refrations.net/>). It can import common file formats used by data producers, especially ESRI Shapefiles and MapInfo files, and is compliant with the Open Geospatial Consortium, OGC’s “Simple Features Specification for SQL” (McKee, 2004; Herring J. R., 2010; Refrations Research Inc. 2011). This standard defines the storage and analysis of geographic data in a “spatially enabled” database. Imagery data from aerial or satellite acquisition and raster results are stored in an exchange format so they can be easily shared.

The data is then made available over the Internet in the form of dynamic map images as well as text-based formats (XML, json, KML...) with the help of GeoServer (<http://geoserver.org/>), and a Web Map Server (WMS) compliant web server.

The data itself is practically useless without its associated metadata. Storing and serving geographic data is half the problem: for users to access that data, they must first find it. Moreover they might need to know who created that data, in what circumstances, what organisation or research program did finance the

author(s), and what restrictions people willing to exploit that data must abide by. All this additional information about the data, the metadata, is essential because it gives the data demonstrable integrity.

We chose GeoNetwork (<http://geonetwork-opensource.org/>) to manage our geographic metadata: records can be stored, retrieved, consulted and edited following standards such as INSPIRE (European Parliament and European Council, 2007), ISO 19115/19139 and Dublin Core. Metadata records are requested through the visual web interface, or the “Catalog Service for the Web” (CSW) compliant web service.

## 2.2. Web interface

Our information system is designed to be accessed via the web (<http://grimm.univ-nc.nc/geoportail>), through a web browser. Users can build entire maps from several layers stacked on top of each other, retrieved from several WMS servers, navigated, and saved as PDF files. The web interface is built in HTML, CSS and Javascript. We use OpenLayers (<http://openlayers.org/>) to handle all the mapping operations such as the displaying of the map and ExtJS (<http://www.sencha.com/products/extjs/>) for the Graphical User Interface (GUI) part. GeoServer and GeoNetwork have their own administration interface, and provide their core services via the web.

We developed a multilingual approach: at present French, English and Bislama (Vanuatu) are available.

## 3. INFORMATION SYSTEM ARCHITECTURE

The main goal of the information system that was built is sharing. Sharing refers to the ability to expose and consume data. To do this, we need an infrastructure that is:

- ▲ Distributed;
- ▲ Interoperable;
- ▲ Standards based.

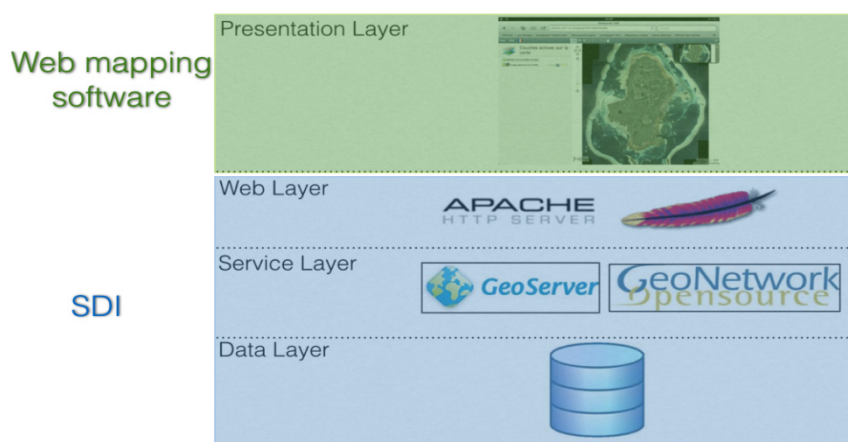
Distributed: partners involved are potentially located in many different places in the same country or around the world.

Interoperable: the distributed nature of this infrastructure requires that each partner exposes their data through common exchange protocols. It doesn't matter how the data is produced or maintained as long as each partner can access others' by using a well know language or protocol that each one agreed on in the first place.

Standards based: to ensure the whole solution can work with others, we rely on mainstream standards defined by well known organizations such as OGC (geospatial data standards) or W3C (web standards).

Networks: at the heart of this infrastructure lies computer networks. All data and communications between the partners' systems go through networks, be they local or global in which case we rely on the Internet.

Let's have a look at how the Spatial Data Infrastructure or SDI is structured at the UNC (fig.1):



**Figure 1.** The structure of our SDI.

Many layers compose the information system:

- ▲ the data layer, where resides all the data in many formats : shapefiles, relational databases or raster images;
- ▲ the service layer, which is the core of the system. There are two main components:
  1. Geoserver;
  2. GeoNetwork;
  3. the web layer is mainly composed of one component which is the Apache HTTP Server. This is the entry point for web based clients;
- ▲ the presentation layer is where all the interactions take place. We have developed our own web mapping software, which allows us to query our own SDI as well as our partners'. It is a central access point giving access to information systems potentially distributed far apart geographically, and belonging to different organizations.

SDIs can communicate. The main purpose of this communication is for metadata exchange which can be in realtime or scheduled to be imported in a local repository. The latter is also known as metadata harvesting. One SDI can also act as a proxy between its clients and other SDIs. For example our SDI enables WMS cascading. This allows us to define layers coming from our partners' SDI in our own SDI and then our web mapping application can query our SDI directly (fig. 2).

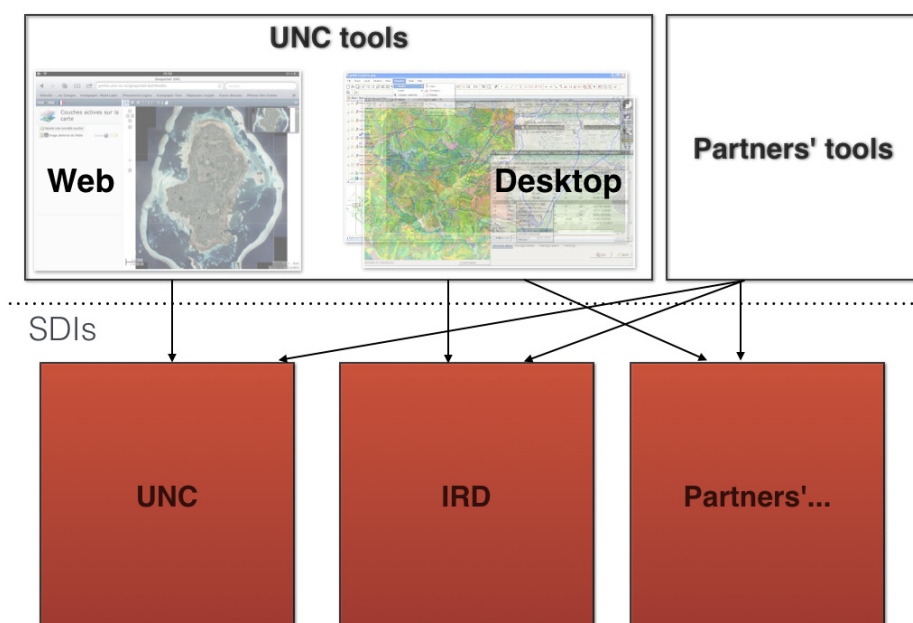


Figure 2. Different client types fetching data from various SDIs.

## 4. DISCUSSION AND CONCLUSION

### 4.1. User feedback

Our SDI has been available for about three years, with additional features and new data added gradually. We have inserted about 150 geographic layers, in both vector and raster formats, including high quality aerial photographs, for the most part coming from several research programs conducted by the researchers at the UNC. It has been especially useful for a PhD student in her research thesis on the islands of Wallis and Futuna, and is of particular interest to our most active partners (a government institute and a private company).

## 4.2. Benefits of our approach

The following benefits contribute to our main goal, which is to make scientific data highly accessible and exploitable for an audience as wide as possible:

- ▲ web applications are easy to distribute. Users do not have to download and install anything, or deal with software dependencies. Our information system can be used on any operating system, all that is required is a compliant web browser and an Internet connection;
- ▲ no identification or subscription are necessary for accessing the data that is publicly available;
- ▲ every geographic layer and every scientific document can be retrieved with keywords, thanks to metadata and search engines;
- ▲ users have access to not only our own data, but also our partners', since interoperability allows us to be part of a greater network of WMS servers;
- ▲ only the most up-to-date data is available with data added, updated and visible to users instantly;
- ▲ industry-leading mapping applications are compatible with the infrastructure, for example ESRI ArcMap, MapInfo or QGIS can fetch data in realtime from GeoServer, people can then use all of the advanced features from experience they have with these tools;
- ▲ the software servers and web interface all support multiple languages, especially English and French, which is important in order to reach a broader audience, and let as many people as possible find, use and reference our data and for now, the user interface only is internationalized, but not the data nor the metadata yet;
- ▲ our SDI is compatible with INSPIRE, the European directive voted to "ensure that the spatial data infrastructures of the Member States are compatible and usable in a Community and trans-boundary context".

We chose open-source over proprietary solutions because they are free of charge, meet most of our needs, and are regularly updated with bugfixes and new features. They are widely used, and online documentation and answers to common problems are thus easier to find. Our choice was also influenced by GeoServer being a Java Servlet application that we host in a Tomcat Servlet container, and Tomcat administration is part of our team's combined skillset.

## 4.3. Difficulties and limits of our approach

Web applications depend on a reliable Internet connection, but Pacific islands tend to have lower connection bandwidths, if at all. This can limit the use of our information system. For instance when working in Vanuatu, the common Internet connection speed we had access to was limited to 32kbits/s.

Interoperability works well when all the nodes of a network are fully compliant with the same standards. Some tools however advertise themselves as being so while they are not. In that case, genericity is broken, the main advantage of interoperability is lost and additional work is required to permit communication with those nodes.

Relying on distant nodes that are out of our control is a trade-off. On the one hand they relieve us from costly, time-consuming installation and maintenance. On the other hand, when one of them fail, not much can be done on our side to solve this problem. Features that are tightly integrated, i.e. relying on multiple external nodes, are vulnerable to multiple potential points of failure. For example, a map made of geographic layers coming from 4 different WMS servers depends on each of them working properly. Should a single WMS server become unavailable, the map would then be useless. Some users, unaware that the data is fetched in realtime from other providers, hold us responsible for these occasional outages.

Compared to their desktop counterparts, web applications are more limited in terms of speed and functionality. Even though modern web browsers are becoming increasingly fast at rendering pages and executing scripts, a web application running in a browser cannot be as fast as a native application. Some limitations make things either more complicated for the development team, for instance setting up a proxy to bypass the same origin policy for AJAX queries, or simply impossible tasks/operations, such as client-side file input / output (although HTML 5 may fill that gap).

To further increase the challenge mentioned in the previous point, web browsers have different levels of web standards compliancy, while older versions are still in use. This forces us to use the lowest common denominator, or spend time developing features accessible to only a fraction of our user base. Fortunately this is largely mitigated by the external libraries we rely upon, such as OpenLayers, jQuery and ExtJS.

Data access restriction is generally not clearly defined in the OGC standards (eg WMS), therefore we use custom solutions, which are suboptimal (http authentication), or not compatible with other clients.

#### **4.4. The SDI for the Government of Vanuatu**

We have been collaborating with the IRD to make our SDI available to the Government of Vanuatu. The SDI was installed in one of their departments in the middle of 2011, for their own needs.

The people working in that department found it difficult to keep their geographic data synchronized with the other departments: it was difficult to find the data itself, to keep track of where it came from, and to keep it up to date. It was the typical situation in which a SDI like ours, i.e. that is centralized and built on standards, would be useful. The fact that our SDI is also compatible with the owner GIS software they use was a real advantage. We translated the web interface in Bislama, the language officially spoken in Vanuatu, to make results transfer easier.

#### **4.5. Future improvements**

Adding more data and of better quality: this is a continuous task, and is paramount to widening our audience.

Right now we add geographic layers in our SDI manually. And when a single geographic feature within a layer has to be updated, we remove the old layer completely, and insert the new one. This workflow is time-consuming, error-prone, and does not scale as the number of data providers increase. Instead, we should let users themselves edit their own data. We plan to use Web Feature Service Transactional (WFS-T), which is another OGC standard for the editing of features within a geographic layer. GeoServer's specific web services will allow layers to be added and edited.

We would like to add simple spatial analysis with the OGC's Web Processing Service (WPS) standard. For example to let our users simulate the rise of the sea level on the map, and compute how many people or how many roads would be affected.

### **ACKNOWLEDGEMENTS**

We acknowledge the financial support from the University of New-Caledonia as well as the collaboration of UNC and IRD researchers for sharing scientific data.

The authors would also like to thank Ni-Vanuatu colleagues for Bislama translation.

### **REFERENCES**

European Parliament and European Council (2007). 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). *Official Journal of the European Union*, L 108/1-14, April 25.

Herring J. R. (2010). OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 2: SQL option. *Open Geospatial Consortium, Inc.*, OpenGIS® Implementation Standard OGC 06-104r4, 111 p.

McKee, L. (2004). The OGC – Organization Offering Unique Benefits. *Open Geospatial Consortium, Inc.*, 1-2.

Refractions Research Inc. (2011). PostGIS 1.5.3 Manual. *Refractions Research Inc.*, 21-23.