The Andean Information System for Disaster Prevention and Relief: A case study of multi-national open-source SDI

Martín Molina[1], Salvador Bayarri[2] and Rubén Vargas[3]

[1] Universidad Politécnica de Madrid, Spain: martin.molina@upm.es
[2] IVER Technologies, Valencia, Spain: salvador.bayarri@iver.es
[3] PREDECAN, Lima, Perú: rdvargas@comunidadandina.org

Abstract

The Andean Information System for Disaster Prevention and Relief (SIAPAD) is an initiative promoted by the Andean Committee for Disaster Risk Prevention and Relief (CAPRADE), providing tools for discovery, visualization and access to data related to disaster risk, available in different organizations within the countries of the Community of Andean Nations (CAN): Bolivia, Colombia, Ecuador and Perú. The architecture of SIAPAD takes advantage and leverages current initiatives in the Andean Community to develop a multi-level Spatial Data Infrastructure.

SIAPAD includes four national GeoRiesgo portals which allow to search and display relevant information to support decision making during the different phases of disaster risk management (prevention, mitigation, preparation, relief). Each national catalogue harvests metadata from about nine organizations in each country, using the CS-W standard. The geoportals, as well as most of the source SDI nodes have been implemented using an open-source software stack.

1. Introduction

Effective decision making and planning in the field of disaster risk management requires information on hazards, vulnerability conditions and the human community living in the target area, including knowledge of the legal framework, information systems, human and material resources, as well as methodological tools to support the implementation of disaster risk management policies. It is a truly interdisciplinary and inter-institutional effort for which the comprehensive SDI approach is well suited for, but that requires a strong coordinated effort.

1.1. Situation in the Andean region

The Andean region presents specific conditions that make information management a priority in the design of strategies for the reduction of disaster risk:

- Recurrence of potentially disastrous natural phenomena (earthquakes, volcanoes, floods, etc.).
- Growing vulnerability as large urban populations expand without appropriate planning.
- Government and other organizations in the region (dealing with civil defence, meteorology, seismology, etc.) collect significant information. However, it is usually difficult to access these sources or even know of its existence. Many institutions lack the policy, equipment, software and human resources needed to implement appropriate

ISBN 978-0-620-42117-1

information systems.

Therefore, it was evident a need to promote the development of information systems in the region to increase the visibility and access to the data. The SDI architecture seemed the best choice because of its building on open services and standards, and its distributed nature. In fact, at the time the SIAPAD project started, some national and regional SDI initiatives had been on planning stage or limited execution for some time, but had not really taken off.

1.2. The CAPRADE and the PREDECAN project

The Community of Andean Nations (CAN) developed the Andean Strategy for Disaster Prevention and Relief as a legal framework implemented by the Andean Committee for Disaster Prevention and Relief (CAPRADE). Within this framework, the creation of an Andean Information System for Disaster Prevention and Relief (SIAPAD) was defined as part of the Thematic Area 2: "Information, research and development". This initiative is supported by PREDECAN, a project financed by the European Union and the CAN, whose purpose is to build capacities for disaster risk prevention in the Andean Community, in direct relation to the Hyogo Framework for Action 2005-2015 (Hyogo, 2005).

Activities regarding the definition of SIAPAD's graphic interface and functional design began in late 2005. A preliminary diagnostic was carried out in order to assess the region's state of development concerning risk management data production and publication. The study mainly looked at existing national or transnational information systems, data availability, and the technology used to serve data over the Internet. The results of the diagnostic were validated in participatory regional events, and used to define the main features of the SIAPAD architecture and access portals, with input from data providers and users.

In the last year, the implementation of the SIAPAD by IVER Technologies has involved capacity building activities (training on software –mainly open source– for publication of web services and metadata cataloguing), software development of the GeoRiesgo national geoportals and hardware provision and installation. The demonstration portal in Perú was presented in February (see Vargas, 2008) and currently the four national geoportals are fully operational and accessible from *http://georiesgo.net*.

2. The SIAPAD architecture

SIAPAD is a distributed network of information nodes belonging to different institutions, according to international recommendations for data interchange (Davies, 2004). It is conceived as a thematic Spatial Data Infrastructure following international standards for geographical information services (OGC 2005a, ISO 2006, INSPIRE 2007), allow interoperability among the nodes and also with other SDIs. For most users, the access points to SIAPAD are the four national nodes called *facilitators* (see Figure 1). These nodes host the GeoRiesgo web applications, conceived as thematic geoportals oriented to different user profiles interested in disaster risk management, or disasters in general.

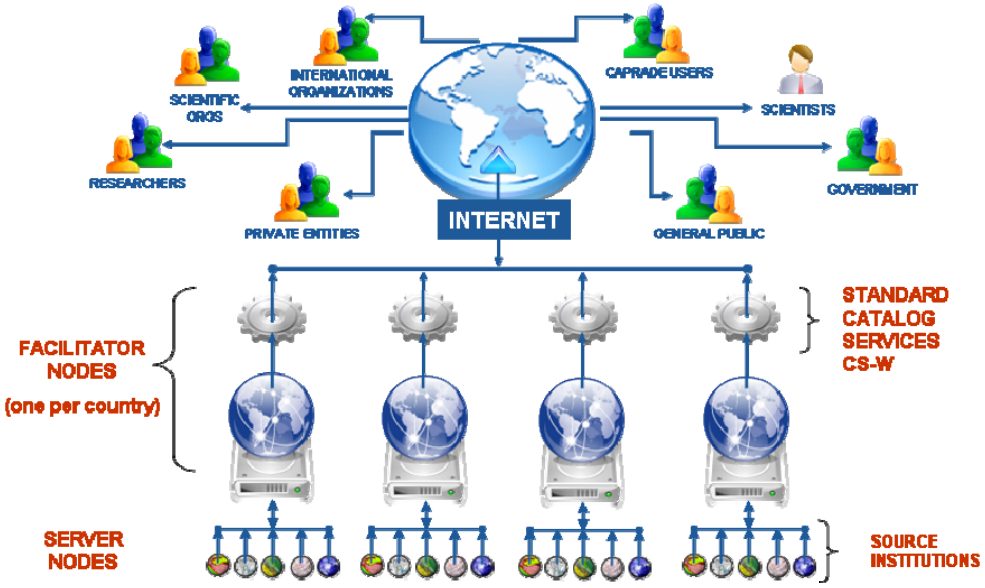

Figure 1. The SDI architecture of the SIAPAD, with emphasis on facilitator and server nodes

As seen in Figures 1 and 2, the national nodes feed from the data and catalogue nodes created in 37 institutions in the CAN. These catalogues are harvested automatically on each national node via CS-W. In turn, these national catalogues are also accessible via CS-W to other SDI clients (see Figure 2). Therefore, facilitator nodes play the role of thematic gateways to the SIAPAD, via the GeoRiesgo portals, but also the role of information integrators at the national level.

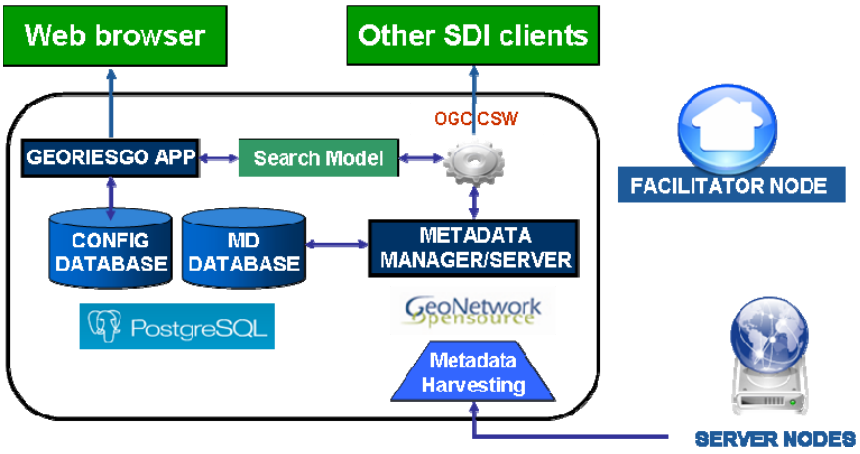

Figure 2. The architecture of a facilitator node of the SIAPAD

The GeoRiesgo application uses the harvested metadata at the facilitator node to quickly perform searches at the national level, although it can also connect to the other facilitator nodes for multi-national searches, and to any other external catalogue configured by the node administrator. It is also possible to perform manual metadata insertion directly in the national node, an option used only in the case that some metadata of interest to SIAPAD can not be hosted and harvested from other nodes.

This design has proven flexible enough to interoperate with the diverse software platforms and data management formats used by institutions in the Andean Community, and has also allowed

linkage with the results of separate disaster-related projects which were not publishing standard map and catalogue services before SIAPAD.

2.1. Open source implementation

The software for the SIAPAD national facilitator nodes has been developed by integrating existing open source tools, strongly compliant with open standards. Geonetwork Opensource (see Geonetwork) is used as metadata manager and CS-W server. Both Geonetwork and the GeoRiesgo portals use the PostgreSQL database for metadata and configuration storage. The map viewer component in the GeoRiesgo portal has been developed by extending components from the Community Mapbuilder library (see Mapbuilder).

Furthermore, because most of the institutions providing the disaster-related information were not publishing standard services before the SIAPAD implementation, a series of national workshops were conducted, and a Moodle-based support platform established to help these institutions build their SDI nodes with the same open source stack used at the national nodes. As a result, there are today 26 instances of fully operational SDI servers implemented with open source software in the CAN.

2.1. Integration of other regional projects

In addition to connecting a large number of institutional servers in the CAN, the standards-based architecture of the SIAPAD has also been key to the integration of the information products resulting from related regional projects. The most important of these are:

DesInventar (Desinventar, 2000). Developed by the Observatorio Sismológico del Suroccidente Colombiano –OSSO– and La Red, a network for social studies in disaster prevention in Latin America. DesInventar has developed methodologies and software to collect, process and analyze data about the effects of diverse types of disasters. The statistical inventory data have been published as MapServer WMS maps, summarizing the disaster impacts by phenomenon and type of effect. These WMS maps have been catalogued and a search service published with Geonetwork opensource.

BiVa–PaD. A network of digital libraries about disaster prevention relief, implemented by the Regional Disaster Information Center Latin America and the Caribbean (CRID). It has compiled, digitized and published a large library of disaster-related laws, reports, educational materials and other documents. The full catalog of this project has been exported to the ISO 19139 metadata schema and imported to Geonetwork servers to be accessible to the SIAPAD and the GeoRiesgo portals.

Geosemantica. Developed under The Andean Multinational Project “Geosciences for Andean communities (PMA:GCA)”, funded by the Canadian International Development Agency, the project has provided geo-scientific institutions in the region with training and a powerful web-based collaborative workspace where users can easily publish documents and WMS maps using the integrated MapServer.

3. The GeoRiesgo portals

The SIAPAD project aims at making risk management relevant information available to a wide range of users (government planners, scientist, educators, journalists, etc.), who in turn can use it to support decision making during the different phases of the disaster risk management process (prevention, mitigation, preparation, relief). The national GeoRiesgo portals (see Figure 3) increase the visibility of the information about disasters and risk management produced by the different institutions, contributing to a more effective execution of their tasks.

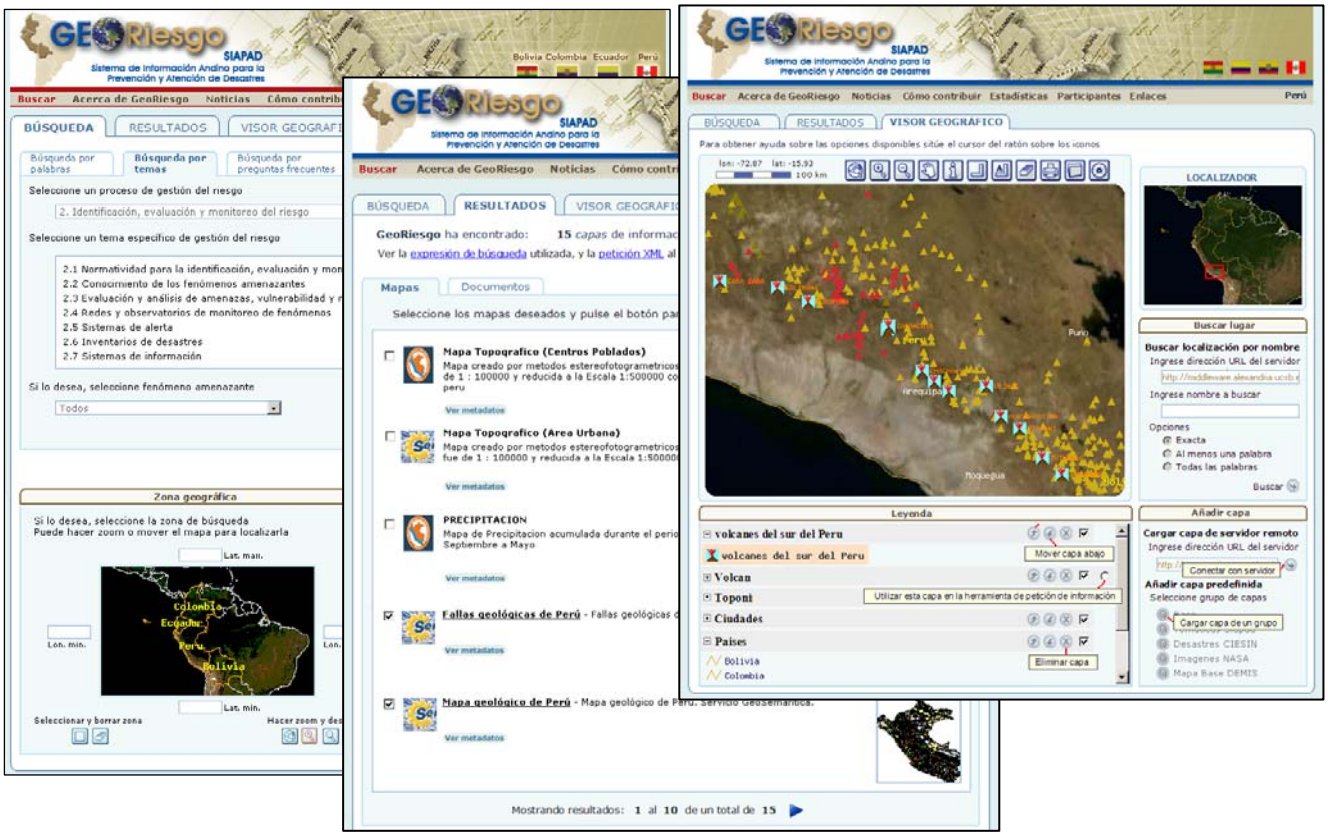

Figure 3. View of GeoRiesgo portal, with Search, Results and Viewer tabs

The conceptual model guiding the GeoRiesgo graphic interface and functional design emphasizes the need to establish a closer link between the users and the producers. It aims at defining not only the type of information product that such a system should deliver but also the most appropriate way that information should be presented to the user. For this purpose, it presents different mechanisms for data searching and retrieval designed to assist different types of users.

- *Word search*. As in conventional search engines, this method allows the user to find metadata containing one or more words. In the case of GeoRiesgo, this search is enriched with the use of synonyms for those terms related to disaster risk management, as it is explained later on.
- *Thematic search*. This method provides the user with a hierarchy of themes related to disaster risk management. The user selects the particular theme of interest and obtains according results. This method is very effective for users who are already knowledgeable about disasters and the risk management processes.

- *Task-oriented search.* This method provides each type of user (researcher, general public, territorial planner, etc.) a set of prefixed typical questions grouped in usual tasks. This is especially useful for users who are not familiar with the vocabulary and concepts of disasters and risk management.

The users can also constraint any of the above searches by providing spatial or temporal bounds, or by restricting the results to a certain phenomenon selected from a standard categorization.

After the search selection, the *Results* tab opens automatically to show the matching information products, grouped in two subtabs, one for map services and another one for documents (images, web pages, anything accessible via a URL link). The user can click on the document links to view them, or select map services to be added to the *Viewer*, which then opens automatically as a separate tab.

One key feature of the portal is the easy navigation between the Search, Results and Viewer tabs (see Figure 3), which allows users to go back to the results or search parameters to select more results and add them to the viewer, or to perform new searches without losing the contents of the current map.

3.1. The SIAPAD search model

The search system is the core of the GeoRiesgo portal and the SIAPAD as an information system focused on a certain domain. SIAPAD's facilitator nodes include a common dictionary of keywords related to disaster risk management and their synonyms. The key guiding idea of the SIAPAD search system is that all user queries are transformed into a search expression built from the keyword dictionary, using parenthesis and Boolean operators (conjunction = AND, and disjunction = OR). This Boolean expression is then passed onto the CS-W search query via the OSG Filter XML (OGC 5b, OGC 07) to be compared with selected fields from the metadata records, most notably the title, abstract and keywords.

The search models must, then, provide a way to generate appropriate search expressions based on the simple user choices in each of the three search methods. This is carried out by using look-up tables containing keywords and keyword-based expressions for phenomena, themes and questions used in the search methods (see Figure 4). In the case of the thematic search, the user can select a specific theme within a risk management process, and an optional phenomenon. Each phenomenon has an associated keyword, given by table P, and each theme has an associated keyword expression given by table T (the expressions in table T and Q are formed using keywords from the common dictionary). The base search expression is then built as a conjunction of the keyword from the chosen phenomenon and the expression from the given theme. A similar sequence is followed in the case of the task-oriented search.

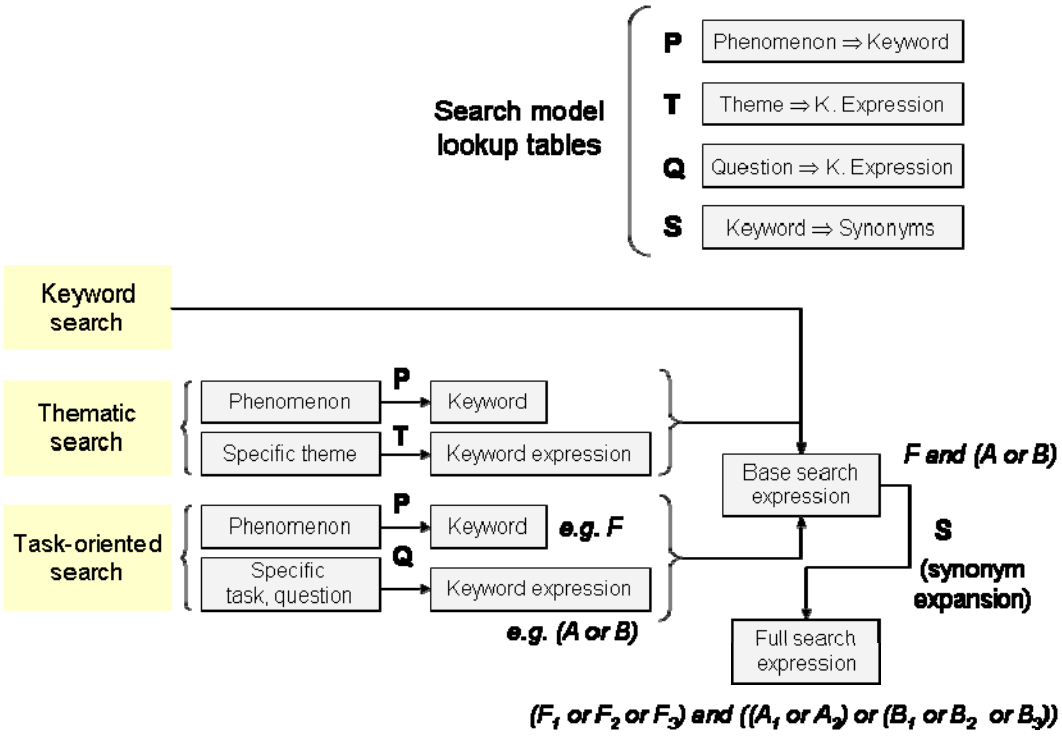

Figure 4. Generation of the search expression in GeoRiesgo

A problem arises in the search process because many keywords have synonyms, which maybe used in the metadata records (for instance, the word “earthquake” has up to six synonym terms in Spanish), and therefore the base search expression may not find many relevant records. To consider synonyms in the search, GeoRiesgo uses a technique known as synonym expansion, which consists in substituting each occurrence of a keyword in the base expression by a disjunction of all its synonyms (this is shown in the figure when “F” is changed into “(F1 or F2 or F3)”, and so on). This expansion is available in some search engines like Lucene, but unfortunately is not directly supported by CS-W or by its implementations on the server side –a GeoNetwork implementation is expected in the future-, which means GeoRiesgo as client application has to take care itself of performing the expansion on the search expression used in the CS-W query. See (Buttcher, 2004) for a similar approach in the biomedical domain.

Extensive testing has been performed to evaluate, for each expression, both the *quality* of the search (the percentage of successful hits in the result set) and its *coverage* (the percentage of information products, significant for each theme or question, that are actually found). These tests have been iteratively used to improve the search terms and synonyms, as well as the Boolean expressions used in the thematic and assisted search options.

4. Results

With four national geoportals and 26 institutional SDI-enabled servers (the total number of participating institutions is 37, since some of them share the hosting of their data and metadata), the SIAPAD system has achieved, in some measure because of its promotion of open-source software, its main goal of increasing the visibility of information useful for disaster risk management in the context of the Andean Community. The development of this system contributes by providing tools

and an information-sharing philosophy to address a problem with a great social and economic impact.

In addition, there is a longer-term impact of using open source software in the Georiesgo application and most of the SIAPAD nodes. This impact is seen first in the sustainability of the software and human resources, with capacity building and technology transfer to the participating institutions having also played an important role, and secondly in the strong leverage provided to regional SDI initiatives and general-purpose data availability, which should have strategic benefits and hopefully coalesce around stronger multinational legislative frameworks similar to EU's INSPIRE.

References

Buttcher S. 2004. 'Domain-Specific Synonym Expansion and Validation for Biomedical Information Retrieval'. *Thirteenth Text Retrieval Conference*. <http://trec.nist.gov/pubs/trec13/t13_proceedings.html>

CRID, BiVa-PaD project web site, Regional Disaster Information Center for Latin America and the Caribbean (CRID). <http://www.crid.or.cr/crid/PREDECAN/index.htm>

Davies R.H 2004. 'Towards an all-hazard, all-source, worldwide disaster information network; surviving and impact'. *Planetary Defense Conference*: Protecting Earth from Asteroids. Orange County, California.

Desinventar, 2000. OSSO, La Red. Project web site. <www.desinventar.org>

Geonetwork opensource. Community web site. <http://geonetwork-opensource.org>

Hyogo Framework for Action 2005-2015: Building the Resilience of Nations and Communities to Disasters. World Conference on Disaster Reduction 18-22 January 2005, Kobe, Hyogo, Japan. < http://www.unisdr.org/eng/hfa/hfa.htm >

INSPIRE 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)'. <http://www.ec-gis.org/inspire/>

ISO 2006, 'ISO/TC 211 Geographic Information/Geomatics' web site. <www.isotc211.org>

Mapbuilder. Community Mapbuilder project web site. <communitymapbuilder.osgeo.org>

OGC 2005a, Open Geospatial Consortium web site. <www.opengeospatial.org>

OGC 2005b. 'OpenGIS Filter Encoding Implementation Specification'. OGC web site. <http://www.opengeospatial.org/standards/filter>

OGC 2007, 'OpenGIS Catalogue Services Specification v2.0.2'. OGC web site. <http://www.opengeospatial.org/standards/cat>

Vargas R. et al. '*The Andean Information System for Disaster Prevention and Relief: SIAPAD an initiative for the development of a thematic SDI*'. Global Spatial Data Infrastructure conference GSDI 10, St. Augustine, Trinidad. February 25-29, 2008.