Open access and open data on natural disasters.

De Giusti, Marisa Raquel, Villarreal, Gonzalo Luján, Nusch, Carlos y Pinto, Analía Verónica.

Cita:

Open access and open data on natural disasters

Marisa R De Giusti[1,2], Gonzalo L Villarreal[2], Carlos J Nusch[2], Analía V Pinto[2], Ariel J.Lira[2]

Abstract

After the flood which devastated the city of La Plata (Argentina) in April 2013, a survey of the region water situation was conducted which revealed that relevant studies and projects related with this area were scattered and insufficiently visible. This prompted academic, scientific, technologic and governmental institutions to take action and work together to compile, organize and disseminate available materials in an unprecedented initiative in the region in response to an environmental crisis. Thus, a dedicated collection was created in the SEDICI, the institutional repository for the Universidad Nacional de la Plata, with the subsequent launch of the Observatorio Medioambiental—La Plata (OMLP), a repository for data on environmental research. This document describes these initiatives, their main characteristics and their current progress.

Keywords

Floods, open acces, digital repositories, environmental observatory

[1] Independent researcher, Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, ARG

[2] Servicio de Difusión de la Creación Intelectual, Universidad Nacional de La Plata, ARG

Corresponding author: Gonzalo L Villarreal, Servicio de Difusión de la Creación Intelectual, Universidad Nacional de La Plata. Calle 49 y 115 SN, Departamento de Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de La Plata, Argentina.

Email: gonzalo@prebi.unlp.edu.ar

1. Introduction

On April 2, 2013, the city of La Plata, capital of the Buenos Aires province, was devastated by the worst flood in its history. A storm of unprecedented intensity in the region took a death toll of 89 human lives, thousands of evacuated residents, a city that was paralyzed for weeks, and economic losses adding up to millions (Cipponeri et al., 2014). This catastrophe had a strong impact on the public opinion and prompted to action thousands of volunteers from across the country, who helped the affected population return to their homes, and later rebuild or recover their lost property. The whole population was shocked by witnessing how a modern city, architecturally planned in the nineteenth century to be a governmental hub for the main province in the country, was devastated in a matter of hours by intense rainfall.

In the aftermath of the disaster, there was a need to take a time to analyze and understand what had happened, and how it had escalated to such a catastrophe. It was concluded that the storm had a volume beyond anything on existing records (Benítez and Scarpati, 2005; Romanazzi, 2015) and that the water infrastructure and crisis management capacity of the city wasn't enough for a storm of such magnitude, not even for strong storms of a much lesser intensity (Facultad de Ingeniería, 2013). It was also noted that every piece of information, data, studies, reports, and all research related to the water and

environmental situation of La Plata was scattered, as most of it was out of reach for the general public, and documents which could be accessed showed a level of disorganization and lack of proper exhibition that made it practically useless for conducting further studies and projects.

One of the most immediate results of this analysis was the creation of a large amount of infrastructure works undertaken by the La Plata municipality, the Buenos Aires province, and the national government, with the aim of improving and strengthening the city's ability to withstand these weather related catastrophes (Romanazzi and Urbitzondo, 2014). As an alternative to the work done by government organizations, a series of initiatives came forth from local academic and research institutions looking to gather, simplify, optimize, and ensure open access to all the information available at that point.

This work describes actions and initiatives created by local academic and scientific institutions, in which SEDICI, the institutional repository of the UNLP, engages as a stakeholder in charge of managing and providing access to knowledge related to water infrastructure, risk management and analysis, crisis response, and environmental studies for the region. It also describes the subsequent creation of an environmental observatory capable of hosting research data gathered from projects related to these topics, with the

collaboration of the National University of La Plata and other relevant institutions on a regional and national level.

2. Water Emergency: a dedicated collection in the institutional repository

In the previously mentioned conditions, the National University of La Plata took as a first measure to task their central repository, the Intellectual Creation Diffusion Service (SEDICI), with the creation of a dedicated space capable of hosting and disseminating works related to local water and environmental infrastructure. For this, faculty members and researchers were asked to compile and store their research material in this new space, designed as a community inside the repository, and called "Emergencia Hídrica [Water Emergency]" (Servicio de Difusión de la Creación Intelectual, 2013a).

2.1 Goals and objectives of the repository

SEDICI is the institutional repository for the National University of La Plata, created to host, preserve, and offer visibility to the production of its own academic units and additional branches. Some of its main goals are: a) to manage, preserve, and disseminate the intellectual production of the university, together with the socialization of the knowledge resulting from the findings of its researches and studies; b) foster the visibility and use of said academic and scientific production; c) analyze and evaluate the impact and excellence of

research financed with public funds, and d) promote the adequate use of the property right of authors over their own work and the rights of the institution over its production (Villareal et al., 2008).

2.2 Organization of the repository

Given the large amount of works and the typological complexity of the different resources, SEDICI organizes its archive in different communities, sub-communities, and collections in order to provide users with fast, organized, and intuitive access the diverse records. The distribution in communities and collections in the repository reflects the structure of academic units in the university and also certain text genres (such as books or journals). Another set of collections is based on specific preservation or digitalization goals to provide access to works from institutions related to the UNLP, e.g. the Academia Nacional de Agronomía y Veterinaria (ANAV) or the Red de Universidades con Carreras en Informática (RedUNCI). Most of these special collections contain distinct works that are prominently non-textual, as is the case of the Open Educational Resources (learning objects), the UNLP Museum Network (in which physical objects are stored, represented by images and sound clips), the Vigo Experimental Art Center, Nanotechnology, and so on.

From a technological standpoint, after a migration in 2012 from a custom system, SEDICI uses the DSpace software as its main digital resource, while

offering a series of services including document digitalization, online dissemination, digital preservation, and support for teaching and interoperability with other systems, among other things (Texier, et al., 2012; De Giusti et al., 2013).

2.3 Open Access policy

SEDICI adheres to the Open Access guidelines for the definition of usage licenses of hosted works, and plays an active role in Open Access training, awareness, and promotion activities both at a local and international level through workshops, postgraduate courses, an online blog, social networks, conferences, and scientific meetings. The repository promotes the creation and dissemination of works under Creative Commons licenses, giving rise to many initiatives, including the UNLP Journals Portal, the UNLP Conferences Portal, and the UNLP Books Portal.

With the new collection called "Emergencia Hídrica", SEDICI opens its self-archiving circuit to accommodate works related to water emergency matters, modifying its online portal to highlight this new collection inside the repository (Figure 1), with a view to assimilating existing works jointly with faculty members and researchers to coordinate efforts and immediately disseminate all the results from previous projects and researches, as well as from those created since the flood.

2.4 Self-archival procedure

The self-archiving circuit is a procedure that allows every professor, researcher, or head of a project to upload works deemed relevant to this collection or another, through the use of a simplified web assistant. This assistant prompts the user in charge of the self-archiving process to input data for the document (authors, title, source organization, abstract, keywords), depending on the type of work, which can be an article, book, conference object, physical object, etc. After this step, the system allows the user to choose a Creative Commons license by answering two questions that will automatically determine which of the six different usage licenses will be assigned to the work.

Once the self-archiving process is completed, repository administrators receive the information and verify that all data is correct, normalize some of the metadata, incorporate additional metadata needed for cataloguing activities to ensure resource preservation, while using the uploaded files to generate versions of the work that are best suited for its preservation (i.e., PDF/A documents). Once this verification is complete, the work is finally published into the repository. The Emergencia Hídrica collection currently contains around 130 works including journal articles, books and book chapters, reports, dissertations, and conference papers.

2.5 Community training

On May 7, 2013, an intensive one-day workshop took place in the UNLP, in which researchers and specialists of different universities gathered to join their efforts and knowledge towards understanding the recent events and preventing similar situations from happening in the future. Part of the staff of the repository attended this workshop to give a series of short presentations, which included a video self-archiving, made available in the repository's YouTube channel (Servicio de Difusión de la Creación Intelectual, 2013b), to show how the repository worked, its collections, and the convenience of self-archiving as a way for researchers to quickly upload their works and later find them in a single location for future research and reference.

After the one-day workshop, the collection gained a large number of works that were incorporated to the existing material in the repository, which, while keeping its place in its source collections (for example, as dissertations) were immediately linked to the Emergencia Hídrica collection without the need of additional uploads. Currently, more and more works are being submitted to the collection, and further works are linked to it whenever an administrator deems their topic to be relevant to the matter. For example, the article by Canevari (2015), published in *Actas de Periodismo y Comunicación*, was immediately added to the collection.

Navegue por nuestras colecciones

Tesis
Tesis de grado, post-grado y otros documentos

Revistas
Publicaciones en revistas científicas

Eventos
Ponencias realizadas en congresos y conferencias

Libros
Libros digitalizados y e-books

Red UNCI
Artículos y ponencias de la Red UNCI

Radio Universidad
Entrevistas y producciones artísticas, entre otros audios

Recursos educativos abiertos
Material educativo de cursos dictados en la UNLP

Red de museos
Material perteneciente a la red de museos de la UNLP

Emergencia hídrica
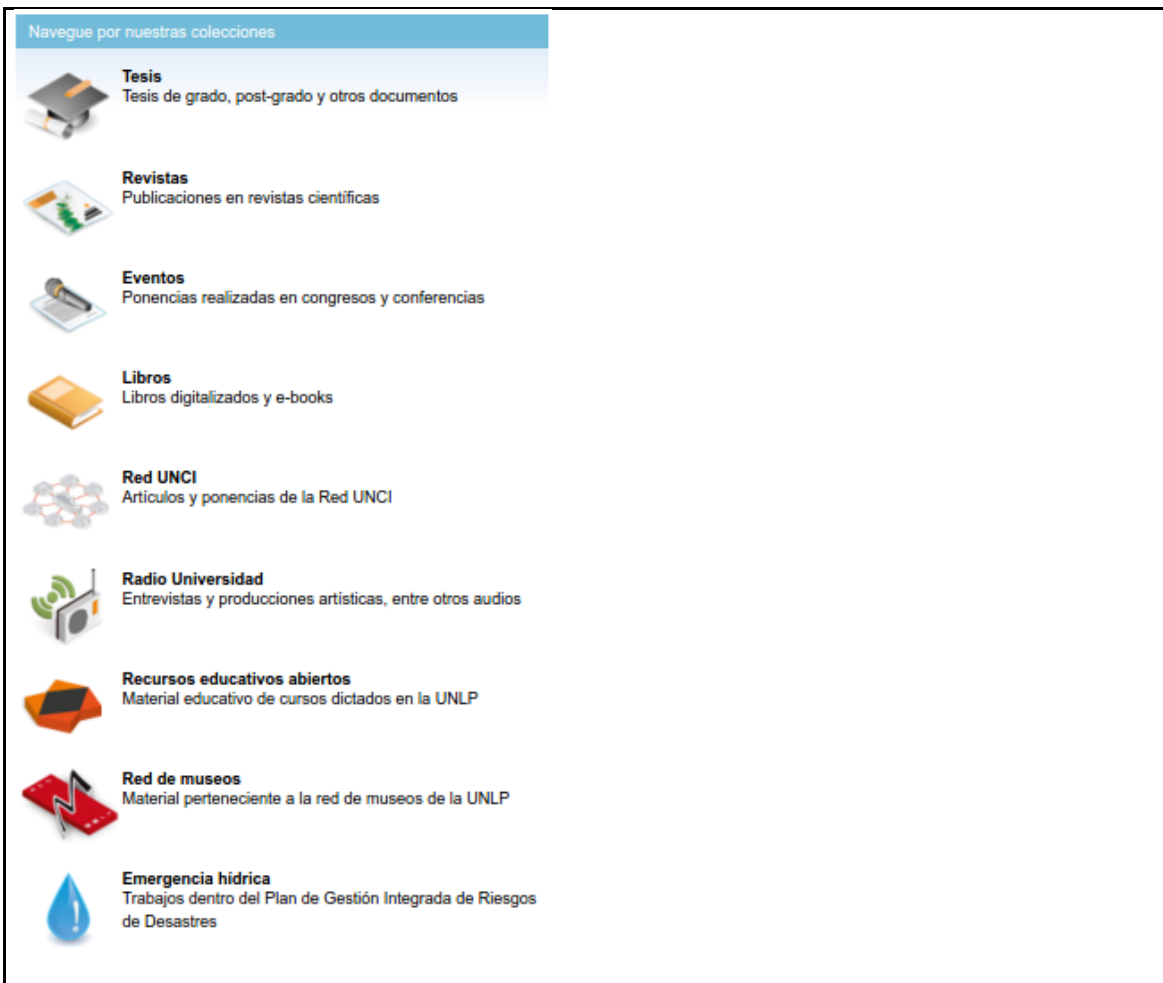Trabajos dentro del Plan de Gestión Integrada de Riesgos de Desastres

Figure 1: List of featured collections in the SEDICI home page, including the Emergencia Hídrica collection. (Servicio de Difusión de la Creación Intelectual, 2016)

3. Observatorio Medioambiental La Plata (OMLP)

Launched in 2015, the Observatorio Medioambiental La Plata aims at disseminating research data and surveys carried out in these areas, driven by the increasing relevance of environmental issues and climate change challenges. This initiative is coordinated by the National University of La Plata and two of the most prestigious research institutions with the largest resource infrastructure in the country: the National Scientific and Technical Research Council (CONICET) and the Commission for Scientific Research of Buenos Aires (CICPBA). The observatory is a shared services center with a Steering Committee in charge of developing the center's activities, including a coordinator and an assistant coordinator as well as a Technical Advisory Commission, for the purposes of influencing decision making on environmental management a national, provincial and municipal levels, and an Advisory Committee for Beneficiary Institutions. This observatory coordinates the efforts and available human and technological resources of multiple national, provincial and municipal entities, universities, scientific and technological institutions and various private organizations which share an interest in environmental issues. Additionally, regarding the collaboration with government institutions, the observatory enables the production of specific reports. The first specific action taken by the OMLP was gathering the results, surveys and raw data obtained

by the Proyectos de Investigación Orientada (PIO), a set of projects related to these issues, as well as offering a platform for their classification, cataloguing and dissemination, providing access to the general public.

3.1 About the PIOs

The PIOs are the result of a special joint call by the Universidad Nacional de La Plata and the CONICET, launched at the end of the first semester 2013 with the aim of financing research projects to solve water challenges in the region. Since then, multidisciplinary work is being done to approach future emergencies considering the relevant technical, urban and social factors. Below there is a list of on-going PIOs to be included in the OMLP at this initial stage:

1. "Construcción de un Sistema Integrado de Gestión del Riesgo Hídrico en la Región del Gran La Plata": This project aims at increasing resilience of the whole environmental, social and territorial system and its capability to face and recover from a threat by generating an integrated risk management system.

2. "Estrategias para la gestión integral del territorio. Vulnerabilidades y Procesos de Intervención y Transformación con Inteligencia Territorial. Métodos y técnicas científicas ambientales, sociales y espaciales": This project covers two scenarios in the area called Gran La Plata, referring to problems and solutions in two extremely vulnerable areas in terms of environmental and social

risk –the most critical areas in the basin of Arroyo Maldonado (La Plata) and the area which surrounds the refinery Refinería La Plata YPF (Ensenada and Berisso).

3. "Evaluación y análisis de Riesgo Ambiental en el Área Gran La Plata": This project's goal is to obtain knowledge so as to correctly assess environmental variables, as well as the risks and consequences caused by the floods by promoting information gathering, monitoring and multidisciplinary studies with a view to improving decision-making and ultimately preserve the environment and people's life quality.

4. "Las inundaciones en La Plata, Berisso y Ensenada": análisis de riesgos y estrategias de intervención. Hacia la construcción de un observatorio ambiental": This project analyzes and explains the characteristics of floods in La Plata, Berisso and Ensenada, which can pose a territorial and environmental threat aiming to build a flood risk assessment matrix. This project's goal is to evaluate and propose action and intervention strategies to prepare for and prevent this type of threats in the short, medium and long term. Another of its goals is to lay the foundations for implementing an open data environmental observatory to provide input for public policies on social and environmental sustainability.

5. "Comunicación y territorio: construcción de mapas territoriales de comunicación": This project aims to creating a georeferencing system from the point of view of complex systems and the research-action-engagement approach (investigación-acción-participación, IAP) to build territorial network maps in parallel with communication network maps, while showing their correlation with relations and practices among communities, organizations and public policies. Its ultimate goal is to systematize information in an online, interactive and intuitive digital platform for governmental, academic and media organizations and institutions.

From the brief descriptions of these projects, their multidisciplinary nature becomes evident, including issues as diverse as mapping communication networks, risk analysis, evaluation of environmental parameters and territorial management. The selection of IT tools that enable the operation of the OMLP is based on this diversity, prioritizing those that provide flexibility and extension capabilities to add new formats, metadata schemas or even customized visualization tools for different data sets.

Selected to support the OMLP, with aesthetic adjustments and functional additions, is an open source software (Free Software Foundation, 2007) that enables creation of open data websites. This tool uses the Python programming language and is being continuously improved both by the development team

and the community of programmers who propose improvements and changes to the software, providing a plugin-based architecture that simplifies and promotes tool extension with new elements and features. Beyond its current features and tools, CKAN offers tremendous potential in terms of technical design and user community, making it a future-proof solution. (Winn, 2013). These features, coupled with clear and complete documentation available online (CKAN Docs, 2013) and the adoption of cataloguing, security and interoperability standards, are the key elements assessed when selecting this platform for the OMLP.

Aesthetically, the adjustments simplify browsing and provide a minimalist interface. The project's interinstitutional nature (CONICET-UNLP-CIC) is also highlighted (Figure 2). Much of the built-in functionality at this initial stage is based on plugins to simplify visualization of data provided by the different projects: *Text view*, for XML, JSON formats and other highlighted-syntax text formats; *DataExplorer*, for easy data access, filtering, charting and mapping in CSV and MS Excel files; *GeoJSON*, to draw points on a map based on JSON files (Figure 3); *Image View*, for PNG, JPEG or GIF files; and *PDF Viewer*, to visualize files within each resource pages.

In the near future, more plugins will be added to enable multiple services and operations in the repository, both for external users and administrators, such as

a plug-in to enable different workflow configurations –implementation of this plug-in is currently under analysis. Furthermore, while CKAN supports OAI-PMH by default, it interoperability will be implemented in subsequent stages.
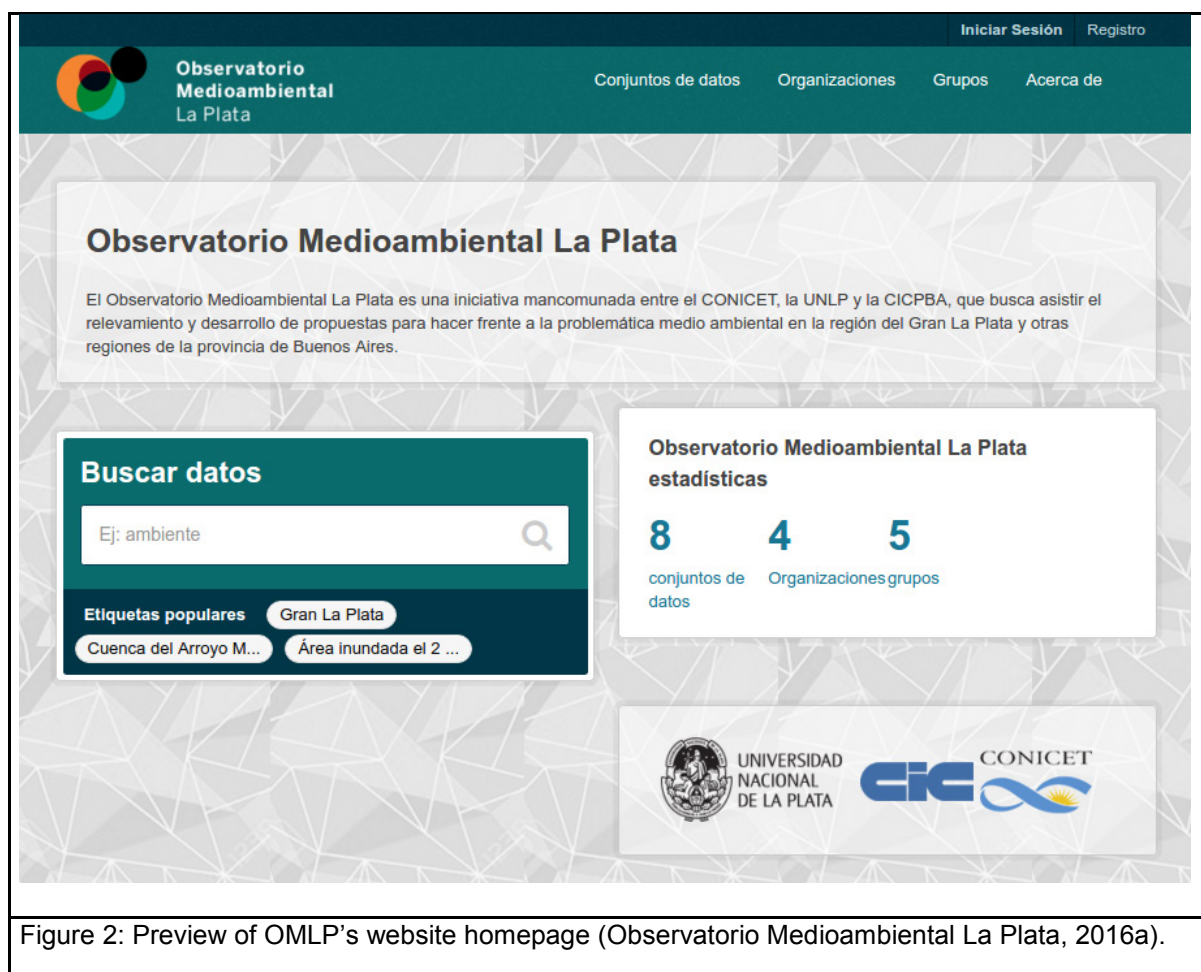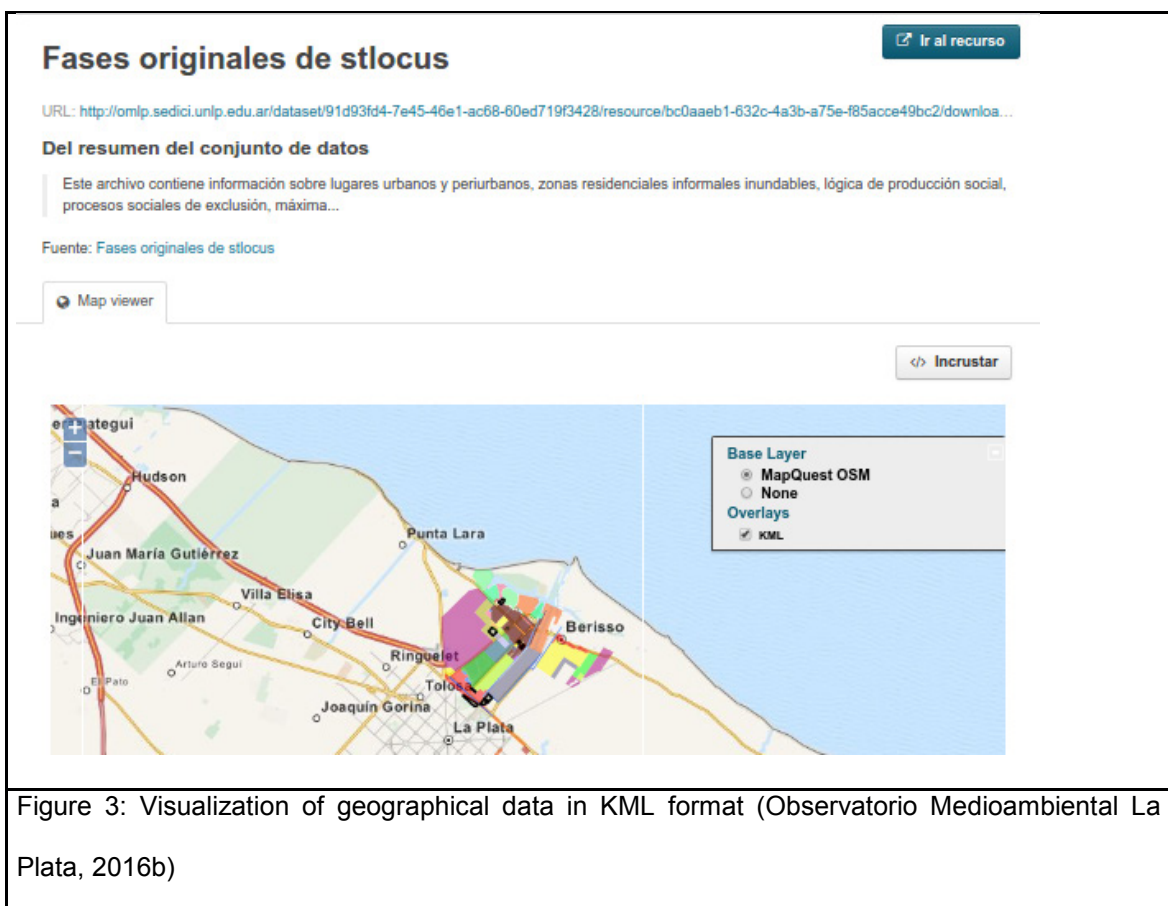

Figure 2: Preview of OMLP's website homepage (Observatorio Medioambiental La Plata, 2016a).

The metadata schema offered by CKAN is used for data cataloguing and implementation of the DCAT standard (Data Catalog Vocabulary, 2014) through a CKAN plug-in is currently being considered. This schema consists of a RDF

vocabulary which enables management of an organization's open data to describe and disseminate them; DCAT standard is the main global standard for a vocabulary designed to simplify data search and interoperability of online data catalogues.

Figure 3: Visualization of geographical data in KML format (Observatorio Medioambiental La Plata, 2016b)

The site for the Observatorio Medioambiental La Plata data repository will initially provide the following services: dataset search and download and data

visualization depending on their format. The contents submitted during this initial stage include GIS layers, high-resolution images, text documents and spreadsheets with survey answers. Initial submittals are being used to determine the most convenient file formats for each scenario and to assess whether they can efficiently be integrated in the web portal –e.g., for map or image visualization. It is worth stressing that the OMLP web portal is not yet accessible to the public; it will be launched once the final stage is completed, including tests, format assessments and specification of a relevant workflow.

4. Conclusions

This document describes the multiple actions taken at the local, provincial and national level in response to a natural disaster. Especially noteworthy is the swift reaction of regional and national academic and research institutions, in a joint effort to prevent the occurrence of events such as the flood which devastated the city of La Plata. The scientific and academic community uses this knowledge via the SEDICI repository, to gather its production and works in related areas in a single Collection accessible to the whole community, with a view to socializing knowledge for early alerts; simultaneously, appointed experts collaborate in multidisciplinary projects funded to give new answers in the area affected by the disaster. Five teams, each including more than fifty

multidisciplinary experts, collect, characterize and systematize a huge amount of information, including primary data; the institutions involved then create the OMLP to meet the special requirements of the authorities and the community concerning the development of statistics and models based on gathered data. With the aim of simplifying visualization tasks and data exploitation, a new technology tool is created –specifically, a data repository where professionals classify, conduct studies, preserve and disseminate the knowledge gained in the area with the ultimate goal of aiding their communities.

Bibliography

Free Software Foundation (2007) GNU Affero General Public License. Available at: https://www.gnu.org/licenses/agpl-3.0.html (accessed 8 September 2016).

Benítez MI and Scarpati OE (2005) Las inundaciones en la ciudad de La Plata Su análisis en relación con las precipitaciones durante las últimas décadas del siglo XX. Geograficando 1: 111-121 Available at http://sedici.unlp.edu.ar/handle/10915/13922 (accessed April 27, 2016).

Canevari T (2015) El campo de la comunicación desde un paradigma científico emergente en contextos de vulnerabilidad socio-ambiental y riesgo hídrico:

Caso La Plata, Berisso y Ensenada. Actas de Periodismo y Comunicación 1 (1) Available at:

http://sedici.unlp.edu.ar/handle/10915/51962 (accessed April 25, 2016).

Cipponeri M, Salvioli ML, Larrivey G, Afranchi AV, and Colli GA (2014) Vulnerabilidad de la población de la ciudad de La Plata (Argentina) a precipitaciones extraordinarias. Aqua-Lac 6(2): 11-20. Available at: http://www.unesco.org.uy/phi/aqualac/fileadmin/phi/aqualac/Numero_6_vol_2 /02_Vulnerabilidad_de_la_poblacion_de_la_ciudad_de_Plata.pdf (accessed April 27, 2016).

CKAN Docs (2013) Overview CKAN 2.6.0a documentation Available at: http://docs.ckan.org/ (accessed April 20, 2016).

Data Catalog Vocabulary (DCAT) (2014) W3C Recommendation 16 January 2014 Available at: https://www.w3.org/TR/vocab-dcat/ (accessed April 20, 2016).

De Giusti MR, Villarreal GL, Terruzzi FA, Oviedo NF and Lira AJ (2013) Interoperabilidad entre el Repositorio Institucional y servicios en línea en la Universidad Nacional de La Plata. In: *PKP International Scholarly Publishing Conferences* (Ciudad de México, México). Available at: http://sedici.unlp.edu.ar/handle/10915/27406 (accessed April 30, 2016).

Facultad de Ingeniería, Universidad Nacional de La Plata (2013) Estudio sobre la inundación ocurrida los días 2 y 3 de abril de 2013 en las ciudades de La Plata, Berisso y Ensenada. Report, Universidad Nacional de La Plata, Argentina, Mayo. Available at: http://sedici.unlp.edu.ar/handle/10915/27334 (accessed April 13, 2016).

Observatorio Medioambiental La Plata, Consejo Nacional de Investigaciones Científicas y Técnicas, Universidad Nacional de La Plata and Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (2016a) Bienvenida - Observatorio Medioambiental La Plata (Forthcoming website).

Observatorio Medioambiental La Plata, Consejo Nacional de Investigaciones Científicas y Técnicas, Universidad Nacional de La Plata and Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (2016b) Lugares: Método Stlocus - Fases originales de stlocus - Observatorio Medioambiental La Plata (Forthcoming website).

Perdomo S (2015) Estimación de parámetros hidráulicos subterráneos a partir de estudios de resistividad eléctrica en 2D. PhD Thesis, Universidad Nacional de La Plata, Argentina. Available at: http://sedici.unlp.edu.ar/handle/10915/46662 (accessed April 27, 2016).

Romanazzi PG (2015) Aproximación a la estimación estadística de la precipitación máxima probable (PMP) para La Plata, Buenos Aires,

Argentina. In: *III Jornadas de Investigación, Transferencia y Extensión de la Facultad de Ingeniería.* La Plata, Argentina. pp. 452-458. Available at: http://sedici.unlp.edu.ar/handle/10915/47844 (accessed April 27, 2016).

Romanazzi PG and Urbiztondo AJ, (2014) Evaluación del desagüe existente y proyectado con un modelo dual: cuenca arroyo del Gato, La Plata, Buenos Aires, Argentina. In: *II Congreso Internacional de Hidrología de Llanuras,* Santa Fe, 22–26 September, 2014 pp.1–10. Available at: http://sedici.unlp.edu.ar/handle/10915/40968 (accessed April 13, 2016).

Servicio de Difusión de la Creación Intelectual, Universidad Nacional de La Plata (2013a) Emergencia Hídrica Available at: http://sedici.unlp.edu.ar/handle/10915/26450 (accessed April 13, 2016).

Servicio de Difusión de la Creación Intelectual, Universidad Nacional de La Plata (2013b) Tutorial de Autoarchivo de Emergencia Hídrica Available at: https://youtu.be/umLtdb5i2LE (accessed April 13, 2016).

Servicio de Difusión de la Creación Intelectual, Universidad Nacional de La Plata (2016) Repositorio Institucional de la Universidad Nacional de La Plata. Available at http://sedici.unlp.edu.ar (accessed 13 October 2016).

Texier J, De Giusti MR, Oviedo NF, Villarreal GL, and Lira AJ (2012) El Uso de Repositorios y su Importancia para la Educación en Ingeniería. In: *World*

Engineering Education Forum (WEEF), Buenos Aires, Argentina. Available at: http://sedici.unlp.edu.ar/handle/10915/22943 (accessed April 19, 2016).

Villarreal GL, De Giusti MR, Sobrado A, Lira AJ, and Vila MM (2008) SeDiCI (Servicio de Difusión de la Creación Intelectual). *D-Lib Magazine*, 14(9/10): 1-11. Available at: http://www.dlib.org/dlib/september08/villarreal/09villarreal.html

Winn J (2013) Open data and the academy: an evaluation of CKAN for research data management. In: IASSIST 2013, 28-31 May 2013, Cologne.

.