APPROACHES TO OPEN DATA FOR SCIENCE IN SPAIN

E Wulff-Barreiro

Marine Sciences Institute of Andalusia, Spanish National Research Council (CSIC),
Campus Universitario Rio San Pedro s/n, 11510 Puerto Real (Cádiz)
Email: enrique.wulff@icman.csic.es

ABSTRACT

As observational data has attained new legal status, allowing their integration into open Internet systems, and experimental data continues to be assembled in common and free platforms, state of the art, easy to access data repositories have been designed in Spain. These repositories have removed many obstacles to re-utilization of GIS and other data. European legislation has also made advances in opening biodiversity data, including a European space in the Latin-American grid infrastructure. Open access biomedical repositories attract commercial attention while astronomical, meteorological, and oncological institutions promote data quality and access. This paper describes recent approaches to open access data for science in Spain.

Keywords: GIS data, Grid architecture, Data policies, Spain, Data repositories

1   INTRODUCTION

The INSPIRE directive adoption into the Spanish law 14/2010, known as LISIGE (Law about Infrastructures and Services of Geographic Information), has created a new scenario for open access to scientific data in Spain. LISIGE mandates central, regional, and local governments to share geographical data and to implement interoperable services. These services include visualization, localization, download, transformation, and enchainment; the first two must be free. If a fee is required, the availability of electronic commerce services must be guaranteed. Administrative procedures, international relationships, intellectual property, etc. are limited in their access rights. A Web Map Service (WMS) is the main standard for the visualization of geo-referenced information. Metadata are also under the schedule of this law; limit dates for its elaboration were established. On December 31, 2010, the reference and basic geographic metadata and services were started, and by December 31, 2013, general cartographic metadata should be established. In Spain, with this legal basis, it could be said that, at least in the case of geographical data, if data collection services are supported by a government body, results must be published.

A recent effort at the national level in Spain has revealed the existence of diverse resources where scientific data can be consulted in a free and no cost manner. This agenda included an information gathering initiative associated with sharing scientific data through repositories or other infrastructures available in the country. On December 2, 2009, the Supercomputing Center of Catalonia organized a workshop on scientific data repositories. Updated information was presented at the recent 4th Plenary Meeting of the Spanish e-Science Network on December 2, 2010.

The data frameworks for e-science that have emerged concern meteorological and oceanographic information in Galicia (Meteogalicia), spatial data infrastructure (IDEE), ES-NGI Spanish National GRID Initiative, CETA-CIEMAT digital repositories infrastructure, data-driven solutions provided to the LHC proton collider supported by PIC, a data initiative in materials and molecular engineering (IMEM), the Spanish Virtual Astronomical Observatory (SVO), and a collection of social sciences datasets of the Fundación Juan March.

Medical geography provided the first case of free scientific data delivery in Spain in the early days of the data explosion through a spatial-temporal approach to surveillance of cancer disparities among demographic subgroups and by analysis of the geographic variations of cancer mortality over annual periods. Profiles of mortality, with particular attention to their follow-up in terms of maps that visualized, analyzed, and identified important issues for health policy, showed that the public domain and open access in digital data were an opportunity for research in Spain. The implementation of the European Community Data Protection 1995 Directive changed the panorama for e-science. Open data, as with open source policy, comes under the auspices of the High Council for eGovernment of Spain (CSAE).

2 THE SPATIAL DATA INFRASTRUCTURE OF THE GENERAL STATE ADMINISTRATION. THE METEOSIX PROJECT, A GALICIAN METEOROLOGICAL AND OCEANOGRAPHIC DATA REPOSITORY

Within the framework of the Infrastructure for Spatial Information in the European Community (INSPIRE) Directive, EU member states shall provide access to their infrastructures through the INSPIRE geo-portal as well as through any access points they themselves decide to operate. Consequently, the Galician regional government is funding for climate science researchers a sensor observation service based on the Open Geospatial Consortium (OGC) specifications for a Meteorological and Oceanographic Serial Digital Interface (SDI) (80 automatic meteorological stations and 4 oceanographic stations). This Meteosix project will launch a new SDI based Geo web site for a wide variety of users to access data through the Internet using OGC and OpenNDAP standards (Viqueira et al., 2010).

In accordance with the INSPIRE initiative, the Permanent Commission of the National Geographic High Council (the Spanish consultative and planning body in cartography) decided on April 10, 2002, that a working group to study and coordinate the spatial data infrastructure (IDEE) in Spain should be established. This initiative to integrate data, metadata, and geographical information provided its potential users with the ability to locate, identify, and access such information.

The current national website for the IDEE provides access to the main node of distribution and screening of data and geographical services in Spain. It has been launched with the nodes and websites of official geographical information under the theme of nature and other equivalent websites that have been established in the Autonomous Communities. Currently 15 out of the 17 regions are involved in the development of IDEE and are building their own regional SDI nodes. In this context in Spain, the development of infrastructures of spatial data (IDE) is carried out in each one of the three administrative levels: national, regional, and local. Furthermore, during the period 2007-2009, this policy, under the umbrella of the European programme INSPIRE, has completed the issuing of reference geoportals to the national and regional levels and begun local level issuance. Similar to that of the Åarhus Convention, the philosophy is to open digital cartography and GI to the citizen.

The INSPIRE Spatial Data Interest Communities (SDIC) are coordinated through periodic meetings of the Geomatic Commission of the National Geographic High Board, the IDEE working group.

The SDIC (Rodriguez et al., 2009) is made up of:

- The Spanish Nature Databank (BDN - Banco de Datos de la Naturaleza);
- EUROPARC Spain;
- The National Assembly of the Land Cover and Use Information System of Spain (SIOSE), which avoids duplication of geographic data and integrates existing data systems to fit EU requirements;
- The Geographical Standards Specialized Commission, SDI of Catalonia, a network of local authorities working under 3.0 specifications (in this context a comparative exercise has been performed between the Catalonia and the Castilla La Mancha IDEs) (Milanović et al., 2010; Diaz et al., 2010);
- The Navarre Territorial Information System;
- the Spanish Railways SDI at Executive Direction of Circulation (ADIF);
- The Engineering Design Group (GID) of Universitat Jaume I;
- Indra Espacio S.A., the leading Spanish company in the Ground Segment for Space Systems and Applications;
- Euskal Herriko Datu Espazialen Azpiegitura (EHDEA); and
- Registra, SL (Capdevila Subirana, 2010).

The IDEE is led by the IGN (National Geographic Institute). The national coordinating body called the Consejo Superior Geográfico (National Geographical High Council) has representatives from national (34), regional (17), and local authorities (6), the IGN, the Hydrographic Institute of the Army, and the cadastre authority. The Commission on Geomatics is the executive board responsible for implementing the IDEE (Spanish SDI).

The ISO 19100 series of standards and OGC specifications are used. The Spanish IDEE Working Group elaborated specific recommendations: a series of WMSs, a Spanish Core Metadata model (NEM v1, (available in ES)), a Spanish Gazetteer Model (MNE v1), and instructions on how to implement WMSs.

WG IDEE organizes three meetings and one technical workshop each year. Since 2010 the technical workshop has evolved into an Iberian (Spain, Portugal, and Andorra) Congress on SDI (Mezcua-Rodriguez, 2009). (See also the ICACI Spanish report http://icaci.org/documents/national_reports/2003-2007/NP_Spain_09.pdf and the public wiki of the Iberian and Latin-American Forum group of the Open Geospatial Consortium, http://external.opengeospatial.org/twiki_public/bin/view/ILAFpublic/WebHome). As a consequence, a new geographical data law was approved in 2010 (LISIGE, 2010). This new policy, which supports the IGN data policy:

- Follows the INSPIRE Directive of 15 May, 2007;
- Establishes free access, free use, and free of charge (with no licensing needed) conditions for the localization of geographical data;
- Fixes accessibility for non commercial uses under conditions of free access, free of charge, and licensing needed with mention of origin and ownership, and
- Provides downloading by Internet, downloading off line, and direct internal use at companies in management systems to promote e-commerce.

Through article 15 of this 2010 law on GI Services in Spain (LISIGE, 2010; Potti Manjavacas et al., 2011), the geoportal IDEAGE integrates the GI data resources from diverse ministries, giving the URLs of WMS, WFS, WCS, and WPS services.

The Ministry of Science and Innovation provides SDI free access to data from the Spanish Council for Scientific Research (CSIC) and the Geological and Mining Institute of Spain (IGME).

CSIC offers:

- SigMayores - a thematic SDI containing social resources for senior citizens in Spain,
- Anthos - an information system on Spanish plants, and
- The Virtual Atlas of the Terrestrial Birds of Spain.

IGME offers:

- SIGECO - a geological data download device for Spanish Cartography,
- SIAS - which makes available and integrates Spanish hydrogeological data,
- WMS services, and
- The Spanish participation in eWater (river basin districts).

Currently, seismic data are not offered free of charge from this institution.

The Ministry of Environment and Rural and Marine Affairs collects data and offers them through a free-data model:

- BDB - the Spanish Biodiversity Databank, which provides a WMS view,
- SIA - a national water information system providing data to the EU,
- IDE-OSE - the SDI system of the Observatory on Sustainability in Spain,
- IDE-Ebro - the geographic data produced by the Ebro River Hydrographic Society,
- IDE-Duero - the Duero River Basin Authority supplied data,
- IDE-CHG - the Guadalquivir River District Infrastructure of Spatial Data,
- CHSIC - viewer information systems of the Hydrographic Confederation of the Segura River, and
- SIGCHG - the GIS viewer of the Guadiana River District.

3   THE SPANISH GBIF NODE

The Global Biodiversity Information Facility (GBIF) in Spain promotes and practices a free-access model data policy and open source free-code software development. The GBIF Data Sharing Agreement is fully in accordance with the Spanish Participant Node. This data provider of free and open access biodiversity data includes records from 151 collections/databases, with a total of 7.306.019 records (http://www.gbif.org/). The log console offers a list of datasets in which the events (harvesting/processing/user/usage/other), log groups and levels, and start/end dates can be specified. A download of these logs follows. The results are grouped into a zip file, whose process time is 30 minutes (a "click the link to download your search results" is prepared, valid for the next 24 hours). The actual event logs can be easily imported into Excel or any other spreadsheet application for analysis. Data resource statistics are provided with the event log download screen. The text can be linked to an Access table that offers the scientific name, the collection code, and other identification fields.

Managed by CSIC, the data search engine (only in Spanish) integrates all the participating data systems. First, the data provider must be selected from a list of biological/botanical classification headings. Then, the taxonomical category and localization specifications are to be selected. Finally, other parameters are required, such as the total number of records to download, visualization conditions (Google Earth), and other personal parameters. Users are supported by tutorial videos, FAQs, and special services offered by the GBIF Spain Coordination Unit (in Spanish).

For structure, contents, and format data sharing standards, the Darwin Core is adopted (Pando, 2008). An evolved version of the 2005 joint CODATA and TDWG (Taxonomic Database Working Group) initiative supported by GBIF is also used. Both the software for the biological collection management and that for the projects management are built in Spain or fully translated into Spanish. Software for botanical and natural history collection programs are Herbar and Zoorbar (http://www.gbif.es/herbar/herbar_in.php and http://www.gbif.es/zoorbar/zoorbar.php). A database application for biodiversity projects management is Bibmaster (Pando et al., 2004).

Two digitalizing software programs are available. From the Canary Islands government, DEMIURGE is a repository to allow fluid, creative use of all the genetic diversity knowledge contained in genotype matrices and other relevant ancillary information This was built by the Viera y Clavijo Botanical Garden CSIC associated unit (Caujapé-Castells et al., 2010). The open-source software Transformer-*3*, for the analysis of molecular population genetic data, was fully developed at the Botanical Garden and simple to download (Caujapé-Castells & Baccarini-Rosas, 2005). The information standard for the data matrices is also that specified by TDWG. In addition, the Vegetation Edition and Analysis Tools (VegAna), an integrated software package oriented towards the storage, management, and analysis of ecological data, is associated with the Catalonian Biodiversity Database.

Other Spanish language initiatives that this GBIF Node promotes are:

- invasIBER (a searchable database on alien species in Spain),
- SIVIM (a cartography of the Iberian corology),
- The Biodiversity Data Bank of the Valencian Region (18.056 species, 1.047.309 citations) (http://www.gbif.es/MasDatos_in.php),
- The Biodiversity Data Bank of the Canary Islands (BIOTA),
- The repository IDBD Biodiversity Data Infrastructure of Navarra,
- The Atlas of the Flora of Aragón,
- The Galician Entomological Inventory, and
- The Flora of Cultivated Fields of Navarra.

4   THE ES-NGI SPANISH NATIONAL GRID INITIATIVE

Grid technologies make it feasible to gain easy access to data accompanied by the requirement for improved security (confidentiality, traceability, integrity, etc.) while data repositories, distributed resource management, and scheduling are topics for the Iberian Grid Infrastructure (see the topics of interest of the 4th Iberian Grid Infrastructure Conference, Braga (Portugal) May 24-27, 2010) (Ibergrid, 2010). The Spanish National GRID

Initiative (ES-NGI) is led by the Institute of Physics of Cantabria (IFCA) and enables access to computing resources from high energy physics to humanities. CSIC is involved although not on an open-access basis.

Data intensive applications with the ultimate goal of reducing the cost of hosting digital repositories on GRID infrastructures and developing applications to exploit them are managed inside ES-NGI by CETA-CIEMAT. Therefore, the European GRID Environment gLibrary/DRI platform Spanish participants CETA-CIEMAT and MAAT International (For a description of MAAT, see http://www.youtube.com/watch?v=GsmrL_ebgMA) are involved in the European project neuGRID. DRI/Mammogrid is the Digital Repositories Infrastructure application that offers a diagnosis support system (DSS) and data federation to support collaborative breast cancer research. A second case of DRI use is the digital repository of the historical archive of Trujillo (Virtual community access only with certificate, http://www.e-ciencia.es/wiki/index.php/Ahtrujillo).

Other ES-NGI repositories available without the usual virtual organization authentication requirement are:

- The Mercurial Grid Configuration Repository of Operation Tools, and
- The Virtual Machines Storage (Réplicas).

Additional Spanish Grid applications in bioinformatics, e-learning, and life and Earth sciences created for online data storage are available from the GISELA (Grid Initiatives for e-Science virtual communities in Europe and Latin America) (Marechal, 2011).

5   PIC, THE SCIENTIFIC INFORMATION PORT DATA STORE

The Spanish Grid for Particle Physics is a tier-1 data centre (PIC, Port d'Informació Científica, Scientific Information Port) that receives all recorded collision events sent from the two proton beams in the LHC at CERN. PIC supports scientific groups working on projects that require strong computing resources for storage and analysis of massive sets of distributed data. The PIC data centre is housed in the Universitat Autònoma de Barcelona. Beyond storage, PIC processes these data to generate calibrated and filtered data that are provided to the seven research institutions (CIEMAT, IFAE, IFCA, IFIC, UAM, UB, and USC) that constitute the Spanish contribution to the LHC Computing Grid (LCG). The data centre, other than just supporting the Spanish particle physics community, has made its computing resources available to the biosciences and Earth sciences communities.

PICNIC is the technological platform of the Medical NeuroImaging Group at PIC that facilitates access to the medical research data of clinicians and the post-processing of magnetic resonance data with free software to download.

MAGIC.PIC is the datacenter for the MAGIC-II experiment designed to study the universe and discover new Gamma-ray sources in the energy range from 50 GeV to 5 TeV, with the largest Cherenkov detectors in the world (located on the Canary Island, La Palma).These data are accessible to all collaboration members through the PIC web site.

6   THE DATA INITIATIVE IN MATERIALS AND MOLECULAR ENGINEERING (IMEM)

A national data resource for materials science has resulted from new computational approaches to provide access to potential energy surfaces of (bio) polymers with a disordered structure. This is NCAD, the Non-Coded Amino Acids Database, a computing resource that has been made freely available. All the data of non-proteinogenic amino acids crucial to control the peptide chains have been assembled and coherently organized in this open access database. The intrinsic conformational preferences of these residues (previously scattered among publications and patents), determined by quantum mechanical calculations, are offered together with the relevant bibliographical information. NCAD was built at the Center for Research in Nano-Engineering (Technical University of Catalonia).

7   ASTRONOMICAL ARCHIVES

TAPAS - Telescope Archive for Public Access System – is an open data archive for the IRAM (Grenoble)/ IAA-CSIC 30m antenna, near Pico Veleta, at Sierra Nevada, Granada (Santander Vela, 2009). It contains more than 200 header variable data for each observational scan, taken between the end of September 2009 and the present. ALMACEN is the IRAM data archive for the 30-m telescope observations taken between 1989 and 2006. AMIGA is a multi wavelength database for a refinement of the pioneering Catalogue of Isolated Galaxies (CIG catalogue, characterizing 1050 galaxies); the data are public and harvested in CSIC (Granada). In TAPAS, CODATA standards are explicitly considered concerning the public release of radioastronomical data.

The Data Archive Unit inside the Villafranca Satellite Tracking Station maintains and develops data archives for the ground-based telescopes (GTC) and space missions (INES, OMC, CoRot). The DSS-63 data archive includes the observations in K-band (18-26 GHz) of the DSS-63 antenna (Deep Space Station of NASA in Madrid) taken during 2006. The data archives of INES (IUE Newly Extracted Spectra), the first space observatory ever launched (INTA, 2000), which ended its 20 year mission in 1996, have been on the Internet since 1998. Its principal centre is operated by LAEFF. Also freely available are the 541802 objects (gamma-ray sources observed from 2002 by the INTEGRAL ESA mission) found in the Optical Monitoring Camera (OMC) Archive. CoRot is the Ground-Based Seismology Programme Archive, which offers public access to the data that are physically stored at IAS (Paris). The GAUDI open repository, a preparatory archive for the COROT mission, contains data and more than 1500 objects gathered since January 1998 in 6 years of observational campaigns.

AXIS public data are available from the Canary Islands IFCA (an XMM-Newton SSC). It contains 319 unique X-ray sources, optically identified to a certain level above 90% by the satellite.

PVOL (Basque Country University) provides a large database containing observations of Jupiter, Saturn, Uranus, and Neptune from the year 2000 onwards contributed by amateur astronomers. It also offers access to IOPW images.

8   THE DATA LIBRARY OF THE FUNDACIÓN JUÁN MARCH

The Center for Advanced Study in the Social Sciences (CEACS) provides integrated access to an extensive data collection for research and teaching. Due to data license restrictions, however, access to most of the datasets is limited to members of CEACS. Nevertheless, to aid their researchers in the curation and sharing of their data, the data librarian has elaborated a Harvard Dataverse data repository. This is a virtual archive in the Harvard Dataverse Network that enables access to everyone signing a licensing agreement (with consent of the researcher).

9   METEOROLOGICAL AUTHORITY DATA POLICY

The State Meteorological Agency (AEMET) in collaboration with CSIC offers, in the framework of the National warning system of African air mass intrusions, an open data device for the period 2004 onwards. Iberonesia is a Network of Brewer Spectroradiometers covering the North Macaronesian and the Iberian Peninsula area. The agency freely offers the calibrations of brewer spectrophotometers for total ozone observations data on the Internet. Additionally, AEMET has an open data policy in accordance with EU directives, thus providing access to the observational and historical data series.

10   NATIONAL CENTER FOR CANCER RESEARCH DATA REPOSITORIES

CARGO, Cancer and Related Genes Online, is a system that presents a list of human cancer genes. As a free service, it facilitates, integrates, and visualizes results from Internet resources, independently of their native access methods. 43,000 unique structures as well as experimental data from more than 350 different biological assays are available in a common chemical and biological repository (CCBR), designed by the drug discovery programs of CNIO (Urbano-Cuadrado et al., 2010). An integral web application (WACBIP) permits querying CCBR. A repository of cDNA clone libraries for human and mouse species is a main goal for the Genomics Unit at CNIO.

11 CONCLUSION

In Spain, public funding of free access to research data has been largely updated by recent legislation. After 2010, geographical data will be published by the public administration under standard services, LISIGE. LISIGE establishes that access for the discovery and viewing of data should be without restrictions and free of charge. Free data download services have recently been proliferating. Also the availability of these data has been improved by providing a central geo-portal on the Internet.

A consensus of opinion at the Spanish National Research Council CSIC considers the Spanish GBIF node to be an important achievement in terms of open data. The work leading up to the biodiversity data sharing practices is seen as advancing international co-operation, and Spain hosted the 2010 European GBIF Nodes Meeting.

The Spanish Grid Initiative shows a consolidated scientific community panorama. e-Science virtual communities close to Latin-America are noted in this Iberian infrastructure. The Spanish Grid for particle physics stores the data distributed by CERN and also makes medical research data easy to access. From a nano-engineering research centre, a non-coded amino acids database has been made freely available. A partially open social science data repository from a private advanced studies institute has been released as well.

Data archives from the astronomical community are constantly being updated and launched. Antennas, missions, and space aircraft units are monitored to be accessed as open repositories. The meteorological authority has a defined open data policy. Finally, diverse repositories are available for cancer research from CNIO.

12 REFERENCES

Capdevila Subirana, J. (2010) *Inspire. Member State Report: España, 2009*. Retrieved October 7, 2010, from the World Wide Web:
http://www.idee.es/SeguimientoINSPIRE/2009/100720_Report_INSPIRE_2009_Spain_v2.pdf

Caujapé-Castells, J. & Baccarani-Rosas, M. (2005) Transformer-3: a program for the analysis of molecular population genetic data. Jardín Botánico Canario "Viera y Clavijo"-Unidad Asociada GBIF & EXEGEN Software, Las Palmas de Gran Canaria, Spain. Retrieved December 10, 2010, from the World Wide Web:
http://www.demiurge-project.org/downloadt3

Caujapé-Castells, J., Sabbagh, I., Medina, D.A., Toledo, J., Ramírez-Pérez, F., Castellano, J.J., Ramos, R., Quintana-Trujillo, F.M., Henríquez, V., Rodríguez, J.F., & Bethencourt, C. (2010) DEMIURGE: an enhancer of knowledge on biodiversity's genetic diversity following GBIF standards of data interchange and structure, Las Palmas de Gran Canaria, Spain. Retrieved December 10, 2010, from the World Wide Web:
http://www.demiurge-project.org/

Diaz, P., Masó, J., & Guimet, J. (2010) Comparative Quality Assessment of Metadata. Two Regional SDI case studies. *INSPIRE conference 2010, 23-25 June,* Cracow, Poland. Retrieved November 25, 2010, from the World Wide Web:
http://www.creaf.uab.es/MiraMon/publicat/oral_p/inspire2010/Comparative_Quality_Assesment_of_Metadata.pdf

Ibergrid (2010) *4th Iberian grid infrastructure conference*, Universidade do Minho, Braga, Portugal. Retrieved December 3, 2010, from the World Wide Web: http://ibergrid.eu/2010/IBERGRID_Agenda.pdf

INTA (2000) *INES, the IUE Archives delivered to the World Scientific Community. An advanced data archive and distribution system for astrophysics*, Noordwijk, The Netherlands: ESA. Retrieved January 2, 2011, from the World Wide Web: http://sdc.laeff.inta.es/ines/Ines_PCentre/ines.pdf

LISIGE (2010) Ley 14/2010, de 5 Julio, sobre las Infraestructuras y los Servicios de Información Geográfica en España (in English). Retrieved September 5, 2010 from the World Wide Web:
http://www.idee.es/resources/leyes/20100706_LISIGE_en.pdf

Marechal, B. (2011) GISELA - Building the future of e-Science in Latin America. *ISGC'2011 Academia Sinica, 21-23 Mar. 2011*, Taipei, Taiwan. Retrieved September 21, 2011 from the World Wide Web:
http://documents.gisela-grid.eu/record/160

Mezcua-Rodriguez, J. (2009) The Spatial Data Infrastructure of Spain as an example of success in Europe. *United Nations Economic and Social Council. Ninth United Nations Regional Cartographic Conference for the Americas, 10-14 August 2009*, New York. Retrieved July 5, 2010, from the World Wide Web: http://unstats.un.org/unsd/ggim/RCC/docs/rcca9/ip/9th_UNRCCA_econf.99_IP16.pdf

Milanović, A., Guimet, J., Rodellas, E., & Bolívar, M.A. (2010) Interorganizational Geo-Synchronization Using Open Geospatial Consortium's (OGC) Technologies to Share and Harmonize Data in Catalonia. *INSPIRE conference 2010, 23-25 June,* Cracow, Poland. Retrieved November 25, 2010, from the World Wide Web: http://www.geoportal-idec.cat/geoportal/cat/documents/articles/Ponencia_INSPIRE_2010.pdf

Pando, F., Gaztelu, J., Rodríguez Vega, D., Pérez Pérez, R., Herrero Nieto, A., & González Talaván, A. (2004) BIBMASTER (version 2.8): A database application for Biodiversity Projects: taxa-level information, nomenclature, literature and specimen management. Retrieved December 10, 2010, from the World Wide Web: http://www.gbif.es/bibmaster/bibmaster_in.php

Pando, F. (2008) *3rd GBIF Workshop on Georeferencing Biodiversity Databases,* Royal Botanical Garden, Madrid, Spain. Retrieved December 10, 2010, from the World Wide Web: http://www.gbif.es/formaciondetalles.php?IDForm=39#prog

Potti Manjavacas, H., Juanatey Aguilera, M., & Abad Power, P. (2011) La LISIGE y el SIG libre. *V Jornadas de SIG Libre. 23-25 March 2011*, Girona, Spain. Retrieved November 25, 2010, from the World Wide Web: http://www.sigte.udg.edu/jornadassiglibre/uploads/articulos/art38.pdf

Rodriguez, A., Abad, P., Alonso, J.A., Sánchez, A., Gonzalez, C., Mas, S., Diez, E., Soteres, C., & Potti, H. (2009) Data and Services Availability in Spanish NSDI. *GSDI 11 World Conference. Spatial Data Infrastructure Convergence: Building SDI Bridges to address Global Challenges 15-19 June 2009*, Rotterdam, The Netherlands. Retrieved November 25, 2010, from the World Wide Web: http://www.gsdi.org/gsdiconf/gsdi11/papers/pdf/161.pdf

Santander Vela, J. de D. (2009) *Integration of tools and radioastronomical archives in the VO architecture,* PhD thesis, Universidad de Granada, Granada, Spain. Retrieved November 25, 2010, from the World Wide Web: http://amiga.iaa.es/FCKeditor/UserFiles/File/radiovo-thesis.pdf

Urbano-Cuadrado, M., Rabal, O, & Oyarzabal, J. (2010) Centralizing discovery information: from logistics to knowledge at a public organization. *Combinatorial chemistry & high throughput screen*, 14(6), 429-449.

Viqueira, J.R.R., Varela, J., Triñanes, J., & Cotos, J.M. (2010) A Sensor Observation Service Based on OGC Specifications for a Meteorological SDI in Galicia. In Trujillo, J., Dobbie, G., Kangassalo, H.;, Hartmann, S., Kirchberg, M., Rossi, M., Reinhartz-Berger, I., Zimányi, E., & Frasincar, F. (Eds.), *Proceedings of the 2010 international conference on Advances in Conceptual Modeling – Applications and Challenges*. Berlin, Springer.

13   APPENDIX

Below is provided the list of Spanish open research data repositories and data systems that were used in this article (as they were on the Internet January 2011).

Acronyms	Spanish Open Research Data Repositories & Data Systems (URL)
SDI	
ADIF	http://inspire.jrc.ec.europa.eu/index.cfm/pageid/42/list/7/id/37804
Anthos	http://www.anthos.es/
Aves	http://www.vertebradosibericos.org/aves/atlas/presentacion.html
BDB	http://servicios2.marm.es/WmsViewer/html/default.htm
BDN	http://www.gbif.es/MasDatos_in.php
BTA	http://www.csg-cnc.es/web/cnccontent/docs/bta/BTAv10.zip
CHSIC	http://www.chsegura.es/chsic/servlet/es.mma.chs.MVCServlet?action= generate&sid=generate&lang=en&predefined=llamadapublico&ancho

	=1200&alto=692
CSG-CNC	http://www.csg-cnc.es/web/cnccontent/en/index.html
EHDEA	http://inspire.jrc.ec.europa.eu/index.cfm/pageid/42/list/7/id/9711
EUROPARC	http://opengis.uab.es/wms/europarc/
GID-Jaume I	http://inspire.jrc.ec.europa.eu/index.cfm/pageid/42/list/7/id/31173
IDE-CHG	http://idechg.chguadalquivir.es/geoportal/en/
IDE-Ebro	http://ide-ebro.chebro.es/gazetteer/indexLayout.jsp?PAGELANGUAGE=en
IDE-Duero	http://www.mirame.chduero.es/DMADuero_09/index.faces
IDE-OSE	http://www.sostenibilidad-es.org/es/actividades/infraestructura-de-datos-espaciales-del-ose-ide-ose
IDEAGE	http://www.ideage.es/web/portal/ideage
IDEC	http://www.geoportal-idec.cat/geoportal/eng/prg/
IDEE	http://www.idee.es/show.do?to=pideep_pidee.EN
IDEE Executive Board	http://www.fomento.es/MFOM/LANG_CASTELLANO/ORGANOS_COLEGIADOS/CSG/Comisiones/comisi_geomatica.htm?lang=en
IDEE Recommendations	http://www.idee.es/show.do?to=pideep_recomendaciones.EN
IDEE Workgroups	http://www.idee.es/show.do?to=pideep_subgrupo_trabajo.EN
IGN	http://www.idee.es/show.do?to=pideep_grupo.EN
Indra	http://inspire.jrc.ec.europa.eu/index.cfm/pageid/42/list/7/id/2129
Registra SL	http://inspire.jrc.ec.europa.eu/index.cfm/pageid/42/list/7/id/2190
SIA	http://servicios2.marm.es/sia/visualizacion/descargas/series.jsp
SIAS	http://www.igme.es/internet/ServiciosMapas/siasweb/sias_general.htm
SIGCHG	http://www.chguadiana.es/consultas/sigchgweb/SIGCHGWEB_pub.jsp
SIGECO	http://cuarzo.igme.es/sigeco/Default.aspx
SigMayores	http://www.sigmayores.csic.es
SIOSE	http://www.siose.es
SITNA	http://sitna.navarra.es/navegar/?lang=en
SITOPO	http://cuarzo.igme.es/sitopo
GBIF	
AEGA	http://www.aegaweb.com/inventario/index.htm
BDBC	http://biodiver.bio.ub.es/biocat/homepage.html
BDD	http://bdb.cma.gva.es
BIOTA	http://www.gobiernodecanarias.org/cmayot/medioambiente/biodiversidad/ceplam/bancodatos/biotaterrestre/Atlantis.jsp
DEMIURGE	http://www.demiurge-project.org/
FLORAGON	http://www.ipe.csic.es/floragon/index.php
GBIF-Spain	http://data.gbif.org/datasets/provider/57
GBIF-Spain Data Portal	http://taray.csic.es:8080/pres/PresentationServlet?action=home
IDENA	http://idena.navarra.es/busquedas/catalog/search/search.page
invasIBER	http://invasiber.org/cerca.php
SIVI	http://www.sivim.info/sivi/
VegAna	http://biodiver.bio.ub.es/vegana/index.html
GRID	
DRI/Mammogrid	http://applications.gisela-grid.eu/application_details.php?l=20&ID=32
Gisela	http://applications.gisela-grid.eu/app_list.php?l=20
Grid CSIC	http://www.grid.csic.es
Mercurial	http://devel.ifca.es/hg/
Réplicas	http://sunsite.rediris.es/replicas.php
PIC	
MAGIC	http://magic.pic.es/
PICNIC	https://neuroweb01.pic.es
IMEM	
NCAD	http://recerca.upc.edu/imem/DATABASE/Install_NCADEXE.zip
ASTRONOMY	
almacén	http://www.iram.es/IRAMES/dataarchive/init.html
AMIGA	http://amiga.iaa.csic.es/DATABASE/Masterswitch.php?body=byname
AXIS	http://venus.ifca.unican.es/~xray/AXIS/publicdata/
CoRot	http://sdc.cab.inta-csic.es/corotfa/jsp/frontpage.jsp
DSS-63	http://sdc.cab.inta-csic.es/robledo/index.jsp
GAUDI	http://sdc.cab.inta-csic.es/gaudi/

INES	http://sdc.cab.inta-csic.es/ines/index2.html
PVOL	http://www.pvol.ehu.es/index.jsp?action=iopw
TAPAS	https://mrt-lx3.iram.es/tapas/
SOCIAL SCIENCES	
CEACS	http://www.march.es/ceacs/biblioteca/datalib/Index.html
CEACS Dataverse	http://dvn.iq.harvard.edu/dvn/dv/ceacs/faces/AdvSearchPage.xhtml
METEOROLOGY	
AEMET	ftp://ftpdatos.aemet.es/
CALIMA	http://www.calima.ws/bd_predicciones.html
IBERONESIA	http://www.iberonesia.net/brewer
CANCER	
CARGO	http://cargo.bioinfo.cnio.es/