eChallenges e-2015 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2015
ISBN: 978-1-905824-53-3

ENERGIC-OD Virtual Hubs: a brokered architecture for facilitating Open Data sharing and use

Paolo MAZZETTI[1], Miguel Ángel LATRE[2], Michael BAUER[3],
Raffaella BRUMANA[4], Stefan BRAUMANN[5], Stefano NATIVI[1]
[1]*Consiglio Nazionale delle Ricerche – Istituto sull'Inquinamento Ambientale (CNR-IIA), Via Madonna del Piano 10, Sesto Fiorentino (FI), 50019, Italy*
Tel: +39 055 5226591, Email: {paolo.mazzetti, stefano.nativi}@cnr.it
[2]*Universidad de Zaragoza, C/ María de Luna 1, Zaragoza, E-50018, Spain*
Tel: +34 976 762331, Fax: +34 976 761914, Email: latre@unizar.es
[3]*GEOkomm Verband der GeoInformationswirtschaft e. V., Große Weinmeisterstraße 3a, Potsdam D-14469, Germany*
Tel: +49 (0)331 27 34 108, Fax: +49 (0)331 2731935, Email: michael.bauer@geokomm.de
[4]*Politecnico di Milano, Department of Architecture, Built Environment and Construction Engineering, P.zza Leonardo Da Vinci, 32, 20133 Milano, Italy*
Tel: +39 02 2399 6534, Fax: + 39 02 2399 6550, Email: raffaella.brumana@polimi.it
[5]*Luftbild Umwelt Planung GmbH, Große Weinmeisterstr. 3a, 14469 Potsdam, Germany,*
Tel: +49 (0) 331 27577-77, Fax: +49 (0) 331 27577-99,
Email:stefan.braumann@lup-umwelt.de

Abstract: Open Data promise to create new market opportunities and social value. In particular, geospatial data have a potentially huge inner value. But to unleash their potential, data need to be not only open but also effectively usable. The present paper describes the concept and implementation of the Virtual Hubs developed in the ENERGIC-OD European project. The Virtual Hubs facilitate the use of geospatial open data, providing seamless discovery of and access to heterogeneous data sources, also providing semantic capabilities and dataset transformations. Virtual Hubs lower the entry barriers for both data providers and users facilitating open data publication and exploitation. Through open Web APIs, they reduce the development time for Web and mobile applications based on geospatial open data, supporting the creation of innovative services and the generation of new market opportunities.

1. Introduction

Open data is a major trend in current information technology scenario and it is often publicised as one of the pillars of the information society in the near future. Indeed, the open data concept encompasses not only technological but also socio-economical aspects promising the birth and growth of new business models, market opportunities and creation of social value [1]. No scientific domain or area of application seems to be immune to the open data wave [2]. In every field, open data are pushed and more and more widely adopted [3]. Relevant initiatives at European and global level have been initiated to support and guide the open data movement, such as the Research Data Alliance (RDA)[1] and the Open

[1] https://rd-alliance.org/

Knowledge Foundation (OKF)[2]. Even several funding agencies now explicitly encourages open data (and not only data) sharing. In particular, the European Commission recognizes the importance of open data in the Digital Agenda for Europe focussing on generating value through re-use of a specific type of data – public sector information, sometimes also referred to as government data [4]. Data processing has been identified as essential for addressing the societal challenges as informed decision making is increasingly depending on analysis of the available data. This brought to an explicit promotion of open access in the Article 18 of the regulation of the recently started Horizon 2020 Framework Programme for Research and Innovation [5].

However, several barriers may act as an obstacle toward an effective sharing and exploitation of open data. Although proper awareness raising and capacity building actions can lower and possibly remove behavioural, legal and technical barriers to sharing, still some problems remain. Indeed, even assuming that openness is achieved, it is not sufficient by itself to assure the expected benefits. What is really necessary is not *openness* but *usability* of data. The adoption of proprietary formats and interfaces, the lack of a long-term preservation strategy and supporting infrastructures, the absence of persistent identifiers, are all examples of obstacles to the practical usage of data which are openly published.

It is therefore necessary to design and develop advanced infrastructures providing services for improving the openness of published data making them really usable, and assuring users the possibility to fully exploit them and to unleash their inner value.

The present paper describes the ENERGIC OD Virtual Hub solution proposed and implemented in the ENERGIC OD[3] (European NEtwork for Redistributing Geospatial Information to user Communities - Open Data) project to facilitate the use of geospatial open data, discussing current achievements and potential benefits. In particular, Section 2 describes the main challenge that the presented activities aim to overcome. Section 3 presents the adopted methodology introducing the concept of ENERGIC OD Virtual Hubs. Section 4 briefly describes the technologies adopted and enhanced to build the Virtual Hubs while Section 5 focuses on the current status of implementation. Section 6 presents the lessons learned from the first phase of implementation of Virtual Hubs and pilot applications. Section 7 discusses the potential business benefits of the proposed solution. Finally, Section 8 draws some conclusions.

2. Objectives

The need of assuring usability of data beyond openness, was recognized since the beginning of the open data movement with the suggestion by Tim Berners-Lee, inventor of the World Wide Web, of a five-star rating system based on the openness characteristics of data [6].

★	Available on the web (whatever format) *but with an open licence, to be Open Data*
★★	Available as machine-readable structured data (e.g. excel instead of image scan of a table)
★★★	as (2) plus non-proprietary format (e.g. CSV instead of excel)
★★★★	All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
★★★★★	All the above, plus: Link your data to other people's data to provide context

Figure 1 - The 5-star classification of open data by Tim Berners-Lee -Lee (from [6])

[2] https://okfn.org/
[3] http://www.energic-od.eu/

According to that classification, the more stars an open dataset has, the more it is usable (see Figure 1). Open data providers should then try to improve the level of published datasets. However, this is not manually feasible due to the increasing amount of available data. This is particularly true when geo-information – i.e. information with an implicit or explicit reference to space and time – is considered, including that from Earth Observation. New remote sensing instruments with high spatial, temporal and radiometric resolution (e.g. the ESA Sentinel missions), and new paradigms for in-situ measurements including crowdsourcing and social media mining, generate a huge amount of open data requiring automated techniques for data management, discovery, access and use.

This outlines a big challenge for the geospatial data sharing world: how is it possible to design and implement advanced systems able to increase the openness of existing and new datasets? The ENERGIC OD project proposed an innovative approach designing and developing advanced infrastructures – the ENERGIC OD Virtual Hubs – lowering, and possibly removing, the existing barriers to a full exploitation of open geospatial data available from disparate (i.e. distributed and heterogeneous) sources. They are designed around the concept of brokered architectures and implemented leveraging the effort done in recent years in the development of Spatial Data Infrastructures, standards, and tools.

3. Methodology

3.1. Identification of target users

Open data sharing is a complex process in itself, and therefore complexity cannot be reduced over a certain extent. However, complexity can be shifted from one actor to another. For example, semantics tools may help data discovery facilitating users, but they require significant effort for design, implementation and maintenance.

The business goal of ENERGIC OD is "the creation of innovative applications and services", thus it considers Web and mobile application developers with no specific expertise in geospatial applications development, as the primary target users. A Virtual Hub should then provide advanced services supporting app developers in creating applications based on open geospatial data according to their skills and preferred formats and interfaces, without the need to become expert in geospatial data technologies.

3.2. Identification of barriers to the geospatial information data sharing

ENERGIC OD assumes that openness is already recognized as a need and a benefit. Although this may be not true in general, ENERGIC OD relies on the many on-going actions that aim to achieve a consensus on the need and usefulness of open data sharing in the geospatial world - e.g. activities of the RDA, the definition of the Data Sharing Principles in building the Global Earth Observation System of Systems (GEOSS) by the Group on Earth Observation (GEO). ENERGIC OD actually builds on the many initiatives contributing to provide geospatial datasets with an open license: the European Union Open Data Portal[4], the Copernicus programme[5] (with particular reference to the space component and the Sentinel missions), the GEOSS Data Collection of Open Resources for Everyone (GEOSS Data-CORE)[6], etc.

However declaring that a dataset can be used does not mean that it is effectively usable, due to many possible barriers raising questions such as: is the open dataset (easily)

[4] https://open-data.europa.eu
[5] http://www.copernicus.eu/main/data-access
[6] https://www.earthobservations.org/geoss_dsp.shtml

discoverable? Once it is discovered, can it be (easily) accessed? Once it is retrieved, which format is it encoded in?

ENERGIC OD recognizes two main categories of technological barriers: a) *interoperability* that is the need to address the syntactic and semantic heterogeneity of geospatial datasets and related sharing services; b) *usability,* that is the need of a user-friendly handling of datasets by the target users.

3.3. The concept of the ENERGIC OD Virtual Hubs

ENERGIC OD aims to facilitate the use of open geospatial data delivering Virtual Hubs providing the needed interoperability and usability functionalities.

The call for pilots that ENERGIC OD answered to did not specify what a Virtual Hub is. The term was just used in a generic way mentioning Virtual Hubs as the instruments for facilitating the use of Open Data: "*The pilots should focus on the development of* *virtual hubs that facilitate the use of open (freely available) geographic data*" [7]. However, the term *virtual hub* is usually adopted referring to the *hub-and-spoke* distribution paradigm, where it means "*Of or being a system of distribution, as of goods, passengers, or data, in which the items being distributed are routed into and out of a central location*" [8]. A software architecture adopting the *(message) broker* pattern is usually referred as a *hub-and-spoke* architecture [9]. Following those indications an ENERGIC OD Virtual Hub is defined as "a virtual node where users can seamlessly access a potentially unlimited amount of datasets by brokering heterogeneous open geospatial data sources".

3.4. Addressing interoperability: brokered System of Systems

An ENERGIC OD Virtual Hub is conceived as a single point-of-access – i.e. a *hub* - to open geospatial datasets. It is *virtual* since it does not physically collect datasets, which are actually kept on their original source according to a System of Systems approach.

In the recent years a wide number of different systems have been developed in order to support geospatial data sharing in different domains such as INSPIRE[7] for European Public Authorities, Copernicus services for environmental and security applications from European Earth Observation data, GEOSS for global Earth Observations, and domain specific infrastructures like the Global Biodiversity Information Facility[8] (GBIF). They differ in architectural and technological solutions, and they have their own governance and mandate needing to evolve autonomously. The notion of "System of Systems" (SoS) and "System of Systems Engineering" (SoSE) emerged in many fields of applications to address the common problem of integrating many independent, autonomous systems. SoSs can be usefully described as follows: *systems of systems are large-scale integrated systems that are heterogeneous and consist of sub-systems that are independently operable on their own, but are networked together for a common goal* [10].

From a technical point-of-view, there are two general approaches for building a SoS: through *federation* and through *brokering*.

In the *federated approach*, the participants agree on a common set of specification (*federated model*). Their choice can range from a loose approach needing just the adoption of a suite of interface, metadata and data model, to a very strict approach imposing the adoption of the same software tools. In any case, participants have to comply with the federated model (specifications or tools) making at least some change in their own systems. Thus, this approach is feasible when:

[7] http://inspire.ec.europa.eu/
[8] http://www.gbif.org/

a) the SoS governance has a strong mandate for imposing and enforcing the adoption of the federated model to all the participants, or the participants have a strong interest and commitment in participating in the SoS
b) the participant organizations have the expertise and skills for implementing the needed re-engineering to make their systems compliant with the federated model

E-Commerce, e-Banking, and e-Government systems are typical examples where the federated approach fits well. In the geospatial world, the Open Geospatial Consortium (OGC) has been historically active in developing standard specifications [11], and the INSPIRE experience is an example where a central authority, the European Union, through a Directive, imposed a set of sharing principles, along with Implementing Rules, for establishing the Infrastructure for Spatial Information in Europe [12].

In the geospatial world, the adoption of a federation based on common software tools is not feasible beyond small cohesive communities. But also the definition of a standard-based federation seems difficult and the attempts of building multidisciplinary federations were not fully successful. Indeed, the definition of a common standard fitting all the possible geoinformation systems is inherently difficult and it has many theoretical obstacles. First of all, even assuming that such a common standard suite – the one standard – would exist in the end, it would not be tailored to very specific but relevant domain needs, making the majority of the adopting systems underperforming. Secondly, the potential one standard, addressing multiple multidisciplinary requirements, would be very complex: some application developers would not have sufficient IT expertise to adopt and implement it. Finally, technologies are continuously evolving, and the evolution of systems would soon produce a misalignment of technologies and architectures.

The *brokered approach* [13] overcomes those difficulties, since no common model is defined, and participating systems can adopt or maintain their preferred interfaces, metadata and data models. Specific components (the *brokers*) are in charge of accessing the participant systems, providing all the required mediation and harmonization functionalities. No (major) re-engineering of existing systems is required, thus lowering the entry barrier for data providers. This approach fits well in situations where the SoS governance does not have a specific mandate, and where the participant organization does not have a strong interest/commitment to be part of the SoS. In the geospatial world, GEOSS is the typical example of an overarching initiative where a third party, the Group on Earth Observation (GEO), has a specific interest in building a SoS collecting existing data systems with their own mandate and governance. The brokered approach is also useful when the participant organization do not have the expertise for complying with complex specifications, as it is common in the Web world. A further advantage of brokering is that introducing a middle tier between clients and servers, the brokers may also add value providing advanced (e.g. semantic) services supplementing server capabilities. The main drawbacks of the brokering approach are the complexity of brokers, and the governance of the brokering infrastructure [14].

In ENERGIC OD, the choice of brokered architectures for the Virtual Hubs is fully justified by three main reasons: a) several data sources of interest are provided through heterogeneous protocols; b) ENERGIC OD has neither the mandate nor the capacity to impose and enforce standards or any federated model to the provider sub-systems; c) advanced brokering tools are already available for re-use.

3.5. Addressing usability: interfaces and APIs

In the ENERGIC OD context, usability specifically relates with the general requirement of facilitating the use of geospatial open data. We can expect that the Virtual Hubs target users (i.e. app developers) have heterogeneous competence and expertise in geospatial technologies. Therefore, ENERGIC OD aims to provide multiple ways of interaction with the Virtual Hubs, ranging from those based on powerful widespread geospatial specifications (e.g. OGC and ISO), to lightweight specifications (e.g. OpenSearch), and finally to simple Application Programming Interfaces (APIs) implemented in JavaScript for rapid development.

Figure 2 depicts the interaction of users with the ENERGIC OD Virtual Hub. A portal handles the human-to-machine interaction; machine-to-machine interaction exploits the available interfaces and Web APIs. Inside the Virtual Hub, a set of brokers support the use of open data providing harmonized discovery and access functionalities, enriched with semantic and transformation capabilities.

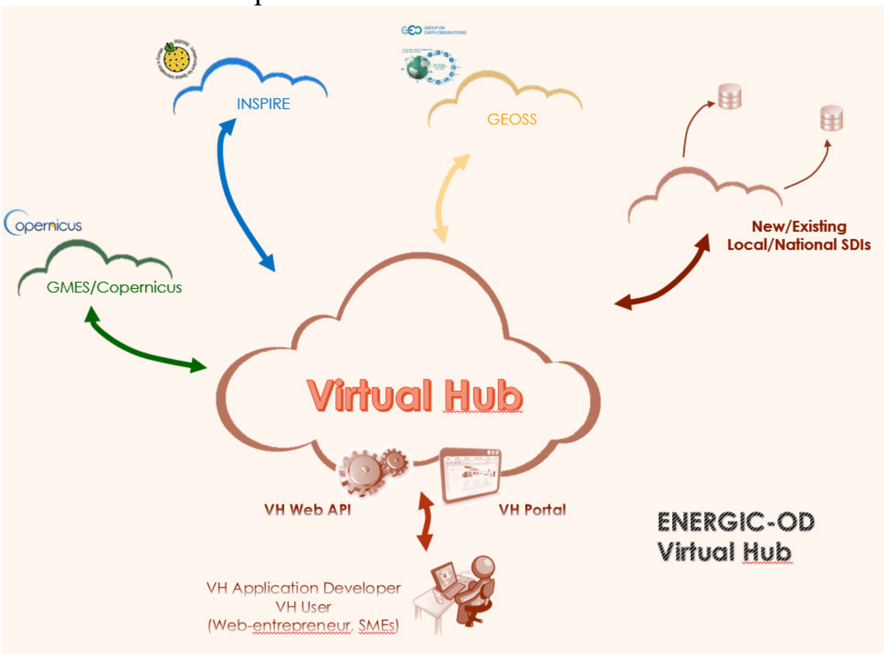

Figure 2 User interaction with the ENERGIC OD Virtual Hub

4. Technology Description

As an innovation project, ENERGIC OD can dedicate only minor effort to research; therefore it adopts an approach for development based on integration of existing tools. ENERGIC OD selected potential technologies through proposal by Consortium partners and through a survey of relevant research projects (mainly funded under the Seventh Framework Programme for Research and Innovation of the European Union) which delivered technological solutions and tools that have now reached a mature status [15].

As described in section 3, ENERGIC OD adopts a brokered architecture to build Virtual Hubs as System of Systems. Thus, brokers play a fundamental role in the implementation of ENERGIC OD Virtual Hubs. The ENERGIC OD proposal was actually built around a suite of geospatial data brokers developed and enhanced in the course of several FP7 projects including EuroGEOSS, UncertWeb and GEOWOW: the GI-suite Brokering Framework [16].

The GI-suite Brokering Framework implements discovery and access brokering towards more than 40 different data source types, publishing more than 10 different interfaces. It supports both metadata harvesting and distributed queries, configurable per data source. It also implements value-added services such as: a) semantic services implemented through query expansion using external knowledge bases (accessible through a SPARQL/SKOS interface and model) [17]; and b) dataset harmonization based on transformations for subsetting, reprojection and format encoding, using internal algorithms or external services (exposed by an OGC Web Processing Service interface) [18].

The GI-suite Brokering Framework is complemented by the GEO API, a JavaScript implementation of Web APIs facilitating the creation of Web and mobile application using the GI-suite Brokering Framework [19].

The GI-suite Brokering Framework is currently adopted in the GEO DAB (Discovery and Access Broker), one of the core components of the GEOSS Common Infrastructure (GCI) where it proved to be able to interconnect many heterogeneous systems and to handle a large amount of geospatial data [20]. The GEO DAB is connected to a semantic service managed by the Joint Research Centre of the European Commission (EC-JRC), which publishes a knowledge base composed of a set of aligned thesauri and ontologies including the GEneral Multilingual Environmental Thesaurus[9] (GEMET) for multilingual support.

During the course of the project, ENERGIC OD will integrate the GI-suite Brokering Framework to other tools and frameworks for supporting other requirements (e.g. data publishing for crowdsourcing apps, access control, etc.). ENERGIC OD will also extend the Web APIs to provide user-friendly access to the new functionalities by target users.

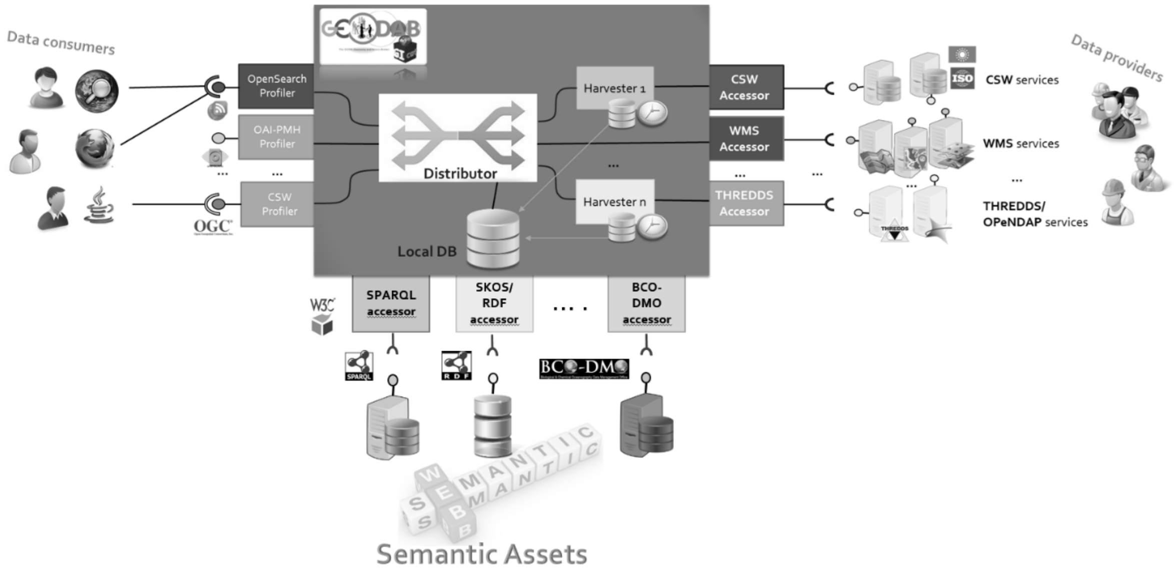

Figure 3 The GI-suite Brokering Framework: discovery broker architecture

5. Developments

The development of the ENERGIC OD Virtual Hubs will span for three years with major yearly releases. The first version, based on the GI-suite Brokering Framework and related extensions, has been delivered in July 2015. It provides the following main functionalities:

- discovery of multiple heterogeneous data sources. Each data source connection can be configured either for distributed query (if supported) or harvesting. Users can perform discovery using any of the exposed interfaces.
- semantic enhancements for discovery accessing external knowledge bases.
- access of multiple heterogeneous data sources. Users can perform access using any of the exposed interfaces.
- data transformations (format, coordinate reference system) using internal or external services
- configuration of data sources and exposed interfaces through a web tool
- test portal including semantic concept browser
- Javascript library implementing APIs for rapid web and mobile app development

[9] http://www.eionet.europa.eu/gemet

The first version of ENERGIC OD Virtual Hub has been deployed on three national-level sites:

- The Italian Virtual Hub (VH-IT) has been deployed on Amazon commercial cloud services. It already connects several geospatial open data sources mainly from Italian Public Authorities (Municipalities, Provinces and regional authorities).
- The German Virtual Hub (VH-DE) has also been deployed on Amazon commercial cloud services. A set of data sources is under selection.
- The Spanish Virtual Hub (VH-ES) has been deployed on University of Zaragoza premises. It currently connects the Zaragoza Open Data Platform and Zaragoza SDI.

Two other national-level sites – the French Virtual Hub (VH-FR) and Polish Virtual Hub (VH-PL) – are under deployment. All current instances are connected to the EC-JRC knowledge base (see section 4) for semantic services.

To assess and evaluate the effectiveness of the Virtual Hubs concept and implementation, ENERGIC OD is developing ten Web or mobile pilot applications. They are characterized by different requirements, allowing to test a wide range of Virtual Hub capabilities. The ENERGIC OD pilot applications will also contribute to the definition of the business models and benefiters of the Virtual Hubs, as they will also have to define their own business models (with some of the apps having already defined them). Indeed, pilot apps have a specific impact on several European and International initiatives like:

1) **EU2020 Societal challenges.** Informed decision making is recognised as crucial for all PAs called to tackle the challenges that contemporary society is facing. For examples, European Commission makes a clear reference to the fact that “intelligent processing of data is essential for addressing societal challenges” in its communication on Open Data in 2011.
2) **European Innovation Partnerships (EIPs).** This new approach to research and innovation is challenge-driven focusing on societal benefits and rapid modernisation of the associated sectors and markets. Some EIPs, such as Smart Cities or Agriculture, are particularly sensitive to advancement in geo-information sector.
3) **GEO Societal Benefit Areas (SBAs).** This last category was considered because of the global impact of GEO and GEOSS initiatives. The use of geo-spatial information at global scale was considered important for ENERGIC OD in terms of an international framework to local challenges.

The development of pilot applications has started in September 2015 and will include two yearly releases, with the first release planned on September 2016.

The preliminary work on the applications, including interoperability tests with the identified data sources and client technologies, allowed to identify broker extensions needed to support required capabilities. For example an extension (accessor) to access SODA (Socrata Open Data API) is required to connect some data sources, and extensions (profilers) for new exposed interfaces are required to provide GeoJSON data to clients and portrayal access through the OGC Web Map Tiled Service Standard (WMTS). Moreover, a coordinate reference system transformation service must be implemented in the Italian Virtual Hub to support the national Gauss-Boaga reference system.

Future works for the second release of the ENERGIC OD of the Virtual Hubs include: a) the development of the identified extensions; b) the integration of sensor observation publishing and access services based on OGC SOS specifications to support crowdsourcing applications; c) the development of the first release of the ten pilot applications. Moreover, following an agile approach it is expected that new requirements will be discovered during the app development phase.

6. Results

The current instances of the Virtual Hubs are accessed both internally in the Consortium and externally by software houses involved in the apps development as subcontractors. During the second year open access to the test portals and public demonstrations are planned. Although no formal validation has been carried out yet, it is possible to present some lessons learned from preliminary feedback and from previous experiences in brokered Systems of Systems adopting the same approach and base technology, such as GEOSS.

On the positive side, the approach seems effectively reaching its objective of facilitating the use of open data. Through any of the many interfaces provided by the brokers, users can seamlessly discovery and access multiple heterogeneous data sources without needing to know complex geospatial data specifications. The provided Web APIs allow the rapid development of Javascript-based applications. They have been successfully used for developing/enhancing community portals for GEOSS, and first reports from app developers confirm their usefulness and user-friendliness. Moreover, the semantic services provide highly valuable services. For example, the support of GEMET thesaurus to expand keywords allows to address multilingualism: the user can search through datasets even if metadata are compiled in a different language, as it was assessed by the project through the survey of existing OD platforms [21]. This is a significant and essential feature in the world of open data which are often provided by public authorities with local/national relevance. However, for more advanced semantic services, such as query by concepts, ontologies tailored to the geospatial open data would be useful. Besides GEMET, the ontologies and vocabularies provided by the EC-JRC service currently connected to the Virtual Hubs, are designed for Earth Observation and they are not fully effective for different use-cases.

An important issue raised during the validation of ENERGIC OD Virtual Hubs is the reliability and quality of data sources. A lot of potentially useful data sources cannot be fully exploited due to poor quality (lack of a catalogue service, incomplete metadata) or reliability (low time of response, service often not available). Brokers can supply some required functionality (for example they provide catalogue interface for harvested metadata), but they cannot do anything in other cases (for example they cannot generate metadata from missing information). Other solutions need to be considered to address these issues. In particular a data source monitoring and validation service is under design to be integrated in the future releases of the Virtual Hubs. Besides technical aspects, policy issues are still a relevant barrier to geospatial open data sharing. For many potential data sources, ENERGIC OD did not found any explicit policy declaration, while others have heterogeneous policies posing problems in their integration and exploitation.

7. Business benefits

We expect that Virtual Hubs can provide significant business benefits supporting the creation of new market opportunities. Lowering entry barriers for both geospatial open data providers and users can deliver direct and indirect benefits. Indirectly, facilitating the connection of new data sources needing no (major) change in their infrastructure, a Virtual Hub may increase the amount of open data available, as well as their usability. At the same time, facilitating the use of open data may increase the number of developers building apps with geospatial open data. This could start a virtuous cycle with increasing data sources which attract developers, and on the other hand, a greater number of developers stimulating providers to publish more datasets. Facilitating the use of open data has also a direct benefit reducing the app development time. Accessing Virtual Hubs, a developer does not need to be a geospatial interoperability expert to create innovative applications. The learning curve is reduced, and the developer can focus on his/her business idea. Obviously nothing comes from nothing: the easier use of datasets depends on the development of complex components – the brokers. However, this complies with the separation-of-concerns engineering principle with different tasks assigned to different responsible actors: interoperability experts can focus on the brokers development, leaving providers and users to dedicate to their main tasks. Application development is not the only market addressed by Virtual Hubs. Service providers can run business in the management of Virtual Hubs.

To realize these expected benefits some actions are required. The first one is to fully bring Virtual Hubs to the market. As an innovation project, ENERGIC OD aims to demonstrate the validity of the Virtual Hubs concept, delivering multiple Virtual Hub running instances and a software package that can be used for implementing further Virtual Hubs. To create a market around the Virtual Hub concept it is necessary to define possible business models sustaining the Virtual Hub management and apps development lifecycle. Experiences from the cloud service providers, app stores, and open data domains may be useful. The second action is the definition Virtual Hub governance aspects, with the identification of actors, roles (Virtual Hub managers, administrators, etc.) and business processes for the Virtual Hub administration, evolution, app development training and support, etc.

ENERGIC OD will explore some of these issues defining an exploitation plan to be released at the end of the project (September 2017). Additionally, the ENERGIC OD pilot applications will also contribute to the definition of the business models and benefiters of the Virtual Hubs, as they will also have to define their own business models (with some of the apps having already defined them).

8. Conclusions

The Virtual Hub concept, as conceived and developed in the ENERGIC OD European project, proposes an innovative approach to geospatial Open Data sharing. Based on previous experiences in the research world, instead of creating yet another sharing system - following the long and complex process of defining new standards and specifications - it leverages the existing systems as they are, focusing on lowering entry and usage barriers. Indeed, ENERGIC OD recognizes that current geospatial and Web technologies allow to address the main technological barriers hampering the usage of geospatial open data. Brokers, knowledge bases, widespread Web technologies can highly facilitate the integration of new data sources, and the development of applications using geospatial open data.

The first release of Virtual Hubs, providing seamless discovery of and access to heterogeneous datasets from disparate data services, already shows that the brokered approach greatly facilitate discovery, access and use of geospatial open data. Future works will focus on providing further capabilities, making them also accessible through simple APIs. The final objective is the deployment of fully-functional Virtual Hubs demonstrating that many complex tasks related to data interoperability and usability can be automated. These ENERGIC OD Virtual Hubs will facilitate the use of geospatial open data, allowing Web and mobile developers to focus on their own business applications, enhancing the opportunities for creating new markets and new business opportunities.

Acknowledgments

ENERGIC OD has received funding from the European Union ICT Policy Support Programme (ICT PSP) under the Competitiveness and Innovation Framework Programme (CIP), grant agreement n° 620400. The project, started in October 2014 and ending in September 2017, is coordinated by the National Research Council of Italy with participation from European research centres, public administrations and SMEs.

References

[1] J. Manyika, M. Chui, D. Farrell, S. Van Kuiken, P. Groves e E. Almasi Doshi, «Open data: Unlocking innovation and performance with liquid information,» McKinsey Global Institute, 2013.

[2] Open Knowledge Foundation, «Value Stories,» 2015. [Online]. Available: http://opendatahandbook.org/value-stories/en/.

[3] A. Donovan, R. Finn, K. Wadhwa, L. Bigagli, G. Vega Gorgojo e M. G. Skjaeveland, «Open Access to Data,» BYTE Consortium, 2014.

[4] European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions,* 2011.

[5] European Parliament and Council of the European Union, *Regulation (EU) No 1291/2013 of the European Parliament and of the Council of 11 December 2013 establishing Horizon 2020 - the Framework Programme for Research and Innovation (2014-2020) and repealing Decision No 1982/2006/EC,* 2013.

[6] T. Berners-Lee, “Linked Data,” 18 June 2009. [Online]. Available: http://www.w3.org/DesignIssues/LinkedData.html. [Accessed 25 May 2015].

[7] «Competitiveness and Innovation Framework Programme (CIP) ICT Policy Support Programme ICT PSP Work Programme,» 2013.

[8] American Heritage® Dictionary of the English Language - Fifth Edition, Houghton Mifflin Harcourt Publishing Company, 2011.

[9] F. J. Chang, Business Process Management Systems: Strategy and Implementation, CRC Press, 2005.

[10] M. Jamshidi, System of Systems Engineering - Innovations for the 21st Century, John Wiley & Sons, 2009.

[11] OGC, «OGC History (abbreviated),» [Online]. Available: http://www.opengeospatial.org/ogc/history.

[12] European Parliament; European Council, *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),* 2007.

[13] S. Nativi, M. Craglia e J. Pearlman, «Earth Science Infrastructures Interoperability: The Brokering Approach,» *IEEE JSTARS,* vol. 6, n. 3, pp. 1118-1129, 2013.

[14] RDA Brokering Governance WG, «RDA Brokering Governance Case Statement,» 2014. [Online]. Available: https://www.rd-alliance.org/group/brokering-governance/case-statement/rda-brokering-governance-case-statement.html.

[15] ENERGIC OD Consortium, «R&D State-of-the-art report,» 2015.

[16] ESSI-Lab, «GI-cat Homepage,» 2014. [Online]. Available: http://essi-lab.eu/do/view/GIcat/WebHome.

[17] M. Santoro, P. Mazzetti, S. Nativi, C. Fugazza, C. Granell e L. Diaz, «Methodologies for Augmented Discovery of Geospatial Resources,» in *Discovery of Geospatial Resources: Methodologies, Technologies, and Emergent Applications*, IGI Global, 2012.

[18] ESSI-Lab, «The GEOSS Discovery and Access Broker APIs,» 2014. [Online]. Available: http://api.eurogeoss-broker.eu/docs/index.html.

[19] S. Nativi, P. Mazzetti, M. Santoro, F. Papeschi, M. Craglia e O. Ochiai, «Big Data challenges in building the Global Earth Observation System of Systems,» *Environmental Modelling & Software,* vol. 68, pp. 1-26, 2015.

[20] GEO, «The GEOSS Common Infrastructure (GCI),» [Online]. Available: https://www.earthobservations.org/geoss.php.

[21] ENERGIC OD Consortium, «Report on survey among project partners and on survey with stakeholders,» 2015.