

RDA Brokering Governance WG

Executive Summary

With the exception of standards, there is little precedence for long-term sustainability of advanced forms of interoperability tools. Effective brokering governance has the potential to support a stable, sustainable middleware capability under a variety of operation and funding models. The Brokering approach or “Framework” provides a series of services such as discovery, access, transformation and semantics support to enable translation from one discipline/culture to another. The translation across cultures is simplified by having a common “technical model” embedded in the broker framework with the translation to and from different disciplines handled by facilitators called “accessors”. To ensure sustainable, stable development and operational environments, an effective model for the governance and reuse of brokering middleware is necessary and will be investigated. The governance structure is the interface and framework between the broker, the repositories and the users. As it will reward and reinforce certain actions, it plays a critical part in guiding the development, use and effectiveness of the broker middleware solutions for interoperability.

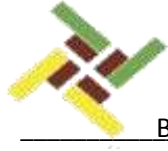
To address the Governance of the brokering framework, three major activities will be carried out:

- 1) Define the scope of governance and priority issues;
- 2) test/evaluation of a governance models and updates defined in the testing; and
- 3) recommendation to RDA for governance options.

Working closely with adopters and engaging with them in use cases, an initial model will be evaluated and refined, leading to a recommendation to RDA and the Community. The expected outcomes of the Brokering Governance WG will be:

- A Position Paper including guidelines and best practices for a **governance approach**.
- **Test (and refinement) of a governance model** to be carried out by the adopters participating in the WG.
- **A recommendation document for the RDA**, including a consensus on paths for adoption of this capability at the international level.

The Working Group will broadly invite participation including experts in the social sciences and humanities. It will interface with other working groups to reach across RDA for inputs and feedback.



Introduction

Multidisciplinary research interoperability is a key challenge for RDA to achieve open research data sharing without barriers.

For disciplinary applications, systems interoperability largely deals with the adoption of agreed technologies, standards, and specifications. However, such a multi-disciplinary approach makes complex demands on the type of systems and arrangements needed to support it. Thus, interoperability among diverse disciplinary systems must be pursued adopting more flexible approaches that reduce the demands on existing and new information infrastructures and that can be both scalable and sustainable. In this context, sustainability deals with many factors including: the ability to adapt to scientific and technological innovation; education and outreach; financial stability.

For multidisciplinary research interoperability (especially in a global dimension), it is unrealistic to expect that all software components or repositories of different disciplines will use the same specification to interoperate. Communities in different disciplines have evolved to support the needs of their own research scientists and users. If the diversity characterizing different Communities can be preserved in moving toward interoperability, the needed evolution toward interdisciplinary interoperability will move more rapidly. Early approaches for interoperability have focused on standards and uniform specifications with the goal of having uniform interfaces adopted by repositories. However, experience shows that standards do not guarantee interoperability because of the differences in interpretation and implementation. Therefore, while there is a drive to adopt common specifications at the disciplinary level, mediation and harmonization are essential to pursue multidisciplinary research in an effective way. Brokers are powerful instruments implementing mediation, distribution, harmonization, and transformation functionalities in a many-to-many context for existing services and components managed by different Communities. These can be architected to scale as N , the number of engaged systems, rather than N squared and thus offer a path to large interoperability networks. The next section on the Broker concept will provide more details.

The Broker Concept and Approach

In an ecosystem of domain infrastructures, multi-disciplinary interoperability has been traditionally pursued on a one-to-one basis or by asking the stakeholders (i.e., both users and resource providers) to be able to utilize the plethora of service buses characterizing the different disciplinary infrastructures. Clearly, this has represented a high entry barrier for developing cross-disciplinary science and applications.

A new approach called the “Brokering “Framework” has been developed that provides a series of services such as discovery, access, transformation and semantics support to enable translation from one discipline/culture to another that may be quite different. The translation across cultures is



simplified by having a common “model” embedded in the broker framework with the translation to and from different disciplines handled by facilitators called “*accessors*”. In this translation, the interfaces used by large information infrastructures are assumed to be stable in the short run. The broker must adapt when any interfaces change. Notification of change is an example of a governance element that enables the broker to function efficiently. Stakeholders are both information users and information providers. Since it is hypothesized that the users will use and access the broker through their discipline portals, the focus is then to connect different discipline infrastructures. The stakeholders envisioned for the governance working-group are the information systems and infrastructures. This hypothesis is, of course, a subject of continuing discussion and will be addressed as part of the governance discussions.

As depicted in Fig. 1, the Brokering approach introduces a new middleware layer of service offerings: the Brokering framework is depicted in the figure as a cloud. This should contain all the necessary existing (and new) components/services such as brokers to implement interoperability among present (and future) service buses of different disciplines. Therefore, a Broker may be defined as an intermediary service dynamically implementing a many-to-many interconnection for a Client-Server framework. This is done by defining and implementing as series of *accessors* (advanced mediators) that translate the discipline infrastructure attributes into a common framework. The current system is in use internationally in programs and initiatives such as GEO GEOSS, ICSU WDS, NSF Earth Cube, IODE.

A major focus in the development of the Brokering approach is to minimize the efforts required for discipline and other infrastructures to participate. To this end, it is based on the following principles [Nativi et al. 2011]:

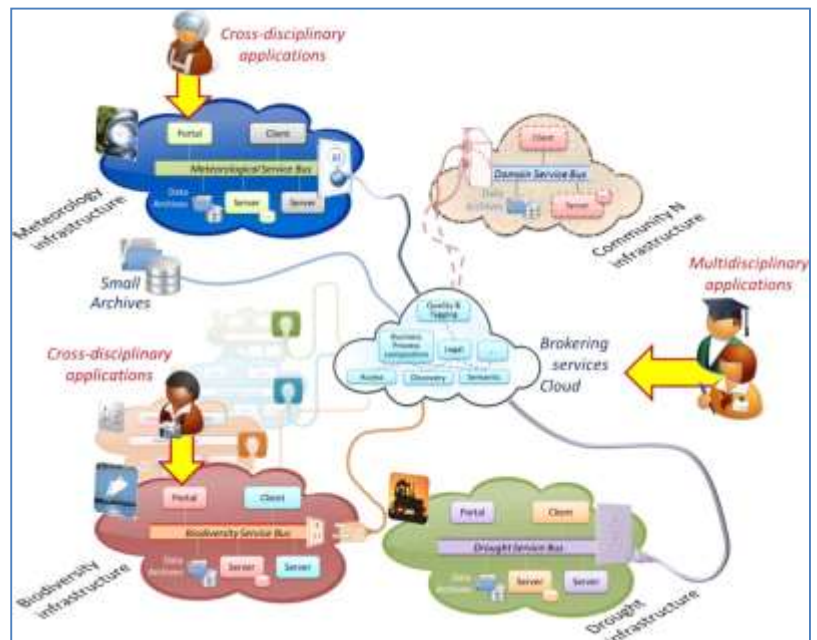
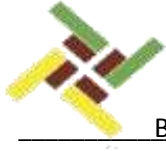


Figure 1 The Broker Framework support interoperability among diverse disciplines

- a) **Autonomy:** Keep the existing disciplinary infrastructures as autonomous as possible, not asking them to implement any “more general” service bus.
- b) **Subsidiarity:** Supplement but not supplant disciplinary infrastructure mandates and governance arrangements by interconnecting and mediating their service buses.



- c) Interconnection: Build incrementally on existing infrastructures and introduce distribution and mediation functionalities to interconnect the heterogeneous service buses characterizing any domain specific or other infrastructure.
- d) Low entry barrier: Minimize the barrier for both users and resource providers of any disciplinary infrastructure.
- e) Flexibility: Be flexible enough to accommodate existing and future information systems and information technologies that will augment the service bus implemented by any discipline.
- f) Scalability: Support the access to growing data resources and interconnected systems without a fundamental change of the overall architecture and supported specifications.
- g) Effectiveness: Address the full range of information exchange needs (discovery, access, semantics, workflow, etc.).

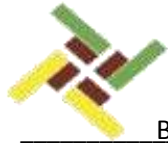
WG Charter

The brokering capabilities discussed in the previous section should be considered in the context of RDA and the science research objectives for open data exchange. In supporting open data research objectives, there should be motivation to go beyond just connecting systems to a capability and environment to connect meaning.

It is quickly obvious that brokers can move in this direction and also that there are other critical elements being examined in RDA from semantics to provenance to policy that will ultimately form a symbiotic environment for effective open data exchange. The larger context then includes technical developments, but must also engage in the more complex interfaces of systems, social interactions and governance.

The RDA WG on Brokering Governance will address selected issues of the larger RDA context with the recognition that any efforts on governance must be focused to provide deliverables in 18 months. In fact, the original concept was to look at middleware operations and sustainability. With the need to support test and stakeholder demonstration as part of the outcome of a working group, the Broker Interest Group proposes to use the current broker framework as developed for GEOSS by the Italian National Research Center as a tool in collaboration with selected adopters/stakeholders to examine aspects of governance and interoperability. The outcome will be recommendations for a broker governance approach based on stakeholder testing and identification of workable, practical solutions for facilitating the exchange of data in large-scale networks.

Governance, in this context, addresses the relations between organizations (institutions, governments or others) that enable a system to be created and to operate. It can include technical operations



agreements, policy or funding paradigms. To do this, the Brokering Governance **WG will start with a governance model described below and, in the process, interface with other RDA WGs and the broader Community to refine how brokering can be sustainably supported to further the larger RDA objectives.** This is addressed more fully below.

Initial Governance Model

In formulating the model, the chairs of the Broker interest group have examined a variety of existing or developing models that are also addressing the objectives of sustainable services. The issue of sustainability, for example, is a key theme of the research community in addressing open data. Previous models supported publication, peer-review and archiving through library subscriptions. With the requirement for free and open data, alternatives for maintaining peer review and archive infrastructure are being discussed with a trend toward payment by authors rather than libraries. On line publications do not have the production costs and this will eventually shape the discussion. For example, the IEEE Oceanic Engineering Journal will no longer be available in printed version starting next months after years of operation in print format. The open software initiatives are working through similar issues.

For the broker, a number of models for sustainable operations have been discussed. The list below is not complete, but illustrates some of the options. Some of the models are not appropriate for the research community such as the Google ad and information sales model. Others may inhibit the community adoption because of labor or cost barriers. It would be disingenuous not to acknowledge that the ideal model is long term funding by government(s). However, the economic cycles and changes in priority for research funding may bring into question whether some mixed model could provide longer-term stability. Looking to the open source community, there are a number of models that offer potential options. Here are some of the models in current use.

Information and Ad sales

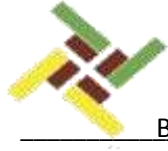
Google is available at no cost for search and for visualization of earth information. Google is supported by advertising and sale of collected information. Facebook has the same model. This is not appropriate for broker middleware.

Product (Document) Sales

- Standards organizations (IEEE, ISO, etc.) sell standards documents and rely on volunteers and corporate participation to formulate standards.

Corporate Support

- OGC has a membership model with fees for participation (different levels are available) and relies on volunteers for standards formulation working groups.



- The Open Source Initiative is moving from a volunteer base to a member/affiliate base. They focus on licenses. The financial base comes from corporate sponsors.

“Software as a Service” (SAS) Model

- Companies provide a mixture of base and enhanced services. WordPress has an open source component (wordpress.org) and a service component (wordpress.com) The latter offers enhanced services for fee. Redhat follows the same model with their web page (<http://www.redhat.com/en>) offering “ open technology, constant improvement, complete control”. This model can work through individual sales or large scale subscriptions. Wikipedia defines a similar freemium model -“Freemium is a pricing strategy by which a product or service (typically a digital offering such as software, media, games or web services) is provided free of charge, but money (premium) is charged for proprietary features, functionality, or virtual goods

Government Funding

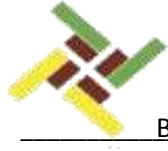
- GEOSS solicits support from governments for their secretariat operations, both in funds and in staff assignments. The activities are supported through in kind contributions of organizations, government and nonprofit.
- SeaDataNet provides an information service based on government grants.

Selection of an Initial Model

The selection of model has certain underlying motivations. First, it would be good to build on a governance/business model that has been used for software and is widespread in the community. There are also preferences for an open source modality that provides for both community support and sustainability. A governance model should also offer an early start in working with adopters – in this case, three major initiatives that have committed to participate as adopters: ICSU/WDS, JRC Danube Project and GEO-Bon (see below).

The model for Broker governance must address certain core requirements: it should: (1) support users and further interoperability; (2) be sustainable; support national and international policies; (3) support core technical capability advancement, be accessible to a wide range of users; (4) create a flexible adaptable framework for incorporation of new developments; and (5) offer a range of services essential to multi-disciplinary science collaborations. This range of services is expected to grow.

As mentioned earlier, the broker framework supports discovery and access, with increasing capability in semantics and business flows (models). When less mature services evolve, the timing when they transition from development to operational capabilities and thus under the operational governance needs to be addressed as part of the governance.

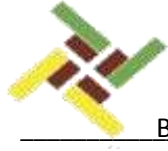


The initial model for operational funding which will be reviewed and tested is a “Software as a Service” Model that is common in the software community and has shown success in a number of software application environments. In this model, a range of basic services (such as templates to develop accessors) would be provided at no charge to users. Where accessors have already been built, they will also be available at no charge. To maintain uniform standards for new accessors, a registration and validation fee could be charged; a test bed could be maintained for performance assessments and validation. For organizations that want enhanced services, e.g., accessors built and other software support, a fee-based service will be available. Additional services such as single point sign-on, access to quality indicators, prioritized discovery, are examples of services that will be considered and tested with the adopters participating in the working group. Service charges could be adapted to single users, but it is anticipated that the majority of parties would be institutions or governments. These could work on a tiered subscription basis that is common among software service providers and needs to be adapted for the brokering class of services. Again, it should be remembered that the preferred model is government-funded operations with free services to the community. We do not believe that this is precluded in any of the working group activities, but alternative models should be examined to provide robust governance recommendations.

Testing of some elements considered for the model have already begun under the U.S. EarthCube broker project (BCube) where templates for accessors are being tested to enable organizations to build their own accessors. The close ties between BCube and the RDA Broker Governance WG will provide additional benefits over the next year.

Other factors than sustainability impact the middleware governance and will be addressed in the Broker WG. There is a need for stability in the translation from local models for data and metadata to the internal broker model. Memoranda of Understanding (MOU) would thus be requested for participation and include agreements for notification of changes in formats for data or metadata or changes in web interface protocols, confirmation of access requirements and release policy. Similarly, restrictions on use or reuse will need to be identified and kept up to date including requirements for sign-on and authorization. Legal aspects such as Intellectual Property Rights and security requirements for data uptake and distribution will need to be included in the MOU. This may become a more difficult issue with the increased hacking observed in recent years.

The question of valuing interoperability is dependent on government policy and user demand. The government is funding capabilities for interoperability through projects such as EarthCube, ODIP, COOPEUS and RDA. Whether Data Infrastructures receive mandates for such interoperability is unclear. The requirements to fulfill such mandates could alter the dialogue.



The above model will be examined in detail during the first two months of the Broker Working Group. This will be done with the adopters in the Working Group and also with other working and interest group internal to RDA and with organizations outside.

To meet the schedule detailed below, the WG will build upon an existing technical capability to address the configuration and strategy for a sustainable and scalable implementation and operation at the system level. It is not clear that one size fits. There will certainly be debates between various philosophies on data management and use. In the longer term, to achieve practical convergence to a recommendation, one aspect of the WG directions involves understanding how to optimize models for Community adoption. For this, we need to define a transparency and documentation strategy for both the system and the software as part of the governance assessment. Much of these particular activities involve the social aspects of interoperability. Thus, the WG will solicit participation of social scientists. This is being done for the EarthCube Program and for the Broker Project BCube, to good effect.

To address the Governance of the brokering framework, three major activities must be carried out:

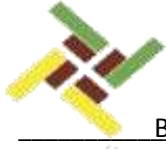
- review and refinement of governance model and priority issues discussed above
- test/evaluation of governance models including updates from tests; and
- recommendations to RDA for governance approaches.

Specifically, the following activities will be addressed:

- Brokering process definition and definition of terms;
- Review of initial governance model; considerations of options
- Stakeholders apply and test a selected governance model;
- Analysis of governance model – examination of updates;
- Develop recommendations for a brokering framework governance approach;
- Review recommendation with a broad stakeholder and RDA Communities; and
- Report writing.

As described in the previous section, brokering philosophy is about accepting disciplinary “diversity” and providing the necessary components to mediate and interconnect. It should be a service to the science community, a service that can be relied upon long-term and is self sustaining.

Adopter Participation in Working Group



Use cases will be utilized to test governance models with adopters. The Brokering IG identified three use cases during their meeting in Dublin in which adopters committed their facilities to participate (including points of contact):

- International repositories: ICSU WDS (Michael Diepenbroek, Mustapha Mokrane)
- Environmental sciences: European Commission Danube SDI (Max Craglia)
- Global Changes: GEO-BON (Wim Hugo)

GEOSS and EarthCube have interest in furthering brokering through grants or agreements and both organizations will be represented in the working group. Two of the co-chairs of the proposed working group are working in EarthCube through and NSF grant called BCube, which is further developing the brokering capabilities considered by the proposed RDA working group. There will be a strong and continuing interface between the working group and EarthCube.

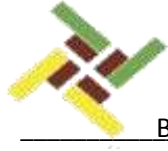
Value Proposition and Deliverables

With the exception of standards, there is little precedence for long-term sustainability of advanced forms of interoperability tools. The Community has many debates on the trades between open source and other forms of software development; the issue is still open as noted in the large variety of governance models described above. For open data exchange, the practices shift in response to different funding and property models, under different architectures. To ensure sustainable, stable development and effectiveness in an operational environment, an effective model for the governance and reuse of brokering middleware must be tested.

The WG will consider and recommend a set of best practices, and a model with related options, for governing and managing brokering middleware to facilitate and enable broader capabilities to be part of an operational framework. The results will be of value not only to architects and to developers (who can plan integrated systems assuming the continued use and support of brokering middleware) but also to system managers and end users. The potential for scaling and expansion of integrated data resources and systems in brokering middleware is of value to increasingly interdisciplinary research work as well as in managing growing big data sets.

The expected outcomes of the Brokering Governance WG will be:

- A Position Paper including guidelines for **governance models**.
- **Test of a selected governance model** to be carried out by the Adopters (see above) who participate in the WG.
- **A recommendation document for the RDA**, including a consensus on paths for adoption of this capability at the international level.



The working group will work with three adopters (ICSU, EC Danube and Biodiveristy), but will reach out to other communities including the social sciences for use cases, with consideration of these later in the first year as part of the evolution to governance model testing.

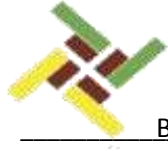
Work Plan

Evaluation of Governance models must involve working with adopters and using their feedback to refine the model through the course of the first year. This will involve an MOU, the selection of an array of data sets for testing and the creation (by the adopter or the service provider (CNR)) of accessors. Of course, some 40 accessors exist and will be used where appropriate. The community portal will have a broker interface introduced to provide users and the adopter with hands-on experience. The NSF BCube test bed will be used to validate performance of adopter developed accessors; this is already being done as part of the EarthCube program in a collaboration between BCube and GEOWS building blocks.

Various aspects of the governance model described above will be iterated with the detailed schedule for testing and evaluation agreed with each of the three adopters during the first task in the work plan. As testing occurs, documentation of issues and lessons learned will be done in the existing tracking system that is part of CNR development and BCube. Early in the work plan, reviews with other working groups will be done, with an emphasis on the RDA Washington DC meeting in November if the working group starts in September (this will allow enough time for more detailed definition of each of the tasks and capabilities with the WG adopters). Recommendations and best practices will be developed based on work with the adopters and the technical CI community. The recommendations will be brought to the RDA community and the broader stakeholder communities for review and comment and a final report will incorporate these at the end of the 18 month life of the working group. The planned schedule for this is provided below.

The work plan considers the following main tasks:

- TB1: Brokering process definition and definition of terms; MOUs with adopters
- TB2: Review of initial governance model; considerations of options
- TB3: Stakeholders apply/test the governance model; document experience
- TB4: Analysis of governance model – examination of updates; testing of updates
- TB5: Develop recommendations for a brokering framework governance approach;
- TB6: Review recommendation with a broad stakeholder and RDA Communities;
and
- TB7: Report writing.



Task	Timeframe (months)	Expected outcome
TB1	M1-M2	Process definition and definition of terms; MOUs
TB2	M2-M4	Review of initial governance model; considerations of options
TB3	M5-M10	Stakeholders apply/test the governance model; experience documented
TB4	M8-M14	Analysis of governance model – examination and test of updates;
TB5	M13-M16	Recommendations for a brokering framework governance approach;
TB6	M14-M16	Stakeholders review of recommendations for a brokering framework
TB7	M16-M18	Final report

Community Adoption Plan

The first step is to engage in dialogue with other RDA working groups and disciplines to refine the objectives and tests for model development. This will engage users and developers.

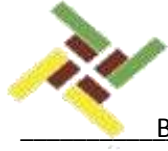
The next step of Community adoption involve the formation of the recommendations and stakeholders reviews These start at M13 and are part of the plan to encourage ownerships of the outcomes by the broader Community.

WG activities and outcomes will be presented at the major conferences and workshops dealing with research infrastructures and interoperability.

Initial Membership

An initial set of significant Stakeholders agreed to participate and support the use cases. These will be expanded. There will be an open call for membership upon the approval of the WG. Initial members of the WG are (alphabetically):

Point of Contact	Organization	Areas
Max Craglia	European Commission – JRC	e-Governance
Michael Diepenbroek	PANGEA/ICSU-WDS, Germany	Digital Library; Scientific Data Systems
Giuseppe Fiameni	CINECA, Italy	Supercomputing Centre
Milena Žic-Fuchs	University of Zagreb, Croatia	Social Science
Wim Hugo	SAEON/ICSU WDS, South Africa	Biodiversity
Bente Lilya Bye	BLB, Norway	Science & Technology; Arctic Data
Mustapha Mokrane	ICSU-WDS	Digital Library; Scientific Data Systems
Stefano Nativi	CNR-IIA, Italy	Information infrastructure; Earth System Science
Francoise Pearlman	J&F Enterprise, USA	Outreach and Public Engagement
Jay Pearlman	J&F Enterprise, USA	Information infrastructure
Roger Proctor	IMOS, Australia	Oceanography Data Systems
Stephen Slota	Univ. of California	Social Science
Tobias Spears	Fisheries and Oceans, Canada	Biodiversity (OBIS, ODIP, IODE)



The WG will be chaired by (alphabetically):

Max Craglia (European Commission –Joint Research Centre);
Stefano Nativi (Italian National Research Council);
Jay Pearlman (J&FE).

Useful References

- The EarthCube Brokering Concept Award Team, "**Brokering for EarthCube Communities: A Road Map**", 2012 [Online]. Available: http://api.ning.com/files/e1f7JW76fC4atLM7lc-wdrsh4bpyUb9knJhLXfEtBqjhhF*rgAdvZNTwzvtnWsqVSXA4EBbE6pNrBO*vpLKR1EiwePStjpnL/BrokeringRoadmapAug171.pdf
- Freemium web site: <http://www.freemium.org/>
- S. J. Khalsa, S. Nativi, and G. Geller, "**The GEOSS Interoperability Process Pilot Project (IP3)**," IEEE Trans. Geosci. Remote Sens., Special Issue on Data Archiving and Distribution, vol. 47, no. 1, pp. 80–91, Jan. 2009.
- S. Nativi and L. Bigagli, "**Discovery, mediation, and access services for earth observation data**," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 2, no. 4, pp. 233–240, 2009.
- S. Nativi, M. Craglia, J. Pearlman, 2013, **Earth Science Infrastructures Interoperability: The Brokering Approach**, IEEE JSTARS, Vol. 6 N. 3, pp. 1118-1129.
- S. Nativi, S. J. Khalsa, B. Domenico, M. Craglia, J. Pearlman, P. Mazzetti, and R. Rew, 2011, "**The Brokering Approach for Earth Science Cyberinfrastructure**", EarthCube White Paper. U.S. NSF [Online]. Available: http://semanticcommunity.info/@api/deki/files/13798/=010_Domenico.pdf
- S. Nativi, M. Craglia, and J. Pearlman, "The brokering approach for multidisciplinary interoperability: A position paper," Int. J. Spatial Data Infrastructures Res., vol. 7, pp. 1–15, 2012. Available: <http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/view/281/319>.
- S. Pearlman, "The Capitalist in the Co-Op: The Art & Science of the Premium WordPress Business", available at: <http://www.slideshare.net/shanepearlman/the-capitalist-in-the-coop-the-art-science-of-the-premium-wordpress-business>
- Redhat website: <http://www.redhat.com/en>)
- M. Santoro, P. Mazzetti, S. Nativi, C. Fugazza, C. Granell, and L. Diaz, "**Methodologies for augmented discovery of geospatial resources**", in Discovery of Geospatial Resources: Methodologies, Technologies, and Emergent Applications. Hershey, PA, USA: IGI Global, 2012.
- Schindler, U, Diepenbroek, M, 2008. **Generic XML-based Framework for Metadata Portals**. Computers & Geosciences 34 (12), 1947-1955. doi:10.1016/j.cageo.2008.02.023
- L. Vaccari, M. Craglia, C. Fugazza, S. Nativi, and M. Santoro, "**Integrative research: The EuroGEOSS experience**", IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 5, no. 6, pp. 1603–1611, 2012.