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

1) definition of governance and priority issues (including the scope for governance models);
2) formulation and test/evaluation of governance models; and
3) recommendation to RDA for governance options.

Working closely with stakeholders and engaging with them in use cases, one or more models will be formulated and evaluated, 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 of a selected governance model** to be carried out by the Stakeholders 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 WG will broadly invite participation including experts in the social sciences and humanities.

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.

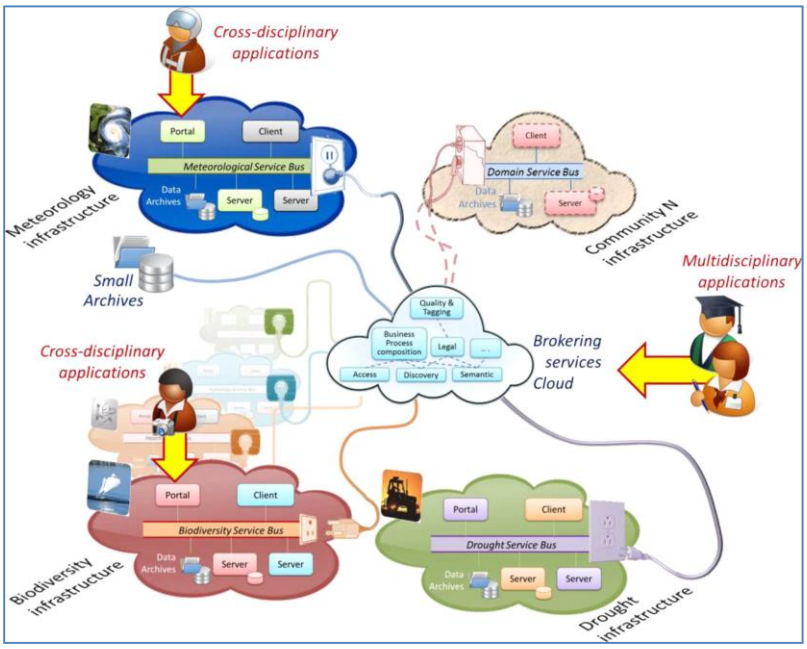

Figure 1 The Broker Framework support interoperability among diverse disciplines

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]:

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 proposed to use the current broker framework as developed for GEOSS by the Italian National Research Council (CNR-IIA) as a tool in collaboration with selected 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 develop a governance approach and, in the process, interface with other RDA WGs and the broader Community to refine how brokering can support larger objectives**. This is addressed more fully below.

In formulating the model, 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. Another is to have stakeholders involved in the process, end-to-end. Thus, the WG will conceive and test a model for guiding broker operations with stakeholders offering specific use cases for the assessment and trades. 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:

1) **definition of governance and priority issues** (including the scope of governance models);
2) **formulation and test/evaluation of governance models**;
3) **recommendations to RDA for governance approaches**.

Specifically, the following activities will be addressed:

1. **Policy and Users agreements** that impact brokering configuration and strategies (at both the overall system and single service level);
2. **Transparency and documentations** of the Brokering system and software;
3. **Community adoption**, **scalability** and **sustainability**;
4. **Bottom-up considerations** for defining one or more governance models through evaluation of use cases and interfacing with other RDA Groups;
5. **Definition of an approach for governance models** incorporating the above;
6. **Test of a selected model through specific use cases** with stakeholders that are part of the Working Group
7. **Community engagement** for inputs to refine the governance approach; and
8. **Recommendations to RDA** for governance options.

In addressing these, there must be a consensus definition of what brokering is and what specific capabilities it offers now and what it should do in the future. As described in the previous section, brokering philosophy is about accepting disciplinary “diversity” and providing the necessary components to mediate and interconnect.

One or more **use cases** will be utilized to recognize options, and potential governance models. They will also be used for test and evaluation. The Brokering IG recommended three preliminary use cases during their meeting in Dublin:

- Global Changes: **GEO-BON**;
- Environmental sciences: **European Commission Danube SDI**;
- International repositories: **ICSU WDS**.

GEOSS might be easily considered as a fourth one: the GEO secretariat confirmed its support to the WG activities. The European Commission and ICSU have confirmed their intent to provide support as stakeholders including through exercising use cases.

The **main objectives** of the proposed governance action are:

(1) **Study existing middleware governance models** and best practices considering significant use cases;
(2) **Define and refine one or more multi-disciplinary governance models** for research data interoperability working with stakeholders and other RDA synergic activities;
(3) **Test a selected model** demonstrating a proof-of-concept;
(4) **Recommend a governance approach to RDA** and the Community.

Value Proposition and Deliverables

Effective brokering governance has the potential to support longer-term development under a variety of operation and funding models. 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 and the issue is still open. 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 agreed upon.

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. These practices will work to ensure future interoperability, access, and use to brokering middleware independent or in light of various developments and funding models to support long-term planning of brokered, integrated systems. These 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 Stakeholders 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 WG will consider at least three use cases (ICSU, EC Danube and Biodiveristy), but will reach out to other communities including the social sciences for use cases, with a selection to be made during the first year as part of the evolution to governance model testing.

Work Plan

The **work plan** considers the following main tasks:

TB1: Brokering process definition and definition of terms;
TB2: Scoping of a "brokering agreement" and stakeholders roles;
TB3: Input to governance framework approaches through the use cases and interface with RDA Groups and Community;
TB4: Evaluation of options for governance approaches and selection of a model;
TB5: Stakeholders apply and test a selected governance model;
TB6: Analysis of governance model;
TB7: Develop recommendations for a brokering framework governance approach;
TB8: Review recommendation with a broad stakeholder and RDA Communities; and
TB9: Report writing.

Task	Timeframe (months)	Expected outcome
TB1	M1-M3	Process definition and definition of terms
TB2	M3-M5	Guidelines for "brokering agreement" and stakeholders roles
TB3	M5-M8	Options for framework governance approaches and model selection
TB4	M8-M13	Applications and testing of the framework governance model
TB5	M13-M14	Stakeholders inputs from test and evaluation
TB6	M14-M16	Recommendations for a brokering framework governance
TB7	M16-M17	Stakeholders review
TB8	M17-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	Oceanography Data Systems; Biodiversity (OBIS, ODIP, IODE)
To be confirmed		Agriculture
To be confirmed		Health

The WG will be initially 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
- 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://semanticommunity.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

[xxv] 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.

[xxvi] 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

[xxvii] 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.