

OntologySummit2008 Communique

OntologySummit2008 Communiqué: Towards an Open Ontology Repository [\(1B3Y\)](#)

- Version: 1.0.0 was adopted and released on 29-April-2008 2:49pm EDT / Gaithersburg, Maryland, USA. [\(1GMG\)](#)
- Current Version: 1.0.3 [\(1H16\)](#)
 - this version has been edited (incorporating non-substantive changes) mainly by [BarrySmith](#) / 2008.05.04 [\(1HM2\)](#)
- Lead Editors: [LeoObrst](#) & [MarkMusen](#) [\(1GOY\)](#)
- Co-Editors: [BarrySmith](#), [FabianNeuhaus](#), [FrankOlken](#), [MichaelGruninger](#), [MichelleRaymond](#), [PatHayes](#) & [RaviSharma](#) [\(1GOZ\)](#)

1. Introduction [\(1GUD\)](#)

Each annual Ontology Summit initiative makes a statement appropriate to each Summit's theme as part of our general advocacy designed to bring ontology science and engineering into the mainstream. The theme this year is "Towards an Open Ontology Repository". This communiqué represents the joint position of those who were engaged in the year's summit discourse on an Open Ontology Repository (OOR) and of those who endorse below. In this discussion, we have agreed that an "ontology repository is a facility where ontologies and related information artifacts can be stored, retrieved and managed." [\(1GUJ\)](#)

We believe in the promise of semantic technologies based on logic, databases and the Semantic Web, a Web of exposed data and of interpretations of that data (i.e., of semantics), using common standards. Such technologies enable distinguishable, computable, reusable, and sharable meaning of Web and other artifacts, including data, documents, and services. We also believe that making that vision a reality requires additional supporting resources and these resources should be open, extensible, and provide common services over the ontologies. [\(1GUK\)](#)

A number of controlled vocabularies and ontologies have been encoded in RDF, OWL, and other knowledge representation languages, but only a fraction of these have fostered significant reuse. While there are many issues that can limit the potential for reuse, a significant contributing factor is the lack of well-specified policies for vocabulary management, metadata, and provenance specification. Several of the most prominent RDF vocabularies currently in use have emerged from a close collaboration between a relatively small community of developers and a larger community of users. The prominence of these vocabularies may be attributed to their utility, but also to the commitment, made by those responsible for developing and maintaining the vocabularies, to accommodating, serving, and working with a community of users. ([ref.](#)) In addition to a lack of policies and metadata, the lack of open and available infrastructure and services to support reuse is an impediment to adoption of these semantic technologies. [\(1GUL\)](#)

The purpose of an Open Ontology Repository is to provide an architecture and an infrastructure that supports a) the creation, sharing, searching, and management of ontologies, and b) linkage to database

and XML Schema structured data and documents. Complementary goals include fostering the ontology community, the identification and promotion of best practices, and the provision of services relevant to ontologies and instance stores. Examples of anticipated services include automated semantic interpretation of content expressed in knowledge representation languages, the creation and maintenance of mappings among disparate ontologies and content, and inference over this content. We believe that the Open Ontology Repository will ultimately support a broad range of semantic services and applications of interest to enterprises and communities. [\(1GUM\)](#)

Achieving these goals will help reduce semantic ambiguity whenever and wherever information is shared, thereby allowing information to be located, searched, categorized, and exchanged with a more precise expression of its content and meaning. The artifacts of the repository will provide a semantic grounding for diverse formats and domains, ranging from the conceptual domains and specific disciplines of communities to technical schema such as WSDL, UDDI, RSS, and XML schema, and of course expressed in standard ontology languages such as RDF, OWL, Common Logic, and others. Perhaps most importantly, the repository will enable wide-scale knowledge re-use and reduce the need to re-invent the wheel when defining concepts and relationships that are already understood. [\(1GUN\)](#)

These goals cannot be achieved all at once, and must track the evolution of best practices as well as technology itself. It is also good system development practice to bound complexity by defining a system in terms of a series of short-term, achievable objectives. For this reason, as for other such initiatives, it is envisioned that the Open Ontology Repository will be developed in a series of phases, proceeding from the simple to the complex, with achievable goals that capitalize on previous experience and the emergence of technology over time. It is important to note that in any given phase, planning and prototyping for subsequent phases is always in progress. [\(1GUO\)](#)

2. Requirements for an Open Ontology Repository [\(1GUP\)](#)

The Ontolog community in the past year determined that the primary technical areas that needed to be discussed and illuminated to make the vision of an Open Ontology Repository a reality were the following: 1) determining the current state of the art in ontology repositories, 2) determining quality and gatekeeping criteria for registering and then provisioning ontologies and their instances, 3) developing an ontology of ontologies that would act as structure and metadata for registering ontologies and supporting the common repository of their instances, data, and services, and 4) developing a sound architecture for the envisioned Open Ontology Repository. Elaborations of these four technical areas together help to provide both specifications of requirements and the ideas and tools that could help to realize them. The remainder of this communiqué thus summarizes the results of the discussions in these four areas. [\(1GUQ\)](#)

3. State of the Art [\(1GUR\)](#)

The purpose of this section is to set out the major design decisions and the technology choices which are important to the creation of ontology repositories. [\(1GUS\)](#)

Ontology repositories support the storage, search, retrieval and interoperation of multiple ontologies. [\(1GUT\)](#)

Ontology repositories support macro-level storage, query and retrieval (across the collection of ontologies) and micro-level operations (within individual ontologies). At each level we would like to support both text search, and semantic search (variously faceted search, SPARQL, ontology and ontology language literate search). Some ontology repositories have used the same technologies for both macro-level and micro-level operations. [\(1GUU\)](#)

A key decision is the choice of a representation of the ontologies. Current practice includes: text, frames (e.g., OBO), graphs (e.g., RDF), and various types of logic, e.g., description logics (e.g., OWL-DL), first order logic (e.g., Common Logic), sorted logics, possibly higher order logic (HOL). Other possibilities include the use of UML (e.g., in the OMG Ontology Definition Metamodel). [\(1GUV\)](#)

Ontologies have been stored in long narrow relations, e.g., "triple stores" of RDF triples (subject, relationship, object), relational databases, customized data stores. Increasingly implementers are using "quad stores" in order to support Named Graphs. "Column stores" such as MonetDB and Vertica have also been used to store ontologies. [\(1GUW\)](#)

For the purposes of ontology interoperability it helps to have all of the ontologies in the repository encoded in a common representation. However, this requires the sometimes difficult and lossy translation of ontologies among various representations into the common representation. Some ontology repositories store ontologies in their native representation, with metadata to identify the representation language. [\(1GUX\)](#)

We also need some way to support ontology interoperability by specifying the mappings among entities, e.g., via relationships such as `same_as`, `is_a`, and `part_of`. Other mapping relationships include: `see_also`, `similar_to`. Some ontology mapping consistency checking tools check that mappings between partially ordered ontologies, e.g., taxonomies, preserve the partial orders. [\(1GUY\)](#)

Many ontology repositories which support partially ordered ontologies (taxonomies and partonomies) may decide to materialize the transitive closure of the partial order relation. This provides faster query evaluation at the expense of additional ingestion costs, storage, and maintenance. [\(1GUZ\)](#)

Provenance of definitions in ontologies is important to the credibility, scientific attribution, and regulatory compliance of ontologies. In particular, many definitions are embodied in legislation, administrative regulations, court decisions, professional society standards. [\(1GV0\)](#)

Provenance and other metadata are distinguishing features of recent ontology repositories. Such metadata ranges from authorship, and creation date, version information, to evaluation and usage reports. Other metadata may include intended use (context). [\(1GV1\)](#)

Modularization support is useful for large ontologies, and for facilitating the reuse and mapping of portions of ontologies. [\(1GV2\)](#)

In a distributed setting, ontology repository developers increasingly are adopting Service Oriented Architectures (SOA), providing access, search, and other capabilities via web services. Two major approaches to SOA are REST and SOAP. REST is built on HTTP, with a small set of operators (GET, PUT, POST, DELETE) and the use of URL (or URI) addresses for all objects of interest. SOAP is based on XML RPCs. REST is much simpler to implement and should be adequate for typical ontology repository functions. SOAP is supported by a wide variety of software tools. Both SOA approaches are currently being used. [\(1GV3\)](#)

Finally, an ontology repository typically facilitates access to a variety of ontology related tools: creation, editors, pretty printers, visualization tools, differencing tools, modularization tools, import / export, version management, access control, inference engines, explanation, summarization. [\(1GV4\)](#)

4. Quality and Gatekeeping [\(1GV5\)](#)

We distinguish between gatekeeping and quality control. Gatekeeping criteria are a set of minimal requirements that any ontology within the OOR has to meet. The latter are intended to enable the users of the OOR to find quickly ontologies that fit their needs; the criteria are not supposed to ensure the quality of the ontologies. [\(1GV6\)](#)

4.1 Gatekeeping Criteria [\(1GV7\)](#)

The ontologies in the OOR have to meet the following criteria: [\(1GV8\)](#)

1. The ontology is submitted in a publicly described language and format. [\(1GV9\)](#)
2. The ontology is read accessible. [\(1GVA\)](#)
3. The ontology is expressed in a formal language with a well-defined syntax. [\(1GVB\)](#)
4. The authors of the ontology provide the required metadata as specified under section 5. [\(1GVC\)](#)
5. The ontology has a clearly specified and clearly delineated scope. [\(1GVD\)](#)
6. Successive versions of the ontology are clearly identified. [\(1GVE\)](#)
7. The ontology is appropriately named. [\(1GVF\)](#)

It is particularly important that the required metadata include information about the process that is employed to create and maintain the ontology. (Is the ontology maintained in a cooperative and transparent process? Can anybody participate in this process?) Further, the metadata has to include information about the license under which the ontology is submitted. [\(1GVG\)](#)

4.2 Quality Control [\(1GVH\)](#)

The community agrees that it is not sufficient for the OOR just to store ontologies, but that it needs to enable the evaluation of the ontologies within it. The OOR will offer functionalities like those on social networking sites which would allow users to comment on ontologies and rank them. Further, the OOR will enable selective views of the repository using tags provided by subcommunities that characterize ontologies with respect to their chosen criteria. For example, such a view might select for ontologies for specific fields of research or industries, or for ontologies satisfying specific quality criteria or levels of organizational approval. [\(1GVI\)](#)

5. Metadata for Ontologies [\(1GVJ\)](#)

5.1 Purpose of the Ontology Metadata [\(1GVK\)](#)

The community agrees that it is not sufficient for the OOR just to store ontologies but that metadata for ontologies are necessary to support the sharing and reuse of ontologies within the repository. [\(1GVL\)](#)

The metadata should allow users to: [\(1GVM\)](#)

- determine whether an ontology is suitable for a user purpose; [\(1GVN\)](#)
- capture the design rationales that underlie the ontology; [\(1GVO\)](#)
- find information about author, author credentials, and source of ontology reference material [\(1GVP\)](#)
- retrieve ontologies for use in domain applications; [\(1GVO\)](#)
- retrieve ontologies to be integrated with other ontologies; [\(1GVR\)](#)
- retrieve ontologies that will be extended to create new ontologies; [\(1GVS\)](#)
- determine whether or not an ontology can be integrated with given ontologies; [\(1GVT\)](#)
- determine whether a set of ontologies retrieved from the repository can be used together; [\(1GVU\)](#)
- determine whether an ontology in the repository can be partially shared. [\(1GVV\)](#)

The discussions surrounding the [OntologySummit2007](#) provide a basis for understanding the metadata for ontologies. [\(1GVW\)](#)

There should be policies for creation and modification of metadata and documentation of ontologies and the management of the persistence and sustainability of ontologies. [\(1GVX\)](#)

Users (including end-users, ontology and repository developers, subject matter experts, stakeholders) should participate in the collaborative ontology development life cycle and in decisions regarding what metadata are suitable for ontologies in the repository. [\(1GVY\)](#)

We can consider logical metadata (logical properties of the ontology independent of any implementation or engineering artifact) and engineering metadata (properties of the ontology considered as an engineering artifact). [\(1GVZ\)](#)

5.2 Logical Metadata [\(1GW0\)](#)

5.2.1 Language [\(1GW1\)](#)

The first logical property is to identify the language used to specify the ontology. [\(1GW2\)](#)

The report "[Evaluating Reasoning Systems](#)" contains a classification of formal languages used to specify ontologies. A formal language has a syntax (logical symbols together with a formally specified grammar) and a model theory (which specifies the conditions under which expressions in the language can be given particular truth assignments). [\(1GW3\)](#)

A formalizable language has a syntax, although it does not have a model theory. Examples of such approaches include Topic Maps and folksonomies (which are written in XML) and ISO 15926 (which is written in EXPRESS). [\(1GW4\)](#)

Finally, some ontologies are only specified in natural language, including Wordnet, taxonomies, and thesauri. [\(1GW5\)](#)

5.2.2 Modularity [\(1GW6\)](#)

A second property of ontologies is based on modularity -- is a particular ontology a monolithic set of axioms, or is it composed of a set of smaller modules? Furthermore, is each module considered to be a separate ontology within the repository? If not, what are the relationships between the modules and which modules of an ontology can be used separately? [\(1GW7\)](#)

For example, the Process Specification Language ([PSL](#)) consists of a set of modules which are extensions of a common core theory PSL-Core. Metadata for each module specifies which other modules must also be included when using the module. [\(1GW8\)](#)

5.2.3 Relationships between ontologies [\(1GW9\)](#)

We can also specify various logical relationships between ontologies within the repository, including mutual consistency, extension, and entailment, and semantic mappings. [\(1GWA\)](#)

5.3 Engineering Metadata [\(1GWB\)](#)

In addition to the logical metadata for ontologies, we need to specify metadata for ontologies as considered as engineering artefacts. This includes [\(1GWC\)](#)

- provenance [\(1GWD\)](#)
- versioning [\(1GWE\)](#)
- existing applications of the ontology (e.g. interoperability, search, decision support) [\(1GWF\)](#)
- domain-specificity (e.g. biology, supply chain management, manufacturing) [\(1GWG\)](#)

5.4 Conclusions regarding Metadata [\(1GWH\)](#)

The Ontology Metadata Vocabulary ([OMV](#)), Dublin Core, ISO 11179, ISO 19763, and other existing approaches to provenance and versioning metadata are all candidates for aspects of the metadata for ontologies in the OOR. [\(1GWI\)](#)

We strongly urge an empirical approach to the identification and evaluation of ontology metadata. We

should begin by collecting ontologies from Summit participants, and test out the different proposals for metadata on these ontologies. We should also develop use-case scenarios that will motivate the use of the metadata with these ontologies and help establish best practices. We especially challenge the participants in the [UpperOntologySummit](#) to create a prototype of the OOR that includes the upper ontologies. [\(1GWJ\)](#)

6 Repository Architecture [\(1GWK\)](#)

The Architecture of a repository for enabling wide-scale searching and sharing of ontologies, must be open and extensible. The Architecture design should be modular in nature and provide for ontology storing, sharing, searching, governance, and management of the repository infrastructure and content. [\(1GWL\)](#)

6.1 Architecture Approach [\(1GWM\)](#)

The core approach for the Open Ontology Repository is a federated, service oriented architecture. This approach provides for distributed ontology storage, repository management and service support. [\(1GWN\)](#)

The overall assessment of the community is to enable open, distributed, federated repositories, and to provide metadata for each ontology registered, as well as providing connections for logical services, inference engines etc. [\(1GWO\)](#)

Those who engage in the federation must include required metadata. This metadata must include any access constraints. [\(1GWP\)](#)

Over the repository will be an ontology that is inclusive of both the metadata of ontologies and the information we need for operational use. [\(1GWQ\)](#)

6.2 Core Requirements [\(1GWR\)](#)

The requirements presented are important to the enablement of wide-scale knowledge re-use. [\(1GWS\)](#)

1. The repository architecture shall be scalable. [\(1GWT\)](#)
2. The architecture shall be optimized for sharing, collaboration and reuse. [\(1GWU\)](#)
3. The repository shall be capable of supporting ontologies in multiple formats and levels of formalism. [\(1GWV\)](#)
4. The repository architecture shall support distributed repositories. [\(1GWW\)](#)
5. The repository architecture shall support explicit machine usable/accessible formal semantics for the meta-model of the repository. [\(1GWX\)](#)
6. The repository shall provide a mechanism to address intellectual property and related legal issues/problems. [\(1GWY\)](#)
7. The repository architecture shall include a core set of services, such as support for adding, searching and mapping across ontologies and data related to the stored ontologies. [\(1GWZ\)](#)
8. The repository architecture shall support additional services both directly within the province of the repository and as external services. [\(1GX0\)](#)
9. The repository should support all phases of the ontology lifecycle. [\(1GX1\)](#)

6.3 Repository Management [\(1GX2\)](#)

An ontology repository requires mechanisms for effective management. The understanding is that as a repository and its infrastructure evolve, more management support mechanisms will be included. [\(1GX3\)](#)

Required mechanisms will provide the capabilities to: [\(1GX4\)](#)

1. enforce access policies [\(1GX5\)](#)
2. enforce submission policies [\(1GX6\)](#)
3. enforce governance policies [\(1GX7\)](#)
4. enforce change management policies [\(1GX8\)](#)
5. control user and administrator access [\(1GX9\)](#)

Highly recommended mechanisms will provide the capabilities to: [\(1GXA\)](#)

1. create usage reports [\(1GXB\)](#)
2. validate syntax [\(1GXC\)](#)
3. check logical consistency [\(1GXD\)](#)
4. automatically categorize a submission [\(1GXE\)](#)

6.4 Service and Application Support [\(1GXF\)](#)

OOR interfaces should support internal and external services and applications including: [\(1GXG\)](#)

- Ontology creation tools [\(1GXH\)](#)
- Ontology editors [\(1GXI\)](#)
- Ontology differencing tools [\(1GXJ\)](#)
- Ontology modularization tools (clustering, etc.) [\(1GXK\)](#)
- Ontology export [\(1GXL\)](#)
- Ontology visualization (e.g., graph visualization) [\(1GXM\)](#)
- Version management [\(1GXN\)](#)
- Access control [\(1GXO\)](#)

6.5 Discovery Support [\(1GXP\)](#)

To facilitate knowledge discovery the repository shall provide metadata capabilities to support search capabilities, governance process, and management. [\(1GXQ\)](#)

The repository should support discovery by, for example: [\(1GXR\)](#)

- domain [\(1GXS\)](#)
- author/creator/source [\(1GXT\)](#)
- version [\(1GXU\)](#)
- language [\(1GXV\)](#)
- terminology and controlled vocabularies [\(1GXW\)](#)
- quality [\(1GXX\)](#)
- mapping [\(1GXY\)](#)
- inference [\(1GXZ\)](#)

7. Conclusion: Toward the Future [\(1GY0\)](#)

We look forward to establishing an open ontology repository in the future that adheres to the requirements set forth above. We endorse an open ontology repository that seeks to honor and implement the following overarching mission requirements: [\(1GY1\)](#)

1. Supporting the Open Ontology Repository (OOR) Initiative that will promote the global use of ontologies, their instance bases, rules, and services, and mappings among these. [\(1GY2\)](#)
2. Enabling and facilitating open, federated, collaborative ontology repositories. [\(1GY3\)](#)
3. Establishing best practices for expressing interoperable ontology work in open

registries/repositories. [\(1GY4\)](#)

4. Enabling and facilitating the development of common services to support the repository and to extend the capabilities available to providers, users, and developers who use the repository. [\(1GY5\)](#)

We believe that creating this kind of infrastructure will facilitate the emerging Semantic Web. [\(1GY6\)](#)

This Communiqué was reviewed, collaboratively edited, finalized and adopted by individuals present at the Ontology Summit 2008. [\(1GY7\)](#)

Endorsed by: [\(1GY8\)](#)

The above Communiqué has been endorsed by the individuals listed below. Please note that these people made their endorsements as individuals and not as representatives of the organizations they are affiliated with. [\(1GY9\)](#)

The following individuals were at the 2008.04.28/29 workshop and contributed to the fine tuning and finalization of the Communiqué: [\(1GZV\)](#)

- [AlexGarcia \(1GYA\)](#)
- [AmandaVizedom \(1GYB\)](#)
- Amy Davidson (BBN) [\(1GYC\)](#)
- [BarrySmith \(1GYD\)](#)
- [BenjaminDai \(1H13\)](#)
- [BruceBargmeyer \(1GYE\)](#)
- [ConradBock \(1GYF\)](#)
- [DeborahMacPherson \(1GYG\)](#)
- [DeborahMcGuinness \(1GYH\)](#)
- Doug Clark (Gard Associates) [\(1GYI\)](#)
- [DougHolmes \(1GYJ\)](#)
- [ElisaKendall \(1GYK\)](#)
- [EvanWallace \(1GYL\)](#)
- [FabianNeuhaus \(1GYM\)](#)
- Faheem Aziz (Northrop Grumman) [\(1GYN\)](#)
- [FrankOlken \(1GYO\)](#)
- Gail Hodge (Information International Associates) [\(1GYP\)](#)
- [GaryBergCross \(1GYQ\)](#)
- [IndraNeilSarkar \(1GYR\)](#)
- Joanne S Luciano (MITRE) [\(1GYS\)](#)
- [JohnSowa \(1GYT\)](#)
- [K_Goodier \(1GYU\)](#)
- [KennethBaclawski \(1GYV\)](#)
- [LeoObrst \(1GYW\)](#)
- [LiDing \(1GYX\)](#)
- Line Pouchard (Oak Ridge National Lab) [\(1GYY\)](#)
- Luis Bermudez (SURA) [\(1GYZ\)](#)
- [MalaMehrotra \(1GZ0\)](#)
- [MarkMusen \(1GZ1\)](#)
- Michael Pendleton (EPA) [\(1GZ2\)](#)

- [MichaelGruninger \(1GZ3\)](#)
- [MichelleRaymond \(1GZ4\)](#)
- [MikeDean \(1GZ5\)](#)
- [NancyLawler \(1GZ6\)](#)
- [NatashaNoy \(1GZX\)](#)
- [PatrickCassidy \(1GZ7\)](#)
- [PeterBenson \(1H00\)](#)
- [PeterYim \(1GZ8\)](#)
- [RamSriram \(1GZ9\)](#)
- [RaviSharma \(1GZA\)](#)
- [SteveRay \(1GZB\)](#)
- [SusanTurnbull \(1GZC\)](#)
- [ThomasBrunner \(1GZK\)](#)
- [ToddSchneider \(1GZD\)](#)
- Xiang Li (NIST) [\(1GZE\)](#)

Individuals who emailed in their endorsement: [\(1GZW\)](#)

- [RickMurphy \(1GZI\)](#)
- [RexBrooks \(1GZJ\)](#)
- [FrankAlvidrez \(1GZL\)](#)
- [KathyLesh \(1GZM\)](#)
- [BonnieSwart \(1GZY\)](#)
- Xuan Zha (Extension Systems International) [\(1GZZ\)](#)
- Richard Lee (Booz Allen Hamilton) [\(1H01\)](#)
- [MatthewWest \(1H02\)](#)
- [SeanBoisen \(1H03\)](#)
- [RonWheeler \(1H04\)](#)
- [EdDodds \(1H05\)](#)
- [AntoinetteArsic \(1H12\)](#)
- [CarlMattocks \(1H14\)](#)
- [OthelRolle \(1H1A\)](#)
- [AnnWrightson \(1H6N\)](#)
- [JohnBateman \(1HSA\)](#)
- [DilvanMoreira \(1IWS\)](#)
- [MarciaZeng \(1K1U\)](#)
- [CameronRoss \(2HMU\)](#)