Semantic Infrastructure Concept of Operations Initiative 1 Repositories

Contents of this Page

Semantic Infrastructure Concept of Operations Links

- Semantic Infrastructure
- Vision
- Background
- Mission
- Objectives
- Overview
- Initiatives
- Stakeholders
- Alternatives

The metadata infrastructure v2 effort will be incremental and iterative in nature, and will ensure seamless transition for all caDSR users. The current metadata infrastructure (caDSR) will remain in production through the development of the new infrastructure and for as long as necessary for complete transition of all caDSR users to the new infrastructure. NCI CBIIT will ensure migration of all existing caDSR data to the new metadata infrastructure. Prototype semantic tools will be developed and reviewed by stakeholders, and feedback on how well the new tools accommodate stakeholders' requirements including their current workflow, will drive iterative development of the tools. Metadata users are encouraged to continue using caDSR for their needs until the new metadata infrastructure becomes available for production use. (from CBIIT Management)
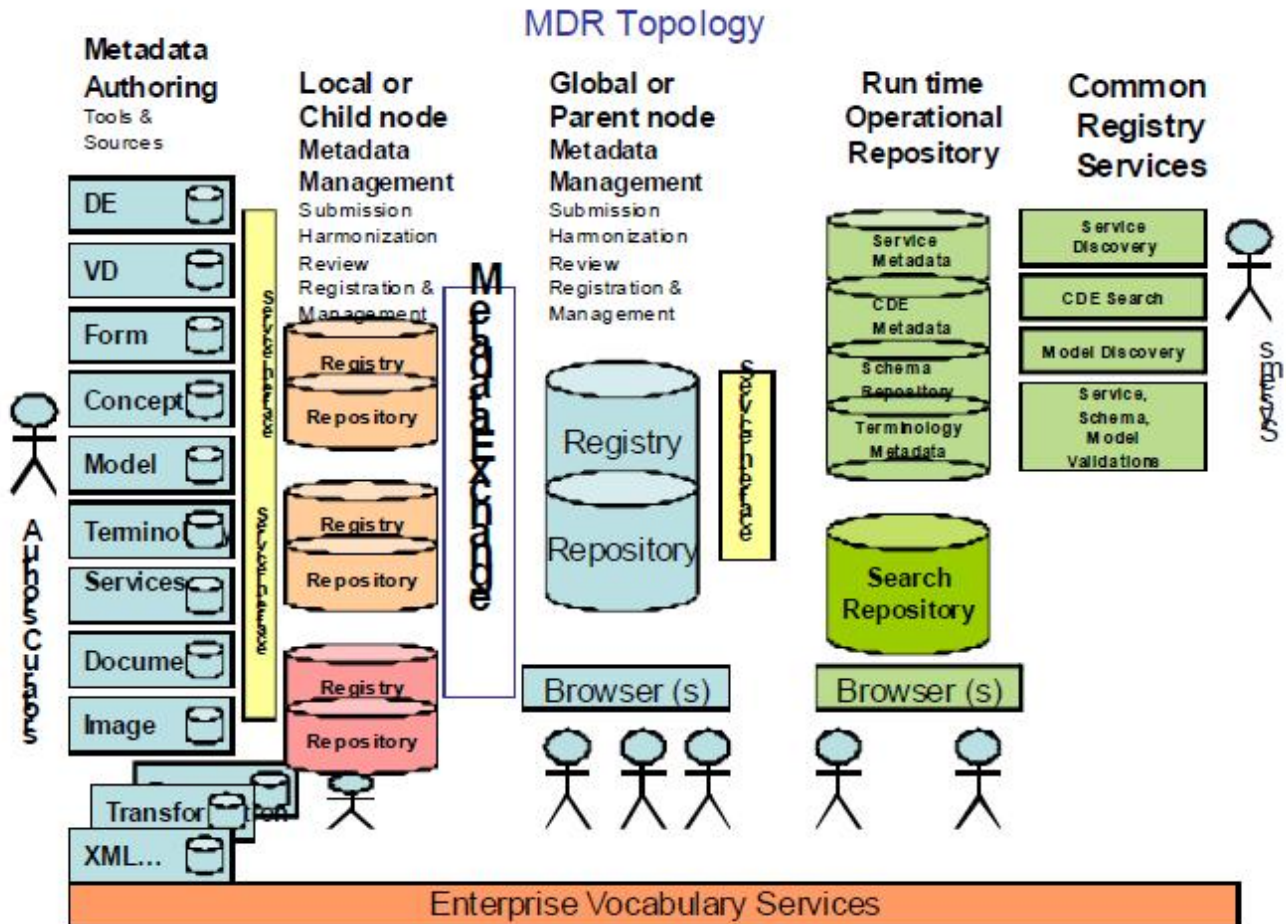
Distributed, federated metadata repositories and model repositories and operations

The architecture for the distributed metadata repository (MDR) will be decentralized in nature, allowing multiple peer repositories to be present at the same time, for sharing of data elements.

Therefore this initiative will include:

- Definition of a platform independent model (PIM) for the metadata repository and
- A production realization of the PIM of the distributed metadata repository including representation of complex data types. The following figure shows a possible topology. This topology does not dictate how each of the metadata repositories are represented, nor which tools are used to author the metadata. Some maybe be IMAGE repositories, some may be OWL, some may be XML Schema, etc.
- Development of a set of user applications or services as needed for creation, management, search and retrieval and of metadata.

Possible Topology for Metadata Repository

The infrastructure will support federated queries to allow users to find potential semantic metadata anywhere in the community, inside or outside caGrid, facilitating sharing across government, industry and national business.

Any one MDR by itself will not fully support the semantics requirements. Building a conceptual view of shared semantics first by the BRIDG community and also by the Life Sciences Domain (LS-DAM) community has been hampered by lack of a suitable repository to point caBIG® participants to. Also, the shared semantic view that should be conveyed is further hampered by lack of tools and knowledge to manipulate such models. Therefore, in addition to the MDR, this initiative will include:

- Exploration of options for model repositories and tooling including storage and sharing of DAMs and of 'lower-level" models derived from DAMs.
- Design of an implementation of model authoring capabilities and a model repository suitable for use with derivative (top down) and compositional (bottom up) models. (Possibly our experience with GForge, Wikis, and UML modeling tools can inform development.)
- Design and implementation of transformation capabilities to allow model repositories to be decomposed into data descriptions that can be reused through the existing infrastructure to support deployment of these semantics in practical end user solutions via software engineering techniques such as forms development and forms generation.
- May also encompass evaluation of semantic profile authoring and services.

Requirements

Initiative 1 section of the Requirements and Initiatives Master List

Forum

Initiative 1 - Distributed, federated metadata repositories and model repositories and operations