Open Ontology Repository

Ken Baclawski
College of Computer and Information Science
Northeastern University

OOR Charter

- Promote the global use and sharing of ontologies by:
  - establishing a hosted registry-repository
  - enabling and facilitating open, federated, collaborative ontology repositories
  - establishing best practices for expressing interoperable ontologies and taxonomy work in registry-repositories.

http://openontologyrepository.org

Why?

- Isn’t the Semantic Web notion of distributed islands of semantics sufficient as a de facto repository?
- If you put it out there, will they come?
- If you build it better and put it out there, will they prefer yours?
- History does not show this laissez faire “field of dreams” is good reality.

Okay, but why a new infrastructure?

The Clickable Web

- The “clickable” web has been very successful in employing a “lazy strategy” for HTML documents.
- However the use and content of the Semantic Web has different characteristics that make it far less tolerant of:
  - change
  - frequent errors which are commonplace on the clickable web.

Okay, but why a new infrastructure?

Distributed Data Sets

- This is also true for increasingly large and complex data sets.
- Generally impossible to judge integrity of data without access to metadata.
  - Increasingly important over time.
- Metadata standards are important.

SemWeb Distinguishing Characteristics

- Machines rather than humans are the primary consumers of content.
- Errors that a human may be able to diagnose and fix (such as a change in location of a document) are likely fatal for machine processing.
- The use of owl:imports creates a strong transitive dependency between ontology documents.

SemWeb Distinguishing Characteristics

- Changes in any imported document (imported directly or through nested import) can cause the resulting import closure to be:
  - inconsistent
  - change its meaning
  - change computational characteristics.
- Ontologies convey a precise meaning with an unambiguous machine interpretation.
- When using this content, careful selection and precise reference is critical.

Open Ontology Repository

- *"An ontology repository is a facility where ontologies and related information artifacts can be stored, retrieved and managed"*
- The registry
- The persistent store
- Value-added services:
  - Ontology sharing, searching, management, etc.
  - Linkage to databases, XML Schemas, documents, etc.
  - Reasoning capabilities.

Open Ontology Repository

All types of artifacts on the ontology spectrum:

- Folksonomies
- Terminologies
- Controlled vocabularies
- Taxonomies
- Thesauri
- Data schemata
- Data models
- OWL ontologies
- Logical theories
- Shared understanding
- Ontological commitments

Open Ontology Repository

Open Access

- compliance with open standards
- open technology (open source)
- open knowledge (open content)
- open collaboration (transparent community process)
- open to integration with “non-open” repositories via an open interface

Purpose and Scope of the OOR Initiative

- Limited to providing an infrastructure that enables ontology and metadata management.
- Formal architecture:
  - Interfaces
  - Required services
  - Enable interoperability among OOR instances.
- Reference implementation:
  - Basic services
  - Basic metadata lifecycle.

OOR Value Added I

- The OOR is reliably available.
- The OOR is persistent and sustainable, so you can be confident when committing to its use.
- The OOR has information about when, why, and how an ontology has changed, so you can be aware of changes that may effect its usability.

OOR Value Added II

- You can find ontologies and metadata easily.
- Ontologies and metadata are registered, so you know who built them.
- Metadata provides the ontology purpose, KR language, user group, content subject area, etc.

OOR Value Added III

- The OOR enables services such as:
  - Ontology search capability
  - Mappings for connecting ontologies to each other
  - Content review and certification, and gauge quality and value by some recognized criteria.
- OOR services allow users to:
  - Use services that others have developed, or
  - Plug in your own services.

OOR Value Added IV

- Ontologies can reuse or extend other ontologies, including common middle and upper ontologies.
- The OOR codebase can be extended to:
  - Domain specific architecture
  - Enterprise specific architecture.

Top Requirements

1. Scalable repository architecture.
2. Optimized for sharing, collaboration, and reuse.
3. Support for multiple formats, languages and levels of formalisms.
4. Distributed (federated) repositories.
5. Explicit machine usable formal semantics.
6. Support for IPR and related legal issues.
7. Allow for extensions and domain specific services.
8. Community involvement and engagement of the ontology lifecycle.

Challenges: Basic

- Inadequacies of the “clickable web” as a basis for the Semantic Web:
  - Need long-term maintenance rather than ontologies maintained in author’s web site.
  - Varying levels of coverage
  - Intellectual property concerns.
- Best practices:
  - Policies and procedures
  - Provenance to enable trust.

Challenges: Metadata and Interfaces

- Ontology Metadata:
  - Dimensions: Expressiveness, Structure, Granularity, Intended Use, Automated Reasoning, Prescriptive vs. Descriptive, Governance
  - Ontology Metadata Vocabulary (mv.ontoware.org)
  - eXtended Metadata Registry (xmdr.org).
- Interface ontologies:
  - Internal APIs for core modules and plug-ins
  - External APIs, especially web services
  - Federation APIs, among OORs.

Challenges: Quality and Gatekeeping

- Quality is a major concern
  - Emphasized at the Ontology Summit 2008
  - Ontologies are similar to standards and publications
  - A mechanism for review is necessary
- Gatekeeping
  - Ensures minimum level of quality
  - Allows for many policies and workflows

Challenges: Management

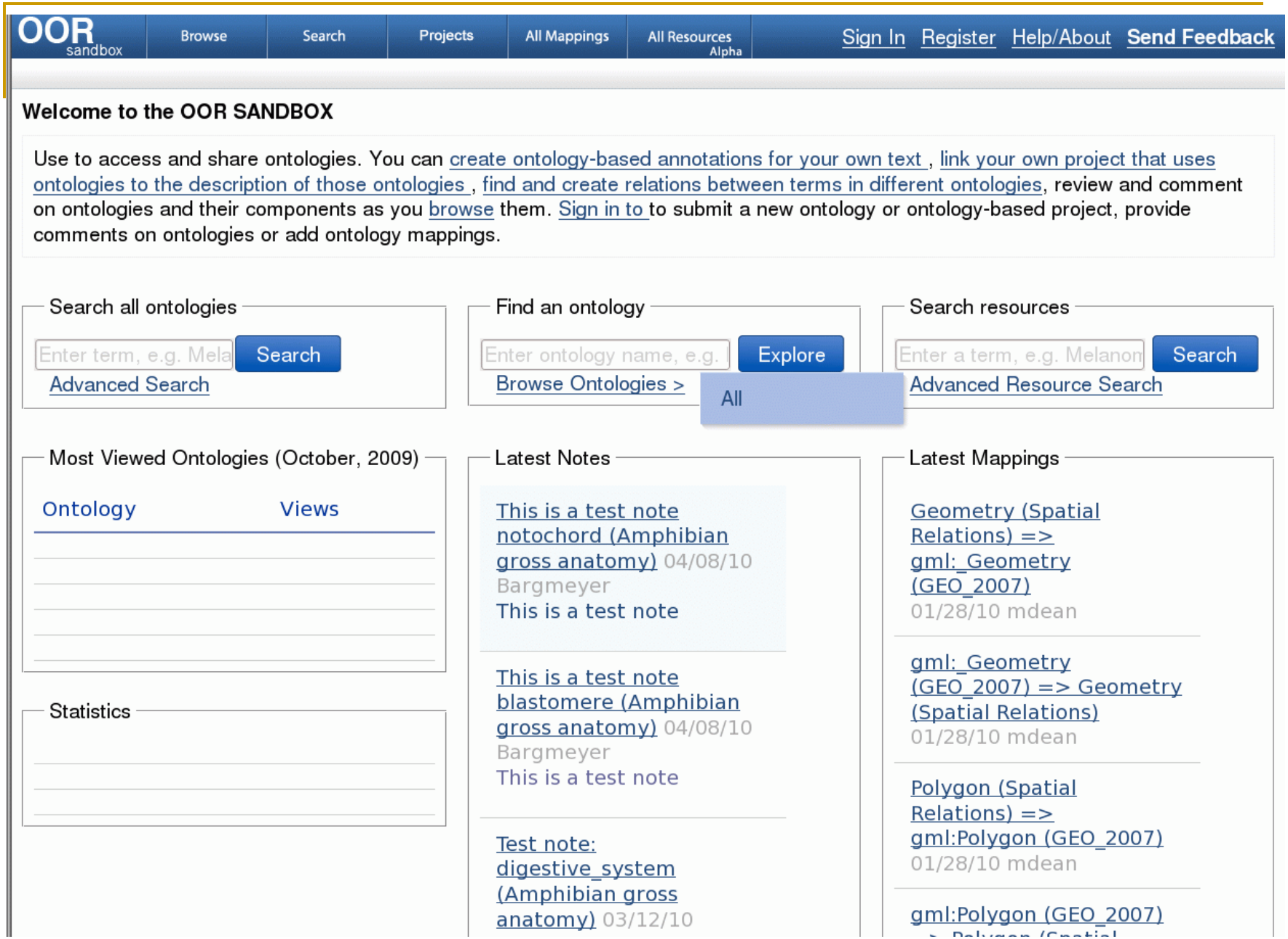
- Version management
- Configuration management
- Provenance and other metadata
- Metadata validation/authentication
  - Related to policy issues
- Ontology ownership and access control
- Federation management

Challenges: Administration and Policy

- Gatekeeping policies and enforcement
  - Should there be overall policies or left to individual repositories?
- Policy-making bodies
  - How are they constituted?
  - How can they interoperate?

OOR Team Efforts

- OOR meetings every other Friday at 12:00 Noon US Eastern Time
- Series of Joint Ontolog-OOR conferences
- ISWC 2009 Paper and Poster.
- Home page: openontologyrepository.org
- OOR Sandbox based on BioPortal
  - http://oor-01.cim3.net


OOR sandbox
Browse
Search
Projects
All Mappings
All Resources Alpha
Sign In
Register
Help/About
Send Feedback
Welcome to the OOR SANDBOX
Use to access and share ontologies. You can create ontology-based annotations for your own text , link your own project that uses ontologies to the description of those ontologies , find and create relations between terms in different ontologies, review and comment on ontologies and their components as you browse them. Sign in to to submit a new ontology or ontology-based project, provide comments on ontologies or add ontology mappings.
Search all ontologies
Enter term, e.g. Mela
Search
Advanced Search
Find an ontology
Enter ontology name, e.g.
Explore
Browse Ontologies >
All
Search resources
Enter a term, e.g. Melanon
Search
Advanced Resource Search
Most Viewed Ontologies (October, 2009)
Ontology
Views
Statistics
Latest Notes
This is a test note notochord (Amphibian gross anatomy) 04/08/10 Bargmeyer
This is a test note
This is a test note blastomere (Amphibian gross anatomy) 04/08/10 Bargmeyer
This is a test note
Test note: digestive_system (Amphibian gross anatomy) 03/12/10
Latest Mappings
Geometry (Spatial Relations) => gml:_Geometry (GEO_2007)
01/28/10 mdean
gml:_Geometry (GEO_2007) => Geometry (Spatial Relations)
01/28/10 mdean
Polygon (Spatial Relations) => gml:Polygon (GEO_2007)
01/28/10 mdean
gml:Polygon (GEO_2007)

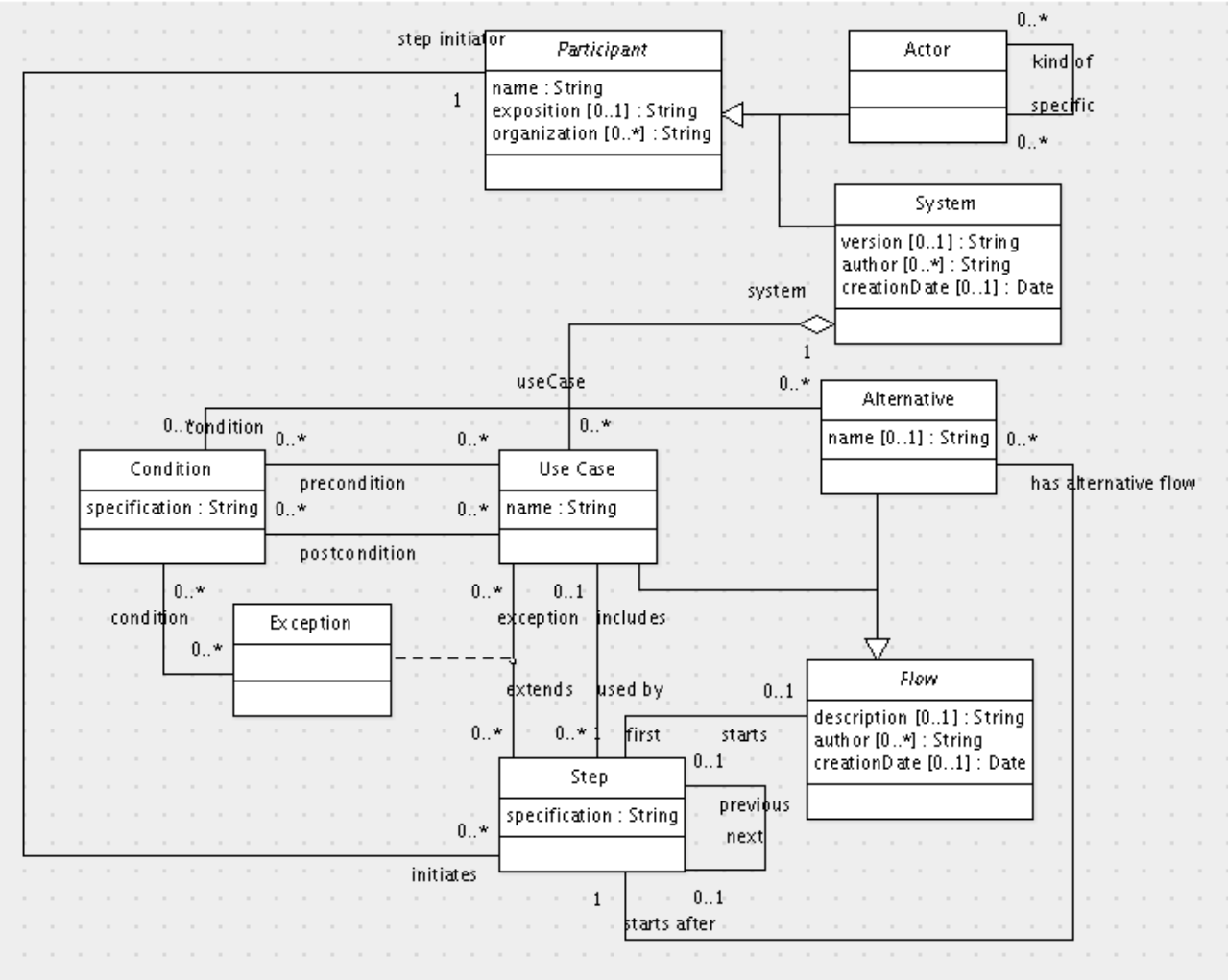
Current Activities

- OOR prototyping is now occurring
  - Several deployed instances
  - Uploaded ontologies
  - Experiments with federation
- Use case descriptions
  - OOR Requirements are on a wiki page
  - Use Case Descriptions are available
- Architecture development
  - Web service definitions

Sources of Requirements and Use Cases

- Ontology Summit 2008
  - Several sessions devoted to requirements
  - Wiki page for requirements and use cases
- ISO 11179: Metadata Repository (MDR)
- XMDR Project: Extended MDR
- BioPortal
- Marine Metadata Interoperability (MMI)

Use Case Description Ontology



Use Case Description Ontology Subscribe to updates

Submit New Version | Edit Ontology Information

Metadata | Views (0) | Projects (0) | Reviews (0) | Metrics | Ontology Widgets

ONTOLOGY NAME:	Use Case Description Ontology
ONTOLOGY ID:	1124
FORMAT:	OWL
CATEGORIES:	
CONTACT:	Ken Baclawski, Ken@Baclawski.com
HOME PAGE:	http://www.ccs.neu.edu/home/kenb
PUBLICATIONS PAGE:	
DOCUMENTATION PAGE:	
DESCRIPTION:	Use cases are a commonly used technique for specifying the functional requirements of a project. This ontology is concerned with the formal representation of the flow of events that occurs in a use case, including alternative flows and exceptional flows. Nonfunctional requirements are specified separately.

Notes

No Notes Created yet

Mappings

No Mappings Created yet

VERSION NUMBER	RELEASE DATE	ONTOLOGY FILE	VISUALIZ
0.2	05/20/2010	Download Ontology	

Powered by NCBO BioPortal | Release Notes | Terms of Use | Privacy Policy

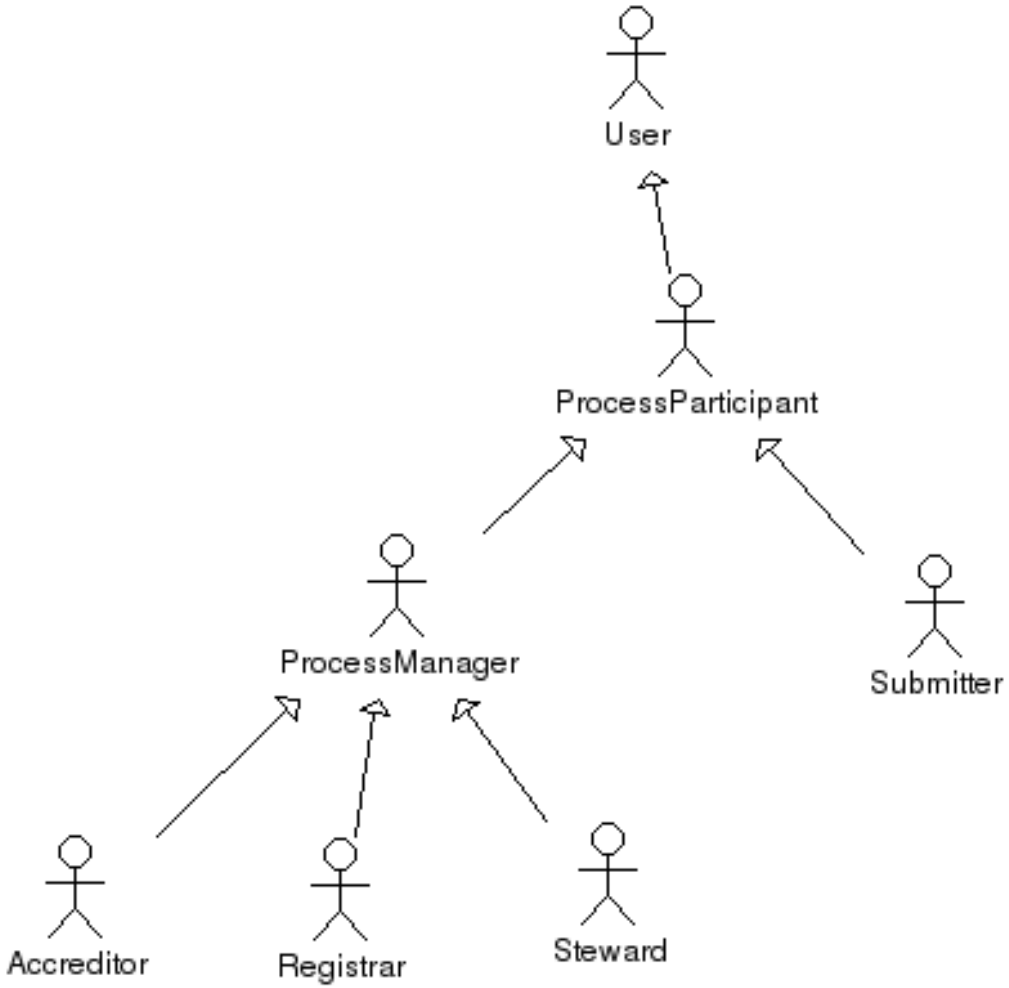
Development of OOR Use Cases

- Consulted existing examples
- Developed Use Case Description Ontology
- Developed OOR use cases as instances of UCDO
- Available at http://www.ccs.neu.edu/home/kenb/ontologies/
- Ongoing effort…

Classification of Use Cases

- Query, retrieval and navigation
- Administer authorities
- Application-oriented use cases
- Registration and validation
- Process definition and workflows
- Review and evaluate ontologies
- Potential use cases

Use Case Actor Hierarchy



System

Open Ontology Repository [OOR]

Name: Open Ontology Repository
Authors: Ken Baclawski - Máximo Gurméndez -
Description: The OOR as a participant in use cases. All communication between any user and the OOR is assumed to occur on an authenticated and secure channel. Activities such as login are therefore not included in any of the use cases. The term item is used generically for any entity that is administered by the OOR with versioning and annotations such as provenance information. This includes ontology modules, mappings, configurations, compositions, process definitions and policies. Configuration is itself a generic term that includes contexts, frameworks and situations.
Organizations: Northeastern University - OOR Initiative -
Creation Date: May 20, 2010

Questions?

Please email authors for questions.

System Actors

Accreditation Authority [Accreditor]

Description: The singular authority that manages the registration authorities of the OOR. One important responsibility of the accreditor is the allocation of top-level names. Consequently, the accreditor is analogous to a domain name registrar. The primary difference is that domain names can be bought and sold, while the names allocated by the accreditor must be permanent and in the public domain.
KindOf: #ProcessManager

Administered Items

- Ontology modules
- Ontology mappings
- Frameworks (contexts)
- Compositions
- Process and policy definitions
  - Handled with their own use cases

Use Cases

Query Administered Item [queryItem]

Description:
A user queries the system to find an administered item. Items include ontology modules, mappings and contexts. Process and policy definitions have their own use case.

Step-by-step Description:

1. [#User] - The user submits a combination subset of query component values. These can include title, release date and status. The user also specifies how to sort the results, and a desired rank range of results. For example: the user might search for "Virus Infection", released before 01/10/2009, with STANDARD status, sorted by release date, and returning the ranked list that goes from entry 20 to 30. Optionally, the specific version can also be included in the query.
2. [#OOR] - The system returns a list of administered items that match the user query and satisfy any access constraints.

Navigate Administered Item [navigateItem]

Description:
An elementary step during browsing. Unlike a query, the user has the item identifier rather than some search criteria. The result of the navigation is a set of attributes of the item that provide a minimal description of it as well as the item identifiers of related items.

Step-by-step Description:

1. [#User] - The user submits an item identifier.
2. [#OOR] - The system returns a list of attributes of the item if the user has the right to access the list of attributes of the item.

Query and Retrieval Use Cases

- Query item
  - Find an item using metadata annotations
- Navigate item
  - Retrieve description and links to other items
- Retrieve item
  - Retrieve the entire ontology module, mapping,…
- Extract from item
  - Retrieve part of an item
- Advertise
  - OOR instance query for the user interface

Administration Use Cases I

- Accredit
  - Creates a registration authority
  - Assigns a top-level name
- Register steward
  - Assigns an intermediate name
- Register submitter
  - Contributor to the community
- Retire process participant
  - Reassign responsibilities to another process manager

Submitter Registration [registerSubmitter]

Description:
A user wants to become a content submitter for a registration authority.

Step-by-step Description:

1. Includes: [#queryAuthorities]
2. [#User] - The user places the request to become a content submitter for a particular steward.
3. [#OOR] - The system notifies the steward that a request has been made for the user to become a submitter.
4. [#Steward] - The steward approves the request.
5. [#OOR] - The system notifies the user that the request has been accepted.

Retire Process Participant [retireProcessParticipant]

Description:
A process participant desires to retire from their system responsibilities.

Step-by-step Description:

1. [#ProcessParticipant] - The process participant submits a request to retire from the system.
2. [#OOR] - The system responds with a confirmation message.
3. Alternative Flow: [#reassignProcessParticipant]

Administrative Use Cases II

- Query authorities
  - Find registrar or steward information
- Query contact
  - Retrieve registrar or steward contact information that is publicly available
- Update contact

Application-Oriented Use Cases I

- Most use cases are focused on creating, querying and updating ontologies and mappings (CRUD).
- Application-oriented use cases support the “compilation” of ontology items for use in applications at “run-time”.
- Frameworks (also called contexts or situations) are collections of items and instance data for a specified purpose.
- Compositions are theory combinations defined by the colimit of a diagram of theory morphisms.

Application-Oriented Use Cases II

- Map items
- Compose items
  - Construct a composition
- Use item
  - Notify the OOR that an item is in use by an application
- Other creation, update and retrieval operations for compositions and frameworks are handled by the same use cases as those for ontology modules.

Registration and Validation Use Cases

- Register item
  - Used for new items
- Update item
  - Update the metadata for the item, not the item itself.
- Validate item
  - Performed asynchronously
- Version item
  - Introduce a new version of an item
  - Optionally one can map from the previous version
- Delete item
  - There is no deletion, items are deprecated instead.

Workflows

- Ontologies are the basis for communication in a community.
- Agreement on an ontology is a standardization process.
- Processes can be trivial or elaborate – It depends on the community.
- Most interactions with the OOR are workflows that are based on BPM process definitions.

Process and Policy Use Cases

- Upload process definition
  - Create a new process definition or policy
  - Default processes and policies are unconstrained
- Process instance form
  - Elementary step in a workflow
- Query procedures
  - Find and retrieve process definitions and policies

Review and Evaluation Use Cases

- Review item
- Retrieve reviews
- Query metrics
  - Item specific metrics
  - Community metrics
  - Registration authority metrics
  - OOR instance metrics

Potential Use Cases

- Federation
- Terminology
- Interpretation
- Reasoning
- …

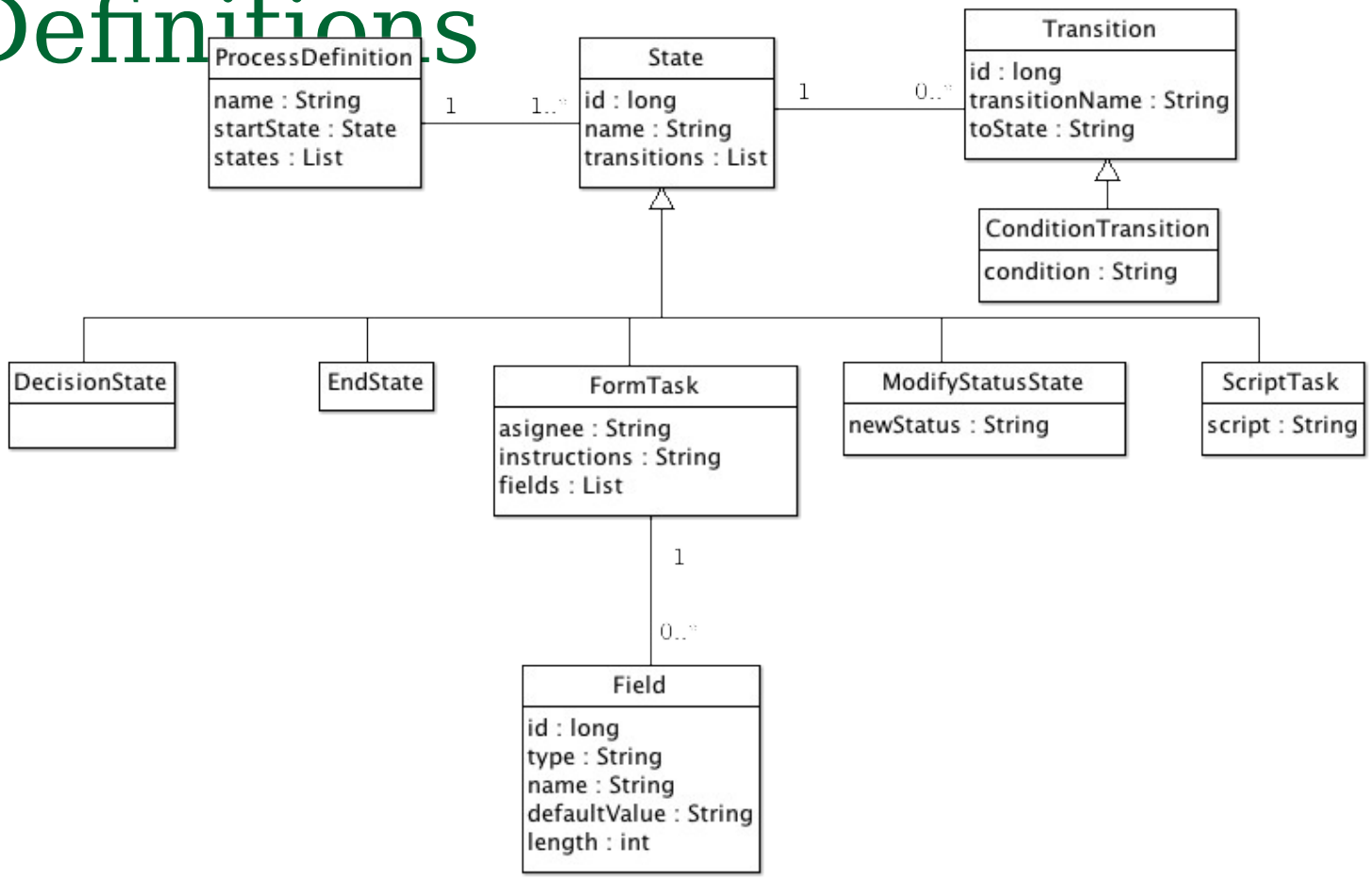
KEEPER

- Gatekeeping prototype for OOR
- Serves as the “boundary” layer of the OOR architecture
- General purpose gatekeeping web service suitable for registries, standards organizations, legislatures, publishers, etc.
- Joint effort with Máximo Gurméndez and other students at Northeastern University

KEEPER Prototype Implementation

- Experimental
- Web Services Based
- ISO 11179 Foundations
- JBPM as Workflow Engine

KEEPER Prototype: Process Definitions



KEEPER Prototype: Sample Scenario

- Registrar uploads process definition:

"Stewards validates model before it becomes a standard"

- Submitter uploads new model (Status=PENDING)
- Workflow Engine notifies Steward (according to process definition)
- Steward Validates Model (Form Task)
- Workflow Engine Modifies Status (Status=STANDARD)

Acknowledgements

- OOR Conveners
  - Mike Dean
  - Leo Obrst
  - Peter Yim
- OOR Team
  - http://ontolog.cim3.net/cgi-bin/wiki.pl?OpenOntologyRepository#nid17YW
- Máximo Gurméndez
- My students at Northeastern University