

Implementation of Prototype Biomedical Registries for PORTAL-DOORS

Carl Taswell, Global TeleGenetics, Ladera Ranch, CA

Abstract

Software implementation of the architectural design for the PORTAL-DOORS cyberinfrastructure system for resource metadata management on the semantic web has resulted in code for prototype registries in various problem-oriented domains: the GeneScene registry for genetics, ManRay for nuclear medicine, BrainWatch for brain imaging and neuropsychiatry, and BioPORT for biomedical computing. These registries facilitate translational bioinformatics by assuring globally unique identification of resources while promoting interoperability and enabling cross registry searches between different specialty fields.

Introduction

The PORTAL-DOORS System (PDS) has been designed as a resource metadata registering and publishing system that addresses the current "cybersilo" problem by facilitating a transition from original web to semantic web, and by enabling cross registry searches between different problem-oriented domains¹. Initial service interface schemas with basic ontologies have been drafted for prototype registries enabling the important example of pharmacogenomic molecular imaging² queries that demonstrate search across multiple specialty domains. However, such XML-based models represent only a piece of the puzzle. A full implementation requires many other components especially back-end databases and front-end browser clients for the PORTAL registries and DOORS directories. This presentation describes the database models now implemented for PORTAL-DOORS including both the original design and a new "bootstrapping" design.

Methods & Results

An iterative process of software development beginning from both the UML perspective and the SQL perspective resulted in code for both the original PORTAL-DOORS design as well as a new alternative design with distinct advantages. All essential design concepts initially proposed¹ have been successfully retained in the software implementations. However, one of the important design principles required resolution of a circular reference: PORTAL registries were designed to restrict registration of resource metadata at each registry to those resources meeting the criteria required for the problem-oriented domain declared for that particular registry. At the same time, DOORS directories were designed to publish the resource descriptions providing the RDF triples and

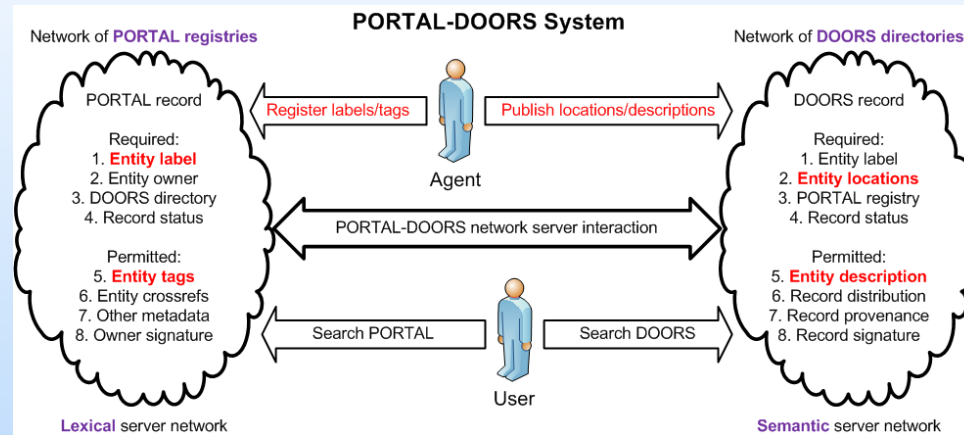


Fig 1: Resource metadata registered and published by agents for search by users in the PORTAL-DOORS server networks.

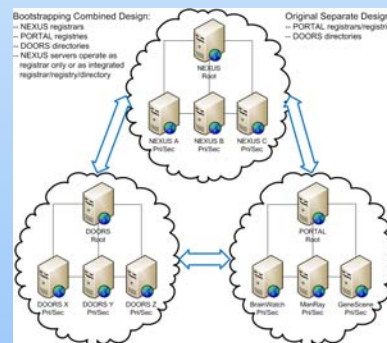


Fig 2: PORTAL-DOORS server networks.

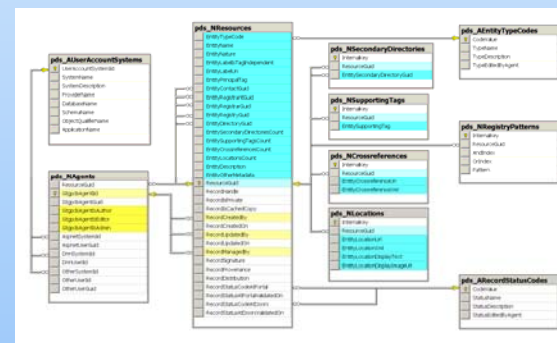


Fig 3: Relational data tables for NEXUS combined design server.

The screenshot shows the GTG Registrar web site interface. It displays a table of resources with columns for Handle, Status Type, Name, Nature, URL Label, Tag, Register, Registry, and Directory. The table lists various organizations and their associated registries and directories.

Handle	Status Type	Name	Nature	URL Label	Tag	Register	Registry	Directory
80170243	V000	Organization American Board of Genetic Counseling	certification in genetic counseling	http://www.genetics.org/abgc	ABGC	GTG Registrar	GTG Registry	GTG Directory
80882546	V000	Organization American Board of Medical Genetics	certification in medical genetics	http://www.genetics.org/abmg	ABMG	GTG Registrar	GTG Registry	GTG Directory
67872284	V000	Organization American Board of Nuclear Medicine	certification in nuclear medicine	http://www.nucmed.org/abnm	ABNM	GTG Registrar	GTG Registry	GTG Directory
47639656	V000	Organization American College of Genetic Epidemiology	physician in medical genetics professional organization	http://www.genetics.org/acge	ACGE	GTG Registrar	GTG Registry	GTG Directory
51948120	V000	Organization American Medical Informatics Association	professional organization for medical informatics	http://www.biomedicalinformatics.org/ama	AMA	GTG Registrar	GTG Registry	GTG Directory
42817636	V000	Organization American Society of Human Genetics	professional organization for human genetics	http://www.genetics.org/ashg	ASHG	GTG Registrar	GTG Registry	GTG Directory
56219171	V000	Organization American Society of Nuclear Cardiology	professional organization for nuclear cardiology	http://www.nucmed.org/asnuc	ASNuc	GTG Registrar	GTG Registry	GTG Directory
58385917	V000	Organization British Nuclear Medicine Society	professional organization for nuclear medicine in the UK	http://www.nucmed.org/bnms	BNMS	GTG Registrar	GTG Registry	GTG Directory
36189623	V000	Organization Board of Genetic Testing in Molecular and Medical Genetics	annual meeting of genetics working in cytogenetics and molecular genetics	http://www.genetics.org/bgtm	BGTm	GTG Registrar	GTG Registry	GTG Directory
58385917	V000	Organization GeneTests	National Library of Medicine for genetic disease and testing information	http://www.genetics.org/genetest	GeneTests	GTG Registrar	GTG Registry	GTG Directory

Fig 4: Screen capture of GTG Registrar web site at www.telegenetics.net for an agent logged in to edit metadata for resources of type organization.

Methods & Results (Cont'd)

thus the information necessary to determine eligibility of the resource for registration in the particular PORTAL registry.

Various solutions for implementations that resolve the circular reference problem include the following:
1) Splitting the resource description into a PORTAL required portion and a DOORS permitted portion;
2) Using record status codes "Invalid", "Pending", and "Valid" exchanged between PORTAL and DOORS;
3) Using PORTAL resource tags instead of DOORS resource descriptions to determine eligibility; and
4) Building an alternative design that combines both PORTAL and DOORS services into a single component instead of separate components.

This new scheme called the "combined" PORTAL-DOORS design can coexist together with the original scheme called the "separate" PORTAL-DOORS design. The combined design offers significant advantages in enabling an efficient self-referencing, self-describing, and bootstrapping process amongst the core system constituents (agents, registrants) and components (registrars, registries, and directories).

Open source software with code for SQL and UML models together with new XML-based schemas and ontologies will be available for download from www.portaldoors.org. Figs. 1 – 4 display design principles, data structures, and an operational web site for PDS draft version 0.5 implemented at www.telegenetics.net now available for registration of resources relevant to the problem-oriented domains of the GeneScene, ManRay, BioPort, and BrainWatch registries.

Conclusion

A new bootstrapping combined PORTAL-DOORS design, together with the original separate PORTAL-DOORS design, has been implemented for the GeneScene, ManRay, BioPORT, and BrainWatch registries. The combined design has many important advantages during early stages of PORTAL-DOORS adoption and use. However, the separate design will become useful when concerns about performance, efficiency, and scalability become more significant.

References

1. Taswell C. DOORS to the semantic web and grid with a PORTAL for biomedical computing. 2008 IEEE Trans Inf Technol Biomed 12(2):191 in Special Section on Bio-Grid (DOI 10.1109/TITB.2008.905861).
2. Taswell C. PORTAL-DOORS Infrastructure System for Translational Biomedical Informatics on the Semantic Web and Grid. 2008 AMIA STB poster 43.