



IMSC
Integrated
Media Systems
Center

INTEGRATED MEDIA SYSTEMS CENTER
A National Science Foundation
Engineering Research Center at the
UNIVERSITY OF SOUTHERN CALIFORNIA

PRINCIPAL INVESTIGATOR
Prof Dennis McLeod

OTHER RESEARCHERS, AFFILIATION

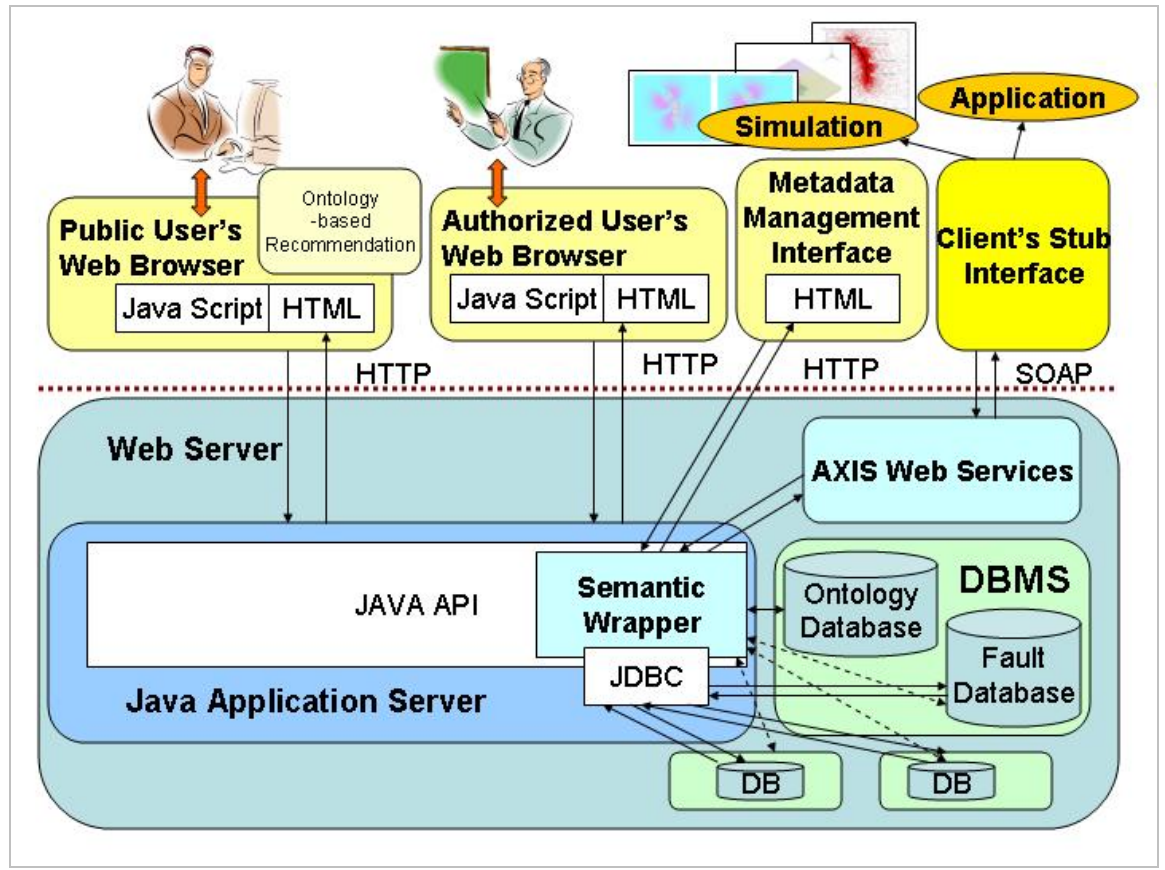
JPL; Indiana University; University of California, Irvine; University of California, Davis

USC STUDENTS (EXPECTED DEGREES):

Anne Yun-An Chen, Sangsoo Sung, Shan Gao (Ph.D.)

Data Management: Interoperability and Semantics for Heterogeneous Earthquake Science Data

In earthquake science and geoscience, scientists typically have their own interpretations and analyses of the raw data from different sources (observations, simulations, etc.). Individual databases are distributed and different data schemas are designed. In order to break the barriers and useful to access the tremendous amount of heterogeneous geoscience data, a semantic metadata management system and wrappers for web services are required to support effective information retrieval and web-based search for data of interest to specific scientists. In this work, we design and build a semantic based system to provide interoperability for heterogeneous data, different applications and database systems, and user-defined packages – for the earthquake science domain.



BRIEF DESCRIPTION OF DEMONSTRATION

The goal of this project is to support interoperability for heterogeneous geoscience data and experiments/simulations using that data. The nature of data in the earthquake science domain is full of variety and the interpretations of it differ from resource to resource, and scientist to scientist. Therefore, the semantic representation of the data and those interpretations are essential to enable access to such federated databases and the integration of those data. To this end, we employ domain ontologies and ontology-interconnection capabilities to support interoperability. Such ontologies are also used to support information discovery and recommendation for users (scientists, students, etc.).

UNIQUE OR DISTINGUISHING CHARACTERISTICS RELATIVE TO STATE-OF-THE-ART

- metadata services – dynamic and intelligence-based-updating ontologies.
- federated database system – ontology-based management of data for interoperability.
- data assimilation – assimilation of remotely sensed data from earth-based sensors, satellites, GPS, etc.
- data mining – schema integration and data interpretation based on semantics (concept, event mining).
- web services - information exchange among different platforms and applications and invocation of remote applications.

APPLICATIONS

- Automated interpretation, integration, and update of metadata of federated databases, including dynamic creation of the domain ontologies.
- Semantics-based information indexing approach for next-generation consumer information presentations in areas such as communication, education and entertainment.

RECENT HIGHLIGHTS, LEVEL OF DEVELOPMENT, UPCOMING MILESTONES

- An initial domain ontology has been developed and three databases created.
- Federated databases of earthquake science are created for the testbed. One of the databases contains processed data of California faults and the sources are journal articles in the field of paleoseismology; another contains data of California layers; the other contains the integrated data from California Geographical Survey.
- A simple basic search with wildcard search engine was developed and a prototype of the ontology-based recommendation system has been constructed. Initially, the web services client stubs are implemented and used for earthquake simulations.
- The user-friendly, easily-accessed, and browse-based authorized interfaces are implemented to manage the databases, and provide access at different levels of abstraction.
- The semantic metadata management system has been designed based on the Classified Interrelated Object Model. This system prototype is under development.

UNDERLYING TECHNOLOGIES

- A subset of Semantic Data Model, the Classified Interrelated Object Model, serves as the core of the semantics wrapping process.
- Rule-based reasoning agents accommodate data-mining, learning and updating capabilities for metadata management.
- Compatibility is assured with web languages and technologies such as XML, XML Schema, RDF, DAML+OIL, and GML.
- Web Services are used for light-weighted protocols and universal data formats.
- Topic mining is being developed to perform event-based detection and tracking.
- The integration and representation of multi-modal data is accommodated by agent-based combination/merge routines.
- Real time dynamic data acquisition (e.g. GPS data, satellite data, and on-site observation data) is supported.
- Data assimilation and 3-dimensional graphic based simulation technologies are employed to convey information to and from scientists and users.

LIST OF PUBLICATIONS, REFERENCES, URLs

- S. Decker, D. Brickley, J. Saarela, and J. Angele, "A Query and Inference Service for RDF", *Proceedings of QL'98 - The Query Language Workshop*, December 1998.
- D. Fensel, F. van Harmelen, I. Horrocks, D. McGuinness, and P. Patel-Schneider, "OIL: An Ontology Infrastructure for the Semantic Web", *IEEE Intelligent Systems*, 16(2): 38-45, March/April 2001.
- J. Heflin and J. Hendler, "A Portrait of the Semantic Web in Action, Intelligent Systems", *IEEE Expert*, 16(2): 54 -59, March-April 2001.
- L. Khan, D. McLeod, and E. Hovy, "Retrieval Effectiveness of an Ontology-Based Model for Information Selection", *The VLDB Journal*, 2003
- G. Aslan and D. McLeod, "Semantic Heterogeneity Resolution in Federated Database by Metadata Implantation and Stepwise Evolution", *The VLDB Journal*, 18(2), October 1999.
- S. Chung, and D. McLeod. "Dynamic Topic Mining from News Stream Data", *Proceedings of ODBASE'03*, November 2003 (to appear).
- M. Pierce, C. Youn, and G. Fox. "Application Web Services" Internal Comm. Grid Lab Report. <http://www.servogrid.org/slide/GEM/Interop/AWS2.doc>.
- L. B. Grant and M. M. Gould, "Assimilation of Paleoseismic Data for Earthquake Simulation", *Pure and Applied Geophysics*, 2003
- T. Bray, J. Paoli, C.M. Sperberg-McQueen and E. Maler, Editors. "Extensible Markup Language (XML) 1.0, Second Edition," World Wide Web Consortium. 6 October 2000. This version is <http://www.w3.org/TR/2000/REC-xml-20001006>.
- H.S. Thompson, D. Beech, M. Maloney, N. Mendelsohn, Editors, "XML Schema Part 1: Structures", World Wide Web Consortium Recommendation, 2 May 2001. This version of XML Schema Part 1: Structures is <http://www.w3.org/TR/2001/REC-xmlschema-1-20010502/>

