2002

NASA FACULTY FELLOWSHIP PROGRAM

MARSHALL SPACE FLIGHT CENTER
THE UNIVERSITY OF ALABAMA

STANDARDS ADVISOR – ADVANCED INFORMATION TECHNOLOGY FOR ADVANCED INFORMATION DELIVERY

Prepared By:	J. Scott Hawker
Academic Rank:	Assistant Professor
Institution and Department:	University of Alabama Department of Computer Science
NASA/MSFC Directorate:	Engineering
MSFC Colleague:	Paul Gill

Introduction

Developers of space systems must deal with an increasing amount of information in responding to extensive requirements and standards from numerous sources. Accessing these requirements and standards, understanding them, comparing them, negotiating them and responding to them is often an overwhelming task. There are resources to aid the space systems developer, such as lessons learned and best practices. Again, though, accessing, understanding, and using this information is often more difficult than helpful.

This results in space systems that

- Do not meet all their requirements,
- Do not incorporate prior engineering experience,
- Cost more to develop,
- Take longer to develop.

The NASA Technical Standards Program (NTSP) web site at http://standards.nasa.gov has made significant improvements in making standards, lessons learned, and related material available to space systems developers agency-wide. The Standards Advisor was conceived to take the next steps beyond the current product, continuing to apply evolving information technology that continues to improve information delivery to space systems developers. This report describes the features of the Standards Advisor and suggests a technical approach to its development.

Product Overview

Figure 1 illustrates the concept for the Standards Advisor.

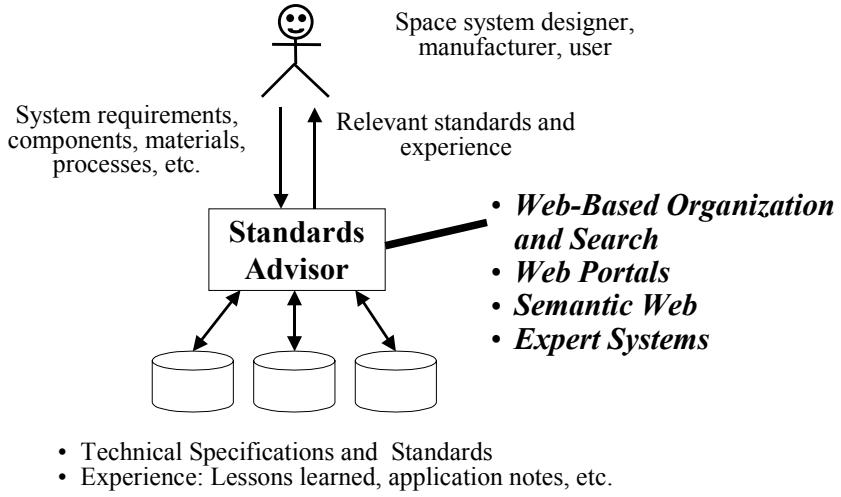

Figure 1: Standards Advisor Concept.

The current standards product provides keyword-based search of extensive standards and experiences databases, as well as organization of documents using various categorizations. The Standards Advisor will extend this capability in three areas:

- **Group/Project-Specific Web Portals.** These portals will provide the specific standards and other documents the group uses for their work, directly in the context of their work.
- **Advanced Document Search and Organization.** Newly available web search technology allows "fuzzy" word search, automatic concept/summary extraction, automatic document categorization, and search of the content of documents not stored in HTML (web text) format.
- **Knowledge Capture, Query, and Reasoning.** Augment the documents with additional data, using the language of the space system developer, capturing information about the meaning and content of the document. This will provide a more natural organization of and access to relevant documents. It will also enable software applications that reason about the documents in the context of the user's needs.

Technical Approach

This section suggests candidate technical approaches to provide the product features.

Web Portal

The primary issue in linking standards documents to web portals is passing user login credentials to 3rd-party servers of documents.

Figure 2 illustrates an approach to resolve this. An intermediate server script running on the NASA web server can front-end the document access, and a client-side script can pass user information as part of the document request.

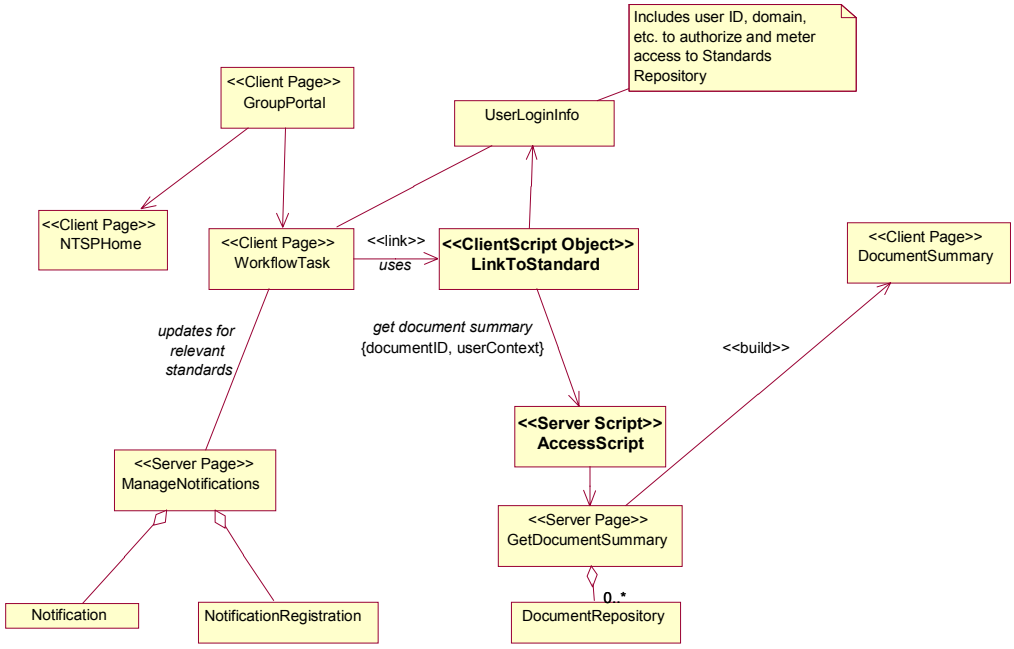

Figure 2: Using an intermediate server script.

Advanced Document Search and Organization

Products from a number of commercial vendors, such as Verity (www.verity.com), Google (www.google.com), and Inktomi (www.inktomi.com), all provide web site developers with tools to embed their search technology into intranets, allowing the Standards Advisor to use existing search technology to crawl, index, access, and display a diversity of data, including HTML and XML pages, application-specific document types (Word, Acrobat, etc.), and ODBC-compliant databases.

To crawl and index a web site or other repository, a crawler must have authority to do so. The crawler must be allowed through firewalls, and Robot META tags and robots.txt settings should be appropriate. Figure 3 illustrates a concept to crawl the sites and databases that get linked to standards documents. This would give the space systems developer a more complete view of standards in context.

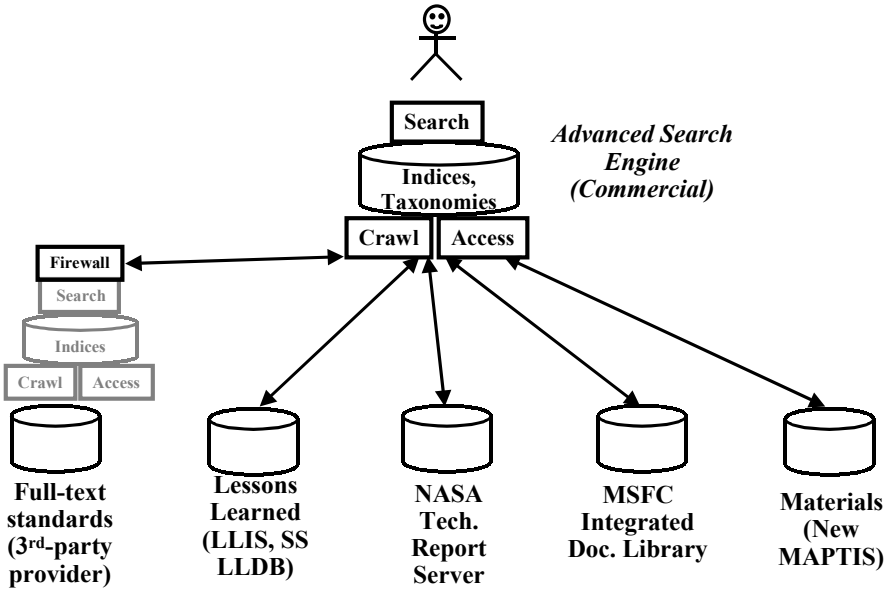

Figure 3: Advanced search across related databases.

Knowledge Capture, Query, and Reasoning

The dominant use of the web is, currently, to exchange and present documents for humans to view. More recently, the web has also become a means of data exchange and integration between software applications. Evolving standards for web services and metadata technologies are allowing web site and web application developers to use the web as a distributed computing environment, application database, and a knowledge base, as well as a document repository for humans. These technologies, collectively called the semantic web, provide an infrastructure to capture the meaning and intent of documents and related information, and to manipulate these to provide better information delivery for space systems developers.

As Figure 4 shows, the semantic web technologies enable capture of increasingly rich layers of data and knowledge, which, in turn, enable increasingly sophisticated processing and reasoning to support end users.

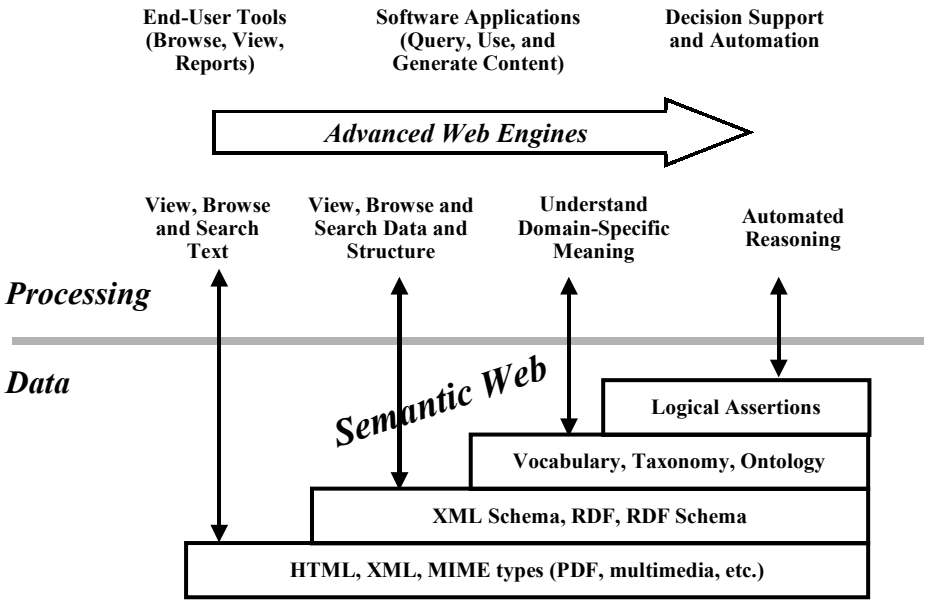

Figure 4: Increasing layers of metadata enable increasingly sophisticated functionality.

There are risks associated with using a technology as abstract and new as that of the semantic web. Users will now have to understand the basics of a new technology, and document developers will now have to create new information: ontologies and metadata.

Recommended steps

The next steps use prototypes to explore the possibilities in each area. Example steps include

- Prototype 1
  - Develop a "Smart Standards" portal for a selected NASA MSFC Engineering Directorate Group or manufacturing group
- Prototype 2
  - Implement an advanced search across the 3rd-party provider of full-text standards
  - Develop a strategy for integrated search across standards, lessons learned, etc.
  - Develop the requirements for and evaluate commercial search engine products
- Prototype 3
  - Use an advanced search engine to extract candidate metadata from focused collections of existing documents – begin to identify taxonomies and vocabularies
  - Build on existing ontologies to create a top-level standards ontology and taxonomy
  - Work with Topic and Discipline Working Groups to develop discipline-specific ontology and taxonomy
  - Add metadata to existing and new documents
  - Exploit the metadata (new search techniques, agents, expert systems, etc.)
  - Provide tools to develop and maintain ontologies and metadata

There are integration and overlap opportunities in related application areas such as workflow management, knowledge management, document management, product data management, and data warehouses.