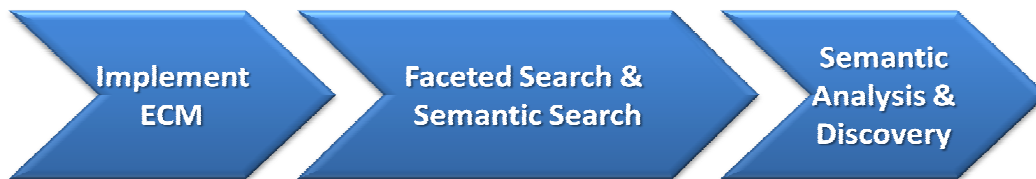


Semantic Business Intelligence System

The United States Air Force Acquisitions office has almost 200 data silos. These information repositories have unstructured and semi-structured information in them with the unstructured information largely consist of text based document such as reports. They recognize a growing need to be able to use and improve their knowledge generation ability with this unstructured information. When a General asks a question they often have to put together a team of analysis and data engineers just to be able to work on answering the question – with an unacceptably long turn-around time and cost.

As with any organization, using and getting the value out of these silos is a timely and costly task. Air Force Acquisitions understood they needed more than the ability to search – keyword search was not enough. They needed a way to connect and explore information relationships between and within information silos. Their goal is to improve knowledge generation and usage and make their vast amounts of unstructured content understandable. They want to build a business intelligence (BI) system for their unstructured content, with the ability to identify relationships between information and domain concepts.

Innovative Query, Inc. (IQI) helped them design a Semantic Business Intelligence system with a data warehouse and IQI’s knowledge management and discovery capabilities. This is a three step process.



1. Implement Enterprise Content Management (ECM) – XML server in this case
 - Data Warehouse, SharePoint, etc.
2. Implement Enterprise Search Capabilities
 - Faceted Search, part of the XML data warehouse
 - Semantic Search, part of IQI knowledge management and discovery system (IQeXplore)
3. Layer-on Semantic Analysis & Discovery -- IQeXplore
 - Dynamic Analysis
 - Discovering Connections
 - Integrate with Knowledge Model (Ontology)
 - Visualization & Collaboration
 - Knowledge Management

Innovative Query, Inc. provides tools and services that enable finding and exploring the Relationships & Facts in unstructured information for the management, discovery and creation of knowledge. We use novel visualization and semantic analysis to allow users to explore information relationships and make connections faster and easier – enhancing their natural work process.

An additional goal of Air Force Acquisitions is to allow analysts to identify and connect warning flags across multiple reports and information sources that may indicate a troubled project. Once these fact patterns are

Case Study: Air Force Acquisitions

identified, we can then monitor for these occurrences in future reports. IQI's tools are the central enabler for this goal.

IQI's tools semantically analyze information (documents and data sources) for entities and facts of interest, based upon an Air Force Acquisition defined ontology. The analysis is stored in industry standard XML/RDF format so other vendors and applications can utilize it also. We then put our IQeXplore semantic search & exploration tools on top of this data, all with the purpose to identify relationships and patterns in the original information to explore in new ways and generate knowledge more quickly.

See video screencast demonstrations on our website, under case studies at www.innovativequery.com for several illustrations of how this works and its value. These include an overview of the Air Force Acquisitions project, its need and solution; semantic search, illustrating fact based search and connecting facts together; relationship graphing, exploring fact relationships and following patterns; document relating by entities, shows relating documents by entities and categories of information; and correlating entities to document, which illustrates connecting entities by their relationship to collections of documents and NLP examples.

Implementation Overview

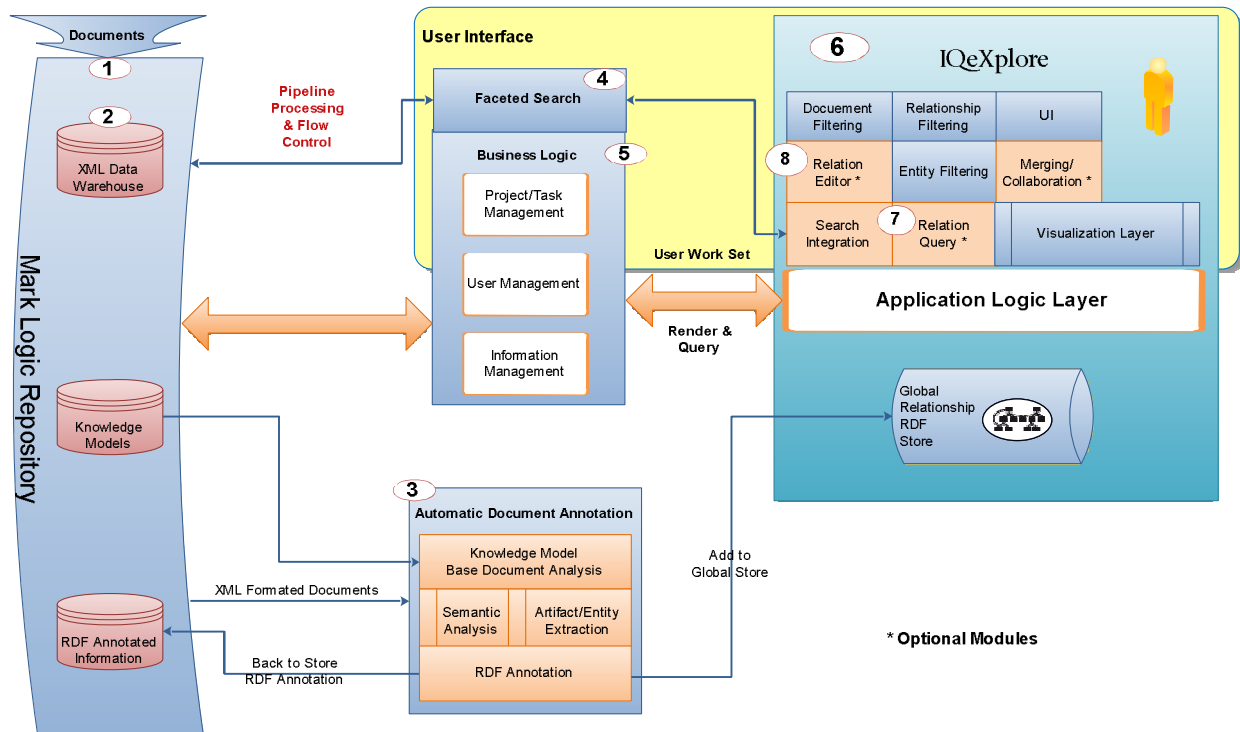
The following discussion refers to the figure below. The ECM system is implemented with an XML data warehouse using a Mark Logic repository (1). All documents and information sources are linked to or stored in the repository. When documents are added to the repository they are transformed into XML format (2) and this triggers the Automatic Document Annotation (3) module. This module semantically analyzes the XML formatted documents using state-of-the-art NLP (natural language processing) to extract entities and relationship based upon the Air Force Acquisition supplied domain knowledge models. Domain knowledge models are stored as OWL ontology files (Web Ontology Language, OWL), which are standards based semantic web formats. The documents are richly tagged with this semantic data that is used for search, knowledge management and information explorations activities.

The ontologies are also used to drive the faceted search engine (4) that is part of the Mark Logic server. The Business Logic (5) and the IQeXplore (6) modules are integrated with the search to control data access and help users create and manage document work sets for business intelligence (BI) exploration and knowledge discovery.

The Business Logic (5) module allows users to create and manage BI related projects for knowledge management and discovery tasks. It controls work flow, has user management, saves and loads work sets for projects and tasks. It is tightly integrated with the IQeXplore (6) module allowing the user to interact with search results (4) and the document repository (2).

The IQeXplore (6) module is the core of the BI and knowledge management system. IQeXplore is a semantic BI system solution that combines information relationship visualization, semantic relationship mapping, categorization and relationship filtering and other information exploration tools to accelerate information discovery and new knowledge generation. IQeXplore performed Ad-Hoc knowledge structuring using relationships and entities identified in the annotation process (3), allowing users to save tremendous time, cost and effort getting value from the unstructured information assets. Integrating this module with repository searches allows the user to add entities and relationship not previously defined in the ontologies. These can be used to add annotations to a document or work set or even add entities and relationships to the project ontologies.

Case Study: Air Force Acquisitions



Once new relationships, information connections and knowledge has been discovered with IQeXplore the user can then display (render) the information sets desired because of the integration with the Mark Logic server and its rendering engine. All information links back to the original content.

The Semantic Relation Query (7) and Relationship Editor (8) modules provide additional capabilities that allow deeper analysis and discovery. Semantic Queries allow deeper relations exploration on an ad-hoc level using a simple drag-and-drop editor to discover and query relationships in the user's work set or entire data warehouse. See the web site for a demo under case studies.

The Relationship Editor (8) module allows the user to create their own relationships (not previously defined or discovered) for even richer semantic analysis, knowledge management and discovery capabilities. The module allow easy knowledge model development for better exploration and knowledge discovery purpose. They can also integrate and feedback easily into master system wide knowledge models (ontology files), with management controls, providing a powerful positive feedback mechanism by driving an even richer semantic analysis (3).