Infosys helps Airlines Major to Build ‘Knowledge Lab’

Infosys GRADIENT ensures low-cost, high-value data integration architecture

The client is a major player in the highly competitive airlines industry. It has to its credit a vast pool of IP and knowledge amassed over years of research. As airlines industry thrives on superior customer service, right information at the right time is extremely crucial for business competitiveness. The client wished to develop a Knowledge Lab (KL) platform to leverage its knowledge capital. It partnered with Infosys to ensure that its IT infrastructure responds to its business requirement and make appropriate knowledge available on demand to its decision makers.
Situation
The airlines industry constantly combats with changes—from regulatory to customer expectations. Quick decision making based on appropriate information is crucial for sustained business competitiveness. The client is home to a large knowledge capital accrued over years of successful customer service and in-depth research and wished to leverage the same to retain its leadership position.

Currently, however, the client’s data is spread across heterogeneous formats and exists in multiple silos within the organization. The absence of a system that ensured a holistic view of the data to facilitate informed decision-making prompted the company to build a “Knowledge Lab”—a platform that leverages the company’s rich knowledge base and makes it available on demand.

Challenges
The current system, with data stored in heterogeneous formats that exist in multiple silos in the organization, failed to provide a holistic view of the data to facilitate informed decision making.

A significant amount of the client’s IP and knowledge was in the form of research data, technical papers, patent documents, design documents, PLM data and so on, and needed to be consolidated without affecting the quality of data.

The airlines major tasked Infosys with building a data integration solution that formed the basis of its Knowledge Lab (KL) with the following features:

- Design and build a knowledge extraction procedure to extract data from heterogeneous data formats
- Develop efficient knowledge transformation techniques to transform the data and store it in database
- Build an ontology of terms (present in the knowledge sources) and their relationships with the help of a domain expert
- Develop an information retrieval mechanism to retrieve documents from its repository of knowledge sources based on user-defined queries.

Infosys solution
Backed by a rich domain expertise in the area of building data integration solutions, we conceptualized a robust architecture for the client’s Knowledge Lab platform. With the help of our own Enterprise Information Integration tool, Infosys GRADIENT, we successfully achieved integration of the client’s disparate knowledge sources.

Infosys GRADIENT (GRid based Access of Distributed Information in the ENTerprise) is an Infosys EII solution that combines two powerful technologies—EII and Grid—to achieve scalable data virtualization. It is a real time data integration solution that integrates disparate data from distributed and heterogeneous data sources on demand and without any data replication.

Infosys GRADIENT leverages grid principles to achieve fast and scalable data integration. It is a lean middleware that lies lightly on your current IT infrastructure and utilizes advanced distributed and parallel query optimization techniques to enable efficient data integration. It is an SOA compliant platform built on open standards and technologies.

In addition to Infosys GRADIENT, other tools and techniques such as Copernicus tool for extraction of keywords, key phrases and document summarization, and Jena for enhancing user queries based on ontology were employed to enable the client to harness its rich knowledge capital.

As a stand-alone data integration solution, it does not directly change/enhance the business processes. The output of the solution can in fact be utilized by decision makers to influence or even enhance their business processes.
The Knowledge Labs solution can also be integrated with other operational systems of client enterprise, though these were outside the purview of the project.

**Implementation**

The key to the success of the project was efficient knowledge retrieval from disparate and distributed knowledge sources of different forms, knowledge engineering was of paramount importance.

We followed a two-phased approach for rolling out the solution to ensure that the business processes are least disrupted. The first phase involved extensive knowledge engineering to extract data from knowledge sources, infer domain ontology with the help of domain experts and create a metadata repository. Subsequently, knowledge extraction/retrieval mechanisms were designed using Infosys GRADIENT to extract relevant data based on ANSI compliant SQL queries.

The result: A light-weight solution that addressed the specific requirements of the client business. Off-the-shelf commercial tools are generic in nature and must be customized, proving to be cost, time and effort intensive. Besides, these tools come up with host of features, which may be redundant to client, making it bulkier.
**Business Benefits**

Our team of engineers took a few innovative routes to ensure that the client got what it had envisioned: using ANSI compliant SQL queries for fetching knowledge stored in various forms and locations is unique and differs from the generic approaches practiced in the industry, like DHT based P2P network. Our approach allows ad hoc SQL queries to be executed against the distributed knowledge repository. Also since our approach brings the benefits of Grid computing to Enterprise Information Systems, the solution is scalable and caters to a humongous distributed knowledge repository. The ANSI approach allows the integration of the platform with various BI tools and other custom applications for reporting.

Here are a few key business benefits of the solution:

- Seamless integration of structured and unstructured knowledge assets dispersed across the enterprise
- Greater flexibility to the end user in searching and reusing the distributed and heterogeneous knowledge assets. 15 different unstructured documents and several structured relational sources were integrated and efficiently searched through the solution.
- Information virtualization without building an expensive data-warehouse infrastructure enabling higher ROI. A data-warehouse solution might cost roughly $4 million to 5 million and take 2-3 years of implementation time, whereas our EII based solution takes roughly 2.5 months for implementation at a fraction of cost — around $0.15-0.3 million.