

Text Mining Hybrid Extraction Market Analysis

ASSYSTEM GROUP IMPROVES ITS VISION OF MARKET WITH WEBLAB FROM OW2

Key Words:

- SOA Integration
- WebLab Open Platform
- Industrial Market Analysis


Organization
 ASSYSTEM
Market

Industrial
Engineering

Status

A major player in
engineering and
innovation
consulting for
over 45 years

Objective

To offer Semantic
Business Intelligence
products

Country

France

OW2 Project

WebLab

The ASSYSTEM Group Chooses WebLab as the Architecture Base for its Market Analysis project called THEMA.

**Business: Industrial Engineering and Innovation
Consultancy**

45 years of experience have made Assystem one of the key partners of many of the world's largest industrial groups. Its partners include Airbus, Areva, Alstom, EDF, EADS, GE, MTU, Peugeot, Renault, Rolls Royce, Safran, and Thales. For example, Assystem combined its experience in electric vehicles and complex infrastructure management to provide assistance in managing the Autolib' project on behalf of the City of Paris. Assystem's engineering teams design and develop new products and services, while building and ensuring the optimum use of their investments throughout their life cycle.

With offices in 16 countries and a staff numbering 10,200, Assystem supports the projects of its customers and partners to provide cleaner energy, lighter aircraft, and electric modes of transport. are among the largest

Why Assystem Chose WebLab

Thema Market Analysis is an Assystem project aimed at providing Business Intelligence products to process open information sources. Given the dynamic and fast-paced nature of such information, the Thema system has to be flexible and adaptable enough in order to change or adapt its behaviour to new formats, sources, services or processing approaches. Assystem was looking for an SOA-compliant platform that would be both highly scalable and able to integrate heterogeneous components.

Assystem chose the WebLab open source platform as the base architecture of its Business Intelligence offering for three main reasons: a) it addresses its initial requirement; b) it provides a coherent architecture including flexible processing, work-flow and a unified presentation and c) it integrates the complete chain from the collection of open sources raw data to its processing to extract meaningful information and retrieval.

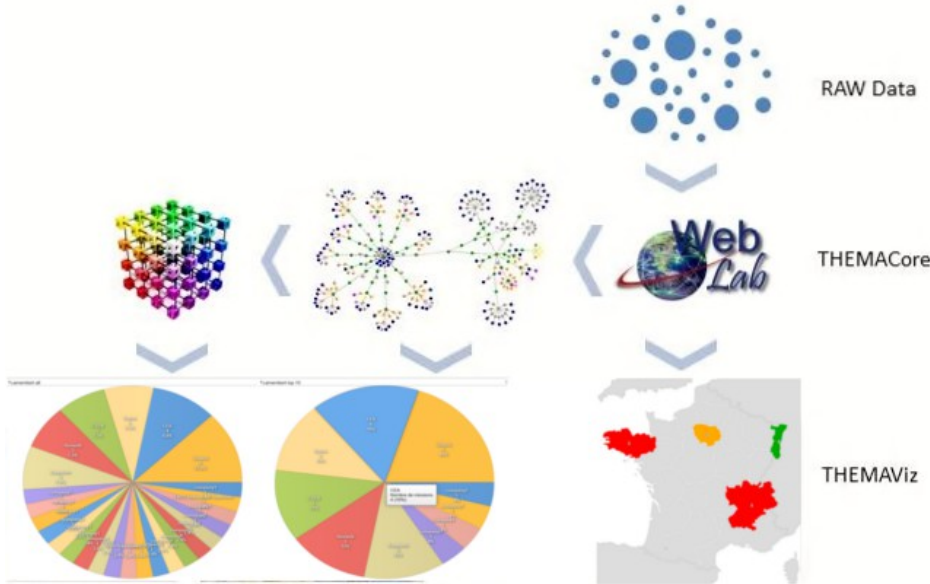
Assystem made WebLab the core platform of its OSINT integrated system out of which offerings such as the Thema project are instantiated. WebLab has enabled the fully integrated Assystem OSINT system to go from unstructured data out of open information sources to structured and actionable knowledge that enhance decision-making as required by projects such as Thema Market Analysis.

Text Mining Hybrid Extraction Market Analysis (THEMA)

The Webservice solution shown in the diagram below is called THEMACore. It is a complete environment extending the WebLab OW2 Web Services platform. Starting with the raw data, THEMACore absorbs company data, market data and human resources data at their entry points. THEMACore then structures the raw information into knowledge using WebLab-provided services and Components-on-the-Shelf that are plugged and orchestrated.

THEMAViz provides an efficient way to search and visualise structured data and Business Intelligence indicators, while providing the interconnection mechanism among the following:

- Raw data and semi structured data
- WebLab platform and BI Projection
- Data in general and meaningful knowledge.



The THEMA platform can be divided into three main steps:

- WebLab Semantic Structuration:**
Raw data (CVs, reports, requests for proposal, market data...) are structured along orchestrated services
- Business Intelligence Projection:**
Structured Data are projected into a BI Tool using a connector developed in a constant discussion with the main BI tool providers.
- Knowledge Visualization:**
Users can search, view and absorb the knowledge in a designed interface called THEMAViz.



Key Architecture Decisions

THEMACore is widely based on the WebLab open source platform hosted on OW2.

- Compared to other platforms, WebLab is the only platform that is open source, Webservice-based, media mining-oriented and semantic-compliant, all at the same time.
- As an open source platform with no license fees, it boasts an active and collaborating community and many already developed, tested and available services at no extra cost.
- Integration spending is reduced to a minimum with the Webservises and the semantic orientation. It allows development teams to easily plug and develop new components in order to maximise the business customization and orchestration of the processing chain.

Key Business Benefits

The powerful semantic web technologies incorporated in the Thema Market Analysis enable Assystem to offer unique value to their customers. The platform's LGPL licences allows Assystem to develop a new business activity by reselling Thema as a product.

About WebLab

WebLab is an open source platform providing business and strategic intelligence solutions and other applications that need to process multimedia data, text, image, audio and video.

Professional Support

User Mailing list: user@weblab-project.org
 Development Mailing list: dev@weblab-project.org
 Contact: Arnaud Saval, arnaud.saval@gmail.com

For more information, visit WebLab at: <http://weblab.ow2.org>

OW2

OW2 Consortium
 112-114 Bd Haussmann
 75 008 Paris, FRANCE
management-office@ow2.org

About OW2

Founded in January 2007, OOW2 is an independent industry community dedicated to developing open source code infrastructure (middleware and generic applications) and to fostering a vibrant community and business ecosystem. The OW2 Consortium hosts some one hundred technology projects, including ASM, Bonita, eXo Platform, JOnAS, JORAM, Orbeon Forms, Orchestra, Petals ESB, Spagic, SpagoBI and XWiki. Visit www.ow2.org