

# WEBSI

IST-2001-35458

## Data Centric Web Services Integrator

### Rationale

Nowadays, organisations produce, during their daily activity, a huge amount of data and information. This information is normally stored in heterogeneous information systems with diverse data formats. These companies are aware of the need of exploiting these intangible assets in order to obtain added value and competitive advantages in the market. They are demanding intelligent tools for searching and classifying information, and lately extracting knowledge from it.

On the other hand, XML has become rapidly the standard format for creating, archiving and exchanging information, as it permits to easily classify and extract knowledge from documents, while Web Services are self-describing, self-contained, modular applications that can be mixed and matched with other Web Services to create innovative products, processes, and value chains. The convergence of these two technologies provides a sound basis for exploiting information and knowledge repositories. To this end, **WebSI** introduces a combination of innovative abstractions, technical solutions, interfaces, components, and design tools to support the business sector in the building of applications as composed Web services, thereby offering to a broad spectrum of enterprises a low entry cost technology, capable of collecting information and composing services from different sources.

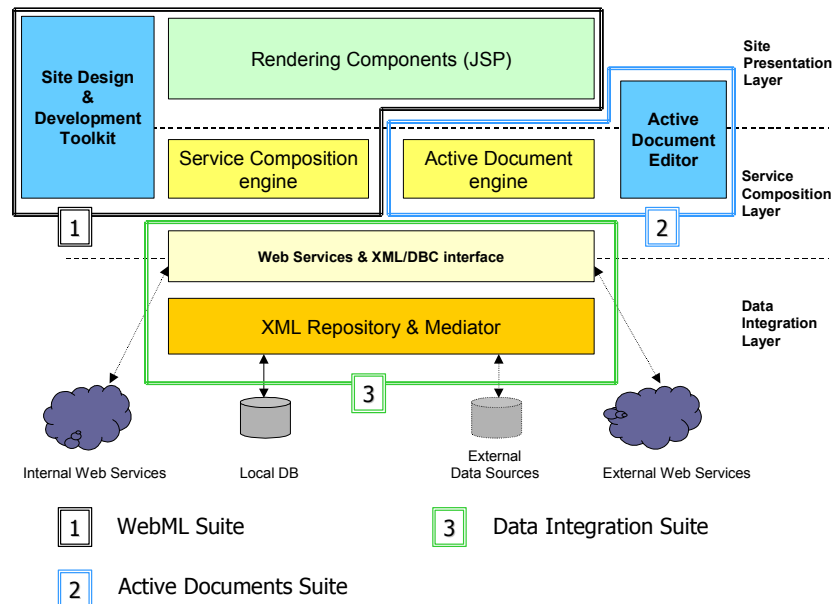
### Project Objectives

**WebSI** focuses on development and demonstration of three suites of tools for developing data-centric Web applications in open architectures, including the ASP (Application Services Provision) framework. It is focused on new low entry-cost technologies: (i) a **design and development toolkit** providing a graphical interface, supporting the visual specification of Web sites, and a code generator transforming visual specifications of hypertexts, web services and workflow primitives into application code; (ii) a **service composition layer** offering the run-time environment for integrating services as specified by the design toolkit, as well as a focused environment for supporting event-driven composition of complex documents from the assembling of pieces derived from queries, service invocations or other existing documents; and (iii) a **data integration service** allowing the gathering of integrated and enriched information from a distributed and secured environment.



## WebSI Architecture

**WebSI** enables fast and easy development of applications and services in open architectures based on integration of heterogeneous data sources and composition of existing and new web services. A typical **WebSI** application is a hypertext of Web pages integrating information access widgets to data sources and Web services. The **WebSI** platform provides tools for quick Web application design through simple data source modelling, visual specification of hypertext navigation, graphical composition of Web services and parameterization of presentations and renderings. The design notation is particularly easy and oriented to non-technical users.



The system is composed of three layers of software, covering the whole cycle of information, from original data sources to their presentation at user interface level, previously integrated and processed as required in the data integration and composition layers. This is illustrated in the figure.

## Main Components

1. The **WebML Suite** includes a graphical interface supporting the visual specification of Web sites in WebML, a high-level visual design language, and a code generator, transforming visual specifications into application code, a set of page templates (e.g. JSP pages) and unit descriptors which enable the execution of the application in the runtime environment.

The initial version supporting the specification of hypertexts and the graphical interface is now extended to allow composition of Web services and incorporates the units needed to specify Web services conversations and workflows, including message queuing services to allow the asynchronous interaction with remote Web services.

2. The **Active Documents Suite** consists of a set of cooperative web services implementing a distributed event-handling architecture and an editor of active documents. Active documents allow composing complex documents from the assembling of pieces derived from queries, service invocations or other existing documents in long-running event-driven stateful sessions.

The core web service in the suite is the Active Documents Manager, which manages the whole lifecycle of active documents. The suite includes additional web services like the Event Manager (for dispatching of XML-based events on the network), the Xscheduler (for registering of schedules and generation of the corresponding time events), and the XMail (an XML-based email server).

3. The **Data Integration Suite** transforms a set of local or remote data sources in an XML view accessible as a Web service through XQuery, the W3C XML query language. The XML/DBC interface offers a set of functions to query the XML view through a Java and SOAP APIs similar to JDBC but designed for XQuery. The suite includes the following main components:

- The XML Repository, which manages persistent collections of XML documents in a native XML file system and provides query access to them. It supports W3C XQuery and XML Schema standards, user access and authentication through Access Control Lists, a notification manager for typical subscription service, and stemming and semantic features for several European languages.

- The XML Mediator (XLive), a data integration middleware receiving XQueries, distributing them to sources and then returning integrated XML answers. Relational data sources are connected to the system through wrappers (XMLizer). The XML Repository can work also as a source underneath the Mediator.



## Innovation

**WebSI** provides an innovative whole-cycle solution based on a threefold paradigm:

- The use of high-level abstractions to design and generate web-based applications
- The composition of Web services as main core to generate these applications
- The integration of heterogeneous and dispersed information sources in XML format

Each one of these dimensions is based on state-of-the-art technology and produces technical innovation. The architecture proposed in **WebSI** explores indeed several novel directions that differentiate it from competitive products:

The most innovative aspect is probably the extension of the WebML conceptual modelling language and toolkit to allow visually specified Web applications to integrate XML data sources, the arbitrary composition of Web services, and the interplay with active documents.

The active document paradigm is also quite unique, as it provides a document-oriented application development paradigm, supported by the proper tools to compose, generate and publish documents as business workflows.

Finally the data integration suite is also an original contribution, as it provides a unified interface available in Java and SOAP called XML/DBC (XML Database Connectivity), supporting XQuery. It also includes a distributed XQuery processor (XML mediator) implementing query canonization and optimization, based on a new and promising native XML technology, and an advanced scalable XML repository.

In addition, all these innovative features will be ready to be deployed in an ASP environment, which is itself innovative

## Business Dimension

### Target market/audience

**WebSI** toolkit is especially addressed to industry sectors, including the service industry, and both the private and public sectors, with a special focus on SMEs with needs for EAI, eBusiness and Knowledge Management solutions.

### Advantages and benefits

In this context, the major business benefits that **WebSI** provides are the following:

It allows quickly design, development and deployment of web services- and XML-based advanced and complex solutions in distributed environments through the use of high-level tools like WebML toolkit and AD editor.

It allows low-cost exploitation of data sources and contents available within organisations, and the building of advance content-based and knowledge-based solutions.

It facilitates a better control over the lifecycle of the contents as well as the personalisation of contents and services made available all over organisations.

It facilitates the development of a flexible and extensible architecture with open interfaces based on web services.

It enables SMEs to catch up on the technology race in IT field, through low investment solutions.

There is a potential market for **WebSI** suites in the Application Service Providers sector, that will find in the three suites the way to build innovative solutions and services (a 25% of ASP's customers are SMEs)

The synergy between ASP and Web Services can give rise to new business models and opportunities, based on a wider vision of the traditional ASP paradigm: Service Providers, where companies will be able to interact directly over the Internet publishing specialised services on information and knowledge that can be engaged in complex applications using external IT infrastructure providers to host software and hardware resources, fostering new activities in a wider scope.

## Consortium Description

### Technological Partners

#### Software Houses



<http://www.ibermatica.com/>

**Ibermática** is a Spanish Consultancy and Software Engineering company with a twofold role in the project as co-ordinator and technology developer. As co-ordinator partner, Ibermática offers a sound experience in European projects management that has been demonstrated for the last eight years in different projects. As a technological partner, Ibermática brings in the project a solid background collected from leading European projects.



<http://www.xyleme.com/>

**Xyleme** is a French company providing breakthrough content solutions that are based on XML and database technologies. Xyleme brings their Zone Server in the project, an XML-native scalable repository, which will be enhanced within the project. They will also take the responsibility for the mediation components.

#### University Research Centres



<http://www.elet.polimi.it/>

**Politecnico di Milano** is the largest technical university in Italy. They participate in the project with the team of the Department of Electronics and Information, a well-known research centre for Computing Engineering. This group bring in the project its expertise in European projects and some results from W3I3 project that will be extended in **WebSI** (a web site design tool along with a web modelling language, WebML).

#### End Users



<http://www.acer.it/>

**Acer** is a world-leading PC manufacturer, offering a broad range of PC products, and a leading Internet-enabler, offering Internet component technologies, devices, services, end-to-end solutions, etc. Acer Italy will play the role of end-user in **WebSI**, developing a business portal application integrated in the new e-business strategy this company is currently involved in.



<http://www.diphuelva.es/>

The **Diputación de Huelva** (County Council of Huelva) is a Spanish local supramunicipal entity in charge of the government and administration of the Huelva province (Andalusia). The objective of Diputación de Huelva in the project is to define and implement an application to access, enrich, deliver and personalize the data sources held by the Regional and Local Public Administrations of Huelva, to give professional advice to new (or existing) entrepreneurs about the profitability of original undertakings.



<http://www.tiscover.com/>

**Tiscover AG** is a leading European and even world-wide provider of Destination Management Systems (DMS). The main product, Tiscover2000, is an ASP oriented information- and reservation system for SMEs within a Destination, based on a direct and distributed content (and availability) maintenance by the suppliers. The objective of this Austrian company in the project is to extend the existing system with comprehensive content services and search facilities.



**Project Officer:** Franco Accordino

This work is partially funded by the European Commission under the Information Society Technologies (IST), RTD Programme, Contract IST-2001-35458.

## Project information

<b>Project No:</b>	IST-2001-35458
<b>Start date:</b>	01.05.2002
<b>End date:</b>	30.09.2004
<b>Duration:</b>	29 months
<b>Project co-ordinator:</b>	José M. Muñoz
	e-mail: jm.munoz@ibermatica.com
	Tel: +34 91 384 9100
	Fax: +34 91 384 9144
	Website: <a href="http://www.ib-ia.com/websi/">www.ib-ia.com/websi/</a>
	Ibermática, S.A. Camino de Hormigueras, 172 E-28031 Madrid

