

# SARE-Bi: Gestor de documentación multilingüe sobre XML/TEI

**Joseba Abaitua, JosuKa Díaz,  
Josu Gómez, Inés Jacob**  
Universidad de Deusto  
Apartado 1 - 48080 Bilbao  
abaitua@fil.deusto.es  
{josuka,ines}@eside.deusto.es

**Garikoitz Araolaza  
Luistxo Fernández**  
Code&Syntax  
BIC-Berrilan, Azitaingo Industrialdea  
20600-EIBAR  
{GAraolaza,LFernandez}@codesyntax.com

**Resumen:** SARE-Bi es un gestor de corpus multilingües que permite la importación, exportación, actualización, alimentación y consulta directa de documentos originales y traducidos en un entorno de web. Los documentos pueden seguir tanto el estándar TEI como el estándar TMX, lo que aporta una gran flexibilidad. Se ha desarrollado con ayuda del servidor Zope y la herramienta Localizer.

**Palabras clave:** corpus multilingües, gestión documental

**Abstract:** SARE-Bi is a multilingual corpus management application that allows importing, exporting, updating, feeding, and browsing of both original and translated documents, with a web interface. The documents may follow the TEI standard as well the TMX standard, with the result of a great flexibility. The system has been developed with the Zope open server, and the Localizer tool.

**Keywords:** multilingual corpus, document management

## 1 Descripción

SARE-Bi es el resultado más destacable del proyecto del grupo de investigación DELi (Universidad de Deusto, [www.deli.deusto.es](http://www.deli.deusto.es)) titulado “XML-Bi: Procedimientos para la gestión del flujo documental multilingüe sobre XML/TEI”, que fue presentado en SEPLN 2001 (Abaitua *et al.*, 2001). Particularmente, el sistema SARE-Bi está siendo desarrollado en colaboración con la empresa Code&Syntax, S.L. ([www.codesyntax.com](http://www.codesyntax.com)).

SARE-Bi es un gestor documental multilingüe que permite la importación, exportación, actualización, alimentación y consulta directa de documentos originales y traducidos en un entorno de web.

El gestor se ha desarrollado sobre la arquitectura Zope. Tanto la importación como la exportación puede realizarse desde los estándares TEI (TEI Consortium, 2001) y TMX (Localization Industry Standards Association, 2002). TEI es más adecuado desde el punto de vista documental; TMX facilita la transferencia

entre SARE-Bi y los programas de memorias de traducción (WordFast, Déjà-Vu, etc.).

El sistema trabaja en estos momentos con documentos en tres idiomas (español, euskara e inglés, elegidos a partir del corpus recopilado durante el proyecto XML-Bi), pero resulta sencillo modificarlo para que maneje otros o más idiomas.

Por su parte, la arquitectura Zope permite una integración perfecta del gestor en entornos de trabajo colaborativo a través de web.

## 2 Características y requisitos

El sistema SARE-Bi funciona en cualquier sistema operativo (Windows, Linux, Solaris) que disponga del servidor web Zope (Zope 2002).

Zope es *open source*, y está especializado en gestión de contenidos, portales y aplicaciones personalizadas. Puede funcionar como servidor autónomamente o utilizar otros servidores ya instalados (como Apache).

SARE-Bi utiliza los servicios de presentación y desarrollo multilingües que

proporciona Localizer (Ibáñez Palomar, 2002), programa añadido para Zope bajo la licencia GPL.

La programación específica del sistema SARE-Bi se ha desarrollado en Zope con el lenguaje Python (Python Software Foundation, 2002).

### **3 Agradecimientos**

El proyecto “XML-Bi: Procedimientos para la gestión del flujo documental multilingüe sobre XML/TEI” está financiado por el Departamento de Educación, Universidades e Investigación del Gobierno Vasco (PI1999-72).

### **Bibliografía**

Abaitua, J., A. Casillas, T. Diedrich, A. Domínguez, A. Garay, C. Isasi, I. Jacob, I. Madariaga, R. Martínez, J.L. Ramírez. 2001. XML-Bi: procedimientos para la gestión de flujo documental multilingüe sobre TEI/XML. En *Procesamiento de Lenguaje Natural 27*: 291-292.

Ibáñez Palomar, J.D. 2002. *Localizer home*. <http://www.j-david.net/localizer/>.

Localization Industry Standards Association (LISA). 2002. *LISA/OCAR-TMX-Translation Memory eXchange Standard*. <http://www.lisa.org/tmx/>.

Python Software Foundation. 2002. *Python Language Website*. <http://www.python.org/>.

TEI Consortium. 2001. *Text Encoding Initiative*. <http://www.tei-c.org/>.

Zope. 2002. *Welcome to Zope.org*. <http://www.zope.org/>.