EASTIN-CL

**Crosslingual and multimodal search
in a Portal for Support of Assisted Living
Project no. 250432**

Deliverable D6.2
EASTIN-CL portal software

Version: V1.0

Date: August 2011

Document Information

Deliverable Number	D6.2
Deliverable title	EASTIN-CL portal software
Due date of deliverable	August 2011
Actual submission date	August 2011
Main Author(s)	A. Agnoletto
Participants	
Workpackage	WP 6
Workpackage title	Integration, test, and evaluation
Workpackage leader	Fondazione Don Gnocchi
Dissemination level	restricted
Version	V1.0

History

Version	Date	Status	Contributions
V1.0	2011-08-31	draft	

Executive Summary

This deliverable presents the new version of the EASTIN portal software. It has been adapted to be able to host the EASTIN-CL language technology components.

Content

1. Introduction

The current deliverable, ‚*D6.2 EASTIN-CL portal software*‘, provides the platform for the integration of the EASTIN-CL language technology components. This was not possible with the previous architecture of the EASTIN portal, and therefore a re-design of this software was required, according to the lines discussed in teh EASTIN-CL specification (‚*D1.1: EASTIN-CL System specification*‘).

The new version of the portal software is supposed to provide the same functionality to the users as the previous version; but in addition it enables the integration of the other EASTIN-CL components.

This deliverable decribes the new portal software. The software is fully operational and can be inspected and tested under the URL
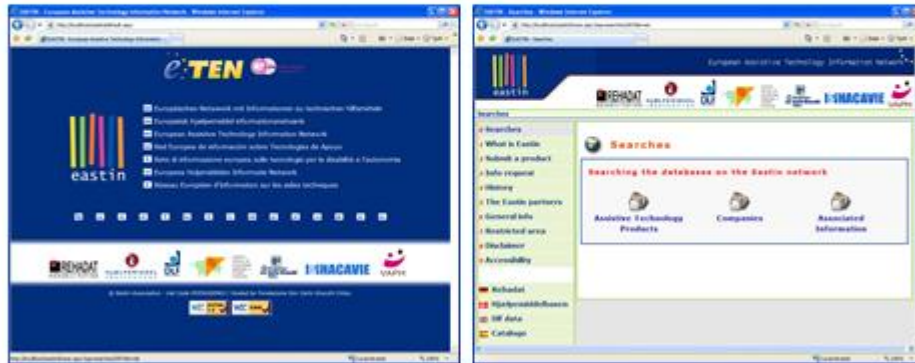
<div align="center">

http://www.eastin.eu

</div>

This software is the real deliverable.

The present document gives an explanation, and argumentation, what has been changed and why. It uses the form of a presentation, to highlight the main aspects of the software.
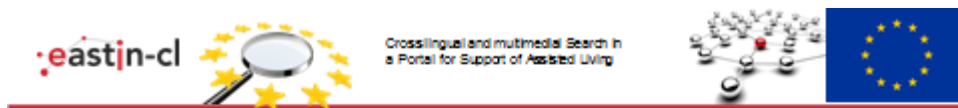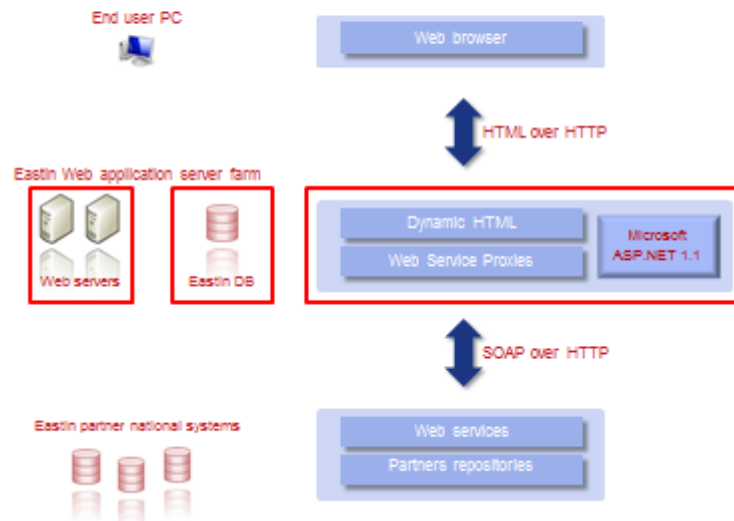
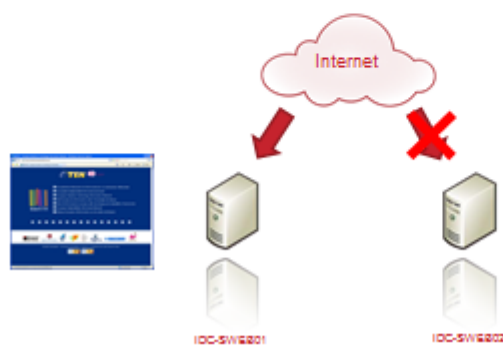## 2. The previous version of the EASTIN portal software

Problems: the Web servers

1. No load balancing. Even if in our server farm there are two Web servers which are capable to evenly split the requests coming from users (load balancing) the old Eastin was installed in one of the two machines only, because of its architectural structure.

Problems: the database

2. Old database engine. Eastin database was hosted by SQL Server 2000; it is a good product but "a bit" too old and this may result in a loss of performance and in maintainability issues.

3. **Not optimal database design.** The old database was designed to support a fixed number of partners and languages. Scalability, that is adding new partners or languages in the Web site, was a heavy issue.
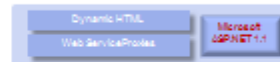

Problems: the architecture

4. **Not compliant HTML.** Eastin site HTML, generated using Microsoft ASP.NET 1.1 technology was not compliant with the last W3C standards and accessibility specifications.

Demo

·eastin-cl

·eastin-cl   Crosslingual and multimedial Search in
a Portal for Support of Assisted Living
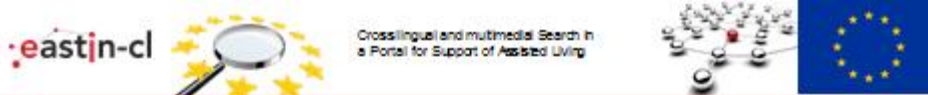
5. Bad indexing by search engines. Several pages of the site could not be correctly indexed by search engines

·eastin-cl   Crosslingual and multimedial Search in
a Portal for Support of Assisted Living
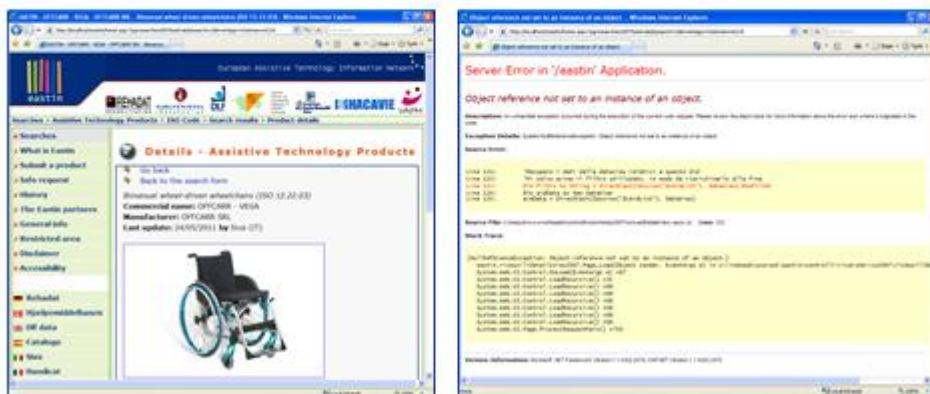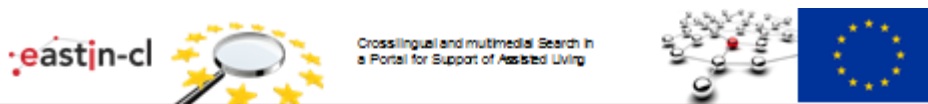
6. Difficult external functionalities integration. Since there was no separation between the logic layer (data elaboration) and the presentation layer (Web pages production) it was very difficult to plug in new functionalities in a modular way.
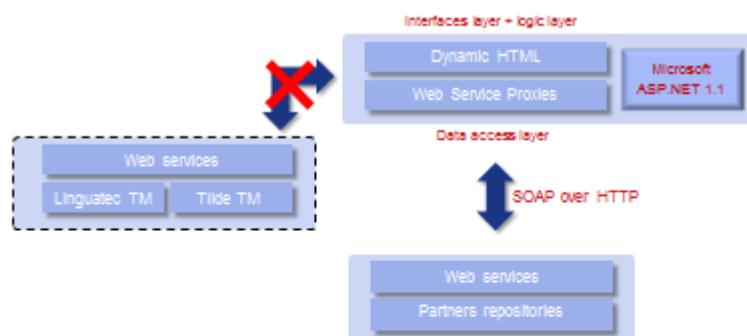
7. Not professional look and feel. Old Eastin site graphic layout gave in some way the idea of an amateur application, which was not true: Eastin reached the top level ranking for its technological design from Eastin project European Commission.

3. The new version of the EASTIN portal software
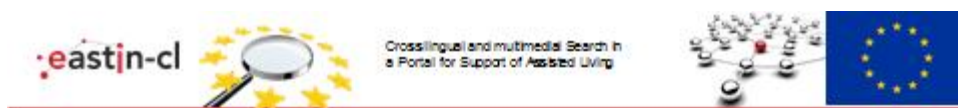
3.1 Web server
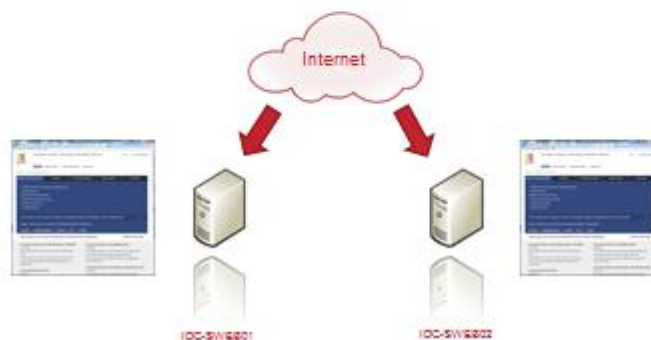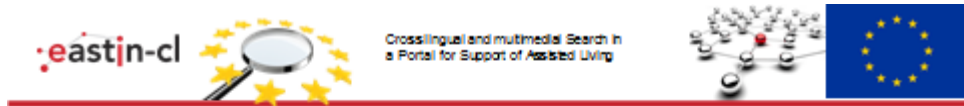
3.2 Database

The new Eastin: database

1. Engine step–up. The database has been migrated from SQL Server 2000 to SQL server 2005. The SQL Server 2005 is a two–machine–cluster designed for redundancy and high performances.

SQL Server 2000 → SQL Server 2005 Cluster
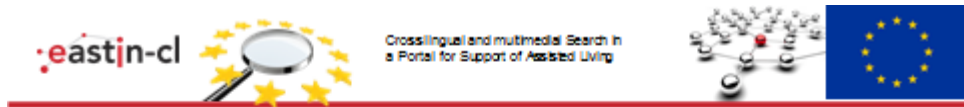
2.1 Database schema step–up. A dedicated software application has been developed to migrate old data from the old database to the new one.
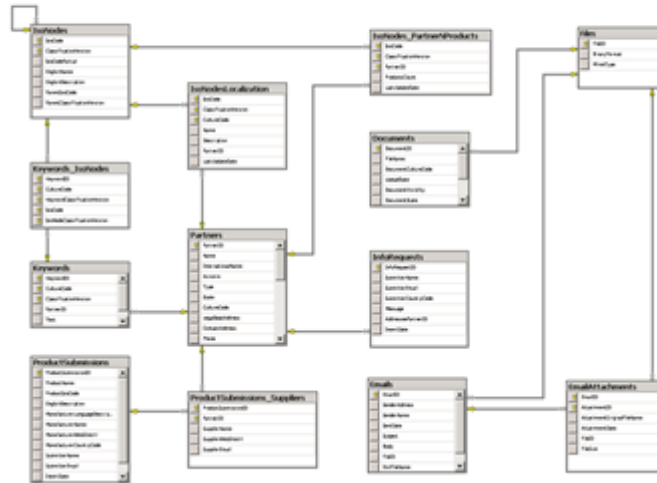
2.2.1 Database schema step–up. Now labels and ISO classification are fully and natively multi-language designed.
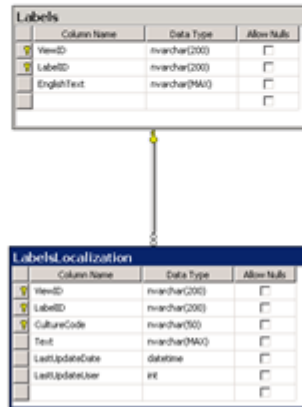
2.2.2 Database schema step–up. LabelsLocalization table.

·eastjn-cl

Crosslingual and mutimedial Search in
a Portal for Support of Assisted Living

2.3 Database schema step–up. Cultures.

Cultures
- Code
- LanguageCode
- CountryCode
- DotNetCultureInfoName
- IsDefault

Countries
- Code
- EnglishName
- Version

ISO 3166-1

Languages
- Code
- EnglishName
- Version

ISO 639-1

Table Cultures

|	Code	LanguageCode	CountryCode	DotNetCultureInfoName	IsDefault
1	bg-BG	bg	BG	bg-BG	1
2	cs-CZ	cs	CZ	cs-CZ	1
3	da-DK	da	DK	da-DK	1
4	de-AT	de	AT	de-AT	0
5	de-DE	de	DE	de-DE	1
6	de-LU	de	LU	de-LU	0
7	el-CY	el	CY	el	0
8	el-GR	el	GR	el-GR	1
9	en-GB	en	GB	en-GB	1
10	en-IE	en	IE	en-IE	0
11	es-ES	es	ES	es-ES	1
12	et-EE	et	EE	et-EE	1
13	fi-FI	fi	FI	fi-FI	1
14	fr-BE	fr	BE	fr-BE	0
15	fr-FR	fr	FR	fr-FR	1
16	fr-LU	fr	LU	fr-LU	0
17	hu-HU	hu	HU	hu-HU	1
18	it-IT	it	IT	it-IT	1
19	lt-LT	lt	LT	lt-LT	1
20	lv-LV	lv	LV	lv-LV	1
21	mt-MT	mt	MT	mt-MT	1
22	nl-BE	nl	BE	nl-BE	0
23	nl-NL	nl	NL	nl-NL	1
24	pl-PL	pl	PL	pl-PL	1
25	pt-PT	pt	PT	pt-PT	1
26	ro-RO	ro	RO	ro-RO	1
27	sk-SK	sk	SK	sk-SK	1
28	sl-SI	sl	SI	sl-SI	1
29	sv-FI	sv	FI	sv-FI	0
30	sv-SE	sv	SE	sv-SE	1

17

·eastjn-cl

Crosslingual and mutimedial Search in
a Portal for Support of Assisted Living

2.4 Database schema step–up. Now files are stored as BLOBs (Binary Large OBjects) into the database directly, with an improvement in download and upload performances (in the old Eastin they were stored in the Web server's file system).

Files
- FileID
- BinaryFormat
- MimeType

Documents
- DocumentID
- FileName
- DocumentCultureCode
- UploadDate
- DocumentVisibility
- DocumentState

Emails
- EmailID
- SenderAddress
- SenderName
- SentDate
- Subject
- Body
- FileID
- EmlFileName

EmailAttachments
- EmailID
- AttachmentID
- AttachmentOriginalFileName
- AttachmentDate
- FileID
- FileSize

18

·eastin-cl

·eastin-cl   Crosslingual and multimedial Search in a Portal for Support of Assisted Living

2.5 Database schema step–up. Search performance can be monitored through the Logs tables.

19

·eastin-cl   Crosslingual and multimedial Search in a Portal for Support of Assisted Living
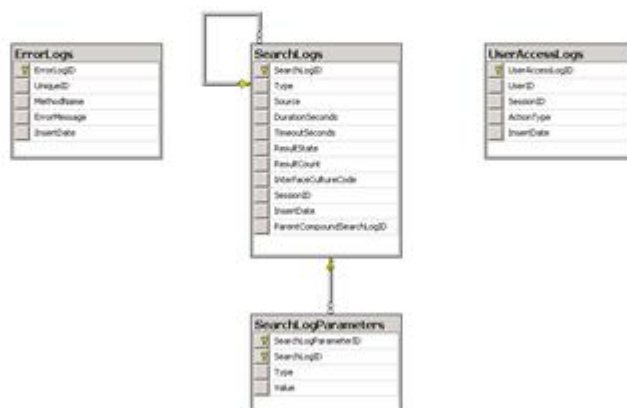
2.6 Database schema step–up. User roles.

	Enum/UserRoleID	Name	Description
1	0	Undefined	Undefined
2	1	Administrator	Administrator
3	2	NationalContentEditor	National content editor
4	3	PrivateAreaUser	Private area user
5	4	TesterUser	Role reserved to Eastin tester users

20

14

2.7 Database schema step–up. Many parameters can be configured in the Settings table.

3.3 Architecture

The new Eastin: architecture

1. Platform step-up. All code has been rewritten from Visual Basic.NET (Framework 1.1 2003) to C#.NET (Framework 4.0 2010)

2.1.1 Application architecture re-design. From ASP.NET Web Forms to **ASP.NET Model ViewController** (MVC). Old architecture.
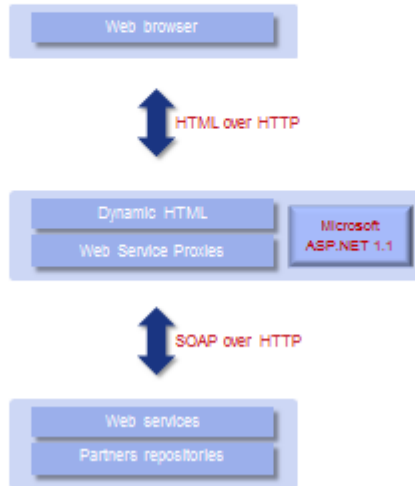
2.1.2 Application architecture re-design. From ASP.NET Web Forms to **ASP.NET Model ViewController** (MVC). New architecture.
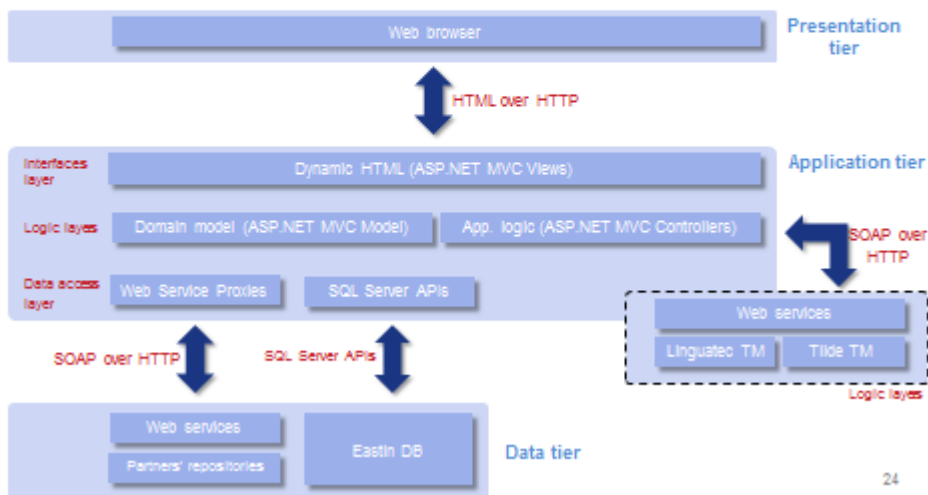
16

2.1.3 Application architecture re-design. Model View Controller architecture encourages a clean separation among three distinct modules:

– the **Domain Model,** or simply **Model**, which is a set of classes (objects) constituting an abstract representation of Eastin entities (Products, Companies, Associated);

– the **Controllers**, which are containers that wrap the application logic; each Controller is responsible for receiving an user request (that is one of the URLs of the site), elaborating data and passing them to the proper View for visualization;

– the **Views**, which are simply templates mixing HTML markup and the data coming from the Controllers.

25

2.1.4 Application architecture re-design. Now let's see the request – response cycle inside the layers of Eastin MVC architecture.
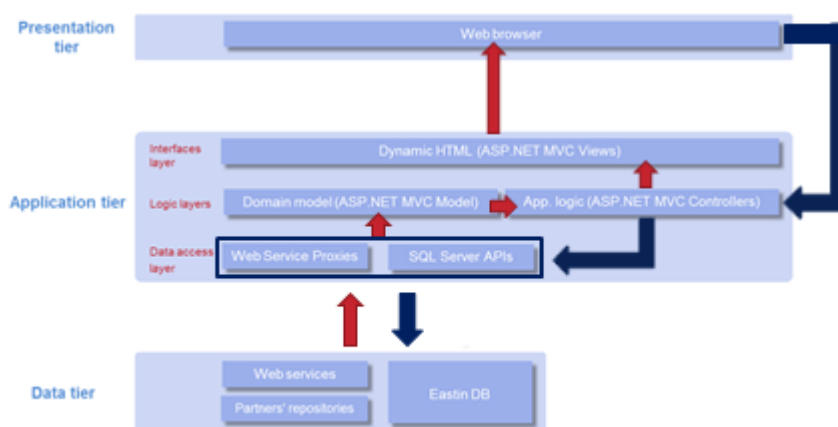
26

17

2.2 Application architecture re-design. Into the Domain Model data are a set of objects and collections. When elaborating data (sorting and filtering) the Controllers take advantage of the new query language provided by .NET 4: **LINQ**.

```
List<Partner> partnersWithDB = (from partner in globalStore.Partners.Values
                                where partner.CanProvideWebServices
                                orderby partner.DatabaseShortInternationalName
                                select partner).ToList<Partner>();
```

Thanks to a careful algorithmic design and to an optimized compilation engine all LINQ queries reach the best possible performances. For example, in a "where" clause, given a collection of m objects and a statement involving n properties of an object the computational complexity of the query execution is $O(m \times n)$ which is the best possible result.

27

2.3 Application architecture re-design. Thanks to this new technology Eastin comply with the most advanced HTML standards and accessibility recommendations defined by the **World Wide Web Consortium (W3C):**

- XHTML 1.1

- CSS 2.1.

- WCAG 2.0 level AA

Demo: the two validators in home page and the Accessibility section.

28

18

2.4 Application architecture re-design. The new Eastin is **fully Search Engine Oriented** (SEO):

1. An URL that is **human oriented** is also Search Engine Oriented

2. URLs must contain **all information** needed to reach the Web page

Note: to avoid the problem of broken URLs all old Eastin URLs has been redirected to the new site using a **permanent redirection** (HTTP 301).

Demo:

1. URL human friendly structure (URL driven application, REST)

2. URL cutting and pasting

29

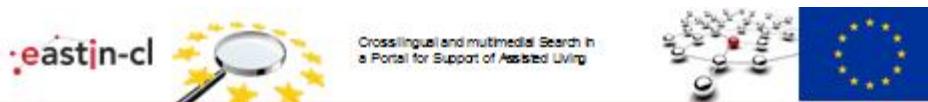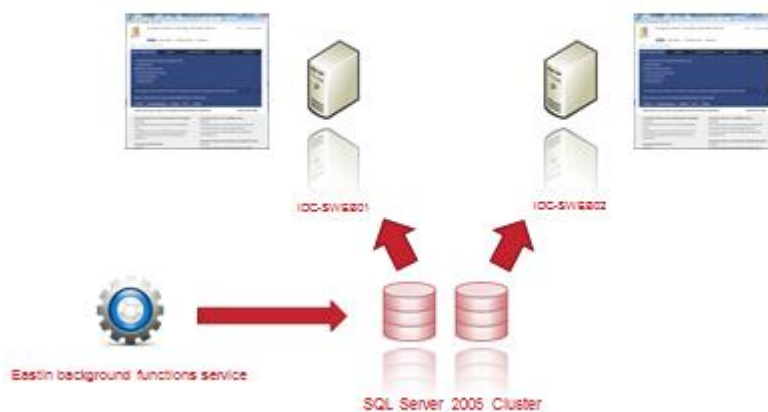**2.5 Application architecture re-design.** The new Eastin search engine takes advantage of the brand new **.NET Task Parallel Library**, which optimizes the computational effort distribution among the multi-core processors of the Web servers.

30

19

2.6.1 Application architecture re-design. The new Eastin background functions service is a stand alone Windows service, updating ISO classification and keywords by night.

31
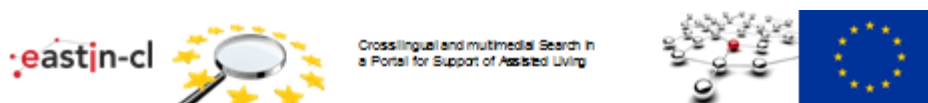
2.6.2 Application architecture re-design. Eastin Background updater performs four actions and their execution is scheduled as follows:

- **ISO Classification description:** the 2nd day of every month starting at 3.00 AM
- **Keywords:** the 3rd day of every month starting at 3.00 AM
- **ISO Classification number of products:** every day starting at 4.00 AM
- **Eastin Short and Extended distribution lists' emails storage:** every hour starting at the first minute (check the presence of email in the mail@eastin.info mail-box).

Eastin Web sites synchronize with the database every day at 5.30 AM.

32

20

3.4 .Net Cache object
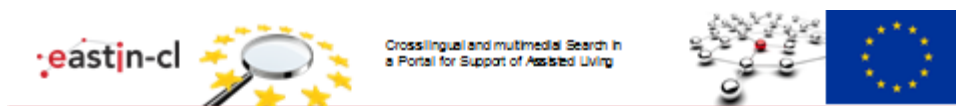
The new Eastin: the .NET Cache object

To improve the performance of Web pages generation, most frequently requested data (labels, ISO classification, keywords, partners) are fetched from the database and stored in the.NET Cache object. Each Web server has its own Cache. This kind of data have a fixed expiration policy; they are re-synchronized with the database every day at 5.30 AM.

Also data coming from searches (products, companies, associated information) are cached but with a different policy: a sliding timeout which can be configured in the Settings table (currently set to 10 minutes) is associated with each cached set of search results. If no user recalls that set of data within 10 minutes the search is executed again.
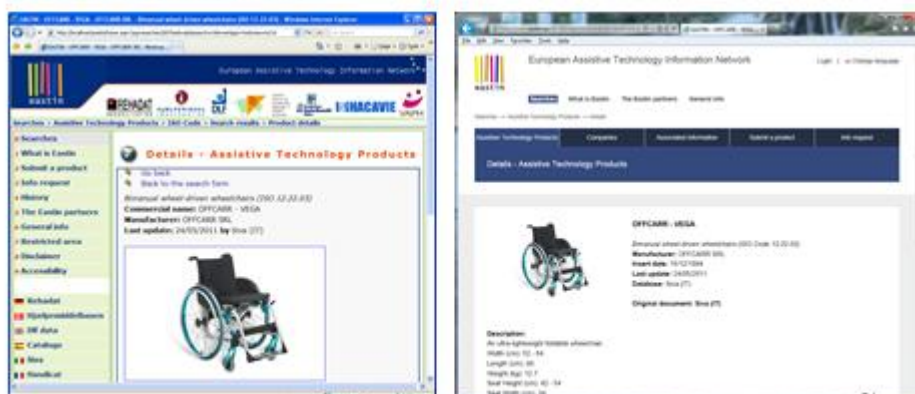
Demo

3.5 Graphic layout
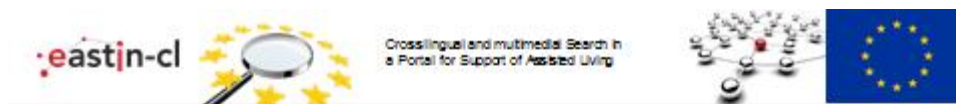
The new Eastin: graphic layout

The new graphic layout gives to the user the look and feel of a professional Web site about Assistive technologies.

3.6 Content Editing
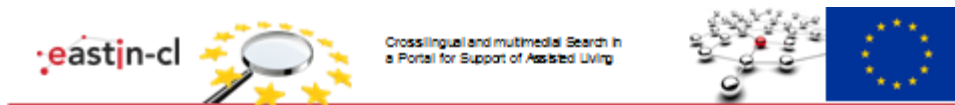
The new Eastin: content editing

A brand new content editing system has been developed to give each who has a content editor role the possibility to modify almost all contents of the site working on the Web site interfaces directly.

In the old Eastin content editing system only a small part of the site contents was editable.

Demo: the new content editor system: two levels of editing. Not editable content (select, form messages, meta text, some form's regions)

35

4 Next Steps

Crosslingual and multimedial Search in a Portal for Support of Assisted Living
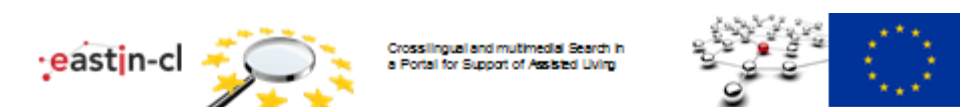
What's coming next?

1. September – October 2011: Eastin CL cross-lingual functionalities integration.

- A Query Processing engine to provide a free text search functionality (Goolge-like search)

- Real time language translation functionalities for some of the contents (products description)

- Text to speech functionality

2. November – December 2011: Eastin-Etna concept. A proof of concept evolution of the Eastin site will be deployed to show possible integrations with other data sources and other external functionalities.

36

Crosslingual and multimedial Search in a Portal for Support of Assisted Living
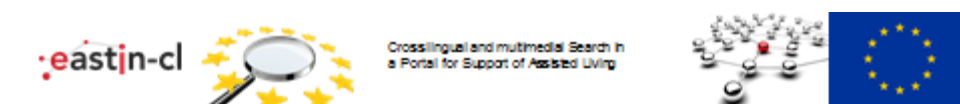
Final thanks to...

1. Don Gnocchi Sistemi. The collegue Enrico Carnelos, IT systems administrator, who assisted us during the processes of database migration and application deployment on Web servers. The Eastin Web servers and database are hosted by **Telecom Internet Data Center** (Milan) on behalf of Don Gnocchi Sistemi.

2. Sobrio. Who designed and tested the new graphic layout.

3. All of us. In particular: all SIVA collegues, Sabrina, Lucia, Renzo, Valerio, Matteo, and Antonio, who gave their precious support for the final testing of the Web site; all Eastin Association's partners, in particular Thomas, Martin, Marc and Dave, for their suggestions on the new Eastin design and for the effort in the deployment and test of the new SEO Web Services for Eastin; the Eastin-CL project, thanks to whom the Eastin site re-engineering has been possible.

37