Semantic Web technologies to increase Web Experience Management during Web Information Portal fruition

Vienna, 17/9/2015

Raffaele Palmieri & Vincenzo Orabona

R&D Projects/Knowledge & Technology Direction
Eustema Spa
Outline

- Company profile
- Initial situation
- Challenges
- Demo
- IT tools
- Faced problems
- KPI & Metrics
- Lesson learned and next steps
- Vision
Company Profile: Eustema

Main skills

- Business Intelligence & Analytics
- Business Process Management
- Enterprise Information Management
- Digital Media
- Enterprise Applications & SOA
- Business Information Security
- System management

25+ Years of ICT Experience

Offices (Rome, Naples, Milan)

3

90+ Italian Clients

530+ people

Research & Development
30 dedicated people
- **Italian funded project MISE n. E01/000812/00/X 17**
  
  Start date: 01/06/2010 - End date: 31/08/2014

- **Motivations:**
  - company has many years of experiences in the development of web portals and **CMS**
  - the need to offer customers new features to improve *knowledge management* within organizations
  - optimize operating costs of *web portal management*

- **Inspirations:**
  - EU funded project: **IKS**
Intrage is one of the first information portals for elderly, giving useful informations that regard employment, pension, taxes, health with attention to the care and needs of elderly and disabled people.

The portal is implemented using Microsoft Sharepoint and it’s shown use case.

- More than 12M visits, unique visitors 5M
- 6M sessions
- About 100K registered users
Intrage initial situation

• The contents are managed by a **team of editors**, each of them is an **expert** in a particular **subject**, monthly **hundreds** of **content** are published.

• Editors don’t **share** effectively the **knowledge**.

• **Knowledge** is all **enclosed** in the unstructured **CMS content**, and cannot be exported outside.

• The contents don’t link external **LOD**.

• Publishers do not use tools for the **semi-automatic annotation** that **optimizes** the **process** of **content production**.

• Users cannot **search semantically** published contents.

• **Visibility** of web site on **search engines** (Google, Yahoo..) isn’t favored by **schema.org markup**
Rethinking WCM to improve **User Experience**

- Make Publishing work **easier**
- Increase **Web Presence**
- Facilitate **Content Search**
- Engage and retain **Users**
- Support **third-party** data sources integration
### How to achieve the targets

#### Content Driven phase *(completed)*
- Catch hidden info into content
- Create your resources to represent your domain
- Analyze your content and extract metadata to enrich your data

#### User Driven phase *(ongoing)*
- Track user behaviour
- Profile user and him/her interests
- Recommend classified content and/or media on user profile basis

#### CAPABILITIES
- Content Classification
- Domain modeling by Ontologies
- Entity annotation & linking
- Content similarity
- Semantic Search
- User profiling
- Content Recommendation
Faced Problems

- Ontology Domain Construction (Domain Experts, time consuming)
- Data Sync (CMS-Semantic Engine & Data Model)
- Italian Language richness (Disambiguation, NLP Light)
- Editor WYSIWYG integration
- CMS Independence (CMIS)
- LOD ETL (LOD Ontologies manipulation, Data Fusion & Quality)
- Personalization Effort (to enable knowledge experience from UI)
## Success criteria: KPIs

<table>
<thead>
<tr>
<th>Capabilities</th>
<th>KPIs</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantic classification of content</td>
<td>Better time-efficiency of editing and publishing content</td>
<td>★★★★★</td>
</tr>
<tr>
<td>Semantic Tagging</td>
<td>Enhance user experience</td>
<td>★★★★</td>
</tr>
<tr>
<td>Web Experience Management</td>
<td>User experience enhancement, enabling geospatial content search</td>
<td>★★★★</td>
</tr>
<tr>
<td>Semantic geo-referentiation</td>
<td>Inspiration of editor's creativity through content enrichment</td>
<td>★★★★★</td>
</tr>
<tr>
<td>Editorial support</td>
<td>Increase of web presence through schema.org tagging</td>
<td>★★★★</td>
</tr>
<tr>
<td>SEO</td>
<td>Search experience enhancement enabling content navigation based on extracted information</td>
<td>★★★★</td>
</tr>
<tr>
<td>Faceted Browsing</td>
<td>Search experience enhancement using natural language</td>
<td>★★★★</td>
</tr>
<tr>
<td>Search semantically expanded</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Success criteria: first KPI

Publishing time decreases through the availability of Classification and Entity Annotation services which assist users in the content creation process.

Better time-efficiency of editing and publishing content

<table>
<thead>
<tr>
<th>$\Delta t$ (sec)</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta t \leq 10$</td>
<td>★</td>
</tr>
<tr>
<td>$10 &lt; \Delta t \leq 30$</td>
<td>★ ★</td>
</tr>
<tr>
<td>$30 &lt; \Delta t \leq 60$</td>
<td>★ ★ ★</td>
</tr>
<tr>
<td>$60 &lt; \Delta t \leq 120$</td>
<td>★ ★ ★ ★</td>
</tr>
<tr>
<td>$\Delta t &gt; 120$</td>
<td>★ ★ ★ ★ ★</td>
</tr>
</tbody>
</table>

Semantic classification of content

Semantic Tagging
Prospects & Recommendations

LESSONS LEARNED
- NLP people and skills
- Time consuming ontology construction
- Not CMIS-compliant web portals
- Required data stewards

NEXT STEPS
- Complete “User driven” phase
- Applying “profiling” to Advertising mkt
- Rethink to new business cases:
  - CRM integration
  - HR
  - Social communication channels
Thank you!

Eng. Raffaele Palmieri
IT System Architect
r.palmieri@eustema.it

Eng. Vincenzo Orabona
Enterprise Architect
v.orabona@eustema.it

Other contacts:
info@eustema.it