

# **VICODI:**

## **Visual Contextualisation of Digital Content**

Edvins Snore

Scientific coordinator

Edvins@ri.lv

### **Abstract**

VICODI is a collaborative RTD project carried out by 7 partners from 6 European countries under the 5FP IST programme of the European Union.

The goal of VICODI is to enhance human understanding of digital content on the internet. This will be achieved by introducing a novel visualisation and contextualisation environment for digital content. The project will also provide an innovative set of multilingual contextualisation mechanisms.

The VICODI contextualisation environment addresses the management of search and retrieval as well as the management of information presentation, by:

- 1) creating a cumulative and open knowledge space that can be enhanced by its users
- 2) providing an innovative user interface that employs SVG for the presentation of information that has been contextualised with respect to LOCATION, TIME, and SUBJECT. European history is the showcase application subject that VICODI has chosen to demonstrate the viability of its innovative information structure.

### **INTRODUCTION**

The mission of VICODI is to enhance people's understanding of the digital content on the Internet. This is reached by introducing novel visualisation and contextualisation environment for digital content. The project will provide a novel set of multilingual contextualisation mechanisms which will be demonstrated as follows:

- context sensitivity will be defined over topics, locations, and time,
- a context sensitive search engine operating on multilingual input, and using latent semantic indexing, ontological markup, and neural classifiers will be developed,
- a context sensitive knowledge structuring engine for multi-dimensional indexing will be developed,
- a graphical contextualisation interface based on Scalable Vector Graphics (SVG) will be developed,
- showcases in the application area of European History will demonstrate the viability of such novel information structuring and visualisation devices.

### **WHY**

By definition, data taken out of context is "information". Information structured in context with other "words" and "ideas" becomes knowledge. This conclusion is at the cornerstone of VICODI, the main objective of which is to develop novel methodology and tools for comprehensive structuring and

graphical visualization of context – by that facilitating creation of new knowledge.

VICODI addresses two key technical aspect for next generation, semantic web interfaces: firstly, the need to provide contextualisation services to large knowledge spaces and secondly, the quest for more appropriate user interfaces for the visualisation of richly structured, contextualised content. Current web interfaces fall short of such expectations, because on the one hand, "context" remains an ill-defined concept, and on the other - visualisation remains media-centered (flash, animated gif, mpeg streams) instead of becoming knowledge-centered.

## **HOW**

### ***Approach***

The VICODI Contextualisation Environment addresses the management of search and retrieval as well as the management of information presentation, by:

- 1) creating a cumulative and open knowledge space that can be enhanced by its users;
- 2) providing an innovative user interface that employs SVG for the presentation of information that has been contextualised with respect to **location, time, and subject**.

### ***Contextualisation - Tools and Mechanisms***

Wurman describes contextualisation as a process that can be circumscribed by the forced acronym **LATCH**: (L)ocation - (A)lphabetical order - (T)ime - (C)ategory - (H)ierarchy .

In VICODI, we shall use ontological markup and knowledge authoring tools to cater for (C) and (H). We will use cartographic maps for locational context, and we will use thesauri for historical information in order to explore temporal contextualisation, in the field of European History.

Despite the above-mentioned specifics (show-case) our environment will be generic and at the technical level, will consist of four major modules:

- 1) A Management System of Knowledge Space (MSKS) that encodes the LATCH-properties of any given content, including attendant authoring tools to add LATCH properties where needed. The information system will require knowledge management facilities in order to support reasoning over ontological markup.
- 2) A multilingual contextualisation module that offers translation services and cross-lingual indexing facilities.
- 3) A generative contextualisation module that selects which LATCH properties are relevant for a given context and which creates contextualised content that is enhanced by relevant knowledge structures.
- 4) A context-enabled, SVG based user interface for presentation of, and interaction with, the contextualised content.

At the user end the VICODI contextualisation environment will perform as follows:

Once the user has found (or written himself) an article, he/she will be able to paste the text (or URL) into the on-line contextualisation tool and retrieve the same text with contextual links in it (as underlined words, phrases, names, years, etc.). By recognizing certain parameters (keywords, years, names, etc.), the system will generate appropriate context both for each hyperlinked word and the whole article. The context will be automatically generated from the resources provided by MSKS, and not taken from the whole web, as in case of ordinary search engines.

The user will receive the contextual data also in the form of interactive graphical pallets/maps. For example, world map - in case of global philatelic community, map of Europe - in case of European History portal, etc. So, if the user will paste (into contextualisation tool) an article about *Elisabeth I*, the system will automatically generate the 16<sup>th</sup> century Europe's map from predefined SVG drawings (embedded in MSKS) and will put links on the map – in places where they belong – individual countries, regions, etc. The user will be able to assess the 16<sup>th</sup> century situation both in England, the whole Europe and in individual countries.

Thereby VICODI will provide visually enriched interface to navigate both in place and time.

### ***Validation of the contextualisation environment***

The objective of VICODI is to develop a generic contextualisation system that can be applied to any knowledge field. Particularly, to those requiring geographical references: history, sports, politics, military, statistics, languages, tourism, municipalities, international organizations, etc.

We have chosen the specific knowledge area – European history - to validate the system. This envisages modeling of web community on European history set up by the means and techniques of MSKS and enhanced by the contextualisation tool. At the first stage the linkage of metadata from the repositories will be carried out by establishing 3 real-size testbeds where the validation of the system will take place. Later, more repositories will be attracted through dissemination.

## **PROGRESS SO FAR**

### **User and functional requirements of VICODI system**

Identifying user requirements was the first major task of VICODI team. The initial set of requirements based on user interface mock-ups was discussed already at the kick off meeting.

Also at the kick-off workshop, the technical group established an initial set of technical platform requirements, including choice of operating systems and implementation languages. During the first four month of the project, partners indicated the importance of each requirement by its ranking, using the scale "must", "should", "desirable", "not relevant".

After identifying user requirements partners draw the functional requirements of the visual contextualisation system. During this phase partners carried out analysis and UML use cases, initial ontology requirements relating to the contextualisation engine, template structure.

### **Selection and assessment of contextualisation data**

Over the period of two last month of 2002 VICODI content partners led by UNEW (leader of the "Panel of Content experts") carried out intensive search for relevant contextualisation data, including XML repositories. It was concluded that VICODI would at this stage have to content itself with existing web based resources and that there were no XML repositories available to us. After discussions between partners, UNEW wrote the scope for the historical content that VICODI will use. Content partners assessed the 174 web sites that had been collected and pooled between the partners. This was done to prepare ground for the assessment of historical content [Deliverable 8.1] in which we shall document VICODI data selection criteria.

### **Development of the VICODI Ontology**

During the final month of 2002 the VICODI ontology development process was commenced. Extensive list of competency questions (required for ontology development) was compiled. Also partners initiated discussions about the potential structure of the ontology.

### **System specification**

Technical partners of VICODI led by SRFG started to describe the high level architecture of VICODI. This task involved creation of appropriate UML component architecture and deployment diagrams (e.g. interaction, activity and state diagrams, collaboration, deployment, etc). Components were identified in order to assigned to them at the later stage. Investigation of application framework and application server resulted in the choice of Expresso.

On the basis of the user and functional requirements partners started to describe the functionalities of all the system components. Also the formats to be used to integrate the different VICODI repositories were drafted.

Partners started to describe the Interface specifications among components, also public interfaces for Contextualisation engine and CETools, Transformation Engine, Annotation tool, helper interfaces for Machine translation and user management.

RIDemo started to build mock-ups for the following validation cycle, including mock-ups for critical user interfaces for tools, MSKS interface and knowledge portal (SVG GUI).

### **Development of Contextualisation Mechanisms**

#### *Contextualisation Engine (CE)*

SRFG drafted Executive Summary Reports for partners, including initial components, architecture, initial context model and approach. Context research and ambiguity resolution research was continued. Initial CE

architecture and design was drafted. Issues relating the native service server to the CE component in the VICODI central server were discussed.

#### *Transformation engines development*

LATCH and SVG support issues were investigated. Also partners carried out investigation into linking and highlighting schemes and use of open source "Velocity" templating package. Initial work on transformation related classes was started.

#### *Annotation Tool*

Investigation of annotation tools, techniques, standards was started.

### **User Involvement, Promotion and Awareness**

During the first four month of the project all the VICODI partners started their dissemination activities aimed at promotion and awareness raising about the novel approach of visual contextualisation developed by VICODI. In month 3 (November, 2002) these activities were coordinated on the basis of Dissemination Plan produced by RIDemo.

VICODI project web site ([www.vicodi.org](http://www.vicodi.org)) was created in October, 2002. Partners put a specific link to the VICODI web site on their own web sites.

Active project promotion activities were carried out via the specific professional fields of VICODI partners. Several partners took the advantage of promoting VICODI in press conferences (SRFG : November, 2002) and professional seminars (NKP : December, 2002; SRFG : November, 2002).

The success of the awareness raising activities was demonstrated by a number of enquiries about the data management/contextualisation mechanisms that will be developed during VICODI. Most of the enquiries originated from outside Europe.

### **Future Work**

Entering 2003, VICODI partnership will focus on finalizing system specification, including final version of UML diagrams, system functionalities and interfaces specifications (among components) for the Contextualisation engine, CETools, Transformation Engine, Annotation tool, part of the Machine translation and web components mock-up descriptions.

Content partners will carry out final assessment of contextualisation data and produce recommendations for its potential use for VICODI needs. We foresee intense discussions and interesting developments while creating VICODI ontology.

A very important task for 2003 is technical development of MSKS (Management System of Knowledge Space), which will be built upon the Resource Description Framework. SVG interface development activities are also scheduled for the coming year. Use of the scalable SVG maps to visualize the content is one of the innovative aspects of VICODI.