ATLAS
Applied Technology for Language-Aided CMS
Project Presentation

www.atlasproject.eu

Ivan Koychev

koychev@fmi.uni-sofia.bg

Overview

- What is the problem?
- What is the proposed solution?
- How do we address the problem?
- Who are the target users?
- Who is in the consortium?
- General project information
- More information

Problem

- Large quantities of multilingual digital content
- Growing demand for:
  - new methods and tools for content creation and distribution
  - online content services that incorporate language-based technologies
  - mechanisms facilitating the process of manipulating heterogeneous multilingual content
  - better content navigation

Proposed Solution

- ATLAS – a software platform for multilingual web content management unifying :
  - automatic annotation of important words, phrases and names;
  - summarization;
  - categorization;
  - computer-aided translation.
- Supported languages:
  - English, German, Bulgarian, Croatian, Greek, Polish, and Romanian.
- Solutions:
  - **i-Publisher**
  - *i-Librarian*
  - *EUDocLib*

i-Publisher

- Adds a visualization layer to ATLAS.
- Provides a powerful web-based instrument for creating, running and managing small and enterprise content-driven web sites.

i-Librarian

- A thematic content-driven web site *built with i-Publisher*, which allows its users to:
    - store, organize and publish their personal works
    - locate similar documents in different languages
    - easily obtain the most essential texts from large collections of unfamiliar documents

EUDocLib

- A thematic content-driven web site *built with i-Publisher*.

- Publicly accessible repository of EU documents, which provides enhanced navigation and easier access to relevant documents in the user's language.
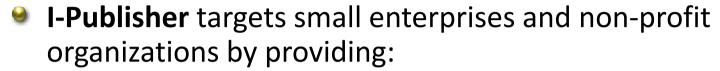
Functionality

- Functionality available in web sites built with i-Publisher:
  - distinction between content and visual presentation, which allows:
    - user interfaces to be created with a variety of technologies
    - content to be displayed on different devices
  - dynamic content propagation – any change of content is instantly reflected across all output channels where the content is visualized
  - multilingual versioning system, which provides:
    - visual comparison of versions of documents in one or more languages
    - versioning of structured content rather than the simple text-based
  - language processing of text content:
    - categorization, summarization, annotation of important phrases, words and names, computer-aided translation and similarity search
    - English, German, Bulgarian, Croatian, Greek, Polish, and Romanian

Target Users

- **I-Publisher** targets small enterprises and non-profit organizations by providing:
  - ability to build content-driven web sites via point-and-click user interface
  - wide set of predefined functionalities
  - automatic processing of textual content – categorization, summarization, annotation, etc.

- The **i-Librarian** web site addresses the needs of authors, students, young researchers and readers by allowing them to:
    - easily create, organize and publish various types of documents
    - find similar documents in different languages
    - share personal works with other people
    - locate the most essential texts from large collections of documents

Target Users cont'd

- The **EUDocLib** web site addresses the needs of users who require easier access to EU documents. Key features are:
  - multilinguality
  - easily find similar documents
  - the summaries of documents
  - extracted important words and phrases.

Consortium

- **Tetarcom IS Ltd – Bulgaria**
- Deutsches Forschungszentrum Fuer Kuenstliche Intelligenz GmbH – Germany
- Atlantis Consulting SA – Greece
- Institute for Bulgarian Language "Prof. Lyubomir Andreychin" at the Bulgarian Academy of Sciences – Bulgaria
- Instytut Podstaw Informatyki Polskiej Akademii Nauk – Poland
- Universitaet Hamburg – Germany
- Universitatea Alexandru Ioan Cuza – Romania
- Sveucilište u Zadru – Croatia
- Institute of Technologies and Development – Bulgaria

http://www.atlasproject.eu/atlas/project/consortium/en

ICT PSP call identifier: CIP-ICT-PSP-2009-3 Theme 5: Multilingual Web
5.3 Multilingual Web content management - methods, tools and processes

More Information

- Web site: www.atlasproject.eu
- Tetracom – project coordinator
    - Email: info@atlasproject.eu
    - Tel: +359 2 495 0444
    - Fax: +359 2 495 0443
    - Contact Persons:
        - Ms. Anelia Belogay – ATLAS CEO
        - Mr. Diman Karagiozov – ATLAS CTO
- ITD – dissemination and user evaluation
    - Contact Persons: Ms. Stanimira Yordanova
    - Email: s_yordanova@fmi.uni-sofia.bg
    - Tel.:   + 359 2 971 35 09
    - Tel./Fax: + 359 2 971 35 43

ICT PSP call identifier: CIP-ICT-PSP-2009-3 Theme 5: Multilingual Web
5.3 Multilingual Web content management - methods, tools and processes