

# Deliverable D 2.3

Software documentation

Grant Agreement number: 250467

Project acronym: ATLAS

Project title: Applied Technology for Language-Aided CMS

Project type:  Pilot A  Pilot B  TN  BPN

---

Deliverable D 2.3    Software documentation    30.12.2011

---

Project coordinator name, title and organisation:

Anelia Belogay, CEO, Diman Karagiozov, CTO,

Tetacom Interactive Solutions

Tel: +35924950444

Fax: +35924950443

E-mail: [anelia@tetacom.com](mailto:anelia@tetacom.com), [diman@tetacom.com](mailto:diman@tetacom.com)

Project website address: [www.atlasproject.eu](http://www.atlasproject.eu)

Authors:

Anelia Belogay, Emil Stoyanov, Diman Karagiozov, Petar Gadjev, Elitsa Marinova, Evgeni Styanov, Kiril Mutafchiev Radostin Surilov (Tetacom)

Project co-funded by the European Commission within the ICT Policy Support Programme
--

Dissemination Level		
P	Public	X
C	Confidential, only for members of the consortium and the Commission Services	

Revision	Date	Author, organisation	Description
0.1	01/05/2011	Anelia Belogay (Tetracom)	First complete outline of the document.
0.2	01/10/2011	Petar Gadjev	Technical Documentation v.0.9
0.3	01/11/2011	Elitsa Marinova	Simple Mode User Guide
0.4	03/11/2011	Petar Gadjev	Advanced Mode User Guide
0.5	01/11/2011	Elitsa Marinova, Diman Karagiozov, Petar Gadjev, Radostin Surilov, Kiril Mutafchiev	Technical Documentation v.1

### Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

## Introduction

The Atlas consortium has delivered three web-based services produced with the multilingual platform Atlas during the first project year:

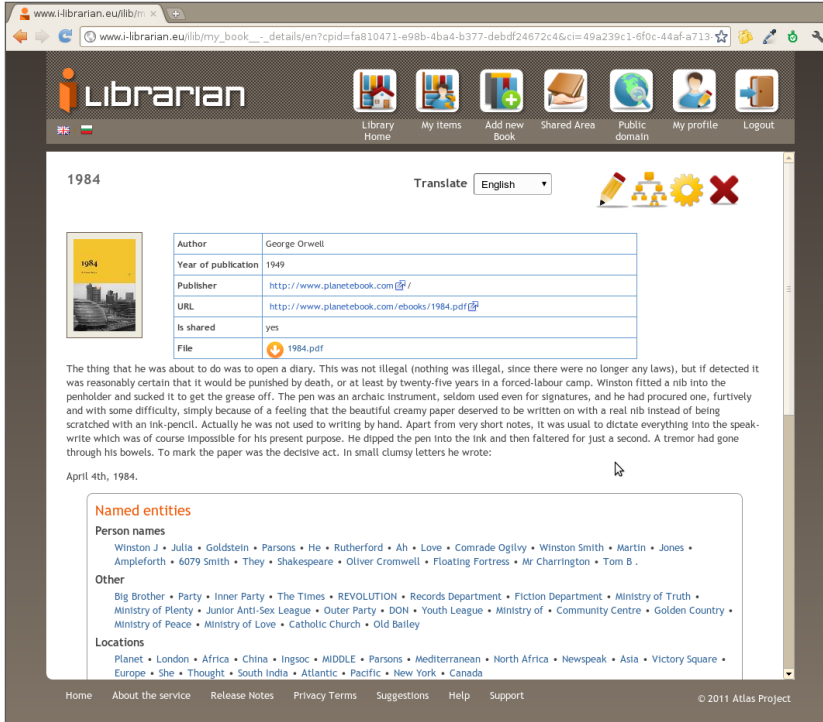
i-Librarian – [www.i-librarian.eu](http://www.i-librarian.eu)

i-Publisher – [i-publisher.atlasproject.eu](http://i-publisher.atlasproject.eu)

EUDocLib – [eudoclib.atlasproject.eu](http://eudoclib.atlasproject.eu)

This deliverable provides the software documentation of the platform and the three services at: <http://i-publisher.atlasproject.eu/atlas/documentation>

### 1. i-Librarian – the intelligent content assistant service



The screenshot shows the i-Librarian web interface. At the top, there is a navigation bar with icons for Library Home, My Items, Add new Book, Shared Area, Public domain, My profile, and Logout. Below this, the main content area displays a document entry for '1984'. The entry includes a table with the following information:

Author	George Orwell
Year of publication	1949
Publisher	<a href="http://www.planetebook.com/">http://www.planetebook.com/</a>
URL	<a href="http://www.planetebook.com/ebooks/1984.pdf">http://www.planetebook.com/ebooks/1984.pdf</a>
Is shared	yes
File	<a href="#">1984.pdf</a>

Below the table, there is a text excerpt from the document, starting with 'The thing that he was about to do was to open a diary. This was not illegal (nothing was illegal, since there were no longer any laws), but if detected it was reasonably certain that it would be punished by death, or at least by twenty-five years in a forced-labour camp. Winston fitted a nib into the penholder and sucked it to get the grease off. The pen was an archaic instrument, seldom used even for signatures, and he had procured one, furtively and with some difficulty, simply because of a feeling that the beautiful creamy paper deserved to be written on with a real nib instead of being scratched with an ink-pencil. Actually he was not used to writing by hand. Apart from very short notes, it was usual to dictate everything into the speak-write which was of course impossible for his present purpose. He dipped the pen into the ink and then faltered for just a second. A tremor had gone through his bowels. To mark the paper was the decisive act. In small clumsy letters he wrote:

April 4th, 1984.

Below the text, there is a section titled 'Named entities' which lists various categories of entities extracted from the text, including Person names, Other, and Locations.

**Named entities**

**Person names**

Winston J • Julia • Goldstein • Parsons • He • Rutherford • Ah • Love • Comrade Ogilvy • Winston Smith • Martin • Jones • Ampleforth • 6079 Smith • They • Shakespeare • Oliver Cromwell • Floating Fortress • Mr Charrington • Tom B .

**Other**

Big Brother • Party • Inner Party • The Times • REVOLUTION • Records Department • Fiction Department • Ministry of Truth • Ministry of Plenty • Junior Anti-Sex League • Outer Party • DON • Youth League • Ministry of • Community Centre • Golden Country • Ministry of Peace • Ministry of Love • Catholic Church • Old Bailey

**Locations**

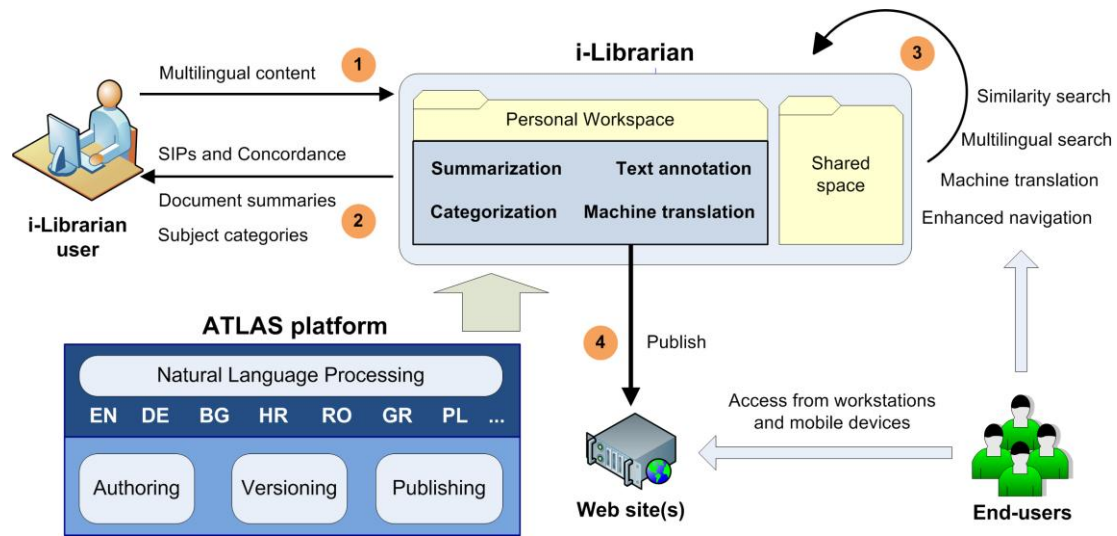
Planet • London • Africa • China • Ingsoc • MIDDLE • Parsons • Mediterranean • North Africa • Newspeak • Asia • Victory Square • Europe • She • Thought • South India • Atlantic • Pacific • New York • Canada

At the bottom of the page, there is a footer with navigation links: Home, About the service, Release Notes, Privacy Terms, Suggestions, Help, Support, and a copyright notice: © 2011 Atlas Project.

### i-librarian.eu

i-Librarian is a web-based content assistant service, which encourages visitors to register and get a personal workspace where they can store, share and publish various types of documents and have them automatically categorised into appropriate subject categories, summarized and annotated with important words, phrases and names. Advanced language-based technology is implemented to help users easily navigate between and access both their personal works and unfamiliar documents. After processing a large collection of unfamiliar texts, i-Librarian displays short summaries and extracted concepts that enable users to easily decide which documents are worth reading and which could be discarded. Furthermore, i-Librarian interlinks all user documents based on the extracted phrases, words and names, and thus improves significantly content navigation. Finally, the service helps users with no previous experience to publish their own content using the power of a modern content management solution but without struggling with the inherent complexity of such systems. The features of i-Librarian will be initially available in seven languages – English, German, Bulgarian, Croatian, Greek, Polish and Romanian. However, as more languages could be easily integrated in the service, the consortium will explore several options to

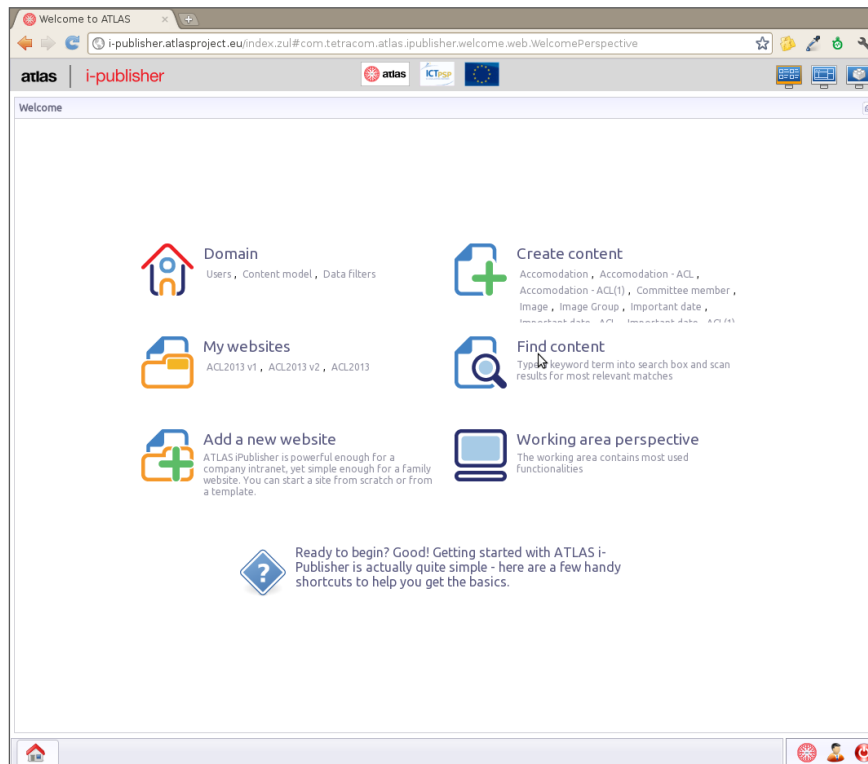
secure the necessary funding after the end of the project in order to enable support for the remaining major European languages.



Some of the main characteristics of i-Librarian are summarized below:

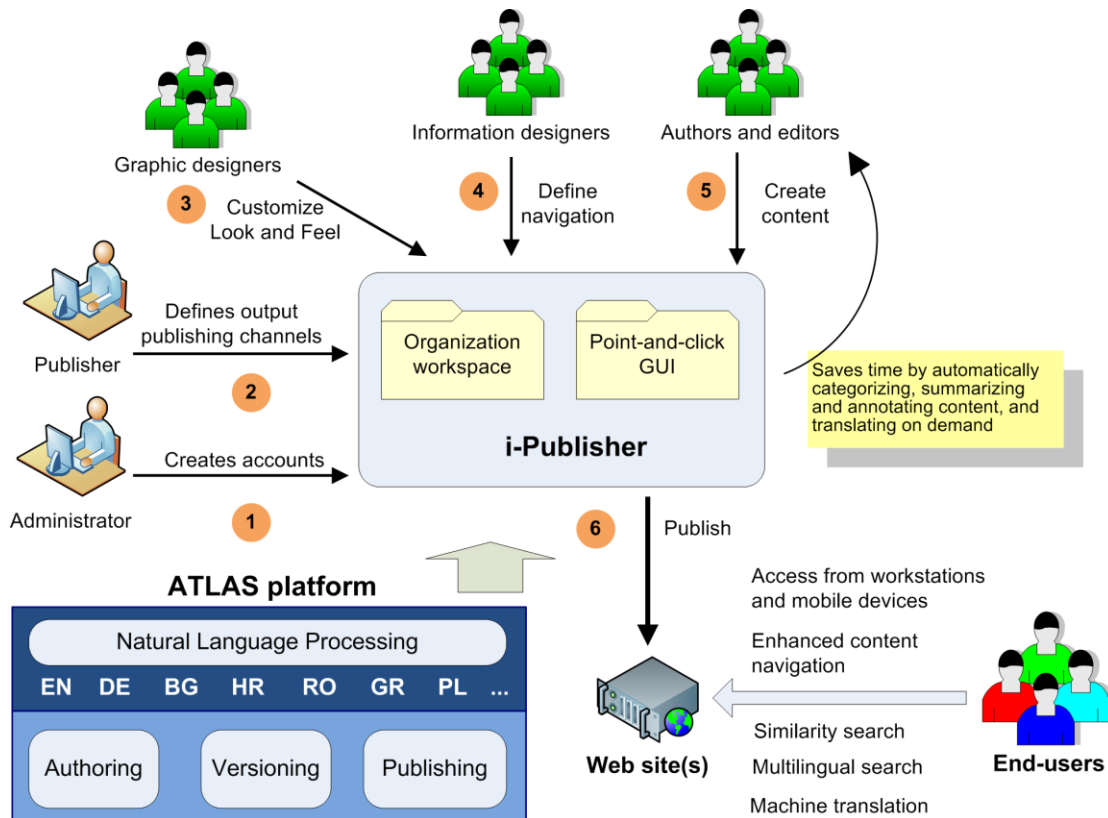
- i-Librarian offers multilingual full-text search inside personal or shared documents.
- i-Librarian provides multilingual similarity search, which enables users to easily locate both personal and shared similar documents in different languages.
- i-Librarian implements powerful instruments for computer-aided translation, automatic content categorisation, summarization, and annotation of important words, phrases and names.
- Users can rate the quality of automatic translations and improve them, which would help the consortium to build better translation models for future use.
- Users can publish heterogeneous multilingual content on a personal web site hosted by i-Librarian or on existing web sites and portals.
- Users can freely annotate documents in their personal workspace and search through the annotations (possibly in different languages) shared by other users in order to find documents of interest.
- i-Librarian is accessible from all browsers, and also features a mobile version for smartphone users.
- i-Librarian includes a mechanism for reporting and removing of materials that violate copyright laws.

## 2. i-Publisher – the online web content management solution



### [i-publisher.atlasproject.eu](http://i-publisher.atlasproject.eu)

i-Publisher is a novel software-as-a-service solution for web content management, which allows both small and large organizations to deploy and manage multilingual web sites without spending time and effort for installing and maintaining a content management system. This service assists organizations in retrieving, unifying, and packaging heterogeneous pieces of content, and dynamically rendering them on multiple web sites. i Publisher fosters collaboration in content creation by enabling authors, editors, and other contributors to work together. It also facilitates the process by automatically categorizing, summarizing, and tagging the newly created content. Furthermore, web sites may be built with i Publisher with a point-and-click graphical user interface by people with different expertise but no programming experience – publishers, information designers and graphic designers. The service leverages the full benefits of the ATLAS platform and becomes an ideal choice for promoting any type of organization on the Web. i Publisher will be available free of cost for non-commercial use in order to promote web standards and encourage language diversity in content creation. Different subscription plans will be available to those who desire more storage space and customer support, or who would like to use i Publisher for commercial purposes.



#### i-Publisher characteristics :

- i Publisher is well-suited to both small and large organizations as it is designed with scalability in mind, i.e. if an organization needs to handle more content and users, additional servers will help address these needs and provide the desired results in terms of performance.
- i Publisher improves content navigation by dynamically interlinking content items based on extracted important words, phrases and names.
- i Publisher utilizes a flexible user access rights system comparable to that of a modern server operating system – security policies may be set for groups and specific users as well as for specific content items or even content item properties.
- i Publisher implements an industrial strength versioning system, which supports the versioning of structured content rather than the simple text-based versioning found in most existing solutions.
- i Publisher allows content to be mass exported to or imported from file systems, databases or file servers.
- Web sites created with i Publisher offer to end-users multilingual full-text and similarity search as well as clustered, summarized and annotated content.

### 3. EUDocLib



#### [Eudoclib.atlasproject.eu](http://eudoclib.atlasproject.eu)

EUDocLib is a publicly accessible repository of EU law documents from the EUR-LEX collection. This web site (online service) provides enhanced navigation and easier access to relevant documents in the user's own language.

- The automatic categorisation tool classifies the documents according to the three classification trees in EUR-Lex.
- The implementation of EUDocLib is done entirely with i-Publisher as a proof of concept
- All documents of the EUR-LEX collection, available in English, that are added up to January 1, 2011 are imported into ATLAS (resp. into EUDocLib).
- The algorithms used for full-text searching and the identification of similar items are improved in terms of performance and so they work within a large pool of documents.

## Three level technical support for Atlas users

### 1. Technical documentation of Atlas and i-Publisher

The first level technical support is provided by the technical documentation of the open-source multilingual platform for content management Atlas, published on the dedicated documentation website. The documentation provides detailed technical information about the Atlas platform, and outlines the most common workflows in i-Publisher. It serves as an informative tool for users who wish to understand the intricate mechanisms of the open-source platform in the view of developing it further.

The document provides the users with:

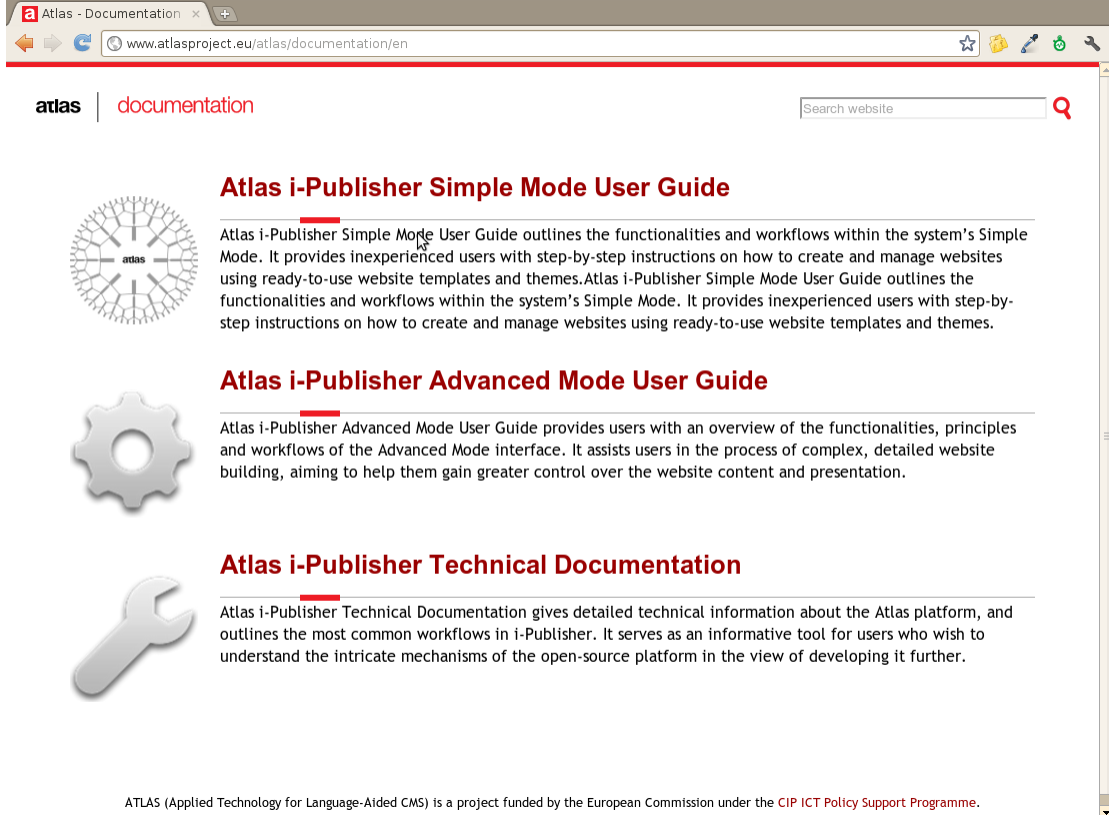
- "Introduction" - an overview of the platform architecture, comprising of a description of the system key features and layers, descriptions of the ATLAS components, the communication channels between the components and the request processing flows.
- "Content management concepts" - a detailed description of the CMS functionalities:
  - o "Domains" chapters including a description of the concepts such as the Content model, Users, Groups and Access Rights, Datasources and Selections and Approval Chain functionalities
  - o "Content items" chapters including a description of the structure and main actions performed on content items, such as content editing, approval and categorisation
  - o "Websites" chapters describe the structure of the websites built with ATLAS, the different types of widgets included in the website pages, navigations, the concepts for human readable URLs (aliases), sitemap, additional site functionalities, and caching mechanisms.
  - o "Export/Import" chapters provides an overview of the concept allowing the users to reuse important system entities such as widgets, content types and even whole websites.
- "Language processing framework" chapters describes how textmining functionalities are interwoven in the processes of content management, and focus on technical implementation aspects of the Bulgarian, German, Greek, English, Polish and Romanian language processing chains.
- "Automatic Categorisation" chapters describe the main workflows related to these functionalities as well as the current status of the implemented categorisation algorithms
- "Database structure" chapters describes the data store design decisions, providing ATLAS with scalable and performant data store.
- "Site using ATLAS" chapters provide technical details on the building of the web services i-Librarian and EUDocLib.
- "Deployment" chapters describe the installation and configuration of ATLAS components and provides an overview of the currently deployed components.

The documentation content will continue to grow and stay updated as more features are enabled. Technical details of the categorisation tool, summarization tool and machine translation infrastructure will be added after the finalization of the working packages for Automatic categorisation (WP 3), Automatic Summarisation (WP 5) and Machine translation (WP 6).



The documentation is accessible at:

<http://i-publisher.atlasproject.eu/atlas/documentation/tech/intro>



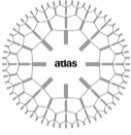
Atlas - Documentation x

www.atlasproject.eu/atlas/documentation/en


atlas | documentation

Search website


### Atlas i-Publisher Simple Mode User Guide

 Atlas i-Publisher Simple Mode User Guide outlines the functionalities and workflows within the system's Simple Mode. It provides inexperienced users with step-by-step instructions on how to create and manage websites using ready-to-use website templates and themes. Atlas i-Publisher Simple Mode User Guide outlines the functionalities and workflows within the system's Simple Mode. It provides inexperienced users with step-by-step instructions on how to create and manage websites using ready-to-use website templates and themes.

### Atlas i-Publisher Advanced Mode User Guide

 Atlas i-Publisher Advanced Mode User Guide provides users with an overview of the functionalities, principles and workflows of the Advanced Mode interface. It assists users in the process of complex, detailed website building, aiming to help them gain greater control over the website content and presentation.

### Atlas i-Publisher Technical Documentation

 Atlas i-Publisher Technical Documentation gives detailed technical information about the Atlas platform, and outlines the most common workflows in i-Publisher. It serves as an informative tool for users who wish to understand the intricate mechanisms of the open-source platform in the view of developing it further.

ATLAS (Applied Technology for Language-Aided CMS) is a project funded by the European Commission under the CIP ICT Policy Support Programme.

## 2. Hint system in i-Publisher

The second level technical support is provided by the elaborate Hint system ingrained in the i-Publisher service. Every dialogue (domain-specific, content-specific or website-specific) within the system contains a detailed description of the functionalities provided to the user on the page. The Hint system acts as a reference guide, providing comprehensive information on every step that is available to the user on the current page; but in order for it to be most effective, some knowledge of the content building process is required. This knowledge can be gained from the Advanced Mode user guide, the Technical Documentation and the Video Tutorials, as well as from experience with other content management systems.

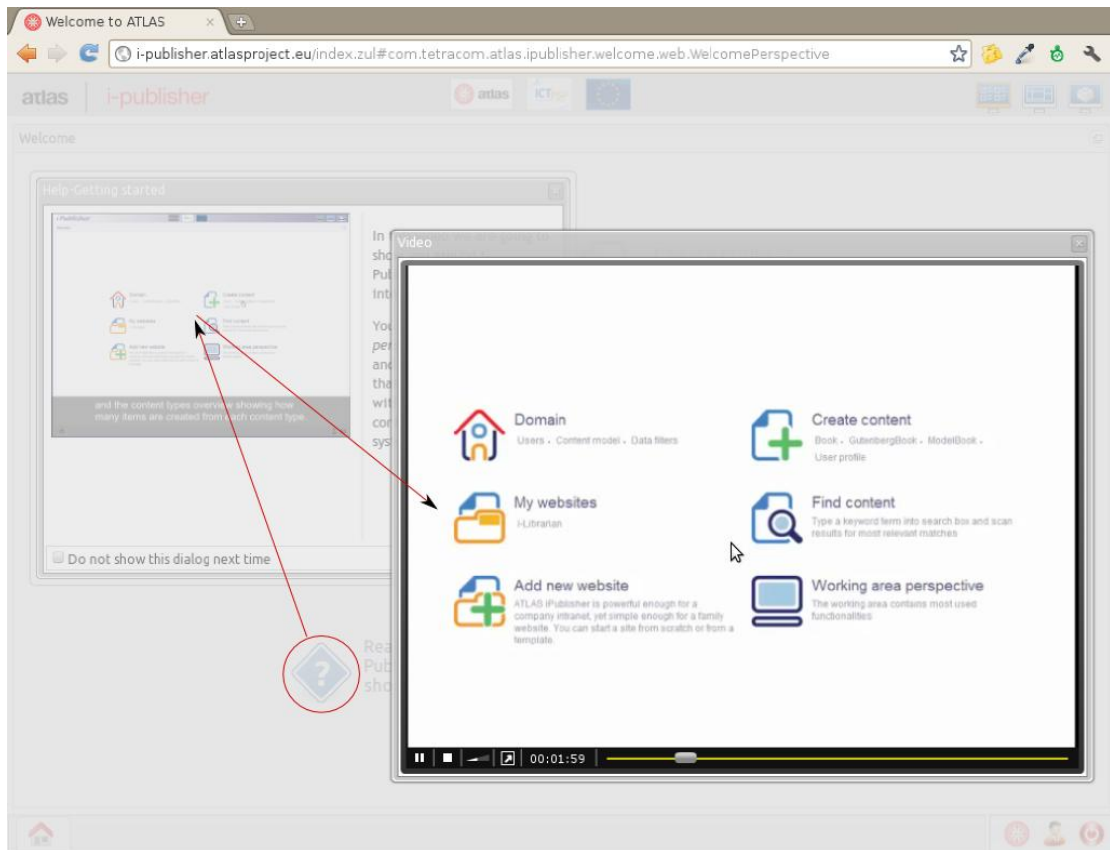
The screenshot displays the i-Publisher web interface for configuring a website. The browser address bar shows the URL: `i-publisher.atlasproject.eu/index.zul#40df8f0b-4e24-4e8f-b120-e5409774e30a`. The interface includes a navigation bar with 'Domain', 'Create content', 'My websites', 'Search', and 'Working area perspective'. The main content area is titled 'Documentation' and contains a 'Properties' sidebar on the left with options like 'Pages', 'Aliases', 'Sitemap', 'Navigations', 'Access log', 'Cache', 'Functionalities', 'Publish', and 'Go offline'. The main configuration area is divided into 'General' and 'Preview' sections. The 'General' section includes fields for Name, Alias, Address, Email, Languages, Theme, and Page renderer. The 'Preview' section shows a website header with 'atlas | documentation' and a search bar. A 'Website Properties' panel on the right contains a list of instructions for configuring the website.

**Website Properties:**

- The website dialog is organised into tabs, accessed from the 'Shortcuts' section on the left. There are ten website features (properties, pages, etc.), as well as three actions that can be used once the website is completed (publish online, go offline and export).
- In this tab you define the general properties of your website, and once you finish you can see a preview of your homepage (provided you have created one) in the 'Preview' field.
- Firstly, give your website a name. Website's atlas is formed by its name. Choose the languages you wish your website to feature by clicking on the icon in the Languages field and then dragging and dropping the desired languages in the field on the left. Define your default language by clicking in the radio button next to it and then click on apply to save your choice. Select a website color theme from the dropdown and choose a page renderer. After you are satisfied with your choices click on 'Save'. When the page is reloaded you will be able to see your website public address(URL).

### 3. Video tutorials

The third level technical support is provided by the video tutorial. The general workflows of how to make a website, a page, how to create a navigation, add widgets are visualized in an informative and easy-to-understand way using short video clips. They are suitable for beginners as a basic help tool.



## User guides of i-Publisher

The users willing to use the web-based instrument for building multilingual websites i-Publisher are advised to gain some basic knowledge on the website building process by reading the comprehensive step-by-step user guides. Pictures, diagrams and short explanations should be enough for an inexperienced user to build a website via the Simple Mode using predefined website moulds, or to build a completely new website with some help from the ready-to-use themes. The advanced users, however, should consult the User Guide together with the three-level technical support in order to build a comprehensive multilingual website.

### Atlas i-Publisher Simple Mode User Guide

The Simple Mode is a simplified layer of i-Publisher, providing inexperienced users with ready-to-use thematic websites and themes. The guide provides i-Publisher users with a general overview of the Simple Mode interface and describes how to create and manage a website and its content. There is also a brief description of some of the system's more advanced functionalities, such as tools that enrich text documents with information on important phrases, names, dates within the text. The Simple Mode User Guide can be used by non experienced users in order to gain a more comprehensive perception of the process of website building with i-Publisher's Simple Mode.

The user guide is accessible at:

[http://www.atlasproject.eu/atlas/documentation/simple/getting\\_to\\_know\\_the\\_atlas\\_simple\\_mode\\_interface](http://www.atlasproject.eu/atlas/documentation/simple/getting_to_know_the_atlas_simple_mode_interface)

The screenshot shows a web browser window displaying the Atlas i-Publisher Simple Mode User Guide. The browser's address bar shows the URL: [www.atlasproject.eu/atlas/documentation/simple/getting\\_to\\_know\\_the\\_atlas\\_simple\\_mode\\_interface](http://www.atlasproject.eu/atlas/documentation/simple/getting_to_know_the_atlas_simple_mode_interface). The page content includes a search bar, a navigation menu on the left, and the main text of the guide. The main text is titled "Getting to know the Atlas Simple Mode Interface" and includes sections for "Wizards" and "Information screen".

**Getting to know the Atlas Simple Mode Interface**

In the next couple of screenshots we will show you how to operate the Atlas simple mode interface.

**Wizards**

The process of creating and managing websites is organized in different work flows. For each work flow there is a different wizard to guide you through the system.

**Information screen**

The i-Publisher's simple mode interface has one main menu, located to the left.

The picture you see below is i-Publisher's simple mode information screen. This is the welcome screen when you log in for the first time. Here you can see a brief review of the main simple mode functionalities. The information screen is accessible through a shortcut in the main menu.

**Navigation menu (left side):**

- Getting to know the Atlas Simple Mode Interface
- Add new website
- Edit your website
- Add new page to an existing website
- Create and edit dynamic content
- Simple page components' wizards
- Navigation blocks
- Components for displaying dynamic content
- Text analysis components
- Translations
- Chapters

**Main content area (Information screen):**

embrace innovation  
i-publisher

i-Publisher is as intuitive as you would like a tool for building websites to be!

- Because we are not only i-Publisher, we'll be there for you 24/7.
- It's easy to manage your website, ready-to-use website designs, themes ready-to-use, localization and translations that save you time.
- It's easy to build your website by choosing a theme, a color scheme and ready-to-use content.
- You can easily manage all your websites, add blogs, photos, lists of items or customize your content.

Select the website you want to work on:

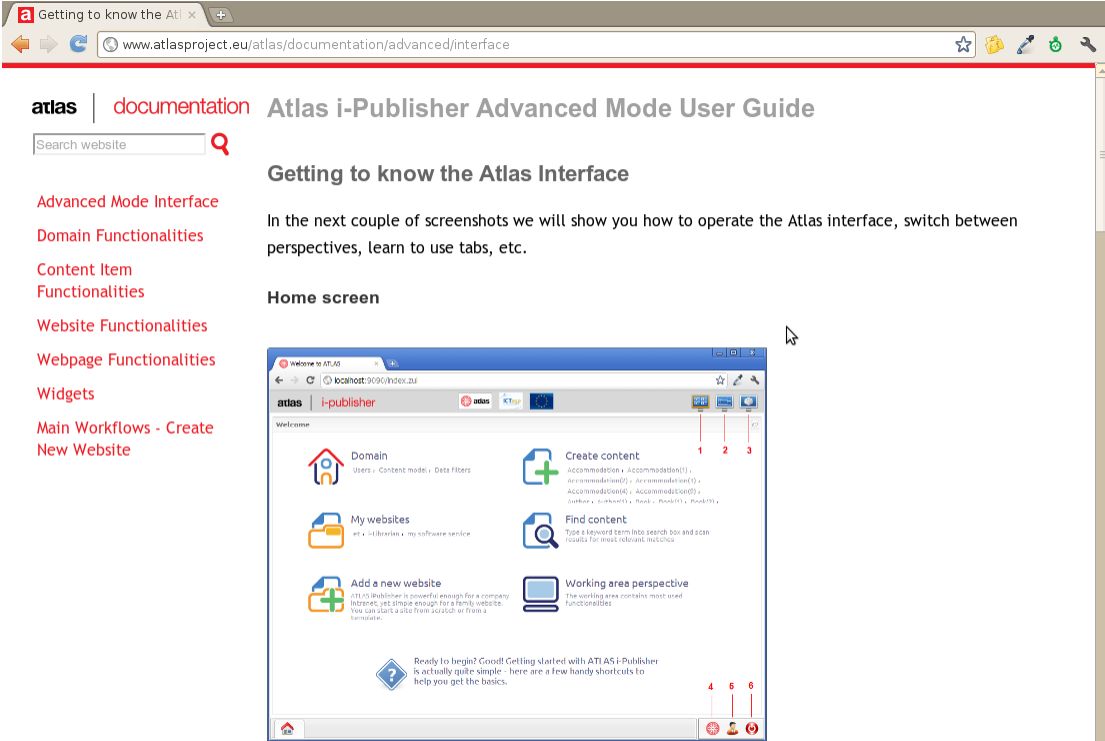
For more information visit [www.atlasproject.eu](http://www.atlasproject.eu)

## Atlas i-Publisher Advanced Mode User Guide

i-Publisher's Advanced Mode guide provides system users with a general overview of the Advanced Mode Interface and functionalities, acting as a complement to the Hints and Video Tutorials. The workflows of managing users, groups and their access rights within a domain, building a content model, managing content, managing a web site and web page and widgets are outlined with graphic representations of the processes, together with an apprehension of the different uses of Advanced Mode's functionalities. Finally, the step-by-step guide on how to make a website supports the user in their attempt to build a complex, multilingual web site using the advanced linguistic technologies that provide automatic classification of content, summarization of content as well as machine translation.

The user guide is accessible at:

<http://i-publisher.atlasproject.eu/atlas/documentation/advanced/interface>



**Getting to know the Atlas Interface**

In the next couple of screenshots we will show you how to operate the Atlas interface, switch between perspectives, learn to use tabs, etc.

**Home screen**

The picture you see above is i-Publisher's home screen. Marked as 1,2 and 3 are the three different perspectives that you can select for your home screen. These perspectives separate i-Publisher in three main