



INFRA-2011-1-284432



## COLLABORATIVE EUROPEAN DIGITAL ARCHIVE INFRASTRUCTURE

**Project Acronym:** CENDARI

**Project Grant No.:** 284432

**Theme:** FP7-INFRASTRUCTURES-2011-1

**Project Start Date:** 01 February 2012

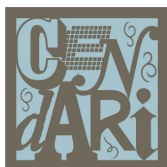
**Project End Date:** 31 January 2016

<b>Deliverable No. :</b>	9.2
<b>Title of Deliverable:</b>	Full Release Portal/VRE
<b>Date of Deliverable:</b>	October 2015
<b>Revision No.:</b>	1
<b>WP No.:</b>	9
<b>Lead Beneficiary:</b>	INRIA
<b>Author (Name and email address):</b>	Jean-Daniel Fekete <a href="mailto:Jean-Daniel.Fekete@inria.fr">Jean-Daniel.Fekete@inria.fr</a> Nadia Boukhelifa <a href="mailto:Nadia.Boukhelifa@inria.fr">Nadia.Boukhelifa@inria.fr</a> Emmanouil Giannidakis <a href="mailto:Emmanouil.Giannidakis@inria.fr">Emmanouil.Giannidakis@inria.fr</a> Carsten Thiel <a href="mailto:thiel@sub.uni-goettingen.de">thiel@sub.uni-goettingen.de</a> Hanna Meiners <a href="mailto:meiners@sub.uni-goettingen.de">meiners@sub.uni-goettingen.de</a> Natasa Bulatovic <a href="mailto:bulatovic@mpdl.mpg.de">bulatovic@mpdl.mpg.de</a> David Stuart <a href="mailto:david.stuart@kcl.ac.uk">david.stuart@kcl.ac.uk</a>
<b>Dissemination Level:</b>	PP = restricted to other programme participants
<b>Nature of Deliverable:</b>	O = other
<b>Abstract (150 words):</b>	Full Release Portal/VRE: This release brings together all elements of the CENDARI development, adding full connectivity through the integration layer developed by WP



INFRA-2011-1-284432

	<p>7.</p> <p>This report provides an overview of the portal used to access all the services and data of the CENDARI research environment. The portal is based on “Liferay” (<a href="http://www.liferay.com/">http://www.liferay.com/</a>), an open-source portal that integrates a large number of social-networking services. The CENDARI portal is available at: <a href="https://portal.cendari.dariah.eu/home">https://portal.cendari.dariah.eu/home</a></p>
--	---



INFRA-2011-1-284432

## Table of Contents

### Contents

Introduction .....	4
Integrating Services into the CENDARI Virtual Research Environment .....	5
Organization of the Portal.....	7
Research .....	7
Standardize .....	17
Lab.....	18
Collaborate .....	18
Extensions of the Portal.....	19



## Introduction

This report provides an overview of the portal used to access all the services and data of the CENDARI project.

As stated in the CENDARI Description of Works, the objectives of the W9 work package are:

1. To provide the user interface and human-oriented services to weave together all the technical resources of CENDARI to allow historians to conduct their research with transnational digital support.
2. To develop a portal that allows integrated discovery of and access to resources held in the participating archives and other web-accessible resources, in coherence with the architecture and functional specifications defined in WP9.
3. To develop functionalities that allow integrated discovery of and access to resources held in the participating archives and other web-accessible resources.
4. To develop an enquiry environment that (i) provides shared access to tools for integrating and processing data; (ii) provides a collaborative space for sharing and discussion; (iii) supports the research lifecycle of the targeted communities.
5. To develop visualisation tools and services to ensure that available information can be maximally exploited by researchers according to his/her research interests.

The portal is therefore an important component of the whole user experience. However, technically, it boils down to being a collector and access point to other technologies developed or used by the CENDARI project. As such, its deployment is essentially a technical issue, and involves less complications than the rest of the CENDARI tools. This is why the CENDARI project has decided to focus more on finalizing the tools themselves before the portal. The tools being under experimentations, the portal can then be deployed.

The portal is based on “Liferay” (<http://www.liferay.com/>), an open-source portal that integrates a large number of social-networking services. Liferay has been chosen because several CENDARI partners were familiar with it; in particular, multiple Dariah projects are hosted on a Liferay platform, making it easier to install, maintain, and possibly adapt for CENDARI. The CENDARI portal is available at: <https://portal.cendari.dariah.eu/home>



## Integrating Services into the CENDARI Virtual Research Environment

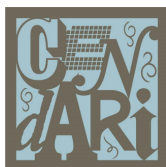
As a web portal, Liferay provides standard services such as a calendar, forums, a wiki, and page management. In addition, it can link to other web-based applications with three levels of integrations:

**Level 1: Link-based Integration:** The simplest mechanism relies on Web links pointing to other applications listed on a Portal page. This mechanism has the advantage of simplicity but when accessing the linked application, users see a brutal change in their user interface, and cannot access the services of the portal unless they leave the portal open in another window.

**Level 2: Frame-based Integration:** another level of integration consists in inserting a CENDARI service inside a portal page using an HTML Internal Frame (iframe). We have experimented with this option with the CENDARI Note-Taking Environment (NTE) and found it unsatisfactory because the NTE uses the whole screen to display multiple components. Integrating it inside a sub-window of the portal reduces its available screen real-estate for no clear benefit.

**Level 3: Content-based Integration:** Liferay provides a rich Application Programming Interface (API), in particular to create web components for Liferay. The main protocol/mechanism for creating these components is called a “Portlet” and we have tried to see if some of our tools could be integrated that way. After multiple tries, we decided not to use Portlets because they did not provide the right level of integration. It is meant to fit tools in the Liferay ecosystem, but not to fit Liferay into an ecosystem of tools. We would have to specialise the CENDARI tools too deeply towards Liferay and decided not to.

In the end, we have decided to simply link our tools from portal pages using static images of the interfaces. This integration is very light and flexible, allowing using clickable static images to present the reachable applications to users; this impact of a brutal change will be reduced through the adoption of a consistent user-interface on linked applications. This flexibility is also important because several services are actually provided by the NTE, such as the faceted browsing, the note-taking application itself, and the Archival Research Guides. The portal provides a nice illustration for each of them, but they actually all link to the NTE with different internal web pages.



INFRA-2011-1-284432

**CENDARI**  
COLLABORATIVE EUROPEAN DIGITAL  
ARCHIVE INFRASTRUCTURE

Research Standardize Lab Collaborate Guides What's new Who we are Sign In

### Archival Directory

- Explore archival resources
- Register new archive
- Enrich the archival directory

### Notes

- Create your research project
- Take Notes
- Tag persons, places, events
- Search through CENDARI resources

### Search

- Search CENDARI data
- Navigate to sources

### Medieval sources

- Find medieval resources
- Go directly to the shelfmark
- Register new databases

### Archival Research Guides

- Define your topic
- Link and describe resources
- Publish your Guide

### Ask Questions

Diagram illustrating the workflow: Cendari data offerings → Linking services → Expert knowledge offerings → Linked Open Data. A central box labeled 'Data extraction & Linking module' connects the first three stages.

Figure 1: The CENDARI Portal

The portal integrates multiple tools and offers multiple services that we list in the next section.



## Organisation of the Portal

The portal organises the CENDARI tools and services into three main groups according to their primary audience: tools and services for the historian conducting research (*'Research'* on portal menu); metadata schemas and ontologies for the information professional wishing to build upon the CENDARI knowledge framework (*'Standardise'* on the portal menu); and resources for developers wishing to build upon the code developed by CENDARI (*'Lab'* on the portal menu). The *Research* page acts as the homepage of the portal as *Standardise* and *Lab* consist of more specialised or experimental tools aimed at a smaller subset of users.

In addition, the portal menu includes:

- *Collaborate* - a dropdown menu of Liferay modules that have been implemented to encourage collaboration between researchers.
- *Docs & Guides* - the documentation and a set of guides on how to use CENDARI tools for researchers and users, developers and administrators
- *What's new* - the news feed for the CENDARI blog.
- *Who we are* - information about the CENDARI project.
- *Sign in* - a login window for registered users.

*Research*, *Standardise*, *Lab* and *Collaborate* and their associated integrated tools are discussed in more detail below.

## Research

*Research* currently consists of links to six tools.

- "Archival Directory": pointing to the list of archives in Europe
- "Notes": pointing to the Notes Taking Environment
- "Search": faceted search over all the documents indexed by the CENDARI project.
- "Medieval Sources": allowing search over the medieval documents via TRAME.
- "Archival Research Guides": pointing to the ARGs project in the NTE



INFRA-2011-1-284432

- “Ask Questions”: pointing to the list of parameterised queries using semantic-web technologies on the CENDARI platform.

**Archival Directory:** The CENDARI Archival Directory is a database of archival descriptions and collections, with a strong transnational focus. It aims to include archives, libraries and other institutions containing resources on the Medieval era and World War One. Special focus is given to “hidden” institutions i.e. those who are not well known by historians or do not have sufficient online presence. Directory is implemented by the AtoM tool, a web-based archival description software that is based on the International Council on Archives (ICA, see <https://www.accesstomemory.org>) standards. AtoM is the abbreviation of "Access to Memory". This software was chosen because it adheres to international standards and comes with a user friendly tool to produce archival descriptions (see Fig. 2). Informational pages of the Archival Directory are available in several languages: English, German, French, Italian, Russian, Polish and Dutch.





The Archival Directory is a large database of archival descriptions and collections and is part of the CENDARI Virtual Research Environment.

It has a strong transnational focus and one of its aims is to include many archives and institutions which are little known or rarely used by researchers. The Archival Directory allows historians to view sources in a rarely seen transnational and comparative view. It is focused on archives and libraries containing resources on the **Medieval era** and **World War One**.

[Searching for records in the CENDARI Archival Directory](#)

[Frequently asked questions](#)

[About CENDARI Archival Directory](#)



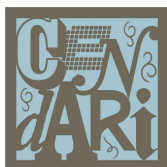
Join Us!

To access and contribute content to the Archive Directory tool, join our [Trusted Users Group](#) (opens new window).

We hope that researchers, historians and archivists will find this opportunity to be of interest. Potential members of the TUG are asked to submit a short statement of interest via [TUG registration form](#) (opens new window) or if you prefer, you may [send a CV instead](#).

Figure 2: The Archival Directory

**Notes:** the **Note Taking Environment** is the main component of the VRE. It is a collaborative online platform that unifies a set of functionalities to support historical research workflow from data collection and analysis to synthesis and publishing. In particular, the NTE allows for *organising, authoring, sense-making, searching* and *sharing* of resources related to historical research. **Organising & Authoring:** Similarly to how historians organise their material on their computers, the NTE allows for creating projects within which resources are organised as notes, documents and transcripts. These resources can be created within the environment using a rich set of text editing tools (using TinyMCE editor), or by uploading existing material such as for high resolution images. **Sense-making:** An important feature of the NTE is the semantic enrichment of resources, either manually or automatically, via a tagging mechanism where strings of text are recognised as named entities (person, place, event etc.) and resolved by associating an entity with a reference URI (e.g. to a DBpedia repository entry). By default, user created files (i.e. notes, documents, transcripts and uploaded images) are private, while the semantic information in the



INFRA-2011-1-284432

form of RDF data contained inside these resources (i.e. the entities) is public and accessible to the CENDARI project and others. This semantic information is also available via dashboard type visualisations within the platform. The NTE supports frequency visualisations for person, place and event entities and an interactive map. Such visualisations provide context for projects and facilitates sense-making of the collected data, for example in terms of time spans, geographical distributions and frequency of events. **Searching & Publishing:** all material created within the NTE is accessible via a faceted search facility and navigation.

**Search:** the Faceted search interface of CENDARI allows searching in all the resources maintained by the CENDARI project, whether created inside the project, or gathered by the project. The interface (Fig. 3) allows full-text search and shows many facets to facilitate the search and provide context to users. Users can use the facets to either get information about the results such as the number of events associated with the query, of geographical locations. The user can also use the facets to further filter the search results, selecting particular values, date ranges, or geographical bounds. Currently, the faceted search allows accessing all the documents entered in the NTE (notes, documents, entities), and all the documents indexed/ingested by the Data API. It also provides several datasets for the convenience of the researchers: all DBpedia (3.6 millions documents), and soon several encyclopaedias of World War One. The faceted search can also forward search query to TRAME, but unfortunately, the results are not faceted yet and appear in another window below the standard search panel.



INFRA-2011-1-284432

Include TRAME ☒  
Limit search to current project ☐

### Facets (reset)

- + Application (2)
- + Artifact (1)
- + Contributor (0)
- + Creator (1)
- + Date ()
  - ? Filter Reset
  - 1676-07-14 2011-12-28
- + Event (98)
- + Format (1)
- + Language (1)
- + Location (7)
  - ? Filter Adapt Reset

92 Results (0/1) [Prev](#) [Next](#)

<https://notes.cendari.dariah.eu/cendari/pavlinabobic/topics/1451/>  
Hindenburg

[http://dbpedia.org/resource/The\\_Hindenburg\\_film](http://dbpedia.org/resource/The_Hindenburg_film)  
The **Hindenburg** is a 1975 American Technicolor film based on the disaster of the German airship **Hindenburg**. The film stars George C. Scott. It was produced and directed by Robert Wise, and was written by Nelson Gidding, Richard Levinson and William Link, based on the 1972 book of the same name, *The Hindenburg*, by Michael M. Mooney.A. A.

[http://dbpedia.org/resource/Hindenburg,\\_Saxony-Anhalt](http://dbpedia.org/resource/Hindenburg,_Saxony-Anhalt)  
**Hindenburg** is a village and a former municipality in the district of Stendal, in Saxony-Anhalt, Germany. Since 1 January 2009, it is part of the municipality Hohenberg-Krusemark.

[http://dbpedia.org/resource/Paul\\_von\\_Hindenburg](http://dbpedia.org/resource/Paul_von_Hindenburg)  
Paul Ludwig Hans Anton von Beneckendorff und von **Hindenburg** (About this sound listen ), known universally as Paul von **Hindenburg** (German: [ˈpaʊl fɒn]) President of Germany from 1925 to 1934. **Hindenburg** enjoyed a long career in the Prussian Army, retiring in 1911.

[http://dbpedia.org/resource/SV\\_Hindenburg\\_Allenstein](http://dbpedia.org/resource/SV_Hindenburg_Allenstein)  
SV **Hindenburg** Allenstein was a German football club from the city of Allenstein, East Prussia (now Olsztyn, Poland).The club was formed in 1921 as Sportvereinigung **Hindenburg** Allenstein and was named for German Field Marshal and Reichs President Paul von **Hindenburg**. Sometime in 1935 it became a military side and played as Standort-Sportvereinigung **Hindenburg** Allenstein.

[http://dbpedia.org/resource/Carl\\_Hindenburg](http://dbpedia.org/resource/Carl_Hindenburg)  
Carl Friedrich **Hindenburg** (13 July 1741– 17 March 1808) was a German mathematician born in Dresden. His work centered mostly on combinatorics and

[http://dbpedia.org/resource/Hindenburg\\_Line](http://dbpedia.org/resource/Hindenburg_Line)  
The **Hindenburg** Line (Siegfriedstellung) was a German defensive position of World War I, built during the winter of 1916–1917 on the Western Front

[http://dbpedia.org/resource/Hindenburg-class\\_airship](http://dbpedia.org/resource/Hindenburg-class_airship)  
The two **Hindenburg**-class airships were passenger carrying rigid airships built in Germany in the 1930s and named in honor of Paul von **Hindenburg** the **Hindenburg** class were widely considered the future of air travel, and the lead ship of the class, LZ 129 **Hindenburg**, established a regular transatlantic service.

[http://dbpedia.org/resource/Preussen\\_Hindenburg](http://dbpedia.org/resource/Preussen_Hindenburg)  
Preußen **Hindenburg** was a German association football club from the city of Zaborze, Upper Silesia in Germany (today Zabrze, Poland).

[http://dbpedia.org/resource/Hindenburg\\_Bridge](http://dbpedia.org/resource/Hindenburg_Bridge)  
[] The **Hindenburg** Bridge was a railway bridge over the Rhine between Rudesheim in the German state of Hesse and Bingen-Kempen state of Rhineland -Palatinate, named in 1918 after Field Marshal Paul von **Hindenburg**, later German President. The bridge was put in service in 1915, destroyed in the Second World War and never rebuilt. Since 2002 the remains of the **Hindenburg** bridge has been the easternmost point of the UNESCO World Heritage Site of the Upper Middle Rhine Valley.

<https://notes.cendari.dariah.eu/cendari/rolandszczok/notes/384/>  
Coercion and Consent Picture to be added© and the copyright or source noted here Author Pavlina Bobic Theme(s) World War I,

Figure 3: The Faceted Search Browser

**Medieval Sources:** TRAME (Texts and Manuscript Transmission of the Middle Ages in Europe) is a web-based metasearch engine which includes a variety of medieval sources, which can be searched by end user. TRAME application provides a layer of interoperability among different digital resources in the Medieval Culture domain.



INFRA-2011-1-284432

**TRAME**  
Text and manuscript transmission of the Middle Ages in Europe

Project Partners Search engine Join us News Contacts FEF

**Sites**

**Site selection**  
☒ Within the TRAME network ☐ All

**External sites**  
☒ Libraries ☒ Projects ☒ Portals

**select all** **clear selection**

- ☒ BAI - Mirabile network
- ☒ BISLAM - Mirabile network
- ☒ CALMA - Mirabile network
- ☒ CANTICUM - Mirabile network
- ☒ CODEX Database: Regione Toscana
- ☒ IRHT – Jonas
- ☒ LIO - Mirabile network
- ☒ MAFRA - Mirabile network
- ☒ MAGIS - Mirabile network
- ☒ Manus - ICCU
- ☒ MATER - Mirabile network
- ☒ MDI - Manoscritti Datati d'Italia
- ☒ MEL - Mirabile network

**Figure 4: The TRAME Search Environment**

Within CENDARI project the list of TRAME resources has been extended with resources from the World War One period for testing purposes. This has demonstrated its general-purpose functionality and applicability to domains other than Medieval.

**Archival Research Guides (ARGs):** Besides entities which are publically available, a special type of notes called *ARGs* are also accessible to the public from within the



INFRA-2011-1-284432

NTE. An Archival Research Guide (ARG) is an introductory guide to archival material, structured around either one or many archival collections. Their main purpose is to offer a research framework for researchers, historians, scholars etc. who are starting a new research project. In contrast to the traditional archival guides which usually focus on one collection, institution or country, CENDARI uses a transnational approach. Depending on the author, each ARG can be divided into multiple notes which are linked to each other and they can differ in structure. ARGs support multimedia content such as images (and in the future video and sound) and a certain level of interactivity. In order for the NTE to support the ARGs, rich editing functionalities were added so that the ARG authors could format their document depending on their needs. In addition, the NTE now supports embedded images and a "Read Mode" view. By offering the ARGs through the NTE authors can tag/resolve entities, and make use of the collaborative environment and the visualisation tools.

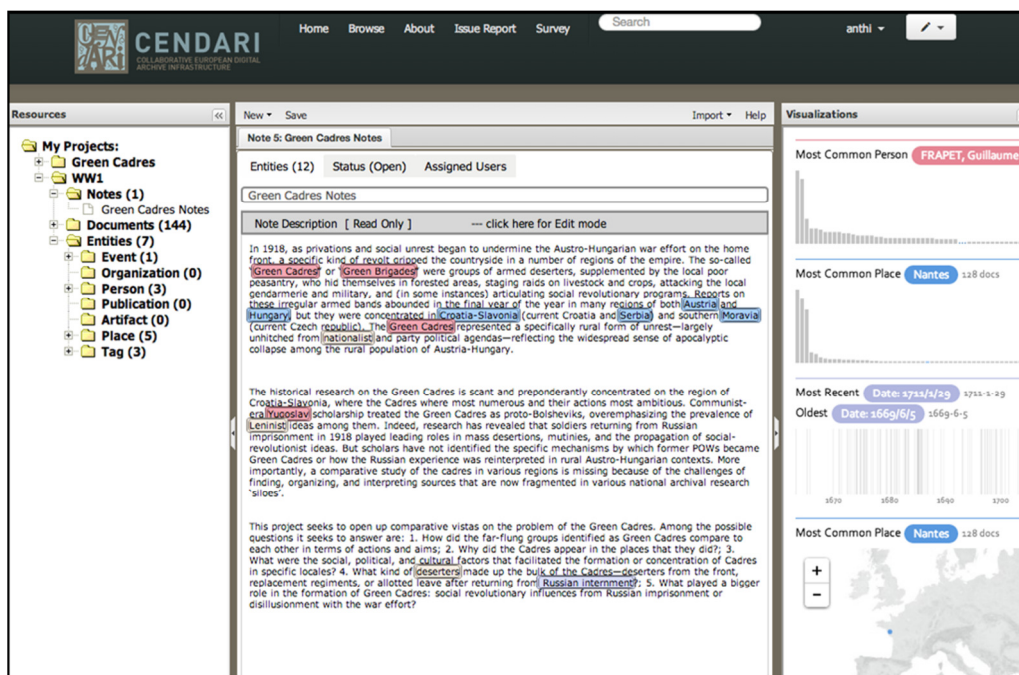
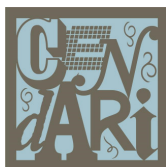


Figure 5: The User Interface of the CENDARI Note-Taking Environment (NTE)





INFRA-2011-1-284432

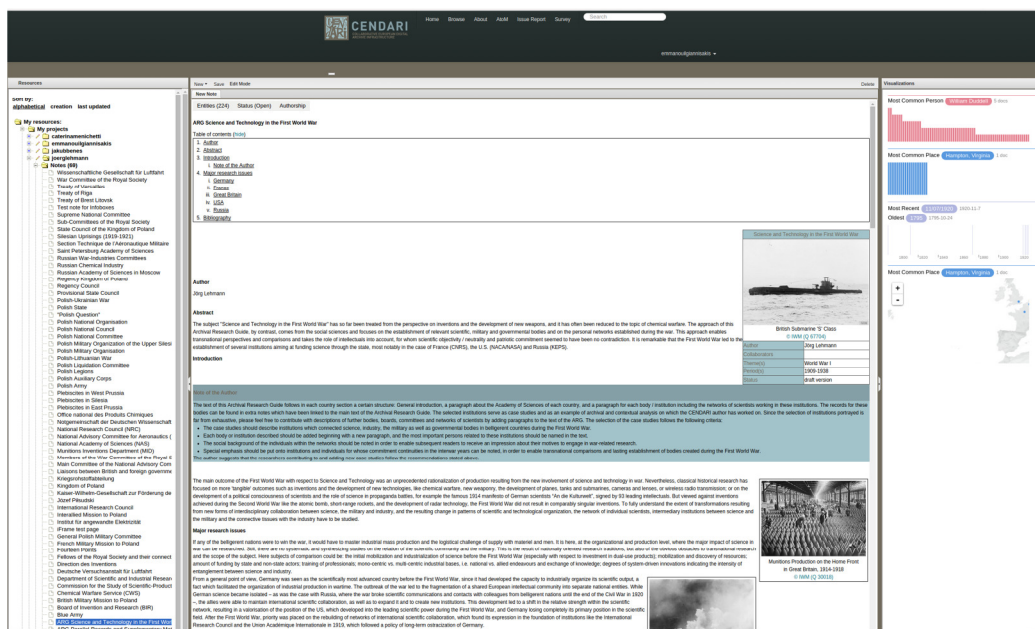


Figure 6: An Archival Research Guide (ARG) created within the NTE

Although the ARGs can be achieved through the NTE, they are a valuable project on their own providing advice for historians and archivists who need to search in multiple archives spread all over Europe and sometimes the world.

**Ask Questions:** Pineapple is a semantic web service, allowing to access and search for various resources from the CENDARI Knowledge Base. While a document might be searchable via the NTE and faceted search, Pineapple understands the document as a “resource” and provides additional information about this resource such as: own properties (i.e. triples extracted from the document), which entities (Persons, Places, Events) are mentioned in a document, which other documents are mentioning this particular resource. This information is populated into the CENDARI Knowledge Base via text-mining and Indexing components (provided by WP7) such as: EAG2RDF Indexer, EAD2RDF Indexer, Named Entity Recognition Service, and NER2RDF Indexer.



INFRA-2011-1-284432



## South African munition workers

**Source** dc-b2bf-47f7-a01f-040035cabdd9  
**Last modified** 19-12-2014 15:46  
**Plain Text**

```
2005
Sat Jan 01 01:00:00 CET 2005
Sat Dec 31 01:00:00 CET 2005
Early 21th century
Начало 21-го века
Mon Jan 01 01:00:00 CET 2001
Sat Dec 31 01:00:00 CET 2033
21..
21??
21e
21er siècle
21-st
21st
21st century
21st century
```

## Mentions

World War  
United Kingdom  
Информация  
Начало  
ShareAlike South Africans Ammunition Automobiles Officers  
Документиране  
2e millénaire après J.  
Attribution  
Фотография

**Figure 7: Results generated by Pineapple**

**Documentation and User Guides:** The Documentation consists of two parts, whereas one part serves as a documentation space for the source code of the





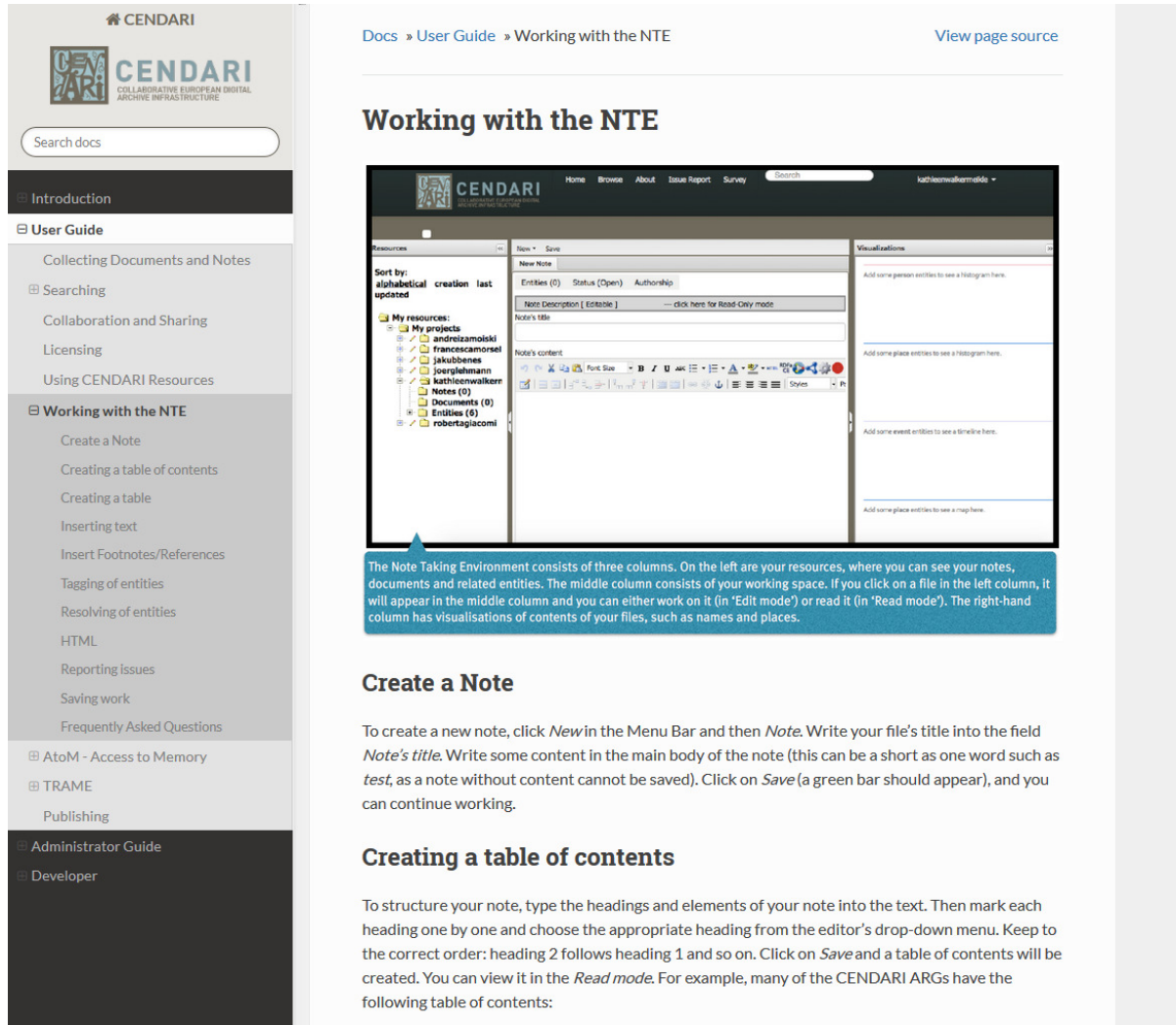
INFRA-2011-1-284432

individual CENDARI components. In this area developers are able to document their work and their code so other technicians may be better able to understand what was done and why.

The other part of the Documentation consists of the User Guides and serves as a manual for the end users of the CENDARI components. In this part there will be detailed descriptions of how to use the single tools, which possibilities were offered in the NTE and how to get to the different tools. Also, there will be FAQs for every single tool which arose from the testing workshops with the Trusted Users Group (TUG). For these workshops, training material was prepared for every single tool and component which was reused for the documentation.

The User Guides shall be accessible via a link on the portal itself.





### Figure 8: The Documentation of the NTE

## Standardise

The 'Standardise' tab will bring together the outcomes of the Knowledge Organisation Framework (WP6) section of CENDARI in one place, including both the work on the ontologies and the common metadata schema. This includes: visualisations of the ontology elements sets in WebVowl; access to the associated



INFRA-2011-1-284432

ontology instances; ontology documentation; and guidelines and schemas for the application of the metadata standard.

## Lab

The CENDARI Lab is envisioned as a practical and a technical corner for all users who would like to learn about and try new methods for data discovery, gathering, processing or visualisation within CENDARI or related developments.

In particular, the Lab allows accessing the source code of all the CENDARI components, for inspection and reuse in other projects.

It is the location where the tacit knowledge about the use of tools, tools development, data acquisition, data transformations, successes, failures and challenges is gathered and shared with the public.

It lists services, scripts and “how-to’s” on particular tasks or problems, and optionally, links to experimental services and applications being developed or deployed for the CENDARI project.

It is our hope that the Lab will become a referenceable exchange and information platform for tools, methods and processes supporting “digital humanities” techniques to help historian researchers in their work.

## Collaborate

In addition to tools linked to from the portal, the portal deploys modules to encourage researcher collaboration. Initially this consists of:

- Discussions
- Calendar
- Blogs

Collaborative modules will be experimented with users and adapted as necessary.





INFRA-2011-1-284432

### **Extensions of the Portal**

We will use the portal as it is until the end of November 2015 to collect feedback from the users. We will then revise the structure of the portal according to the feedback, and possibly add or integrate new services offered by Liferay if our users ask for it.

We hope that historians will use the available tools to change their practice that remains essentially personal to a more social one with the CENDARI tools.

