INFRA-2011-1-284432

Project Acronym: CENDARI
Project Grant No.: 284432
Theme: FP7-INFRASTRUCTURES-2011-1
Project Start Date: 01 February 2012
Project End Date: 31 January 2016

Deliverable No. :	4.3
Title of Deliverable:	Expert Seminar Proceedings
Date of Deliverable:	January 2016
Revision No.:	Final
WP No.:	4
Lead Beneficiary:	UOB
Author (Name and email address):	Pavlina Bobic (p.bobic@bham.ac.uk) Emiliano Degl'Innocenti (emiliano@sismelfirenze.it) Kathleen Walker-Meikle (kathleen.walker-meikle@cerl.org)
Dissemination Level:	PU=Public
Nature of Deliverable:	R = report
Abstract (150 words):	This report sums up the results of three one-day Expert Seminars organized and presented in conjunction with the following WP4 partners: FUB, SISMEL and CERL. The report provides with an overview of seminar activities and addresses methodological challenges for historical research across time periods along with the impact of digital technologies on knowledge production. It discusses the value and use of CENDARI VRE and explores the links between methods in historical research and DH. The report also brings forth the potential that CENDARI infrastructure holds in comparison to other similar digital projects with the intention of strengthening the cooperation between archives and researchers and setting the frame for long-lasting scholarly connections.

SEVENTH FRAMEWORK PROGRAMME

Table of Contents

Deliverable 4.3

Expert Seminar Proceedings

Deliverable description (from DOW): *These volumes will allow the insights on the methodological challenges exposed in the CENDARI expert seminars to be captured and shared among archives and researchers. They will reflect the challenges discussed 1) across CENDARI periods 2) between CENDARI stakeholder groups and 3) focusing on interdisciplinary approaches to European experience - History, Social Sciences and Cultural Critique. They will capture the use case studies as well as more theoretical considerations, and a consideration of the role of the digital humanities in domain research practices.*

Description of expert seminars (from DOW): *The fourth phase of the work package's activities will centre around the nascent integrated collections themselves, with presentations prepared by the participants on the basis of data sets coalescing in the CENDARI infrastructure. In particular researchers who have taken advantage of the trans-national access offered by CENDARI will be invited to attend and present their findings. For these three seminars (two domain, one digital humanities) participants will be asked to prepare their investigations beforehand and submit them in a written form for publication afterward. The insights gained will further feed in to the usability data required by the technical development, while also surfacing possibly unexpected implications or lacunae in the provision of data. These seminars will greatly strengthen the links between the partners and the archives, feeding into a robust infrastructure and long term strengthening of the scholarly networks. The results will be prepared for publication, and thereby disseminated beyond the immediate participants as well.*

1. Executive Summary

1.1 Aim and Context

The report on methodological challenges presented in the CENDARI expert seminars feeds into the deliverable D4.3 ("Report on Expert Seminar Proceedings") in accordance with Description of Work (DOW) of the project "Collaborative European Digital/Archival Infrastructure". It centres on the results of expert seminars and synthesises them into a useful report for archivists and historians. It also explains place and purpose of expert seminars within the CENDARI workflow.

1.2 Method

The report integrates notes taken by Cendari partners during seminars, structured feedback forms completed by seminar participants as well as position papers prepared by seminar participants.

1.3 Structure of the report

The report is divided into two main parts. Both rely strongly on the results of the Expert Seminars and the discussion of research issues or latest developments pertaining to each domain. The first part consists of an overview of one-day ES organized and presented by three WP4 partners with the help of UOB: SISMEL, FUB, and CERL. The reports on MM and DH ES were contributed by SISMEL and CERL, respectively. The second part addresses methodological challenges for historical research across time periods and the ensuing impact of digital turn in knowledge production. The report also draws on the results of the previous

reports (D4.1, D4.4) while assessing the value of digital methods in research practices of Medieval, WWI and DH studies.

2. Brief overview of one-day Expert Seminars

In accordance with the requirements set in DOW, the three Expert Seminars addressed the value and future use of VRE, interconnectedness between methods in historical investigation and other disciplines in humanities (including digital humanities), and the potential that CENDARI infrastructure holds compared with similar online projects. The seminars were organized with the aim of enhancing the links between the partners and the archival institutions, and in the hope of establishing long-lasting scholarly networks.

2.1 Medieval Studies Expert Seminar (SISMEL)

2.1.2 Participants

The expert seminar on Medieval Studies was held in Florence on 30 October 2015. The participants were engaged in the presentation, discussion and evaluation of the digital tools developed by SISMEL and FEF, with a strong focus on their integration within the CENDARI environment.

The seminar was divided into three main sessions followed by discussion. Presenters and discussants of the papers were as follows:

Session I: The CENDARI Ontology in the context of Medieval Research

Presentations by:

- Lucia Pinelli (SISMEL): Scientific requirements for the Ontology: An overview

- Emiliano Degl'Innocenti (FEF/SISMEL): Extending the DM2E model: the mdv Medieval Extension
- Fabrizio Butini (FEF): Technical requirements for the Ontology
- Maurizio Sanesi (University of Florence): Modelling Semantic Medieval Data with DM2E and the mdv Medieval Extension

Discussants:

- Giovanni Fiesoli (University of Florence)
- Gabriella Pomaro (University of Florence)
- Roberto Gamberini (SISMEL)

Session II: Integration of Discoverability tools and Research Data for Medieval Studies in CENDARI

Presentations by:

- Alfredo Cosco (SISMEL - Zeno Karl Schindler Foundation) *Extending TRAME. Towards Data Mining and Knowledge Extraction from external resources*
- Daniele Bologna (University of Pisa - Zeno Karl Schindler Foundation) *TRAME come strumento per la formazione nelle Digital Humanities*
- Zdenko Vozar (University of Prague - Zeno Karl Schindler Foundation) *Building a crawler for Medieval Scholarly Resources*
- Roberta Giacomi (University of Siena) *Integration of reference resources: Authority lists of Medieval Authors' names*
- Vinicio Serafini (University of Siena): *Integration of reference resources: Authority lists of Medieval manuscripts' shelf-marks*

Discussants:

- Agostino Paravicini Bagliani (SISMEL - Université de Lausanne)
- Fancesco Santi (SISMEL - University of Cassino)
- Lino Leonardi (FEF – CNR-OVI, Opera del Vocabolario Italiano)

Session III – Hands-on

Participants: Emiliano Degl'Innocenti (SISMEL/FEF), Maurizio Sanesi (University of Florence), Silvia Nocentini (SISMEL), Roberto Gamberini (SISMEL), Gabriella Pomaro (SISMEL), Lucia Pinelli (SISMEL), Roberta Giacomi (University of Siena), Vinicio Serafini (University of Siena), Zdenko Vozar (University of Prague - Zeno Karl Schindler Foundation), Agostino Paravicini Bagliani (SISMEL - Université de Lausanne), Fancesco Santi (SISMEL - University of Cassino), Lino Leonardi (FEF – CNR-OVI, Opera del Vocabolario Italiano)

This session was followed by

Emiliano Degl'Innocenti – Maurizio Sanesi, *Hands-on session with live Data in the CENDARI Medieval Triple Store*
and
Emiliano Degl'Innocenti, *Report on the Trusted User Group test on TRAME and CENDARI*

The seminar concluded with discussion on currently available digital tools and their integration into CENDARI research infrastructure.

2.1.2 Activities

The seminar started with an introductory session hosted by Emiliano Degl'Innocenti and Lucia Pinelli, presenting two resources that are among the most relevant components in the medieval section of CENDARI: TRAME and MIRABILE. The former is a set of technologies fostering the discoverability of relevant digital resources for the study of medieval culture. The latter is one of the biggest repositories on the web providing research data on the same subject.

Emiliano Degl'Innocenti presented the TRAME development from a stand-alone meta search tool for medieval manuscripts to a web-based application intended to provide a layer of interoperability among different digital resources in the Medieval domain. It implemented most of the features listed in the user requirements for the CENDARI Medieval Prototype (cfr. the CENDARI Medieval Prototype requirements[1]) on a number of selected scholarly digital resources (more than 80).

Lucia Pinelli presented MIRAbile - Digital Archives for Medieval Culture: the knowledge management system for study and research on medieval culture developed by the SISMEL and the FEF ONLUS of Florence, in partnership with other research institutions[2] in Italy and Europe. MIRAbile represented one of the most relevant data providers for the medieval section of CENDARI, providing descriptions of manuscripts (more than 115.000), bio-bibliographical information on medieval authors and their works (more than 15.000 authors)

1 https://wiki.de.dariah.eu/display/CENDARI/TRAME+-+Texts+and+Manuscript+Transmission+of+the+Middle+Ages+in+Europe; https://wiki.de.dariah.eu/display/CENDARI/06+-+TRAME%3A+Texts+and+Manuscript+transmission+of+the+Middle+Ages+in+Europe; https://wiki.de.dariah.eu/display/CENDARI/TRAME

2 http://www.mirabileweb.it/ricerca_semplice.aspx?lingua=en

and other relevant contextual data, ranging from VI to XVI c. and covering all Western countries (including USA).

After the two presentations of TRAME and MIRABILE the seminar continued with three main sessions devoted to different aspects of the development of the medieval section of CENDARI. The first was focused on the ontology building and data modelling process, the second described the integration process of TRAME and MIRABILE within the broader CENDARI architecture, and the last provided a live demonstration of the new semantic tools developed by FEF and SISMEL by testing them against the research questions of the scholars and discussants.

Session I: Ontology, MIRABILE and TRAME

In the first session, Lucia Pinelli, Emiliano Degl'Innocenti, Fabrizio Butini and Maurizio Sanesi focused on several aspects of the development of the CENDARI semantic framework, paying particular attention to the scientific and the technical requirements that guided the ontology[3] building process: the selection of classes and properties needed to match the scientific requirements expressed by the research community[4], the extension of the classes and properties available in the DM2E with the elements required to properly represent the complexity of the medieval scientific domain, the integration of other sources (i.e.: controlled vocabularies, other DBs, authority lists etc.).

At the end of the three presentations, the discussants – for the most part involved in traditional research activities – were asked to provide feedback. They believed the following priorities for digital research infrastructures were important:

3 http://www.digitalhumanist.net/wiki/images/9/9c/7_09_2015_ONTOLOGIA.pdf

4 http://www.cost.eu/domains_actions/isch/Actions/IS1005?parties

1. exploit the richness of the available digital resources (i.e.: MIRABILE and other external datasets) to allow a wide range of research questions that are not currently addressed by traditional relational databases: e.g. relations between medieval authors and properties/attributes belonging to different lists, i.e. authors, title of works, monastic orders, places, etc.
2. develop tools for enrichment and integration of traditional relational DBs' contents
3. develop tools for better human machine interaction, i.e.: interfaces to give suitable results with keywords for browsing etc.

During the discussion, several tools that have been developed within CENDARI[5] and specific tools for Medieval Studies[6] were showcased, including:

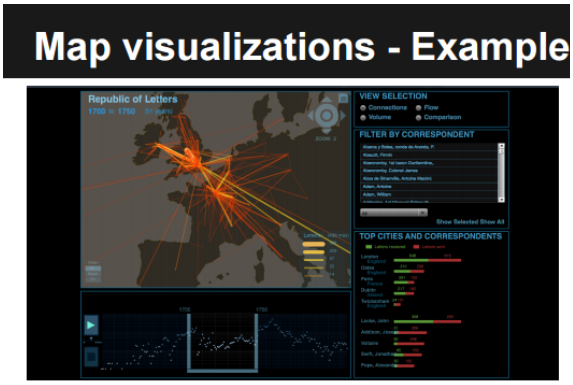
- Triple Store Management System (Virtuoso): to manage the medieval Knowledge base, resulting from the conversion of the available medieval research data into a semantic form
- Sparql endpoint[7]: to allow users and external services query over the semantic Knowledge base and/or build services to support digital research processes;
- a set of candidate tools for data visualization, providing spatial and temporal representations of available information[8] (cfr. below): maps, timelines, networks;

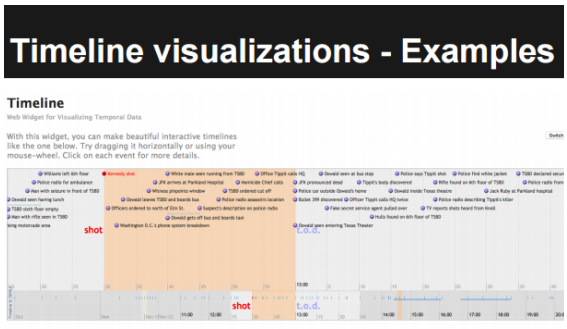
5 https://portal.cendari.dariah.eu/welcome

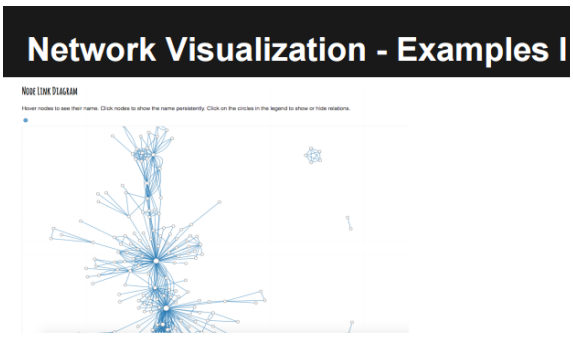
6 http://www.digitalhumanist.net/wiki/images/a/a5/L%27ecosistema_di_TRAME.pdf

7 http://git-trame.fefonlus.it/sparql/

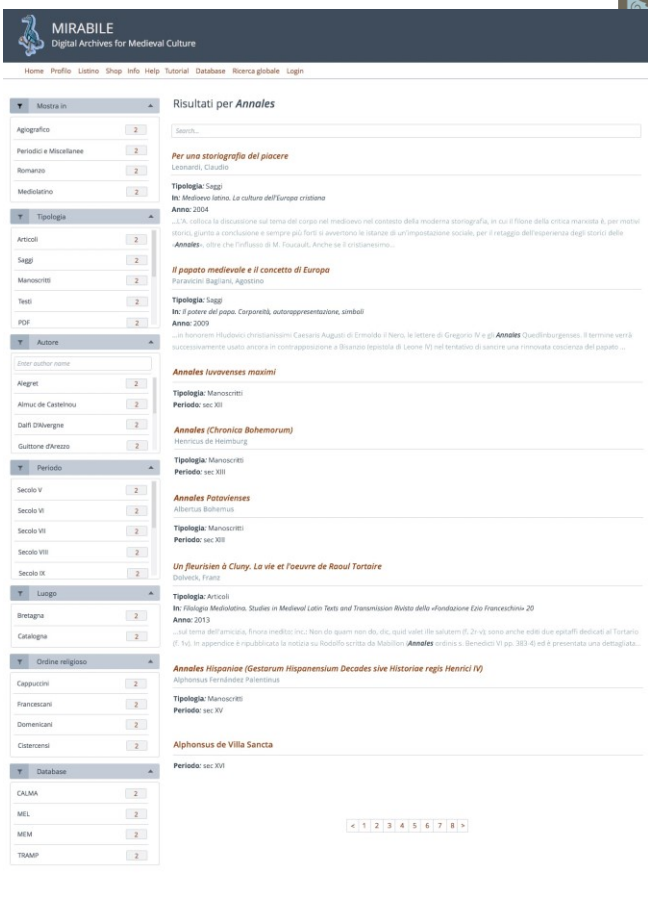
8 http://www.digitalhumanist.net/wiki/images/7/7b/Visualizzazione_delle_informazioni.pdf



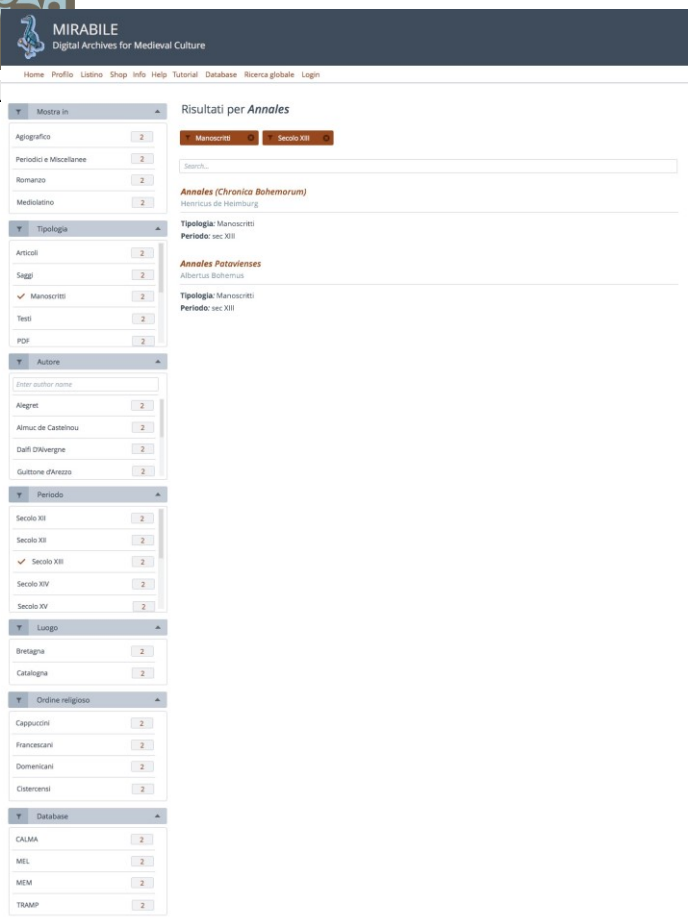




Scholars and discussants were also asked to provide feedback on an application for semantic data search and browse, built on top of the CENDARI semantic framework, following the requirements expressed by the research community (i.e.: surveys and interviews submitted within COST IS1005 meetings and workshops). Because the application was still in beta development phase at the date of the Expert Seminar, the feedback of the participants was crucial to improve the overall quality of its features (i.e.: add/remove facets, improve the design of the results visualization layout, fix bugs in the logic of the application, improve the design of the advanced search page, add more filters etc.).

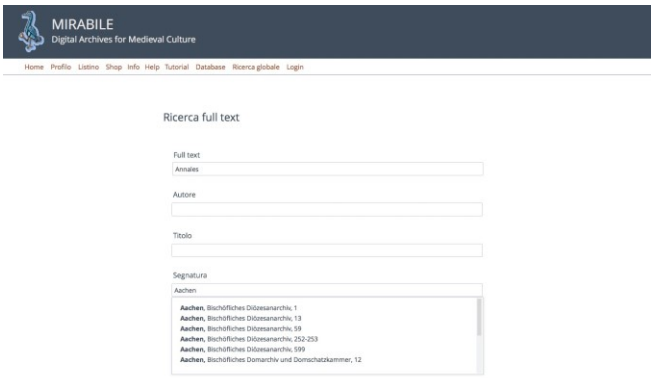


Faceted search example

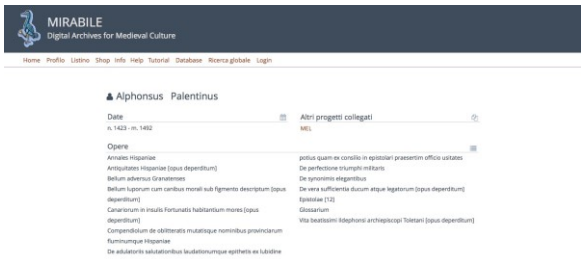


Full text search with auto-complete example

Furthermore, in session III (see *infra*), the participants were asked to formulate their research questions during a hands-on session, in order to test also the scientific relevance of the results provided by the tool. See the examples of the test below.



Shelf-mark search with auto-complete example



Data visualization example: authors

Session II

In this session, the focus of the seminar concentrated on the technological dialogue among CENDARI, TRAME and MIRABILE as parts of a coherent digital ecosystem for medieval digital research.

The first presentation, describing the new release of TRAME, TRAME 2 and its relation with CENDARI[9], was delivered by Alfredo Cosco. Among many new features, the participants were interested in the possibility to build a custom workspace: saving searches and databases preferences, exporting results etc. Cosco also presented a new tool for data analysis to use some technical metadata (i.e.: anonymous) to feed other components (custom data mining agents and crawlers). The final goal of this component is to provide for the knowledge extraction modules developed by CENDARI, a list of relevant resources to be considered, leveraging on users preferences.

Zdenko Vozar (awarded with a TRAME - Zeno Karl Schindler fellowship in Digital Humanities) focused his presentation on the development of the knowledge extraction module mentioned in the previous paragraph. Vozar presented the module as "basically a tool to *search and save* particular pieces of information from the world wide web [...] like taking notes from literature during the traditional research process. However, advance of the computational methods made it possible to automatize the whole process and take it to "a level never reached before using traditional means". Vozar then presented possible risks connected to this process: "important information, or their structure could be lost in the process [...] risk of infringement of the intellectual and institutional property of digitized information, etc.". Vozar also presented a working beta of the tool.

[9] http://www.digitalhumanist.net/wiki/images/3/3b/TRAME-VRE.pdf

Davide Bologna (former TRAME - Zeno Karl Schindler fellowship in Digital Humanities) delivered a presentation about the didactical potential of TRAME in the Digital Humanities[10].

The second half of this session was devoted to data integration: Roberta Giacomi described the process of elaboration of a shared authority list based on three different projects dealing with names of medieval authors: *BISLAM Bibliotheca Scriptorum Latinorum Medii Recentiorisque Aevi*, SISMEL project, *Repertorium Geschichtsquellen des deutschen Mittelalters*[11] and *Narrative Sources*, a repertory on medieval chronicles of the Low Countries[12].

Vinicio Serafini described a similar process aiming at the creation of a cumulative authority list of manuscripts shelf marks based on the DBs maintained by SISMEL, IRHT-Institute pour la Recherche et l'Histoire des Texts, and the *Scriptorium* journal.

Session III

In the third session, Maurizio Sanesi and Emiliano Degl'Innocenti hosted a hands-on session on some of the semantic tools developed by CENDARI, briefly introduced at the end of session I. The session focused on the use of the CENDARI semantic infrastructure (Triple Store, SPARQL endpoint and semantic browse and search application) to solve the research questions of the scholars. A number of questions, dealing with the textual tradition of given authors and works as well as related to the manuscript transmission, were considered. Goals included testing the scientific relevance of the results provided by the tool and

[10]http://www.digitalhumanist.net/wiki/images/d/d5/Sismel_Fellowship_-_Daniele_Bologna_%28versione_definitva%29.pdf
[11] http://www.geschichtsquellen.de/index.html
[12] http://www.narrative-sources.be/about_en.php

measuring the level of improvement made possible by semantic technologies over the traditional relational databases.

Among the most interesting results, scholars considered:

- the ability of the system to make inferences, like having given properties (i.e.: the relation with a certain religious order) inherited from an entity (e.g.: an author) to another (e.g.: a manuscript);
- the possibility of submitting queries involving data coming from different contexts (e.g.: persons, manuscripts and organisations) putting them in relation with spatial and temporal coordinates, without the limitations of traditional relational models;
- coordinates, without the limitations of traditional relational models.

nome_opera	nome_autore	nome_ordine_autore	nome_ente	nome_ordine_ente	info_ente
"Epistolae"	"Stephanus Tornacensis episcopus n. 18-2-1128, m. 11-9-1203"	"CanR"	"Breuil-Benoît, abbazia OSB"	"OSB"	"Francia Haute-Normandie Évreux (Eure) B "
"De amore sponsi ad sponsam"	"Hugo de Sancto Victore n. 1096 ca., m. 11-2-1141"	"CanR"	"Breuil-Benoît, abbazia OSB"	"OSB"	"Francia Haute-Normandie Évreux (Eure) B "
"Quomodo sermo Dei unus est"	"Hugo de Sancto Victore n. 1096 ca., m. 11-2-1141"	"CanR"	"La Trinité, abbazia OSB"	"OSB"	"Francia Haute-Normandie Fécamp (Seine-I abbazia OSB "
"Epistolae"	"Stephanus Tornacensis episcopus n. 18-2-1128, m. 11-9-1203"	"CanR"	"La Trinité, abbazia OSB"	"OSB"	"Francia Haute-Normandie Fécamp (Seine-I abbazia OSB "
"De assumptione beatae virginis Mariae"	"Hugo de Sancto Victore n. 1096 ca., m. 11-2-1141"	"CanR"	"St.-Pierre, abbazia OSB"	"OSB"	"Francia Haute-Normandie Jumièges (Seine abbazia OSB "
"De statu interioris hominis"	"Richardus de Sancto Victore m. 10-3-1173"	"CanR"	"Notre-Dame, abbazia OSB"	"OSB"	"Francia Haute-Normandie Lyre (Eure) Not
"Miscellanea"	"Hugo de Sancto Victore n. 1096 ca., m. 11-2-1141"	"CanR"	"Notre-Dame, abbazia OSB"	"OSB"	"Francia Haute-Normandie Lyre (Eure) Not
"De archa Noe"	"Hugo de Sancto Victore n. 1096 ca., m. 11-2-1141"	"CanR"	"Notre-Dame, abbazia OSB"	"OSB"	"Francia Haute-Normandie Lyre (Eure) Not
"De sacramentis legis naturalis et scriptae"	"Hugo de Sancto Victore n. 1096 ca., m. 11-2-1141"	"CanR"	"Notre-Dame, abbazia OSB"	"OSB"	"Francia Haute-Normandie Lyre (Eure) Not
"De essentia et substantia Dei et de tribus eius personis"	"Guillelmus Campellensis magister n. 1070 ca., m. 1122"	"CanR"	"Mont-Saint-Michel, abbazia OSB"	"OSB"	"Francia Basse-Normandie Mont-Saint-Mich Michel, abbazia OSB "
"De corpore Domini"	"Guillelmus Campellensis magister n. 1070 ca., m. 1122"	"CanR"	"Mont-Saint-Michel, abbazia OSB"	"OSB"	"Francia Basse-Normandie Mont-Saint-Mich Michel, abbazia OSB "
"De sacramento altaris"	"Guillelmus Campellensis magister n. 1070 ca., m. 1122"	"CanR"	"Mont-Saint-Michel, abbazia OSB"	"OSB"	"Francia Basse-Normandie Mont-Saint-Mich Michel, abbazia OSB "
"De verbis Isaie: In die illa nutriet"	"Richardus de Sancto Victore m. 10-3-1173"	"CanR"	"Mont-Saint-Michel, abbazia OSB"	"OSB"	"Francia Basse-Normandie Mont-Saint-Mich Michel, abbazia OSB "
"De eruditione interioris hominis"	"Richardus de Sancto Victore m. 10-3-1173"	"CanR"	"Mont-Saint-Michel, abbazia OSB"	"OSB"	"Francia Basse-Normandie Mont-Saint-Mich Michel, abbazia OSB "
"Tractatus super psalmum «Afferte Domino»"	"Richardus de Sancto Victore m. 10-3-1173"	"CanR"	"Mont-Saint-Michel, abbazia OSB"	"OSB"	"Francia Basse-Normandie Mont-Saint-Mich Michel, abbazia OSB "
"De praeparatione animi ad contemplationem"	"Richardus de Sancto Victore m. 10-3-1173"	"CanR"	"Mont-Saint-Michel, abbazia OSB"	"OSB"	"Francia Basse-Normandie Mont-Saint-Mich Michel, abbazia OSB "
"De exterminatione mali et promotione boni"	"Richardus de Sancto Victore m. 10-3-1173"	"CanR"	"Mont-Saint-Michel, abbazia OSB"	"OSB"	"Francia Basse-Normandie Mont-Saint-Mich Michel, abbazia OSB "
"De meditandis plagis quae circa finem mundi evenient"	"Richardus de Sancto Victore m. 10-3-1173"	"CanR"	"Mont-Saint-Michel, abbazia OSB"	"OSB"	"Francia Basse-Normandie Mont-Saint-Mich Michel, abbazia OSB "
"Causam quam nesciebam"	"Richardus de Sancto Victore m. 10-3-1173"	"CanR"	"Mont-Saint-Michel, abbazia OSB"	"OSB"	"Francia Basse-Normandie Mont-Saint-Mich Michel, abbazia OSB "
"Liber poenitentialis"	"Robertus de Flamesburia v. 1205, m. ante 1234"	"CanR"	"Mont-Saint-Michel, abbazia OSB"	"OSB"	"Francia Basse-Normandie Mont-Saint-Mich Michel, abbazia OSB "

Example of semantic query involving religious orders, authors, texts, place and time

Emiliano Degl'Innocenti also presented a new application for semantic data search and browse, leveraging the CENDARI semantic infrastructure already presented. The goal of the application is "to let users access the rich data in the Triple Store without interacting with a SPARQL endpoint […] to offer traditional tools for regular users and extra services for power users […] to provide a simple and effective user experience. The participants were asked to test the tool, to submit a research question and provide feedback on the available functionalities:

- full-text with autocomplete
- advanced search with autocomplete
- faceted browsing

The participants provided valuable feedback to measure the level of improvement eventually brought by the new application. The experts pointed out that the proposed interface helped to avoid one of the limits of some traditional relational systems (like MIRABILE) in which some cross searches involving multiple search fields (e.g. full-text and title) are not allowed. The aim is to cross more fields using Boolean operators allowing to search for a given textual entity, mixing the full-text search function with operators, increasing the modularity of the resulting query (e.g.: "I want to visualize all the entries with the words 'Annales' and 'Alegre'"). A similar feature is currently not available in the global search of MIRABILE.

Experts were also asked to test a faceted search tool, based on filters and lists (e.g. document type, author, place, time, religious order, etc). Once the facet has been selected, the results will be visualized and updated automatically, providing a dynamic result set. Furthermore, all the entries of the browsing lists will be updated. In this case the participants seemed to be not totally comfortable with

the faceted browse mechanism and asked to make several changes to the number of the facets, and their behaviour.

The hands-on session ended with a quick overview of the reasoning technology applied to the data in the Triple Store.

2.1.3 Medieval Studies, CENDARI Trusted User Group (TUG) and digital research practices

The final part of the Expert Seminar was centred on the activities of the Trusted User Group (TUG) in relation to the medieval section of CENDARI. SISMEL invited a group of international medieval scholars from the COST Action IS1005 to join the TUG activities. Furthermore, a local task force, comprised of Silvia Nocentini, Lucia Pinelli, Roberto Gamberini and Gabriella Pomaro, was also established with the aim of providing feedback for the development of the medieval section of the CENDARI infrastructure. A couple of sessions of the CENDARI TUG were devoted to testing the tools and services developed by SISMEL and FEF with a strong focus on TRAME. Emiliano Degl'Innocenti briefly summarized the results of the tests.

The users in the TUG pointed out some pros and cons of TRAME: the large majority of the participants appreciated the contents made available in TRAME but nevertheless some of them highlighted missing elements, such as library catalogues and information on digitized collections of manuscripts. Many reported on the noise in the results and stressed the necessity of filtering the results of the queries. The duplication of results is also an issue.

The TUG also highlighted some positive elements of TRAME, in particular:

- The ease-of-use of the user interface
- The presence of different search and browse channels (e.g.: simple search, browse by shelf marks etc.)
- The availability of a vast number of sources and repertories

The TUG also proposed some possible improvements for the TRAME user experience:

- Elaboration of interactive tutorials on available search functions
- Refined advanced search mask
- Additional new sites

The Expert Seminar ended with a closing wrap-up session.

2.2 WW1 Expert seminar (UOB, FUB)

2.2.1 Participants

The WWI Expert Seminar was organised at FUB with the help from UOB team. It was held on 24 September 2015 and attended by 9 invited scholars: these included 4 former participants of the CENDARI Summer School (Berlin 2014, and members of the Trusted User Group), 2 TA fellows and 3 humanities scholars with substantial knowledge of various digital research platforms. One participant in particular (Clemens Neudecker) utilized his computer sciences background for the provisioning of digital cultural heritage and related technologies to scrutinise digital resource; his position paper was therefore especially insightful in terms of reviewing CENDARI VRE from a technical rather than merely historical perspective.

Participants were as follows:

- Jenny Sprenger-Seyferth (FUB)
- Anne Baillot (Institute of Literature at Humboldt University)
- Tim Buchen (School of History at the University of Edinburgh)
- Alessandro Salvador (University of Trento)
- Frank Grelka (Europa-Universität in Frankfurt/Oder)
- Deniza Petrova (FUB)
- Clemens Neudecker (Berlin State Library)
- Dennis Mischka (DARIAH-DE fellow, University of Stuttgart)
- Julian Nordhues (Hybrid Publishing Lab, Centre for Digital Cultures, Leuphana Universität Lüneburg)

2.2.2 Activities

Before they came to the seminar, the participants were asked to focus their contributions around a series of questions relating to the CENDARI VRE and its associated tools. Presentations by TA fellows and other invitees were organized into 3-4 panels (by research theme or discipline/profession or career stage) of 3 presenters. Scholars were asked to address the following issues:

1. Their experience with CENDARI infrastructure
2. The importance of digital datasets for their personal research or professional practice
3. The extent of their engagement with digital methodologies in research or professional practice
4. The potential for enhancing research and overcoming fragmentation of archival sources using the CENDARI platform (esp. search functions, VRE, and ARGs)
5. How using CENDARI worked in comparison with other tools/projects.

Value of VRE

The participants mostly stressed the gravity of the (growing) CENDARI Archival Directory, the purpose of which is to identify, describe and catalogue archival holdings and other collections crucial for comparative and transnational historical research.

The participants believed the archival directory enables an efficient search of vast resources and is further enriched by Archival Research Guides (ARGs). They thought it was important that the ARGs were designed to address specific up-to-date research issues and offer fresh perspectives on the questions at stake along with the selected (or suggested) archival information on primary sources compiled in the archive directory. In this regard the VRE was recognized unanimously as a good starting point for examining recent WWI topics, which, with the help of ARGs, was said to build a conceptual "bridge" to the NTE (Note Taking Environment). On a practical level, the seminar participants described the advantages of NTE as being concentrated on "the many options" in the middle window provided for the note taking. This allows a "well structured output when the note is complete", but still requires a "considerable time investment" from researchers to be able to employ all the options as a new note-taking user. According to a critical observation of one of the discussants, the VRE scheme still lacked a clear overall concept and failed to provide a more fitting link between NTE and archival directory. Many of the tools thus seemed convincing on a micro level, while, in view of the same critic, the entire picture "lacked coherence". The NTE, on the other hand, helped the user to benefit from DH techniques to organize, structure, and analyze sources pertaining to specific research project.

Another value of the VRE was attributed to its applicability to broader research in the field of humanities. In response to one of the lead questions, the participants recognized the possibility of organizing texts and documents in a VRE as especially welcome for literature scholars and philosophers. This suggests that the VRE has potential beyond the historical discipline itself. Moreover, the ARGs and the archival directory were acknowledged as "the best possible tools for creating a bridge between historians and archivists and other humanities scholars". In the view of Dennis Mischke, ARGs might "introduce new ways of inserting orientation and consistency into growing archives by establishing novel ways of access and knowledge extraction from existing historical data". While CENDARI VRE has an ambitious collaborative focus, the prospect of visualizing certain features of the data offers new "pathways to the analysis and understanding of archival resources" and has a potential to be further extended. Furthermore, given its transnational character, the future NTE could (and should) become available in multiple languages. The WWI researchers also stressed the value of digitized historical documents and photographs available as online resources, which at the same time poses an additional challenge for CENDARI in that its system has not been developed to support systematic digitization of archival material. Again, this shows the challenges involved in navigating between expectations of users and capacities of virtual research infrastructures like CENDARI.

The WWI presenters attributed the most convincing strength of CENDARI VRE to its ability to connect "collective work with the resources of the digital archive". Moreover, the NTE aspires to "enrich the archive with a sleek tool to categorize, sort, analyze and visualize" (Dennis Mischke) selected archival items. The system's further potential is the categorization of entities and subsequent visualization and in backing the interest of both historians and other humanities researchers. During the discussion the participants asserted that the mapping

tool in particular could be used in a "spacial-historical analysis of narratives, texts and entities without coding skills or elaborate GIS systems." The GIS are increasingly significant in the humanities, while the value of this tool and the archival research environment could sharpen up the historians' new way of viewing sources.

Suggestions on future features in VRE

Presenters proposed and discussed the following key points on how to familiarize new users with the VRE:

- **Guidelines**

a) Exemplary use cases like the introduction of the notion of an ARG by demonstrating the functionalities of the NTE and Archival Directory;

b) Introductory guides to ground concepts (how the developers imagined a certain tool to be used);

c) Screencasts and/or video presentations on best practices;

d) Training users for an efficient use in the further steps of any project (e.g. data analysis options);

e) How-to's for instructors using Cendari in the teaching environment focusing on standard workflows established in academia;

f) Best practices in international multidisciplinary collaborative research processes.

◆Collaborative features in VRE

a) Request for a differentiated rights management was suggested together with the possibility to define groups and to provide (read-)access to non-registered users. It was further suggested to enable sharing of link lists, research data and texts within a group. Discussants likewise requested the possibility of offering "Research Projects" to an open group of users.

b) Project Management Tools (PMT). The idea was to assign roles and tasks to other users and allow for a more managerial perspective on the workflow of the researchers or group (thus tailoring VRE also for teaching). Project management tools would likewise enable notification on assigned tasks or activities of co-workers in a shared project or creating and sharing research schedules. PMT would conjointly assume integration of a feedback tool in order to enable feedback when working collaboratively on a project.

c) Establishment of a mentoring system within the VRE, with mentors or experts being visible in the NTE or Portal.

From the viewpoint of usability, the VRE, especially "accessing NTE", was evaluated as "not entirely user friendly". As some scholars stressed in the discussion, this degree of user-friendliness, depended on who exactly the user community would actually be. Proposals by the experts for help and moderation features included example cases like the introduction of the notion of an ARG by showing the functionalities of the NTE and archival directory. From this perspective, users need to be presented with how to use CENDARI in classes concentrating on standard workflows established in academia.

Alessandro Salvador's cited the role of teaching with regard to CENDARI. According to Salvador, accessing NTE should be made easier especially for fresh undergraduate students with no previous familiarity with digital research tools. In this regard it would be very helpful to be able to rely on guidelines or "how to's" to tackle specific questions, while in the same vein providing examples on how to employ those tools in daily research work. From the perspective of Salvador, CENDARI should invest a lot of effort to "correctly train the users" to ensure a more productive use in later stages of research such as "data analysis" or how to "draft research outputs". He also proposed using "example cases" that showed how CENDARI could be used for a particular research project from start to end. The latter suggestion is in fact best exemplified by the existing domain specific prototype projects, which were produced on the basis of real and complex research scenarios with the addition of specific technical functionalities required to implement them. The WWI prototype in this sense represents a good example of a project that it is hard to pursue in the current digital environment and shows a number of obstacles posed by fragmented, undigitized, multilingual sources in a range of institutions.

CENDARI VRE is designed to share scholarly information on sources with the intention of expediting international collaboration and transnational research. Suggestions on how the system itself could be further improved centered on visualization, which should be "expanded", for instance, allowing visualizing events as a connection between people and organizations, places and content. New categories like "space" or "concept" were suggested as helpful in the semantic analysis in historical projects. The implementation of other popular programs such as zotero, citavi, endnote, etc., could be used to further incline researchers to share literature or annotations made while off-line.

These recommendations also show that large-scale infrastructure projects like CENDARI face a number of challenges. The end users have a vast choice of programs and online resources to strengthen their research projects. What these users require is a system adjusted to their own research practices. However, historians are not traditionally trained to see the technical complexities behind digital infrastructures and what scientific efforts are involved in tailoring them to the humanists' needs. This also represents one of the major pitfalls of an age in which digital practices have only been partially embraced by historical community.

These insightful proposals were also reflected in the broader discussion on the impact of CENDARI's newly envisioned research methodology on both domain research practice and consequential dissemination of knowledge. Along with CENDARI's digital platform DH in general shape a revolutionary outlook on further research in humanities. New digital tools are specifically developed and introduced to ease and, above all, hone the queries and understanding of academic thinkers.

2.2.3 Impact of CENDARI and DH on domain research practices and knowledge production

The major capacity of CENDARI VRE is its attempt to facilitate collaborative and transnational research but also to strengthen the "benefits this entails for transparency and reproducibility in science" (Clemens Neudecker). These have been recognized as the areas in which CENDARI's VRE can leave the strongest mark on current and future scholarly practice. It is namely becoming evident and increasingly crucial that research is being conducted through digital technology, which influences both the epistemologies and ontologies that guide academic projects.

The seminar participants in Berlin exposed visualization tools in the CENDARI VRE as a promising feature fostering new ways of knowledge production and data mining for historians and other humanities scholars. Owing to visualization of data, the VRE could "modernize" people's perception of the past by way of contributing to the "construction and mediation of historical knowledge" (Tim Buchen) and ultimately lead to revolutionizing teaching methods. Visualization of data could gradually gain its prominence due to disclosure of semantic connections, or, as Tim Buchen asserted, "processes of simultaneousness" that will in his opinion stand beside text narratives, particularly in global and international history.

In view of the discussants, the possibility of drawing on research and preexisting work of fellow historians through ARGs likewise bears the potential to greatly enhance the archival research practice and significantly help scholars with lesser experience in archival work. As long as there are common research topics with transnational tendencies as well as clearly defined methodological steps, CENDARI's VRE has the capacity to become "a frequently used tool in historical investigation and [to] motivate historians to become involved in a large community of scholars" (Frank Grelka). The act of sharing of information on sources and research agendas will inevitably support international collaboration as well as widen international horizons and approaches. A highly functional structure (i.e., highly developed ontologies assisting researchers to link names, dates, places, with controlled entities and actors, along with digital interfaces including images of original documents, maps, etc.) could eventually bring about a significant move forward in academic teaching.

The benefits of the NTE and the archive directory are linked to their integration within the same VRE. However, researchers may be utilizing other existing

virtual tools for disseminating or sharing their scholarly work, which, as already mentioned, may prove to be a significant challenge for CENDARI.

Due to persistent source fragmentation, the main task for transnational and comparative WWI historians remains to locate, connect and organize sources in various countries, repositories and formats. Our previous reports (D4.1, D4.4) showed that WWI historians strongly emphasized an advancing historiographical trend towards transnational topics, while recognizing the need to concentrate on the hitherto much neglected zones in Eastern and South-Eastern Europe. Some of the problems in this regard represent the inaccessibility of Russian archives and the often-dispersed primary material in East European archives. The discussion at the ES in Berlin underlined the significance of CENDARI's archival directory, which has built a solid base for further enquiry in the lesser-known repositories in Eastern Europe. These are, however, often largely unattainable to Western historiographers due to language issues. The WWI historians attending the ES were acutely aware of the difficulties posed by multilingual sources. In line with the previous workshop discussants (London 2013), they repeated the need for an array of digital tools to help researchers transcend the barriers imposed by research in foreign languages, fragmentation of sources across national and institutional borders, and further required efficient tools for visualizing collected data.

The shift from analogue to digital is a major paradigm change for the humanities, especially for modern historians. The availability of digital resources and tools offer an overabundance of new options for researchers yet the overall implications this shift will have on research methodology is still debatable.

Digital aggregation of information and archival (meta)data are enabling a number of innovative ways of knowledge production. For instance, Europeana,

which is a rich digital archive, is a good example of how complicated and varied research data can be. According to one of the seminar participants in Berlin, CENDARI VRE might become the first step to manage "the fussy and quickly growing quantities of humanities research data".

The "digital turn" in humanities (and history more specifically) is most evident in the use of functions of Google, Google Books, JSTOR, digital newspapers and online archival finding aids. These functions have opened up an immense new perspective in modern history. For instance, the capacity to find all the mentions of a name in millions of pages of OCR-scanned papers allows historians to analyze the lives of the people in a unique fashion. At the same time historians have access at their disposal to secondary sources about the context of a person, event or organization, which would have previously required multiple trips to archives, including in foreign countries. Still, for historians of the First World War, as many attendees of this workshop indicated in the discussion, the mass of material is such that long research visits in archives remain important, especially given reluctance of many major European archives to allow the use of digital cameras. Thus, this expert seminar made clear that these researchers were looking for a research infrastructure like CENDARI to enhance research practices already in place as opposed to overturning such practices.

The discussion also revealed that historians in particular were interested how digital methodologies could be used to break down the container-like national histories that have dominated the 20th century and still structure archival practice to some extent. It was clear that there was also great hope that the CENDARI infrastructure would help historians to more quickly and easily peer over these national divisions. Several participants noted in the discussion that this was particularly valuable in the case of Eastern Europe where national barriers remain quite strong, yet World War I historians of Eastern Europe study

a series of empires (Germany, Tsarist Russia, the Habsburg Empire, and the Ottoman Empire) not bounded by national spaces of today. In this respect, it is clear that CENDARI, with its explicitly transnational focus, offers real potential to the historians of the First World War.

2.3. DH [Digital Humanities] Expert Seminar (CERL)

2.3.1 Participants

On the 29 October 2016 CERL organised a DH Expert Seminar at CERL that focused on the CENDARI infrastructure, the strategic decisions made in building it, and the challenges presented by large scale digital research infrastructures for digital humanists and the field more broadly in light of CENDARI's experience.

The following scholars participated at the seminar:

- Joris van Zundert (Huygens Institute)
- Marian Lefferts (Consortium of European Research Libraries)
- Marco de Niet (DEN Kenniscentrum Digitaal Erfgoed)
- Emiliano Degl'Innocenti (Società Internationale per lo Studio del Medievo Latino/Digital Research Infrastructure for the Arts and Humanities Italy)
- Kathleen Walker-Meikle (Consortium of European Research Libraries)
- Stijn Van Rossem (Consortium of European Research Libraries)
- Alastair Dunning (Europeana)
- Marnix van Berchum (Data Archiving and Networked Services)
- Ingeborg Versprille (Consortium of European Research Libraries)
- Jonathan Gumz (University of Birmingham)
- Sally Wyatt (University of Amsterdam)
- Henk Wals (Internationaal Instituut voor Sociale Geschiedenis)

- José de Kruif (Universiteit Utrecht)
- Catherine Jones (Centre Virtuel de la Connaissance sur l'Europe)
- Pim Huijnen (Universiteit Utrecht)
- Roxane Wyns (KU Leuven)
- Tom Gheldhof (KU Leuven)

2.3.2 Activities

The Expert Seminar in The Hague focused on the role of the user in large-scale digital infrastructures, of which the CENDARI project is a prime example. The seminar began with five presentations of scholars and project managers, with each of them outlining the challenges or opportunities in connecting users with their digital infrastructures. Each of the speakers had been asked to provide a provocative statement, aimed to generate a discussion. The statements were added to the programme. After the presentations, Marco de Niet led an interactive group discussion and brainstorming session. The starting point for discussion was the statement: "Digital infrastructures for research data and data analysis are not properly aligned with research practices and the needs of individual researchers".

As the field of digital humanities evolves, it still suffers from a split between large-scale digital infrastructures and the future researchers who will use these created digital humanities tools. These infrastructures are often developed by institutions or consortia at an international level. Some of the more ambitious recent projects, such as the CENDARI project itself, were created under the 7th framework program of the European Commission.

In this Expert Seminar the organizing team at CERL tried to draw attention to the role of the user in large-scale digitisation projects and digital infrastructures.

Who decides which tools are created or which projects are funded? Are users involved in that process? Can the target audience be accurately defined? Are the users of these databases and infrastructures involved in the creation, testing, and adjustments? What can we learn from current projects and the way they may or may not involve the user from the initial stage to the digital infrastructure launch?

In the first part of the seminar, Kathleen Walker-Meikle and Stijn van Rossem (project officers for CERL on the CENDARI project) presented the CENDARI platform and discussed how users were involved in different ways in the course of CENDARI's development. At the initial stage, several participatory design workshops were organized in order to gauge the needs of different kind of users for regarding a Virtual Research Infrastructure. The targeted users were medieval scholars, First World War scholars, archivists and librarians). This was followed by creating user stories based on the needs of individual users, so that system functionalities could be defined more clearly. In the project's final year, CENDARI invested heavily in the creation of a Trusted Users Group, which was used to test if created tools were working properly and suited research needs. They presented successes, failures and lessons learnt.

Continuing, Joris van Zundert, a researcher and developer in Computational and Digital Humanities Huygens ING, expressed the desire to have future infrastructures be more open and versatile, and less concerned with standards and point and click interfaces. As research is always evolving and new questions will always emerge, 'closed' infrastructures cannot keep up with developments. Although large-scale infrastructures are important, data should be offered to the user with care. Infrastructures should be opened up, so that the user can utilise the data as desired.

Setting off a provocative question of whether users do or do not need to know about large-scale infrastructures, Alastair Dunning presented Europeana Cloud's new approach of focusing on agile development and minimum valuable products, rather than opening one large infrastructure to the user community in one go. By releasing small but controllable tools every few months, users can provide feedback on a regular basis at all stages of development. Europeana Cloud will be run as a series of services, released in stages, in order to allow maximum interaction and modification. In the discussion after this presentation, Jonathan Gumz pointed out that this agile model for development was at least partially achieved by the formation of the TUG within CENDARI. Future large scale infrastructure projects should look to having something like the TUG built into the early portions of the project, not just the later phases.

On the other end of the spectrum, Emiliano Degl'Innocenti elaborated on the future role DARIAH could play by connecting individual projects and data resources. DARIAH could play an active role in ensuring that tools are used, perhaps even outside of the context they were first developed in. Furthermore, DARIAH needs to also become a political framework, with the ability to talk at the highest level with policy makers and set the agenda.

The seminar was concluded by the contribution of Marnix van Berchum, head of data services at DANS, presented DANS and the conflict between the preservation of data and the access of data institutions such as DANS. DANS hopes to encourage researchers to reuse data that has already been collected and is available as open access, but many researchers still prefer to collect their own datasets.

2.3.3 Digital infrastructures and research practices

Presentations by domain experts were followed by discussion in which the presenters were asked to comment on the statement: *The digital infrastructures for humanities research and data analysis are not always properly aligned with research practices and the needs of individual researchers*.

Each expert was asked to write down positive and negative aspects in the relationship between users and digital infrastructures, based on their own daily practices.

Whereas the abundance of available data and the creation of communities of users were felt as very positive aspects, four clusters were pinpointed on the negative side:

- Data: interoperability, dark data, not always transparent data sets;
- Use: lack of technical skills on the side of the user;
- Vision: monolithic, generic, emphasis on standards, not specific enough;
- Sustainability: doubtful.

The experts chose the following two issues as the most urgent matters that need to be addressed in the next five years:

1. Firstly, conceptual infrastructures. The currently fractured landscape, which is partly due to the funding model, has led to the opportunistic development of infrastructures that do not necessarily fit well together. Jonathan Gumz pointed out that as the Cendari experience showed, this also relates to the question of sustainability of the infrastructures which in turn relates to concerns expressed

by archives when turning material to a "project" based infrastructure like CENDARI.

2. Secondly, there is a great need to invest in user skills. Digital humanities should be a methodology rather than a separate field, which means that all humanities scholars should be familiar with these methods and be able to apply them if needed in the course of their research. Questions were raised in the discussion whether or not disciplines like history are ready for such a large scale change in how they train future practitioners of the discipline. Some discussants raised the notion of recognizing specific footholds that can be gained for DH within these disciplines and acting on them. In this sense, Cendari's extensive research on user practices (undertaken through WP 4 and WP 8) should be invaluable for future large scale infrastructure projects.

2.3.4 DH and Research Practices: Between User and Large-Infrastructure Projects

In concluding remarks of the DH Expert Seminar, Marco de Niet pointed out four areas of tension that exist when attempting to synchronize user needs and large infrastructure projects:

1. Infrastructures as a practical implementation (short term user needs) and infrastructure as a framework/concept (the 'vision');

2. Communities (disciplines, pre-formatting based on group profiles) and individuals (allow for idiosyncrasies; this includes coders as well as non-users);

3. Perspective from researchers (e.g. Dariah / Cendari) and perspective from professionals in memory institutions (e.g. Europeana);

3. Support through products (tools, software, 'infrastructure visible as a building') and support for processes (e.g. persistent identification, licensing; 'infrastructure as a service', invisible as a sewer);

4. Context of infrastructures (politics, governance, funding) and context of research (methodology, domain, funding).

3. Methodological and interdisciplinary challenges for historical research across time periods

3.1 Comparison of key datasets across domains

The Expert Seminar in Medieval Studies showed that the domain historians, who were mostly conducting their research in a traditional way, recognized a number of advantages offered by digital resources. Rich material is available to them through (external) databases, such as MIRABILE, which allows researchers to analyse subject matter otherwise not available in traditional databases (for instance, relations between medieval authors and properties or attributes coming from different lists, i.e., authors, title of works, monastic orders, places, etc.). Another key dataset infrastructure in the MM section of CENDARI is TRAME. The latter is composed of a set of technologies furthering the discoverability of relevant digital resources for the study of medieval culture. It represents one of the biggest online repositories providing research data on medieval history. Notwithstanding, medieval scholars wanted to work with digital tools intended for enrichment and integration of traditional databases. In addition, they expressed the need for improved human-machine interaction.

In the medieval domain, digital resources are becoming increasingly available to scholars. This means that research in medieval studies is more and more reliant

upon acquisition of digital data, tools, and methodologies. These are also valuable for modern historians in that new research techniques (available in a digital enquiry environment) significantly expedite historical investigation. The latter is particularly important in the view of enormity of primary resources available to WWI researchers who are invariably faced with the problem of tracing down sources in the (often) dispersed repositories. This is especially relevant in the early stages of defining the project, storing and organising data. While medievalists have a broader array of digitized materials (objects) available for their research, WWI historians only have very limited amount of online research data available. However, the need for a digital landscape with interoperable datasets to connect different research contexts and different sorts of information binds both medieval and WWI domain in a strongly correlated digital research environment.

The DH Expert Seminar at CERL emphasised that future digital data infrastructures need to become more accessible and flexible. Researchers should be encouraged to employ "open" data while the archival data providers should, on the other hand, allow infrastructures to make use of digital data. Ensuring that the appropriate digital tools are designed to exploit both is important and CENDARI offers clues on how to move in that direction. The ES on DH also stressed the importance of data preservation and how research data could be re-used via available open access online infrastructures, while still enabling researchers to gather and work on their own datasets.

3.2 Comparison of engagement with digital tools across domains

In the medieval domain, experts are often able to rely in their scholarly enquiry upon digitally available study material. To this end, a number of digital tools have also been developed within CENDARI. Virtuoso, triple store management

have been created to sustain the medieval knowledge base. Sparql endpoint provides users and external services an enquiry over the semantic knowledge base or build services to support digital research processes. Digital infrastructures for medievalists likewise currently provide a set of tools for data visualisation, providing spatial and temporal data. The MM digital sphere has been traditionally viewed as a highly fragmented "knowledge arena", facing with the problem of multilingual resources, access to them, and a huge number of non-interoperable sources using a variety of standards and technologies. An early phase of medieval research is a combination of traditional and digital scholarly methods. The latter brings medievalists closer to the ways in which WWI historians employ their methods to clarify their research objectives.

In addition to dealing with a vast volume of primary material, which is frequently dispersed across national repositories, the WWI historians have to overcome a number of difficulties posed by multilingual sources. The latter issue led the CENDARI workshop participants (London 2013) to require a number of digital tools to help WWI experts transcend the research obstacles, and further required efficient tools for visualizing collected data. The CENDARI infrastructure currently incorporates many of these tools.

In the Expert Seminars, scholars working in DH emphasized the role and the place of the user in large-scale digitisation projects and digital infrastructures. It has been suggested that the digital era has both improved and exacerbated problems in the scholarly research process (D4.4). Many large (national) archives and libraries have responded to the "digital turn" with a strong stimulus to digitise finding aids, catalogues, special collections and record series as well as by the creation of portals and hubs to provide searching of dispersed sources.

3.3 Assessing CENDARI platforms' usefulness across domains

Due to fragmentation of scholarly material stored in traditional databases, medieval historians need to have them transferred into semantic networks and ready for academic re-use. WWI historians, in the meanwhile, require a dynamic virtual research space to nurture their projects at various points in the research process, not necessarily a research infrastructure that is there at all points of the project.

Both medieval and WWI history are defined by transnational, multilingual and interdisciplinary research, though the degree of linguistic fracture is greater for the First World War. To this end, CENDARI infrastructure has addressed some of the key technical aspects to assert itself as one of the leading digital infrastructures for conducting virtual research in the two domains. Both seminars noted that the archival directory together with NTE and ARGs are the prime examples of this achievement.

CENDARI's infrastructure is built partly with the aim of bridging the persistent divide between the archival and conceptual work of historians. The CENDARI enquiry environment acts as a tool to virtually integrate the dispersed sources by situating them in historical contexts while simultaneously accommodating and acknowledging the "knowledge populating" structured archive and library metadata (cf. D4.4). The dynamism and flexibility are inherent features of any digital research infrastructure; both characteristics are on the other hand sustaining metadata and ontologies.

3.4 Archival knowledge production vs. historical knowledge production and CENDARI's potential to bridge this divide

The innovative process in DH is reciprocal: advancements in digital methodology have greatly changed (and keep changing) archival historical research, while guiding towards discovery of the meaning of crucial resources and creation of original knowledge. According to our ES discussants as well as to our previous report (D4.1), historians across all periods regard the digitisation of online finding aids and of documents, of the possibility of full-text search, and the availability of online catalogues as a major asset of "digital turn". However, all seminar participants agreed that this new way of finding (as well as seeing) sources and producing historical knowledge is not diminishing the importance of traditional method. Moreover, as Dennis Mischke put it, ARGs in particular have the potential of bringing "the collective intelligence of experts and archivists to fellow researchers, students, early career researchers in related fields". At the same moment the question of whether there is a clear-cut distinction between "archival and researcher-produced data" uncovered a new potential for developing particular visualising function for the benefit of the wider community of scholars.

Archives and libraries "create" knowledge in the way of structuring, describing and classifying their collections. They are required to provide researchers with transparent information as well as with finding or obtaining documents. The "archival knowledge" is inhabited in finding aids, catalogues, classifications, and in the expertise of the archivists and the librarian. On the other hand, knowledge in humanities is derived from a hermeneutic procedure, which claims sound argumentation and a reservoir of previously gained insights in the discipline (D4.4).

CENDARI as a digital "ecosystem" and knowledge "producer" is separating its infrastructure from the model of the infrastructure as a "portal" to archival and library sources. Moreover, it is defining its research model as a meeting place that reflects both flexibility and generation of new scholarly knowledge. Current digital developments seem to point at a deep divide in relations between archivists and historians for their specific approaches to sources.

Throughout all the expert seminars, the advantages of a research infrastructure that allows historians, digital humanists, and archives to exist in dialogue with one another through the research process were acknowledged. As CENDARI evolves, the shape and nature of those dialogues and the precise sites within the infrastructure where such dialogues form, will be of use to future digital research infrastructure projects.

4. Conclusion

In general, the expert seminars in the medieval domain and the First World War domain demonstrated the adaptability of the CENDARI infrastructure when it came to successfully engaging scholarly communities with different levels of exposure to digital humanities tools in their respective research processes.

The medieval expert seminar participants appreciated Cendari's novel approach to collaboration and knowledge production, especially its close connection with TRAME. Especially when it came to medieval scholars at the workshop, there was a higher level of familiarity with the tools within research infrastructures like CENDARI and a greater awareness of the potential tools that exist for research. The WWI ES participants' evaluation of the CENDARI VRE was positive overall, but they raised a series of important questions with regard to the

features and functionalities not yet available (but envisaged) by CENDARI. The traditional practicing First World War scholars expressed doubts about the readiness to share their data in a system that was by default open (which CENDARI is not, since scholars have to consent to have their research data made public). In addition, the points at which the First World War scholars believed that CENDARI would be useful to their research tended to be at the beginning stages of research and after return from the archives. The desire on the part of World War I scholars to have CENDARI "test" cases available for demonstration which show how CENDARI was involved in the research process from beginning to end, underscores the relative lack of familiarity among World War I scholars with the potentials inherent in a research infrastructure like CENDARI. In the end, however, the fact that both domains viewed CENDARI as a potentially positive environment for future research projects reveals its potentially wide application within the historical community.

The DH expert seminar revealed, for obvious reasons, a much broader perspective on CENDARI and digital research infrastructures in general. The participants had long experience with digital research infrastructures, which made them concerned about the long-term sustainability of CENDARI, but also noted that CENDARI being brought under the DARIAH framework offered promise from this perspective. Some of the experts noted the importance of the diffusion of a certain degree of technical knowledge within the historical profession and frustration at the difficulty of effecting this change. In the end, CENDARI is, for these DH experts, an important link in the chain of what will be a long term project to anchor the digital humanities, the research infrastructures that they have produced, and the tools associated with such infrastructures to further the research process, in the broader humanities fields in general, especially that of history, one of its most popular and yet methodologically conservative fields.