



DELIVERABLE

Project Acronym: Europeana Cloud
Grant Agreement number: 325091
Project Title: Europeana Cloud: Unlocking Europe's Research via The Cloud

D3.1 - Personas, scenarios and use cases

Revision: 2.7

Authors:

Erik Duval, KU Leuven
Gonzalo Parra, KU Leuven
Marnix van Berchum, DANS-KNAW / Utrecht University
Hein van den Berg, Stichting VU-VUMC
Anja Jentsch, Open Knowledge Foundation Deutschland
Daniel Dietrich, Open Knowledge Foundation Deutschland
Nikos Manouselis, Ariadne Foundation
Giannis Stoitsis, Ariadne Foundation
Akis Tsekouras, Ariadne Foundation



Project co-funded by the European Commission within the ICT Policy Support Programme	
Dissemination Level	
P	Public
C	Confidential, only for members of the consortium and the Commission Services

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise.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

Revision History

Revision	Date	Author	Organisation	Description
0.1	17 July	Erik Duval	KU Leuven	First integration of working documents
0.2	19 July	H. vd Berg	VU	Adding background, rewriting, editing
0.3	20 July	Erik Duval	KU Leuven	Editing, adding context, adding section on next phase, restructuring appendix.
0.4	21 July	H. vd Berg	VU	Minor comments, additions.
0.5	22 July	N. Manouselis, G. Stoitsis	ARIADNE	Added context and section on process, restructured appendix, reformatted text
0.6	22 July	G. Parra	KU Leuven	Minor edits on scenarios 3.5 and 3.6.
0.7	22-7	H. vd Berg	VU	Substantial change to scenario “Arianna, Hein, the ARIADNE Finder and Pundit”; suggestions to rewrite certain portions of section 5 (in comments); general comments and minor edits.
0.8	22 July	G. Parra	KU Leuven	Final draft for internal review.



0.9	27 July	H. vd Berg	VU	Rewritten scenarios 3.9 and 3.10.
1.0	28 July	G. Parra	KU Leuven	Editing and rewriting of content based on the external reviewers suggestions.
1.1	29 July	Anja Jentsch	OKF DE	Readability improvements. Spell checking. Scenarios rewritten.
1.2	29 July	G. Parra	KU Leuven	Final version.
2.0	14 July	G. Parra	KU Leuven	Outline draft to include work from the year 2 cycle.
2.1	27 July	H. vd Berg	VU	Added section on methodology for year 2 cycle.
2.2	28 July	A. Jentsch	OKF DE	Reviewed Hein's section. Added year 2 part on conclusion. Reorganized tool section and added some tools.
2.3	28 July	G. Parra	KU Leuven	Integration of scenarios and personas. Editing of sections to reflect the work on Year 2.
2.4	30 July	H. vd Berg	VU	Rewriting of sections to reflect the work on Year 2. Finished conclusion, and section on methodology and activities for year 2 cycle.
2.5	31 July	A. Tsekouras	ARIADNE	Updated appendix with list of tools, small modifications in methodology from Y1.
2.6	31 July	G. Parra	KU Leuven	Final draft version.
2.7	5 August	E. Duval	KU Leuven	Finalisation.



Executive Summary

This document describes the personas, scenarios and use cases that WP3 has focused on in the first six months of the first year of eCloud (M1-M6) and the first six months of the second year (M13-M18). This work has been undertaken in collaboration with the DM2E project (and more specifically the Wittgenstein archives at the University of Bergen), the Axiom philosophy group at the VU University Amsterdam (in year 1), and a research community of musicologists (in year 2).

Through the personas and scenarios, we have identified a set of core problems and classes of tools and services that can address these problems for our target communities of Humanities researchers that work with Europeana content.

In the follow-up phases for the first and second year of eCloud, we configure, integrate and evaluate specific tools from the classes identified so far. In this deliverable, an appendix lists the candidate tools for these phases. This follow-up work for the first year of eCloud has been reported in Deliverables 3.2 and 3.3. For the second year of eCloud, the follow-up will be described in updates to these Deliverables, due in M21 and M24.



Table of Contents

1	Background and Context	7
2	Structure of the document	12
3	Year 1: History of Philosophical Ideas	13
3.1	Personas and Scenarios	13
3.1.1	Anita and Muse	13
3.1.2	Anna and Annotateit.....	13
3.1.3	Paul and Pundit.....	14
3.1.4	Pam and GlaMMap	14
3.1.5	John and TiNYARM.....	15
3.1.6	John and More!.....	16
3.1.7	Hein and ARIADNE Finder + Textus.....	16
3.1.8	Pam and Textus	18
3.1.9	Jeroen, ARIADNE Finder, GlaMMap, Pundit.....	19
3.1.10	Arianna, Hein, the ARIADNE Finder and Pundit.....	20
3.1.11	John and Researchr.....	21
3.2	Selected scenario, tools, and workflow	23
4	Year 2: Early Music Tools	29
4.1	Personas and Scenarios	29
4.1.1	Mike and Doug working on Orlandus Lassus and Philippus de Monte.....	29
4.1.2	John and Music-Mapper.....	30
4.1.3	Johnny and Music21-Online	30
4.1.4	Karin and Ariadne Finder.....	31
4.1.5	Maarten and Aruspix.....	31
4.1.6	Erik, Aruspix, and the Europeana Virtual Transcription Library	32
4.1.7	Klaas, Classifier, and MusicRadar	32
4.1.8	Hugo working on Ottaviano Petrucci.....	33
5	From Personas to Tools.....	35
6	Conclusion.....	41
7	Appendix: Tools for researchers in the Digital Humanities.....	42
7.1	Search tools	42
7.1.1	ARIADNE Finder	42
7.2	Visualization tools.....	43
7.2.1	TimeMapper	43
7.2.2	RelFinder.....	44
7.2.3	Muse	45
7.2.4	MappingPhilosophy/GlaMMap.....	46
7.3	Awareness tools.....	47
7.3.1	TiNYARM	47
7.3.2	More!.....	48
7.4	Annotation tools	49
7.4.1	Annotateit.....	49



7.4.2	TEXTUS.....	49
7.4.3	Pundit	50
7.4.4	OpenAnnotation	51
7.4.5	DocumentCloud	52
7.4.6	Researchr	52
8	Appendix: Additional tools specialized for musicologists.....	54
8.1	Search tools	54
8.1.1	IMSLP Petrucci Music	54
8.1.2	Digital Image Archive of Medieval Music (DIAMM).....	54
8.2	Creation of music sheet.....	55
8.2.1	MuseScore	55
8.2.2	CMME.....	57
8.3	Optical Music Recognition.....	58
8.3.1	Aruspix	58
8.4	Classification tools	59
8.4.1	WEKA	59
8.5	Music Analysis.....	60
8.5.1	Music21	60
8.5.2	The Humdrum Toolkit	61



1 Background and Context

The objectives of WP3 include (p.13 of 38 in the Description of Work):

- Development of services and tools that leverage Europeana content for use by researchers.
- Follow an iterative design process to identify typical personas and scenarios of thematic use of content that the Europeana Cloud tools and services as well as the Europeana Research Platform would be expected to support.

In order to reach these objectives, task 3.1 focuses on Personas, Scenarios and Use Cases:

In close alignment with the tasks in WP1 on user needs and existing tools, KU Leuven will lead on developing personas (descriptions of typical researchers that we address with this project), scenarios and use cases that describe in detail what kind of tool a researcher would typically use, and how that use would fit in his typical workflow. This will result in Deliverable 3.1 [M6, M18, M30]

This document is v.2 of deliverable 3.1, due in M18. It will be updated to reflect progress in our thinking as the project evolves again in M30. The basic idea is that we report here on personas and scenarios that we have developed in order to elaborate and consolidate our thinking on tools and services that leverage Europeana content for researchers.

A really nice, but rather lengthy, introduction to the use of personas and scenarios for this kind of purpose can be found in:

Nielsen, Lene (2013): Personas. In: Soegaard, Mads and Dam, Rikke Friis (eds.). "The Encyclopedia of Human-Computer Interaction, 2nd Ed.". Aarhus, Denmark: The Interaction Design Foundation. Available online at <http://www.interaction-design.org/encyclopedia/personas.html>

From that source:

Common understanding is that the persona is a description of a fictitious person [...] In the design process, we begin to imagine how the product is to work and look before any sketch is made or any features described. If the design team members have a number of persona descriptions in front of them while designing, the personas will help them maintain the perspective of the users. The moment the designers begin to imagine how a possible product is to be used by a persona, ideas will emerge. Thus, I maintain that the actual purpose of the method is not the persona descriptions, but the ability to imagine the product.

A shorter introduction, with some examples, is:

<http://www.slideshare.net/pennyhagen/introduction-to-building-and-using-personas-and-scenarios-in-design>



The gist of the idea behind personas and scenarios is that they help us to focus on the intended *user* of the tools and services, rather than on the technical challenges of how to develop those tools and services. Thus, the likelihood that we develop tools and services that are actually useful for and usable by the target user increases.

The main aim of WP3 is to develop tools and services that allow researchers to make optimal use of Europeana content. In the first year, we have focused on developing tools that benefit researchers in philosophy and digital humanities. To this purpose, we have worked more closely with two concrete communities of researchers interested in using Europeana content:

- The DM2E project¹ is also looking at “new tools and services for the re-use of Europeana data in the Digital Humanities”. Of particular interest for us is the Pundit² a ‘semantic annotation tool’ that enables the sharing of annotations between researchers. We have discussed (online and face-to-face) collaboration with this project and will focus in a first phase on a small community of researchers that focus on the Wittgenstein Archives at the University of Bergen³.
- The Axiom group of history and philosophy of logic, semantics and axiomatics⁴ at the VU University of Amsterdam studies the ‘Semantics and Axiomatics from Bolzano to Tarski against the Background of the Classical Model of Science’. Also with this group, represented directly in eCloud through the participation of Stichting VU-VUMC, we have interacted both online and face-to-face.

On the basis of this early interaction with the DM2E and Axiom Group, we have identified three core problems that many of these researchers face:

- (i) problems with navigating and identifying relevant (digital) content and with building corpora;
- (ii) a lack of user-friendly computational tools for conducting fine-grained textual research;
- (iii) a lack of appropriate tools and infrastructure that allow members of research groups to work collaboratively rather than individually.

In our work so far, we have identified

- (i) search tools to find relevant content, and
- (ii) visualization tools (geographical maps, timelines) that assist in overcoming problems (i).
- (iii) annotation tools (see below), incorporating information extraction methods and Linked Data techniques, that help solve given problems (ii),
- (iv) and social awareness and discovery tools (see below) that help tackle (iii).

¹ <http://dm2e.eu>

² <http://www.thepund.it>

³ see http://wab.uib.no/wab_nachlass.page/

⁴ <http://axiom.vu.nl/>



The appendix lists the relevant tools that we have considered. These tools have been further developed, linked to each other, and linked to Europeana content. This work has been extensively described in deliverables 3.2 and 3.3. As such, WP3 provides researchers in the Humanities with tools that assist them with various aspects of their work, ranging from the identification of textual resources to the analysis of these resources, and that allow them to fruitfully exploit content provided through Europeana.

In the second year of our work (M13-M18), we have focused on developing personas, scenarios, and tools for researchers working within musicology, specifically with Early Music. In the first phase of the second cycle (M13-M15), our aims were to understand the research practices and workflow of musicologists, and to identify digital tools and content employed by these researchers. To this purpose, we conducted desk research and organized numerous exploratory online meetings with Marnix van Berchum, a musicologist working on musical culture of the 15th and 16th century, and Marian Lefferts, a specialist in digital library content and cultural heritage. Both are involved in eCloud through activities in WP6 and WP4 respectively.

In the second phase of the second cycle (M16-M18), members of WP3 and Marnix van Berchum have organized three online meetings with a research community of musicologists interested in using digital tools and Europeana content. This community consists of the following researchers:

- Eliane Fankhauser (Utrecht University; working on late medieval polyphonic music from the Low Countries).
- Peter van Kranenburg (Meertens Institute; working on the Tunes & Tales project⁵).
- Laurent Pugin (Swiss office of RISM⁶; works on Optical Music Recognition (OMR)).
- Reinier de Valk (City University, London; working on the computational extraction of polyphonic structures from sixteenth century lute tabulatures)
- Frans Wiering (Utrecht University; works on Music Information Retrieval (MIR)).

In our first and second online meetings (12 and 14 May 2014), we discussed the workflow of the musicologists, as well as computational tools and content they use. After these meetings, we identified a number of challenges these researchers face when conducting their work⁷. These challenges include:

⁵ <http://www.ehumanities.nl/computational-humanities/tunes-tales/>

⁶ <http://www.rism.info/>

⁷ In our work, we have significantly profited from Frans's Wiering's work on user needs and challenges in Digital Musicology. See: <http://www.staff.science.uu.nl/~wieri103/presentations/WieringLondonDigitalMusicLabFinal.pdf>.



- (i) *Problems involving data creation.* Many musicologists require transcriptions of (historical) musical sources to modern scores in order to conduct their research. Although there are computational tools that allow users to create transcriptions and critical editions of musical sources, most transcriptions are done completely by hand. This is a laborious and time-consuming process.
- (ii) *Lack of digital corpora of music scores.* Although there exist a number of initiatives aiming to create standardized digital and symbolic representations of music scores (e.g., <http://music-encoding.org/home>), as well as music libraries offering access to symbolic and encoded representations of music scores, many musicologists still lack appropriate digital corpora of scores.
- (iii) *Information exchange and linking of data.* Musicologists often need to match and compare the original (historical) sources of music or music scores to existing transcriptions of these sources. For example, databases such as Europeana provide images of (historical) music sources (typically in PDF format). These databases typically do not, however, provide digital, symbolic, and encoded representations of music sources. Such representations are provided by other online databases (e.g., <http://josquin.ccarh.org/>). Linking these different types of data and data sources is a non-trivial task.
- (iv) *Retrieve and analyse contextual information.* For contextualizing the music that researchers work on, information has to be gathered from multiple other sources, besides the representations of the original sources or transcribed scores. For example, apart from music sources and scores, musicologists require access to general bibliographical and historical information. It is difficult to process, compare, and analyse these different and heterogeneous information sources. These activities are, however, core to the work of many musicologists

On the basis of this interaction, we identified tools that we thought would aid musicologists in their research, and we constructed personas and scenarios that describe how these tools could fit their workflow. The descriptions of these personas and scenarios have been evaluated by the research community of musicologists in a third online meeting (19 June 2014), and have been rewritten and improved on the basis of this evaluation.

Through our collaboration with the musicologists, we have identified the following kinds of tools as particularly relevant to their work:

- (i) *Optical Musical Recognition tools* that allow users to create transcriptions and critical editions of musical sources, and *music notation software* that allows users to create and edit sheet music. These tools assist in overcoming challenges (i) and (ii) above.
- (ii) *Search tools, digital repositories and archives, and visualization tools* that allow users to identify and aggregate (metadata of) digital and symbolic representations



of music, and that can be used to link these representations to image files of music sources. These tools assist in overcoming challenges (ii) and (iii).

- (iii) *Annotation tools, music analysis tools, and classification tools* that allow musicologists to analyse, process, classify, and enrich their data. These tools assist in overcoming challenge (iv).

Several of these tools will be further developed, linked to each other, and linked to Europeana content in the WP3 activities in the second part of the second year of eCloud (M19-M24).



2 Structure of the document

In sections 3 and 4, we list the personas and scenarios that we have developed in dialogue with our target user community of humanities researchers from the first and second year, respectively. Personas and scenarios are listed in no particular order and there is quite a bit of overlap in the scenarios: this is useful, as it helped us to identify the common issues mentioned in the previous section. Those common issues inform the next phase of WP3: the configuration, integration, and evaluation of specific tools. We conclude the sections by presenting the tools that will be used in the final phase of each year. In Section 5, we further analyse the process to follow in order to develop and adapt tools. This analysis is based on the scenarios presented in Section 3 and 4. Finally, Section 6 concludes this document and the appendix lists the candidate tools analysed for the different humanities researchers in the first and second year of eCloud.



3 Year 1: History of Philosophical Ideas

3.1 Personas and Scenarios

3.1.1 Anita and Muse

Persona

Anita is a post-doctoral researcher on Technology-Enhanced Learning in a computer science department of a European university. She is collaborating with colleagues from different European research units, in a Framework 7 EU project. She regularly travels to project meetings, workshops, and conferences throughout the EU. Anita is pretty tech savvy and works in a department that is well equipped with smartphones, laptops, large screens, and even a few multitouch tabletops.

Scenario

Anita is assisting her supervisor with the preparation of a new EU proposal. There are some partners that her team regularly collaborates with and they decide to invite three of those partners to join the proposal preparation. They then turn to the multitouch table and use Muse to explore whom else they have worked with, and, more importantly, whom their colleagues have worked with before. Together with her supervisor, she thus discovers a team that focuses on Computer-Supported Collaborative Learning (CSCL) in Norway. Anita remembers that she met one of members of that team earlier in the year. Through the visualisation, Anita and her supervisor discover that the Norwegian team has a close collaboration with the team in Lausanne that they had already invited for the new project. They thus decide to ask the Lausanne team about the Norwegian partners first. Moreover, her supervisor will travel to Madrid next week for a conference. With the tabletop visualization, they remember that they had previously worked with two research units in Madrid: Anita's supervisor sends an email to both of them to set up a meeting to discuss the new project proposal.

3.1.2 Anna and Annotateit

Persona

Anna is a research associate at the department of philosophy in Cambridge, UK. She is preparing a thesis on Shakespeare's eighteenth-century reception in England and France. She is interested to explore and apply ICT-based methods and tools to enable new research questions and is eager to exchange knowledge, expertise, methodologies, and practices across domains and disciplines. She is an active member of the Open Humanities Working Group of the Open Knowledge Foundation that meets online weekly to discuss Shakespeare's work and to explore new ICT-based research practices and tools for the digital humanities.

Scenario

For an essay on Shakespeare's thoughts on urbanism and the city Anna needs to scan a lot of documents against the topics of interest for her essay and is therefore also interested in the



thoughts of like-minded colleagues. She and her colleagues have used the Annotateit tool to share annotations on online documents before, and she now invites her colleagues to help her scan through a series of relevant texts for the topics of interest for her essay and share their annotations, comments and tags with her.

3.1.3 Paul and Pundit

Persona

Paul is a post-doctoral researcher on philosophy at the Wittgenstein Archives at the University of Bergen (WAB) in Norway. The focus of his research is on the linguistic aspects of Wittgenstein's work. WAB has published 5,000 pages of the Wittgenstein Nachlass Open Access on the Web and developed a Wittgenstein domain ontology for these texts. WAB is part of a Framework 7 EU research project, called Digital Manuscripts to Europeana (DM2E), where Paul is involved in testing new technologies for collaborative work with semantic annotations online.

Scenario

Reading Wittgenstein's work using Pundit allows Paul not only to annotate the texts, but also to look up various concepts and persons occurring in the texts.

Pundit allows Paul to add and share simple comments as well as semantic links to the Web of Data (as Freebase or DBpedia) for fine granular cross-references and citations using the WAB's Wittgenstein domain ontologies. Its named-entity recognition features allow for disambiguous linking of real world entities mentioned in texts to their entries in Linked Data sets and thus deriving more information on this entity as well as the possibility to further traverse the Web for more fine-grained information.

Furthermore the annotations help finding other relevant texts on a specific topic or concept in Wittgenstein's work.

3.1.4 Pam and GlAMMap

Persona

Pam is a 25 years old PhD-student at the department of philosophy at the University of Utrecht, the Netherlands. She has received her Master's degree at the University of Groningen, where she has written a thesis on Kant's use of the concept 'transcendental logic' in the second edition of the *Kritik der reinen Vernunft* (1787). Pam has enjoyed a classical training in the history of philosophy, which means that she has learned to provide highly detailed analyses of a very small number of texts. At Utrecht, Pam is a member of an NWO Vidi-project in which she is supposed to investigate how the concept of 'logic' is used within German philosophy books published around 1800.

Pam is ambitious, hard working, and desperately wishes to write a good dissertation. Like many of her colleagues at the philosophy department, she is an avid user of technology. She is



a member of various social media sites, she owns a smartphone and a laptop and she is constantly online. Pam is eager to use and explore ICT-based methods to push further her research and has recently developed an interest in digital humanities. Nevertheless, her knowledge of ICT-based methods and tools is quite limited. When conducting her research, Pam mostly uses Google Books and Google Scholar.

Scenario

Having just started research for her dissertation, Pam is overwhelmed by the number and complexity of the philosophical and scientific texts that she needs to read and understand. Although her knowledge of the works of Kant is adequate, she lacks an overview of the works on logic published around 1800 that are relevant for her research. In order to identify publications that are important for her research, Pam uses GlaMMap. Using this tool, she obtains an interactive geo-spatial visualization of bibliographic metadata of logic books published in Europe between 1795 and 1805. She uses the tool to identify and order various (types of) books that she needs to study for her dissertation. She identifies publications on ‘transcendental logic’ that have been published in Jena and Leipzig, she identifies traditional books on term logic published in Halle, and she identifies works on psychology, anthropology, and logic published in Heidelberg. Through exploring the visualization, Pam is also able to quickly distinguish between the relatively small number of books published by famous philosophers in this period, and the relatively large number of textbooks and lecture scripts concerning logic. Since these latter books are easier to read and have also been little studied, Pam decides to start her research by focusing on these books.

3.1.5 John and TiNYARM

Persona

John is a PhD candidate in the doctoral school of human sciences at the Vrije Universiteit Brussel. After finishing his master in arts, he just started to work pursuing a PhD degree in philosophy and history of education. During his master thesis work, he got familiar with tools used by researchers, such as: references managers (Mendeley and BibTeX) and digital libraries.

Scenario

John has a clear idea of his dissertation topic and started with the literature review about the subject. As he is new in the research community and he doesn’t know which colleagues from other European universities may be working on the same topic. Colleagues from his research team suggested him to use TiNYARM. His research group employs this tool to share publications read by them, to know what other research groups are reading, and his advisor uses it to suggest articles to his PhD students. As soon as he joined the tool, John saw a recent suggestion from Anna (a researcher from Germany) to one of his colleagues; this article was quite interesting for his research. He also opened the TiNYARM profile page of his advisor and skimmed the different papers his promoter read over the past months. John found some



articles that were relevant for his research. Looking at his group activity, John got an idea of what everybody is interested in, based on the different papers they read or skimmed.

3.1.6 John and More!

Persona

John is a PhD candidate in the doctoral school of human sciences in the Vrije Universiteit Brussel. He is doing a PhD in philosophy and history of education. During his master thesis work and his first year in his PhD, he got familiar with tools used by researchers, such as: references managers (Mendeley and Bibtex), digital libraries, SlideShare, and academic social networks (LinkedIn, Academia.edu, and ResearchGate).

Scenario

After his first PhD year, John and his advisor get a paper accepted in an international conference. This will be his first time attending a conference. On the first day, John gets the conference programme and tries to figure out which sessions to attend. While creating the list of sessions, he notices that a QR code is available in the conference programme. Using his smartphone, John scans this code and is redirected to the More! mobile web application showing different links to social networks that the presenter is actively participating in. After seeing the presenter profile and checking his publications and recent tweets, John gets a good idea of what the presenter worked on before and which are his current research interests. John decides to meet the presenter later and sends him a mail using More!. After using this application for a few presentations, he is a bit frustrated with some of the other sessions that do not provide a More! profile of the presenter and that he actually has to 'google' them for some minutes in order to obtain the same information as he gets from More!.

3.1.7 Hein and ARIADNE Finder + Textus

Persona

Hein is a post-doctoral researcher working on the history of philosophy, the history of biology and their interplay. In particular, he is interested in 18th century philosophical ideals of science and their influence on 18th century life sciences. His research has been recently focused on debates in philosophy of history and historiography, where he investigates and needs access to philosophical document and digital media corpora and their visualisations.

Scenario for ARIADNE Finder

Hein participates in a group of post-doctoral researchers and academics interested into the history and philosophy of logic, semantics, and axiomatics. This group explores the concept of semantics and axiomatics in the history of philosophy, exploring ideas coming from philosophers like Bolzano and Tarski against classical models of science. Hein has already been using some interesting visualisation tools for relevant document and literature corpora that are working over a bibliographic metadata aggregator/network that his group has set up.



He would like to enhance the web site of this group with a search interface that will allow the group to discover and navigate relevant digital resources that are coming from other large aggregators, such as: material on the history of philosophy collected by the Europeana aggregator (<http://www.europeana.eu>), relevant literature that has been indexed and aggregated from various European national libraries through The European Library (<http://www.theeuropeanlibrary.org>), as well as presentations, notes, slides, and other relevant educational resources that come from a global network of educational repositories called GLOBE (<http://globe-info.org>).

Hein would use the ARIADNE Finder technology to carry out the following steps:

- Define the metadata facets that his group would like to use in order to search and browse through the various digital resources on the history of philosophy, and define the desired properties of a local metadata index that will collect periodically relevant material that can be found in the various aggregators.
- Design the information architecture of the content search pages using a simple pre-existing template that will allow him to integrate the search functionalities into the look and feel of the group's web site.
- Run the tool that will set up and create the local metadata index (also with a mechanism for periodic re-harvesting/ingestion and re-indexing from the various aggregators) running over a cloud-hosted virtual machine through a simple web-based interface. The tool will also allow him to select which thematic keywords or specific collections he would like to include when retrieving descriptions of content from the three external aggregators.
- Get as an outcome of the tool a simple HTML code that he can easily embed into his group's web site in order to get the new ARIADNE-powered search pages easily set up and running with no further technical effort.

Scenario for Textus

Hein would like to incorporate in the web site of the group a functionality that would allow the post-doctoral researchers to select digital documents discovered through the ARIADNE Finder and to carry out further discussions and analyses using Web-based open annotation features - in a similar way that the OpenPhilosophy.org group (<http://beta.openphilosophy.org>) is working with public domain philosophy texts. For this purpose he incorporates the Textus tool into the web site in a way that allows the group to:

- Select/bookmark some interesting digital resource coming from one of the aggregators like Europeana or The European Library and add it to a shared common space.
- If its license scheme allows it, saving a local copy of the digital resource and opening it in the Textus viewer.
- Providing annotations on various parts within the text under analysis and saving them into the shared space together with the digital resource.



- Generate as a separate new document all the produced annotations in a digital format, including a reference to the original document into its metadata description - saving it in the group's shared document repository.
- Enhance the metadata record ingested from an external source with the generated annotations of the group and saving the new version of the record into the group's shared document repository (either as a new, enriched version of the original authority record or as a new metadata record that stores only annotations and other user actions using some scheme like CAM).

3.1.8 Pam and Textus

Persona

Pam is a 29 years old postdoc working at the department of history and philosophy of science at the University of Cambridge (UK). She has a Master's degree in Computer Science and in Philosophy, both obtained at the University of Edinburgh (Scotland). In her doctoral dissertation Pam studied the philosophy of mathematics of the eighteenth-century mathematician Leonard Euler. At Cambridge, Pam is part of an ERC-Starting Grant Project that aims to provide a comprehensive overview of the history of 18th and 19th century philosophy of mathematics. For her research, it is essential that Pam collaborates intensively with philosophers, historians, and mathematicians from different European research units. She and her partners require easy access to historical documents and wish to collaboratively study such documents. Apart from doing her research, Pam also teaches an annual course on the history of mathematics at Cambridge.

Pam knows a lot about existing ICT tools and methods that can help with her research. Thanks to her efforts, her research group and various European partners use project management software (Basecamp) to coordinate projects, they use reference managers (Mendeley) to create shared bibliographies, and they compile large databases using digital resources obtained via Europeana, Google Books, archive.org, and other online repositories.

Scenario

Pam and one of her colleagues working at the University of Duisburg-Essen (Germany) decide to write a joint paper on the reception of Euler's mathematical works at the end of the 18th century, focusing in particular on the work of Immanuel Kant. They decide to use TEXTUS in order to share, study, and collaboratively annotate historical documents that are relevant for their paper. Through their contacts at Duisburg-Essen, they obtain plain-text files of Kant's published writings (<http://www.korpora.org/kant/>). After uploading these writings, they use TEXTUS to identify, annotate and discuss passages within the Kantian corpus that concern mathematical topics. Through sharing their annotations with other project partners and interested researchers, they further obtain valuable information and references that help them in writing their paper.



Apart from using TEXTUS for research purposes, Pam also uses the tool while teaching her undergraduate course. Students following her course are supposed to study historical texts by collaboratively commenting on these texts. They also have to correct and grade each others annotations. Pam notices that by working in this manner her students find it easier to read and understand scientific and philosophical texts.

3.1.9 Jeroen, ARIADNE Finder, GlaMMap, Pundit

Persona

Jeroen is a 25 years old PhD student at the department of philosophy at the Vrije Universiteit Amsterdam. He has Master's degrees in philosophy and intellectual history, both obtained at the University of Groningen. In Amsterdam, Jeroen is a member of an NWO VICI Project that aims to rewrite the history of eighteenth-century philosophy of biology. Jeroen and his colleagues need to identify many unknown historical documents and they wish to collaboratively study such documents. The members of Jeroen's research group are pretty tech savvy. They use Whatsapp to communicate, they use Document management & Intranet software (Papyrs) to coordinate projects and write papers, they use reference managers (Zotero) to create shared bibliographies, and they compile large databases using digital resources obtained via Europeana, Google Books, archive.org and other online repositories. Jeroen is keen to use ICT tools while conducting his research and he hopes that they enable him to work in a more efficient manner.

Scenarios

Compiling and organizing primary sources (step 1)

Jeroen and several other PhD students have to write a paper on methodological debates within eighteenth-century natural history. In order to write this paper, they need to compile a database of relevant primary sources i.e., books on natural history published in the eighteenth-century. When compiling his database of primary sources, Jeroen uses the ARIADNE Finder, a search interface that allows him to discover and navigate digital resources that have been collected by the Europeana aggregator, The European Library, digital resources compiled by Google Books, and resources contained within other repositories.

For his paper, Jeroen decides to search for the term "Natural History" from 1700-1800 within the European Library. He obtains over 9000 metadata records of books on natural history published in the eighteenth-century. Using the ARIADNE Finder, Jeroen exports these records to Zotero and compiles a large database of primary sources.

Navigating the primary sources (step 2)

It is difficult to navigate 9000 works. After having obtained his 9000 metadata records, Jeroen downloads these records and feeds them to GlaMMap in order to obtain an interactive geo-spatial visualization of these bibliographic metadata. He uses the tool to navigate the metadata and he orders books on the basis of time and place of publication. By doing this he can



identify various (types of) books that he needs to study for his paper. After ordering his books, he decides to write a paper on the species concepts of Linnaeus and Buffon.

Identifying secondary sources (step 3)

Having identified his primary sources, Jeroen needs to identify relevant secondary sources, i.e., works on the history of biology published in the twentieth-century. He again uses the ARIADNE Finder in order to aggregate metadata of contemporary (20th century) books discussing Buffon and Linnaeus. He collects metadata from the Europeana aggregator, the European Library, Pubmed, Worldcat, Google Scholar, and other possible sources. He selects all the relevant literature and exports his selections to Zotero. He has now created his database of secondary sources.

Navigating the secondary sources (step 4)

Having obtained metadata records of his secondary sources, Jeroen visualizes these records using GlaMMap (the procedure is identical to step 2). He obtains an interactive geo-spatial visualization of his bibliographic data. He now has two visualizations: one of his primary sources and of his secondary sources. Jeroen is able to compare both visualizations and is able to find several interesting relations between them.

Studying the relevant literature (step 5)

After identifying the specific literature that is to be studied, Jeroen and his colleagues decide to use Pundit to collaboratively study a number of selected texts. They obtain texts of Buffon and Linnaeus and semantically annotate them by relating sections of these documents to Linked Data resources such as DBpedia and Wikidata. They use named-entity recognition in order to identify names and concepts occurring in the text, and link these items to important online resources. Thus, for example, fragments of Buffon's Natural History are related to relevant Wikipedia articles, to more technical articles on the history of biology in the Stanford Encyclopedia of Philosophy (e.g., evolution – see <http://plato.stanford.edu/entries/evolution/>), and other educational resources (e.g., history of biology – see <http://www.ucmp.berkeley.edu/help/topic/history.html>). In this manner, Jeroen and his colleagues can easily navigate and study numerous quite technical texts.

3.1.10 Arianna, Hein, the ARIADNE Finder and Pundit

Persona

Hein is a post-doctoral researcher working on the history of philosophy, the history of biology and their interplay. He is interested in 18th century philosophical ideals of science and their influence on the 18th century life sciences. Arianna is professor of philosophy of language and conducts research on the history of logic, metaphysics and digital humanities. Arianna and Hein are members of a research team investigating the history of logic and science. Both use a variety of digital tools to enhance their research, such as reference managers (Zotero), WhatsApp, and various other tools.



Scenario

Arianna and Hein have, for research purposes, scraped over 7,000 bibliographic records from Wilhelm Risse's *Bibliographica Logica* (<http://www.olms.de/>). As a result, they have obtained bibliographic metadata of over 7,000 books concerning logic published between 1700 and 1940. Both are currently attempting to obtain a comparable database of logic books by aggregating metadata from Worldcat, Europeana, The European library, and other repositories.

Arianna and Hein wish to write a paper on logic books that have been little studied by researchers on the history of logic. In order to write this paper, they need to compare the historical data from Risse (or similar data obtained via other large aggregators) to bibliographic data and references contained in philosophy repositories (e.g., Philpapers – see <http://philpapers.org/>), entries on history of logic published within Wikipedia, entries in the Stanford Encyclopedia of Philosophy (<http://plato.stanford.edu/>), the Internet encyclopedia of Philosophy (<http://www.iep.utm.edu/>), and possibly results provided by Google Scholar and Worldcat. Arianna and Hein use the ARIADNE Finder to obtain bibliographic metadata of thousands of logic books published from the 18th to the 20th data. Using Pundit, they link this data (e.g., name of author) to resources found within Wikipedia, the Stanford Encyclopedia of Philosophy, the Internet Encyclopedia of Philosophy, Wikidata and other Linked Open Data sets. In this manner, they are able to reconstruct which logic books are often mentioned in popular resources and which logic books are rarely mentioned. This allows them to give a reasoned estimate on which logic books have been little studied by contemporary researchers.

3.1.11 John and Researchr

Persona

John is a PhD candidate in the doctoral school of Human Sciences at the Vrije Universiteit Brussel. After finishing his master in arts, he just started to work pursuing a PhD degree in philosophy and history of education. During his master thesis work, he got familiar with tools used by researchers, such as: references managers (Mendeley and Bibtex) and digital libraries.

Scenario

John has to start doing his literature review about the subject. He has downloaded a lot of PDFs (usually with generic or unidentifiable names) and started to take notes using different software and on papers. After realizing the challenges of organizing these, he got really interested in how to improve the management all the PDFs, citation metadata, clippings, notes, and ideas. He started to use Researchr, which was an aggregation of different tools such as: BibDesk (to manage citations), Skim, (to make comments and highlight text), DokuWiki (to manage notes), and scripts; and use it as his personal open publications management system. John uses the uses different scripts of Researchr (via keyboard shortcuts) to: import citations from the browser directly to BibDesk; or send notes, highlights, and images from Skim to DokuWiki.





3.2 Selected scenario, tools, and workflow

As mentioned in section 1, we have identified three core problems for the researchers working on the history of philosophical ideas:

1. problems with navigating and identifying relevant (digital) content and with building corpora;
2. a lack of user-friendly computational tools for conducting fine-grained textual research;
3. a lack of appropriate tools and infrastructure that allow members of research groups to work collaboratively rather than individually.

In the first year of eCloud we aimed to address these problems through a combination of the following tools (a more technical explanation of these tools is present in the Appendix):

1. Search tools to find relevant content: a specifically configured version of the ARIADNE Finder is one candidate that we have started to experiment with – see illustration below. The Finder lets users search content and browse the results of different repositories, hiding protocols and standards from them.

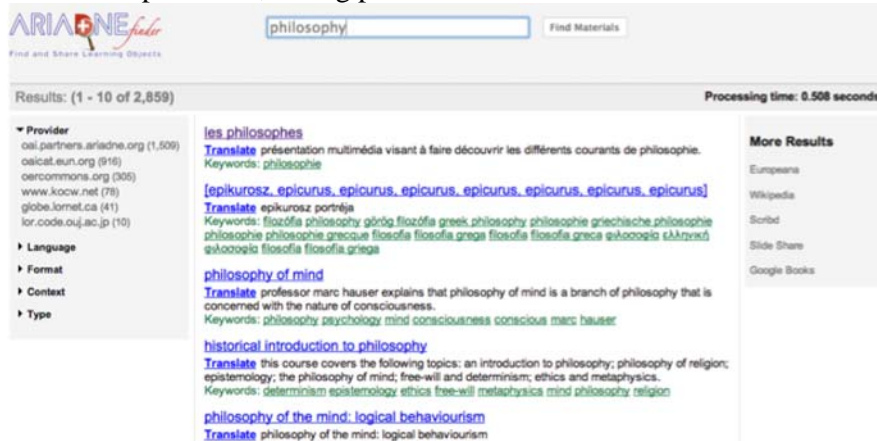


Figure 1: The ARIADNE Finder

2. Visualization tools (geographical maps, timelines) that assist in exploring the search results. For this purpose, we have experimented with geo-spatial visualizations, such as Muse and GlaMMMap; which present co-authorship on a multitouch tabletop and bibliographic metadata from books in logic from 1700-1940 on an interactive map, respectively. Also, we experimented with a timeline visualization of a term or word. Early experimentation screenshots are shown below.

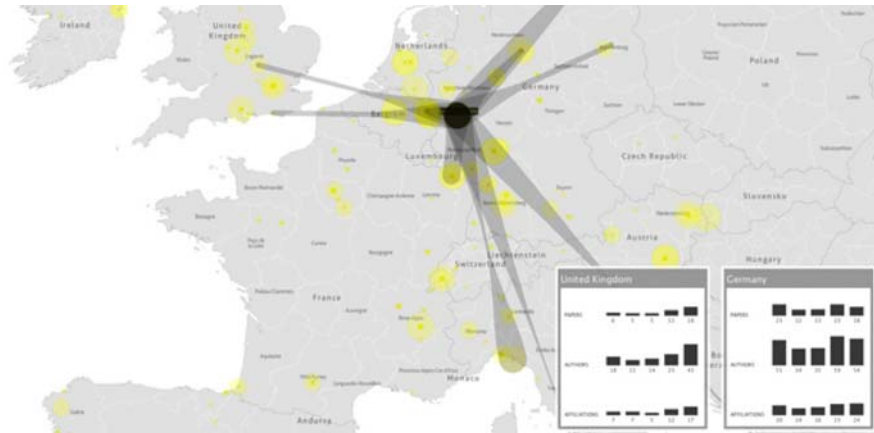


Figure 2: Muse

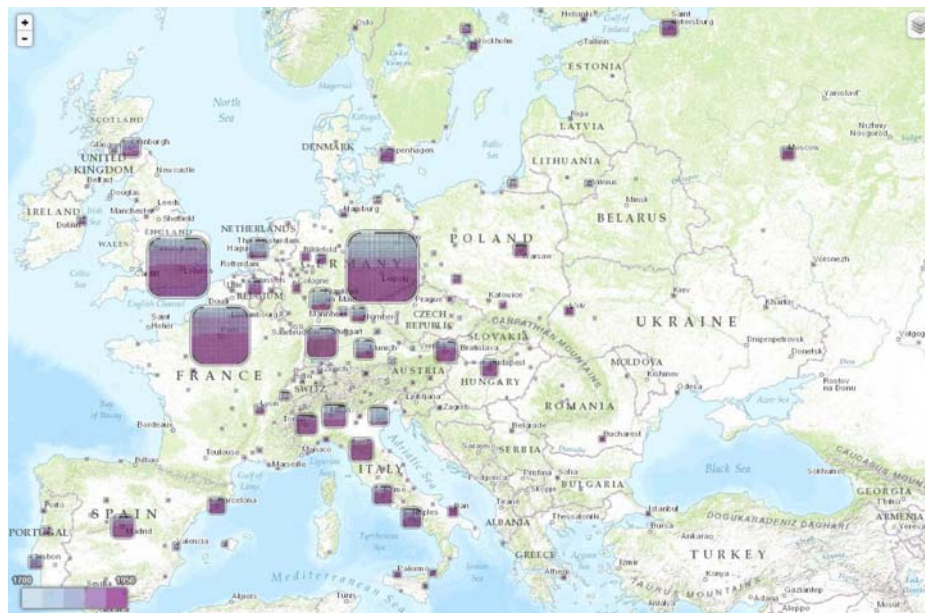


Figure 3: GlaMMap



Wittgenstein Nachlass Timeline

Welcome to the Wittgenstein Nachlass Timeline.
You can search for a word (e.g. 'Derin dies') and see when these words appear in Wittgensteins' work over time.

This app is created by [K.U.Leuven HCL Lab](#) and is powered by [MongoDB](#) and [Node.JS](#).
The data comes from [the Wittgenstein Archives at the University of Bergen \(WAB\)](#).

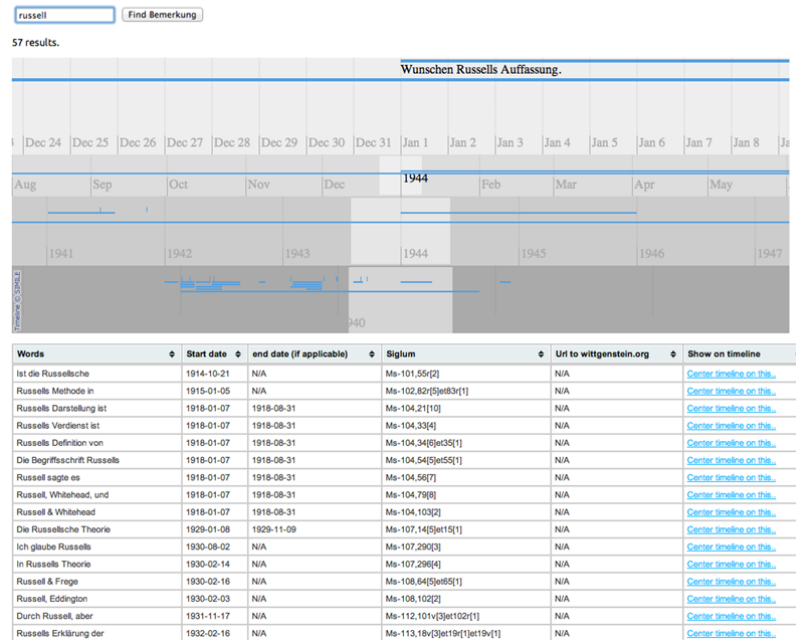


Figure 4: Wittgenstein Nachlass Timeline visualization

3. Annotation tools, incorporating information extraction methods and Linked Data techniques, in order to help researchers to share detailed comments and annotations. We have experimented with AnnotateIt and Pundit, which are web applications that allow users to easily annotate web content. As an extra feature, Pundit also allows the creation of semantically structured data that will enrich the Web of Data. Finally, we also evaluated TEXTUS, a web platform that allows the creation of shared and collaborative annotations between researchers. Again, we include some screenshots of our early experimentation in this area.



Annotator

[Code](#) / [Download](#) / [Wiki](#) / [Issues](#)

The Annotator Demo

Annotator is a Javascript shim that you can insert into any page, allowing you to select and annotate text, images, or (almost) anything else.

How do I use it? We think you'll get the hang of it pretty quickly. It's running right now on this page — try it out by selecting some text and clicking on the note icon. Enter a note and press "Enter →"; you should see your annotation as a highlighted piece of text. Hover over the highlight with your mouse to view, edit and delete the annotation.

What happens to my annotations? The annotations on this page are being sent to [AnnotateIt](#) an online web service for storing annotations. However it's simple to set up your own store. Check out the [Storage](#) and [Annotation Format](#) wiki pages for more details.

Figure 5: AnnotateIt: the Javascript annotator

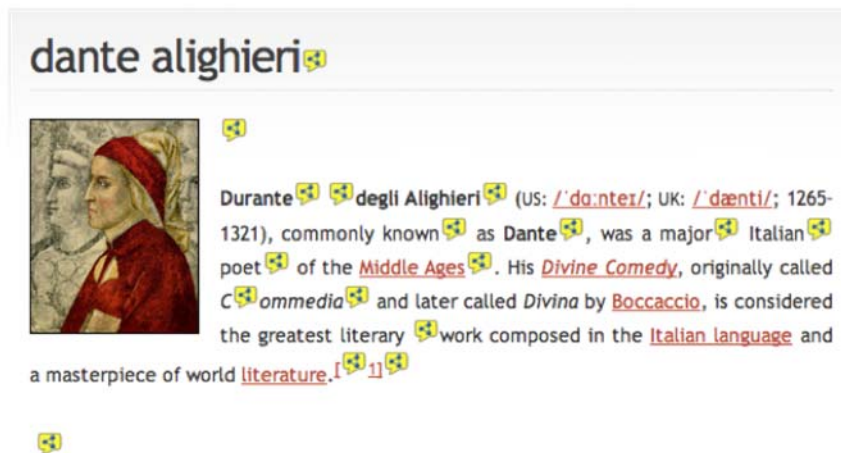


Figure 6: Pundit



OpenPhilosophy.org Home All Texts Login / Register

Index

THE PRINCE by Niccolò Machiavelli (translated by William K. Marriott)

DEDICATION

To the Magnificent Lorenzo Di Piero De' Medici:

Those who strive to obtain the good graces of a prince are accustomed to come before him with such things as they hold most precious, or in which they see him take most delight; whence one often sees horses, arms, cloth of gold, precious stories, and similar ornaments presented to princes, worthy of their greatness.

Desiring therefore to present myself to your Magnificence with some testimony of my devotion towards you, I have not found among my possessions anything which I hold more dear than, or value so much as, the knowledge of the actions of great men, acquired by long experience in contemporary affairs, and a continual study of antiquity; which, having reflected upon it with great and prolonged diligence, I now send, digested into a little volume, to your Magnificence.

And although I may consider this work unworthy of your countenance, nevertheless I trust much to your benignity that it may be acceptable, seeing that it is not possible for me to make a better gift than to offer you the opportunity of understanding in the shortest time all that I have learnt in so many years, and with so many troubles and dangers; which work I have not embellished with swelling or magnificent words, nor stuffed with rounded periods, nor with any extrinsic allurements or adomments whatever, with which so many are accustomed to embellish their works; for I have wished either that no honour should be given it, or else that the truth of the matter and the weightiness of the theme shall make it acceptable.

Nor do I hold with those who regard it as a presumption if a man of low and humble condition dare to discuss and settle the concerns of princes; because, just as those who draw landscapes place themselves below in the plain to contemplate the nature of the mountains and of lofty places, and in order to contemplate the plains place themselves upon high mountains, even so to understand the nature of the people it needs to be a prince, and to understand that of princes it needs to be of the people.

Take then, your Magnificence, this little gift in the spirit in which I send it; wherein, if it be diligently read and considered by you, you will learn my extreme desire that you should attain that greatness which fortune and your other attributes promise. And if your Magnificence from the summit of your greatness will sometimes turn your eyes to these lower regions, you will see how unmeritedly I suffer a great and continued malignity of fortune.

CHAPTER I -- HOW MANY KINDS OF PRINCIPALITIES THERE ARE, AND BY WHAT MEANS THEY ARE ACQUIRED

All states, all powers, that have held and hold rule over men have been and are either republics or principalities.

Principalities are either hereditary, in which the family has been long established; or they are new.

The new are either entirely new, as was Milan to Francesco Sforza, or they are, as it were, members annexed to the hereditary state of the prince who has acquired them, as was the kingdom of Naples to that of the King of Spain.

Activity Stream:

- GenAssem Test
- prason.sharma@systematixindia.com Lovely
- danieldesposito@huridocs.org [revisable] ping = man this is cool
- danieldesposito@huridocs.org oh lala j'aime bien ca
- Alexandre Ubaldo Testando esse sistema.
- Andrew Magiozzi Test annotation to see how this works. I like the interface generally.
- prason.sharma@systematixindia.com hello
- daniel.p.dietrich@gmail.com this comment is public.
- djames.breuer@gmail.com test
- pepper [revisable] whoa, that was new

Previous Next

Figure 7: TEXTUS and OpenPhilosophy.org

4. Social awareness tools: enable researchers to be more aware of the research activities of their colleagues and peers – see below for an early experimentation screenshot with the use of activity streams listing the behaviour of users.

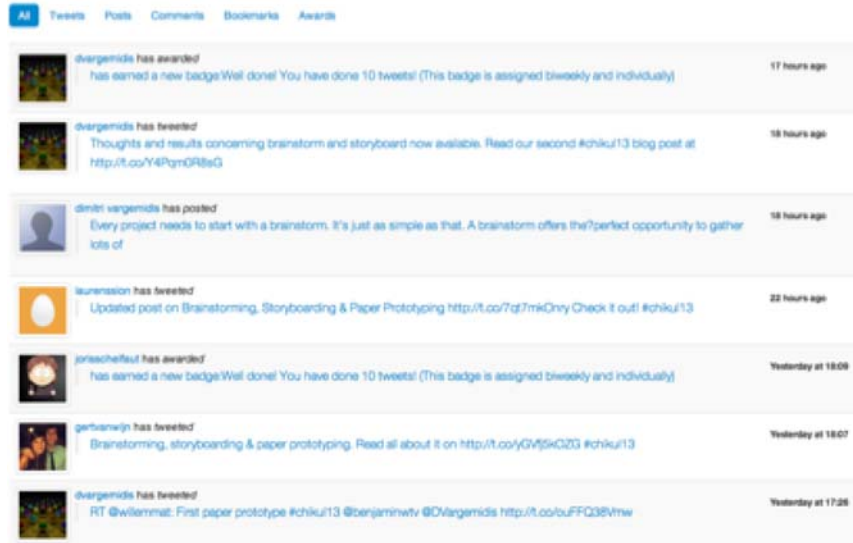


Figure 8: TiNYARMs' Awareness Stream

A generic workflow for these tools could be the following: a Humanities researcher wants to find more about 'Bolzano and Tarski' and uses the ARIADNE Finder to get content to be studied. After that, he tries to visualize them on a timeline or using a interactive map (or other visualization tools) to further filter the content and get a better overview of the different resources found on Europeana. After selecting the most interesting resources, these could be annotated using Pundit (or other). Finally, all the different actions made during this process (search, visualize, explore, annotate, download) are captured and presented via the Activity Streams tools.

In the second part of the first year of eCloud (M7-M12), we have linked several of these tools to each other, and to Europeana content. More specifically, we have created a mash-up of the ARIADNE Finder (a search tool), TimeMapper (a visualization tool), and Activity Stream (a social awareness tool). This mash-up was positively evaluated and was taken to assist researchers with (i) accessing *known* relevant research content or resources; (ii) with discovering *new* research content or data; and (iii) with working *collaboratively* rather than individually. We eventually did not include annotation tools in our mash-up since the philosophers often lacked corpora that allow for annotation by means of tools such as Annotator and Pundit. Work on the mash-up is described in more detail in deliverables 3.2 and 3.3, and in a forthcoming publication⁸.

⁸ H. Van den Berg, G. Parra, A. Jentzch, A. Drakos, and E. Duval. "tudying the History of Philosophical Ideas: Supporting Research Discovery, Navigation, and Awareness". In: Proc. of i- Know '14. Graz, Austria: ACM, 2014. Accepted.



4 Year 2: Early Music Tools

4.1 Personas and Scenarios

4.1.1 Mike and Doug working on Orlandus Lassus and Philippus de Monte

Persona

Mike is a musicologist who is working as a post-doc at the University of Leuven. He is particularly interested in Orlandus Lassus and Philippus de Monte. Doug is a doctoral student at the University of Utrecht who is focusing on the same two composers. Mike and Doug both participate in a small EU project with additional partners in Spain and Poland.

Scenario

Mike has been working on a manuscript he discovered recently in the library of Bologna. He has a hunch that the manuscript may be an unknown piece by Lassus or De Monte or a contemporary. The manuscript is just a fragment of a large piece. Mike has obtained high-resolution photos of the manuscript while he was in Bologna for a research stay.

This morning, Mike uploads the photos to the eCloud environment. A first rough transcription into musicXML⁹ is created automatically and Mike inspects the transcription through MuseScore¹⁰. Mike then searches for similar transcriptions in the AriadneFinder. The results are displayed on a map, on a timeline and as a Google-like list of results. Mike filters the results to the time period that interests him and discovers that there are three existing manuscripts that may be of relevance, one of which was created in Munich, where Lassus lived.

Mike uses Aruspix¹¹ to compare the different original manuscripts as well as their transcriptions. For two manuscripts, a sound recording is available: listening to these recordings helps Mike to get a clearer mental picture of the music represented in the manuscripts.

While he is doing this, he receives a message from Doug, who had discovered what Mike was doing through the eCloud activity stream. Mike is excited about his findings and sends a message back to Doug about his hypothesis that three of the fragments may be part of the same original piece. In a brief videoconference, they discuss how they could validate this hypothesis.

Mike needs to wrap up as he has a lunch appointment. He uploads the Bologna manuscript to IMSLP¹², so that his peers can also make use of the results of his work.

⁹ <http://www.musicxml.com>

¹⁰ <http://musescore.org/>

¹¹ <http://www.aruspix.net>

¹² <http://imslp.org/>



4.1.2 John and Music-Mapper

Persona

John holds PhDs from the University of Edinburgh in Artificial Intelligence and in Musical Composition. His research career covers computational linguistics, computational logic, computational modeling of music perception and cognition, and computational creativity.

Scenario

During his PhD, John created a personal collection of his transcriptions from 15th century music using CMME-XML, as his favorite tool to work with was the CMME software¹³. Currently, he wants to collaborate with other researchers and provide access to his collection, but they only use the Music Encoding Initiative (MEI) schema¹⁴. Thus, he decided to try a tool suggested by one of his colleagues, the Music-Mapper, which will transform his CMME-XML files (using the up-to-date XSLT transformations) to the MEI schema. By transforming the XML files, he discovered some missing features, and wrote the appropriate XSLT templates and contributed his code to the Music-Mapper.

4.1.3 Johnny and Music21-Online

Persona

Johnny is a PhD candidate in the School of Music from the University of Victoria. In the past month, he finished his master and now he just started to work, pursuing a PhD degree in musicology. During his master thesis work, he completed a thesis on hexachordal segmentation of Italian madrigals. His current research interests include musical segmentation and categorization, form, and schema theory.

Scenario

For his PhD, Johnny is trying to analyze music in symbolic form, obtained from his PhD promotor. While his colleagues use Music21¹⁵ for such task, he does not have a strong background in programming languages, such as Python. One of his colleagues mentioned that there is a Music21 online project that he may want to try. Music21-Online provided him with the possibility to just upload his XML files via the tool and try different Music21 scripts already created by the community of musicologists using the tool. By trying a couple of scripts, he noticed that, in order to find modal and hexachordal patterns within polyphonic structures, he needed a combination of three available scripts, so he decided to merge them and create his own, which he later submits to the database of available scripts, so that colleagues can take advantage of his work.

¹³ <http://imslp.org/>

¹⁴ <http://music-encoding.org>

¹⁵ <http://web.mit.edu/music21/doc/about/index.html>



4.1.4 Karin and Ariadne Finder

Persona

Karin is a 20 year old BA student of musicology at the University of Amsterdam. She is interested in historical musicology, especially in the history and development of baroque music. Like most students, she owns a laptop, a smartphone, and she is a member of various social media sites, such as Facebook and Academia.edu. Nevertheless, Karin has little knowledge of ICT tools and methods that may help her with her studies and research. Her use of ICT tools when studying is often limited to standard word processors (Microsoft Word), reference managers (Zotero), and Google, which she uses to find books and articles.

Scenario

Karin is conducting a research assignment on baroque music. For this assignment, she needs to identify and transcribe a number of eighteenth-century music scores of works by composers such as Bach, Händel, and Scarlatti. Karin needs to identify historical prints of works composed in the eighteenth century, but she also wishes to know whether the works she is studying have been included in modern scholarly or critical editions. Identifying historical prints of works and matching these prints to works contained in critical editions is a time consuming affair. Karin uses the Ariadne Finder, a personalized micro site that can be used to search and discover resources, to help her with these tasks. The Finder contains predefined search categories linking to (i) historical music sheets of works by the baroque composers she is interested in (records contained, e.g., in Europeana or RISM), and (ii) links to critical editions containing works of these composers (records provided, e.g., through Google Books and Oxford Music Online). By comparing the search results of (i) and (ii), Karin can match scholarly editions to historical prints more quickly than she could before.

4.1.5 Maarten and Aruspix

Persona

Maarten is a PhD student at Oxford University and specializes in musical culture of the sixteenth and seventeenth century. He obtained his Master degree in digital musicology at King's College London and has only recently started work on his PhD thesis. In Oxford, Maarten is part of a research group that studies written renaissance music. Maarten is tech savvy. He has expert knowledge on popular tools and toolkits used in musicology, such as Humdrum¹⁶ and Music21, and he is familiar with programming languages such as PHP and Python.

Scenario

In order to apply computational tools to written renaissance music, Maarten requires appropriate digital corpora. Such corpora are, however, hardly available. Historical music sources are provided online as digital images, and these images are often of low quality.

¹⁶ <http://www.humdrum.org/>



Maarten uses Aruspix to build his corpus of renaissance music. Aruspix allows Maarten to upload images of music scores, to clean these images, to apply OMR to the cleaned image, and to quickly correct errors that have arisen during the OMR process. As such, Maarten quickly obtains a digital version of the music scores he needs.

4.1.6 Erik, Aruspix, and the Europeana Virtual Transcription Library

Persona

Erik is a professor in digital musicology at the University of Leiden, The Netherlands. He has just received the “spinozapremie” from NWO, and wants to start a large scale transcription project of masses and motets written within the renaissance period. This will allow him to apply a broad range of computational tools to a large corpus of music scores and to detect musical patterns over large periods of time.

Scenario

For his research project, Erik needs to transcribe more than one thousand written music scores. Erik wants to use Aruspix to build his corpus, which allows him to upload images of music scores, to clean these images, to apply OMR to the cleaned image, and to quickly correct errors that have arisen during the OMR process. Erik does not, however, have the time to transcribe all this work himself. Inspired by the Virtual Transcription Laboratory¹⁷, Erik decides to build a virtual transcription laboratory for music scores, which uses Aruspix for OMR. In this virtual environment, musicologists from around the world can upload images, OCR these images, and edit the resulting transcriptions. All transcriptions, including transcriptions in progress, are saved and stored within this environment, and users can also edit and manually enrich the metadata of projects they have been working on. By using this environment, Erik is able to quickly build a high quality corpus, to easily collaborate with research partners abroad, and to share and disseminate research products.

4.1.7 Klaas and MusicRadar

Persona

Klaas is a Postdoc focusing on early music by Josquin des Prez and others, who tries to classify music, find related work and patterns, and share his insights with his fellow musicologists.

Scenario

Klaas wants to find out if a motet by Josquin des Prez he found has been described by other researchers. He therefore first tries to identify relevant scores on ISMPL. As he doesn't find the

17

<https://confluence.man.poznan.pl/community/display/WLT/Introduction+to+Virtual+Transcription+Laboratory>



score for his motet in that repository, he develops the encoding in MusicXML himself. Once he has finished the transcription of the motet, he uploads it to IMSLP, and then runs a classification job from his own music software library. The classifier has access to not only the motet chosen by Klaas but also his other music, the IMSLP, Europeana and other libraries. Internally, the classifier uses one or more components in Weka¹⁸, which Klaas trained before. He finds works by Franchinus Gaffurius and Serafino dall'Aquila and the classifier suggests him clusters of works. Klaas then tags the groups and chooses to share his findings. He furthermore adds some manual tags and notes to the results. His software converts his annotations into an EDM metadata blob and writes it to the Europeana API. As such, his software, which he calls *MusicRadar*, allows musicologists to share and see metadata with provenance information (author, tool, etc.) from the various tools and research tasks.

4.1.8 Hugo working on Ottaviano Petrucci

Persona

Hugo is a 'traditionally' trained musicologist, teaching Early Music at Utrecht University. He has a special interest in music prints from the first decades of the sixteenth century. Hugo recently discovered the richness of the Internet in doing his research. In the past, he mainly used the physical books of his institution library (e.g. the volumes of the Grove music encyclopedia, printed facsimiles of sources of early music and the editions of the *Corpus Mensurabilis Musicae*), but now he turns his attention more and more to digital resources and tools. Hugo doesn't consider himself to be very technical, so a friend of his – a bio-informatics professor – helps him in assessing the resources and tools he find.

Scenarios

Metadata mash up (step 1)

A colleague of the media department pointed Hugo at the Europeana portal, where he searches on one of the first printers of music, Ottaviano Petrucci. On the results page, Hugo sees a lesser-known print: the *Motetti de la corona – Libro tertio*¹⁹. He takes a closer look at the metadata provided and starts combining it with biographical and bibliographical information (titles of pieces/composers, printers surrounding, handbooks) he gathers from the Oxford Music Online²⁰, the digital edition of Stanley Boorman's Ottaviano Petrucci:

¹⁸ <http://www.cs.waikato.ac.nz/ml/weka/>

¹⁹

http://europeana.eu/portal/record/03486/urn_resolver_pl_urn_urn_nbn_de_bvb_12_bsb00077_434_2.html?start=13&query=petrucci+music&startPage=1&rows=24

²⁰ <http://www.oxfordmusiconline.com>



Catalogue Raisonné²¹, the open data of RISM²² and several smaller databases on Early Music (like CMME²³, the Josquin Research Project²⁴ and the Motet Database Online²⁵). He gathers, arranges, and saves all the information in a spreadsheet in Google Drive. In a later stage, he plans to make the document available for his students.

Transcription & analysis (step 2)

After finishing his bibliographical research on the Motetti de la corona print, Hugo continues working on the musical content of this source. He downloads the image files of the print²⁶ and feeds them into Aruspix, recently installed on his computer by his friend. The software produces rough editions of the music, which Hugo analyses making use of the tools of Music21 (which makes a comparative analysis of the styles of the composers involved: how dissonance is treated, distribution voices, the use of pre-existing models). He notices that one of the pieces attributed to Johannes Mouton has features not complying with the style of this composer. He adds his findings to his own spreadsheet, and via the API of Europeana uploads his editions and annotations to the cloud.

Encoded music files matching

A colleague of the music library pointed Hugo at the Europeana portal, where he found a lot of digitizations of early music sources. Hugo recognizes several compositions he saw earlier on the websites of the Josquin Research Project (JRP) and CMME projects. His bio-informatics friends helps him to harvest the files from both projects and advises to contact the informatics people at Utrecht University, who work on Music Information Retrieval (MIR) research. Together with a professor and PhD student of the Interaction Technology department, Hugo is able to extract titles, composer names and incipits from the JRP and CMME files and matches them with the metadata from Europeana, referring to the source images. He uploads the matched encoded files to the Europeana Cloud, linking them with the image files.

21

http://books.google.nl/books?id=xgW4C5Xp2SgC&printsec=frontcover&dq=boorman+petrucci&hl=nl&sa=X&ei=lwq9U662I-XI0wW-i4CIBQ&redir_esc=y#v=onepage&q&f=false

²² <http://opac.rism.info/index.php?id=8&id=8&L=1>

²³ <http://www.cmme.org>

²⁴ <http://josquin.ccarh.org>

²⁵ <http://www.arts.ufl.edu/motet/default.asp>

²⁶ <http://stimbuecher.digitale-sammlungen.de/view?id=bsb00077434>

5 From Personas to Tools

The main objective of this deliverable is to identify the personas, analyze the scenarios, and present specific use cases for the tools that will be developed or adapted in the context of WP3. To achieve this goal, a specific process is followed in order to connect the work described in this document with the development and adaptation of the tools. This process, also called user-centered design, involves the continuous communication and collaboration of both, developers and researchers from the humanities.

The first two steps in this process are described in previous sections. These steps involve the workflow identification of (specific) researchers in the humanities and the preparation of scenarios where the different tools can be used. After writing the scenarios, a process is followed in which, on the one hand, developers continuously update the scenarios, create mock-ups and use cases for the modelled tool, while, on the other hand; researchers continuously provide feedback on their needs and requirements. This continuous communication between the two groups is needed in order to cover the needs of the researchers and final users of the tool. The final steps are the development of a pilot and a demo of the tool. A schematic representation of this process is shown in the figure below.

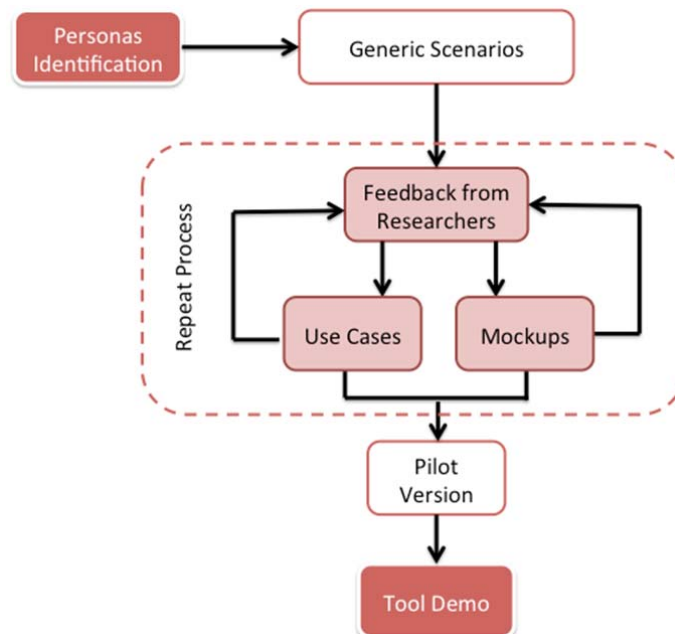


Figure 9: The main workflow followed for the development of the tools

To further analyse this process we will describe the work that has been conducted in the first year of eCloud (M1-M6) for the development of an ARIADNE Finder for the Axiom group of



the VU University Amsterdam. This work is based on the scenario “Hein and the ARIADNE Finder” presented in the previous section.

The first step in our process is persona identification. In our case Hein, a member of the Axiom Group, is a post-doctoral researcher working on the history of philosophy, the history of biology, and their interplay. After analysing the profile of the persona, we proceed with a generic analysis of a scenario for the tool to be developed: Hein could use the ARIADNE Finder technology to embed a search tool in the site of his research group to quickly search and browse through the various digital resources on the history of philosophy and biology. After the creation of the first generic scenario, feedback from Hein (as the Axiom group representative) was requested to better understand his needs and requirements.

Based on the persona and the described scenario, the following step is the creation of the first use case scenario. In our case, the following use case description was developed.

Use Case Description	
Name	Axiom Group – ARIADNE Finder
Description	The finder helps researchers in history and philosophy of science to find material that they can use in their research.
Legal foundation(s)	The copyright license of the content should be respected during the aggregation of metadata and the use of material.
Pre-condition	Metadata for different type of content should be triplified and provided through a single API.
Flow of Events - Basic Path	
Step 1	A researcher is visiting the axiom group site
Step 2	He is entering a search query or using the browse functionalities
Step 3	A results set is shown to the researcher
Step 4	The researcher can use specific facets to narrow down the results set
Step 5	The researcher clicks on specific results to show more information and to access the resource
Post-condition	All the metadata of resources are aggregated, indexed and provided through a single API
5.1 Actors	
End-users	Researchers
Information provider(s)	Europeana, European Library, GLOBE, YouTube, Flickr
5.2 Expected added value	
Expected value added	- Researchers will be able to search simultaneously in many external sources with different type of content e.g. books,

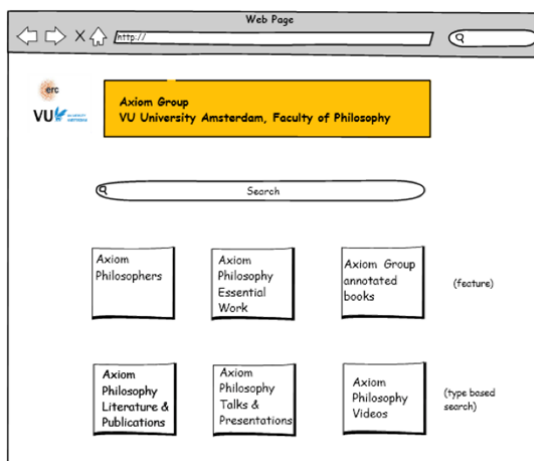


<p>publications, maps</p> <ul style="list-style-type: none"> - the results will include only relevant resources to the user query - the user will be able to make complex queries - discovery of relevant resources with very good performance that will improve the user experience

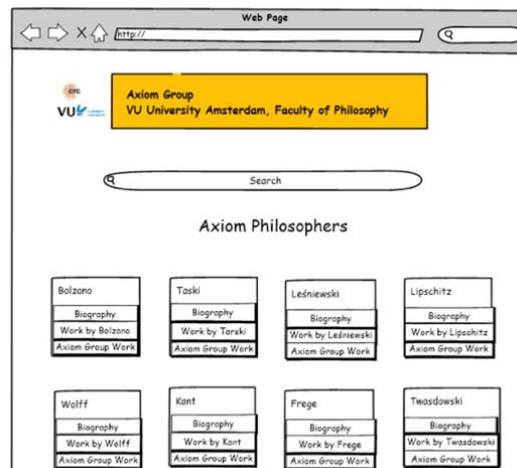
Table 1: The use case description for the scenario of “Hein and the ARIADNE Finder “

The following step required the development of the first mock-ups, shown in the figures below. The mock-ups provide an illustration of what the final tool could look like based on the analysis of the scenario. Since the ARIADNE Finder is a technology that is embedded in an existing site, the first version of mock-ups included the following pages:

- A main page, shown in Figure 10a, where users can search the repositories for data. Additional buttons are included to quickly access predefined searches, specific resources or the local repository of the group.
- A second page, shown in Figure 10b, with a number of the philosophers the group investigates, links to a search for their biographies and work, and access to the group’s work on a specific philosopher,
- A search result page, shown in Figure 10c, with a faceted search to quickly filter the results with a number of different criteria,
- A view more information on a specific resource page, shown in Figure 10d, where all the metadata of the resource are shown and the user has the option to either view the original resource or save a copy of the resource for annotation (enabled based on the copyright).



(a)



(b)

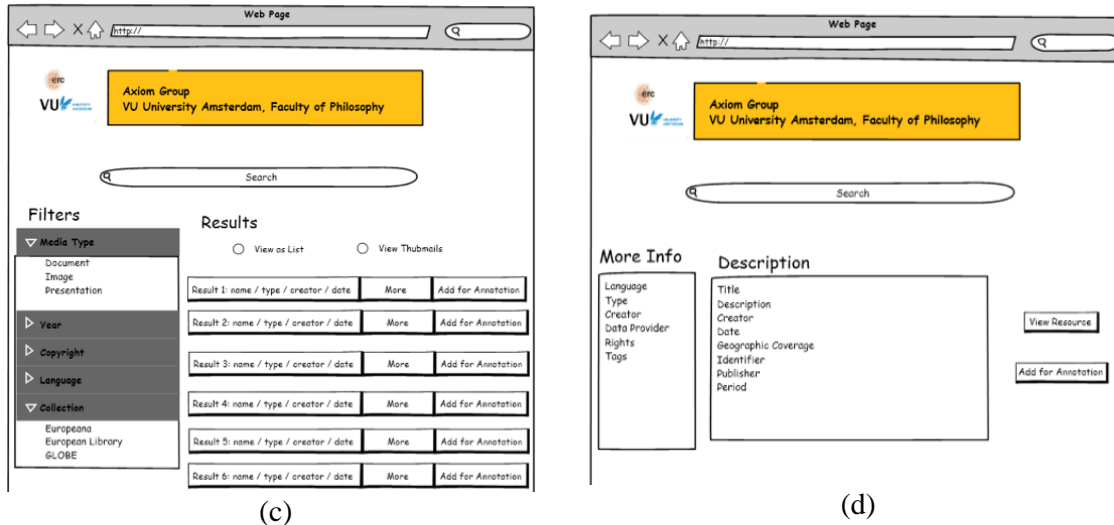


Figure 10: The first version of the 4 pages mock-ups for the Hein and the ARIADNE finder scenario

This set of first mock-ups was distributed among the eCloud WP3 group and an online meeting, open to all members of the WP3 group, was arranged with Hein, the main persona, to discuss them. The feedback gathered from this meeting led to a second version of the mock-ups and to respective changes in the use case description. Apart from comments made by the user-researcher Hein, feedback was also given by other members of the WP3 group, mainly on the technical side for the tool to be developed. The second version of the mock-ups, shown in Figure 11, had the following additional features:

- A bar to quickly add a time period in any search,
- A button to search for the group's work in Mendeley or Zotero,
- A button to view specific and selected resources (essential work),
- More collections to search,
- When viewing a resource, the options to download, export, or view related resources in Mendeley or Zotero are available.

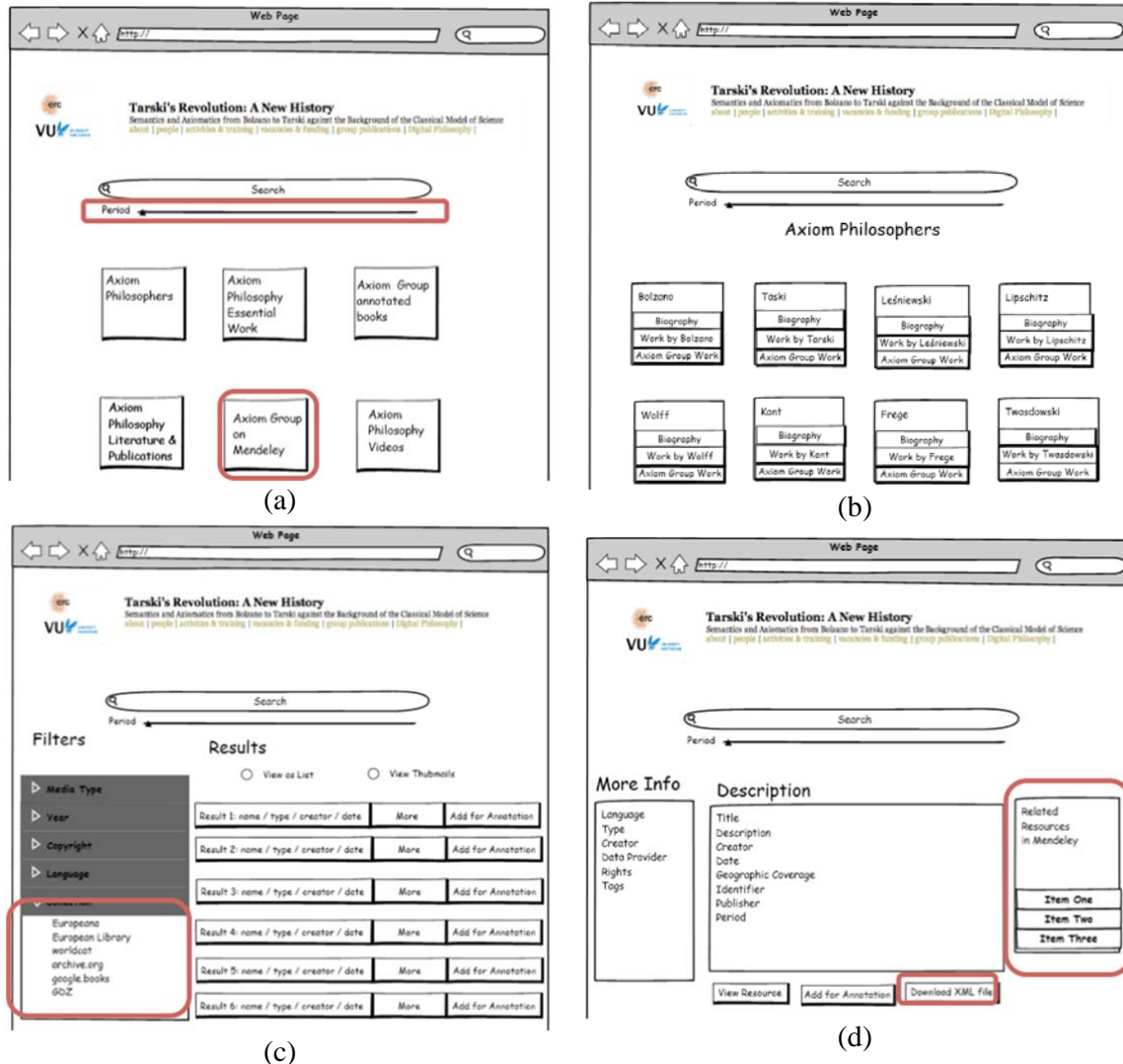


Figure 11: The second version of the 4 pages mock-ups for the Hein and the ARIADNE finder scenario with some of the changes highlighted

As explained earlier, we continuously gather feedback from users in order to update the use case scenarios and mock-ups until we reach a desired outcome and the pilot can be developed. During the first days of July 2013, the eCloud WP3 group arranged a meeting with Hein and the axiom group to discuss, among others, the ARIADNE Finder demos. The mock-ups were presented to the group along with a number of questions to guide the users in providing feedback for the technical team through a presentation that was later distributed to all the members of the WP3. Hein and the rest of the group members decided to gather additional



feedback from other members of the group not present in the meeting before continuing. When all the feedback was gathered, the third and final version of the mock-ups was developed and discussed with members of the Axiom Group.

The following table summarizes all the steps that were followed for the above example in the first seven months of eCloud.

Step	Description	Date
Step 1	Persona identification	28th March 13
Step 2	Initial scenario development	15th April 13
Step 3	Feedback from Axiom group representative	27th May 13
Step 5a	Use case development	13th Jun 13
Step 5b	Initial mock up	13th Jun 13
Step 6	Interview with Axiom group representative and WP3 members	19th Jun 13
Step 7	Revised mock up	5th July 13
Step 8	Presentation of mock up to the whole Axiom group	8th July 13
Step 9	Final mock up & specs for implementation	31st July 13
Step 10	Prototype development – First version	15th September 13
Step 11	Prototype development – Final version	15th December 13

Table 2: Summary of the work for the case of “Hein and the ARIADNE Finder” in M1-M7

After presenting the first prototype of the ARIADNE FINDER to the Axiom Group, the philosophers provided monthly feedback on the tool. This feedback was considered for the next iterations. The same procedure was adopted for the testing and evaluation of the mash-up of tools (described in section 3). Once the philosophers had access to the tools of the mash-up, they were asked to provide expert feedback on whether and how these tools could support their research and workflow. A more detailed description of the mash-up of tools and its evaluation is provided in deliverables 3.2 and 3.3, and in a forthcoming publication²⁷.

²⁷ H. Van den Berg, G. Parra, A. Jentzch, A. Drakos, and E. Duval. “tudying the History of Philosophical Ideas: Supporting Research Discovery, Navigation, and Awareness”. In: Proc. of i- Know ’14. Graz, Austria: ACM, 2014. Accepted.



6 Conclusion

WP3 aims to develop services and tools that leverage Europeana content to be used by researchers. In this deliverable, we focus on the first step before the development of the tools, which is to understand and analyse the user needs by developing personas, scenarios and use cases; and to present a number of different existing tools that can be combined, expanded, and integrated to cover the needs of researchers. In addition, we introduced the process to be followed in order to connect these scenarios with tools development. This document is the second version of the Deliverable 3.1, which will be updated further during the project lifetime to reflect progress of our thinking as the project evolves.

During the first eCloud year, we have identified three core problems for the researchers in philosophy and digital humanities:

- problems with navigating and identifying relevant (digital) content and with building corpora;
- a lack of user-friendly tools for conducting fine-grained textual research;
- and a lack of appropriate tools and infrastructure that allow members of research groups to work collaboratively.

We have addressed these problems through a combination of search, visualization, and awareness tools and the integration of those in the Europeana environment. The results are presented in Deliverables 3.2 and 3.3.

During our second year, we have identified 4 core problems for musicologists:

- Problems involving data creation.
- Lack of music scores.
- Information exchange and linking of data.
- Retrieval and analysis of contextual information.

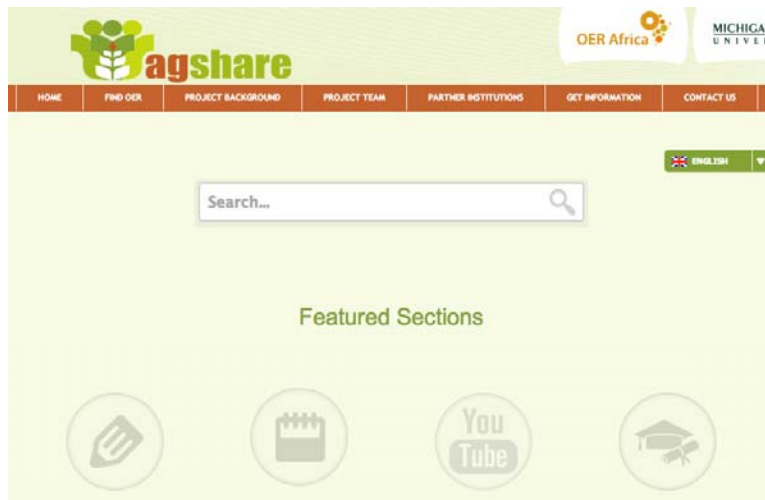
In the remainder of Year 2 (M19-24), we will address a number of these problems through a combination of tools that are already in use by musicologists as well as a set of new tools. This work will be reflected in updated versions of Deliverables 3.2 and 3.3.



7 Appendix: Tools for researchers in the Digital Humanities

7.1 Search tools

7.1.1 ARIADNE Finder



<i>description of functionality</i>	Lightweight search interface (implemented as an HTML page with some Javascript) that works on top of a Solr index. The Solr index is build using a number of metadata facets that can be used to navigate metadata records aggregated through different sources. Ingests metadata records of relevance to the specific users, by a number of ingestion mechanisms/APIs that are creating the Solr index behind the particular search page.
<i>URL</i>	http://ariadne-eu.org/wiki/
<i>contact</i>	nikosm@ieee.org stoitsis@ieee.org
<i>licence</i>	LGPL for older versions, still undefined for current/working ones (but will be a xGPL one)
<i>programming language</i>	HTML and Javascript
<i>current use</i>	at least 5 deployments of beta versions, with estimated usage by >200 users http://ariadne.cs.kuleuven.be/finder/ariadne/ [looking into the GLOBE aggregator]
<i>deployment</i>	http://laflor.laclo.org [looking into the Latin America learning repository aggregator] http://greenlearningnetwork.com/ [looking into the Green Learning Network OER aggregator]



<http://www.greenlearningnetwork.com/organicedunet/> [looking into the organic collections of the Green Learning Network OER aggregator]
<http://agsharedemo.agroknow.gr> [looking into the Africa-related OER collections of the Green Learning Network]

7.2 Visualization tools

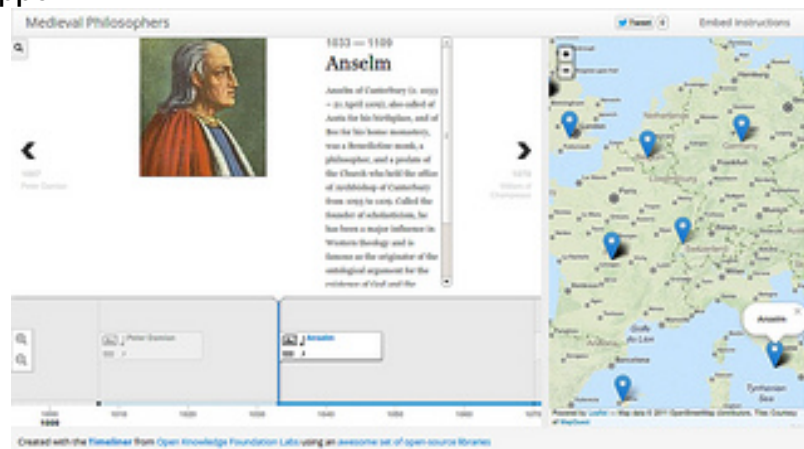
A good overview on visualisation methodologies / tools by functionality:

http://www.visual-literacy.org/periodic_table/periodic_table.html

Data visualization guidelines, by Gregor Aisch:

<http://schoolofdata.org/2013/04/26/data-visualization-guidelines-by-gregor-aisch-international-journalism-festival/>

7.2.1 TimeMapper



description of functionality

A visualization tool making timelines and timemaps using Google spreadsheets.

URL

<http://timemapper.okfnlabs.org/>

contact

info@okfn.org

licence

MIT licence

programming language

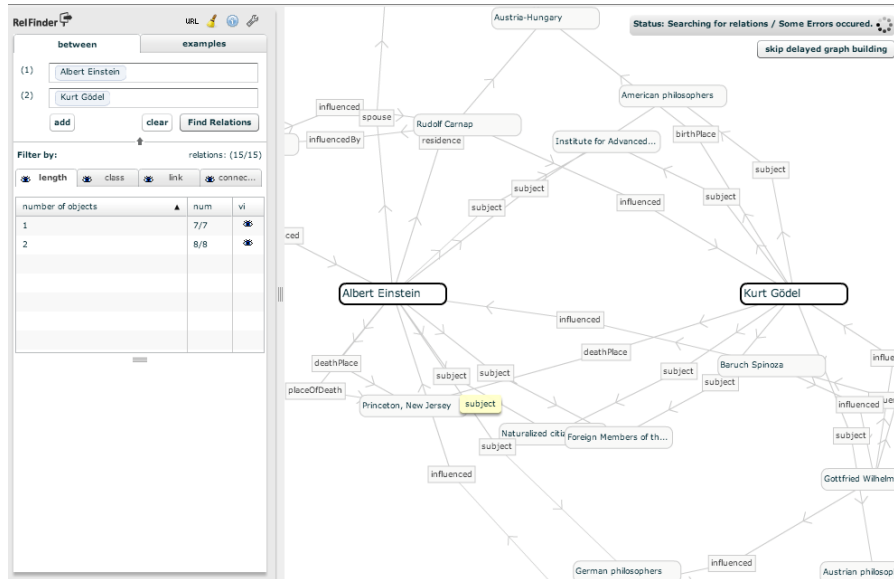
Java Script and other open-source components including TimelineJS, ReclineJS, Leaflet, Backbone and Bootstrap

deployment

Example of an implementation on Medieval Philosophers:
http://timeliner.okfnlabs.org/view/?url=https://docs.google.com/spreadsheet/ccc?key=0A16mO9_3Hr2PdGZnRjEwUWxOekhreTNNZEFEMWRZbke_-2

On the Open Parliament Declaration
<http://www.openingparliament.org/about>

7.2.2 RelFinder



*description of
functionality*

The RelFinder extracts and visualizes relationships between given objects in RDF data and makes these relationships interactively explorable. Highlighting and filtering features support visual analysis both on a global and detailed level. The RelFinder is based on the open source framework Adobe Flex, easy-to-use and works with any RDF dataset that provides standardized SPARQL access.

URL

<http://www.visualdataweb.org/relfinder.php>

contact

contact@visualdataweb.org

licence

GNU General Public License

*programming
language*

Adobe Flex

University of Leipzig,

deployment

<http://catalogus-professorum.org/tools/relfinder/RelFinder.swf>

Ontotext

<http://linkedlifedata.com/relfinder>



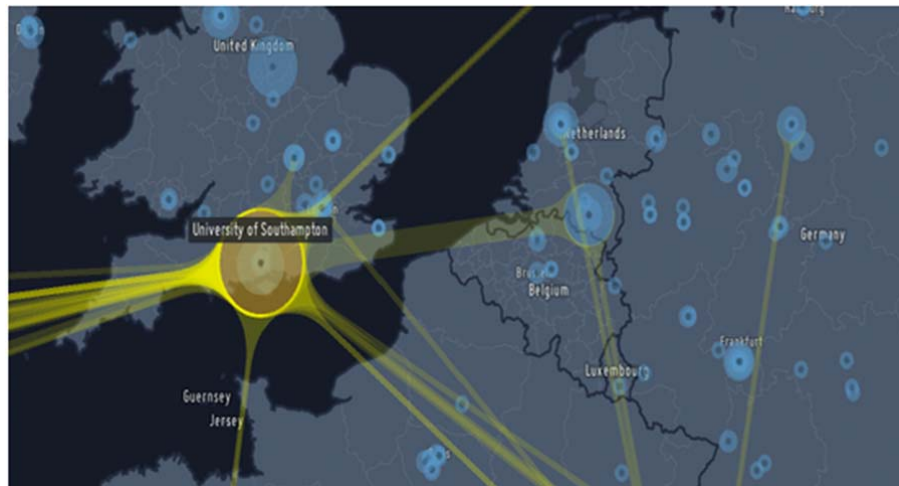
bibliography

Philipp Heim, Steffen Lohmann and Timo Stegemann: *Interactive Relationship Discovery via the Semantic Web*, in Proceedings of the 7th Extended Semantic Web Conference (ESWC 2010), volume 6088, series LNCS, pages 303-317. Springer, Berlin/Heidelberg, 2010.

Philipp Heim, Sebastian Hellmann, Jens Lehmann, Steffen Lohmann and Timo Stegemann, *RelFinder: Revealing Relationships in RDF Knowledge Bases*. in: Proceedings of the 4th International Conference on Semantic and Digital Media Technologies (SAMT 2009), pages 182-187. Springer, Berlin/Heidelberg, 2009.

Steffen Lohmann, Philipp Heim, Timo Stegemann and Jürgen Ziegler. *The RelFinder User Interface: Interactive Exploration of Relationships between Objects of Interest*, in: Proceedings of the 14th International Conference on Intelligent User Interfaces (IUI 2010), pages 421-422. ACM, New York, NY, USA, 2010.

7.2.3 Muse



description of functionality

geo-spatial visualization of co-authorship on a multitouch tabletop

URL

<http://tillnagel.com/2010/11/muse/>,
<http://tillnagel.com/2011/10/interactive-exploration-of-geospatial-network-visualization/>

contact

erik.duval@cs.kuleuven.be

current use

evaluated at ECTEL2010 and Hypertext2011

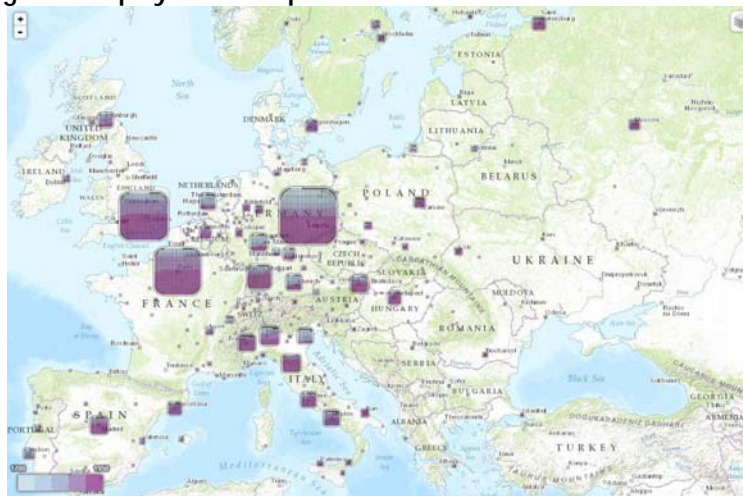
Bibliography

Nagel, T., Duval, E., Vande Moere, A.: *Interactive Exploration of a Geospatial Network Visualization*. CHI 2012, May 5-10, Austin, USA.
Nagel, T., Duval, E.: *Interactive Exploration of a Geospatial Network*



Visualization(Poster). VisWeek 2011, October 21-28, Providence, USA.
Nagel, T., Duval, E., Heidmann, F.: *Visualizing Geospatial Co-Authorship Data on a Multitouch Table*. Smart Graphics 2011, July 18-20, Bremen, Germany.
Nagel, T., Duval, E.: *Muse: Visualizing the origins and connections of institutions based on co-authorship of publications*. Science2.0 for TEL workshop at EC-TEL 2010, Barcelona, Spain.

7.2.4 MappingPhilosophy/GlaMMap



description of functionality

Geo-spatial visualization of bibliographic metadata (place of publication, author, title, year) from books in logic from 1700-1940 on an interactive geographical map of Europe

URL

<http://axiom.vu.nl/MappingPhilosophy.html>

contact

ariannabetti@gmail.com
hein.van.den.berg2@gmail.com

programming language

HTML5, CSS, and JavaScript

current use deployment

early prototype (internal use)
no users apart from developers

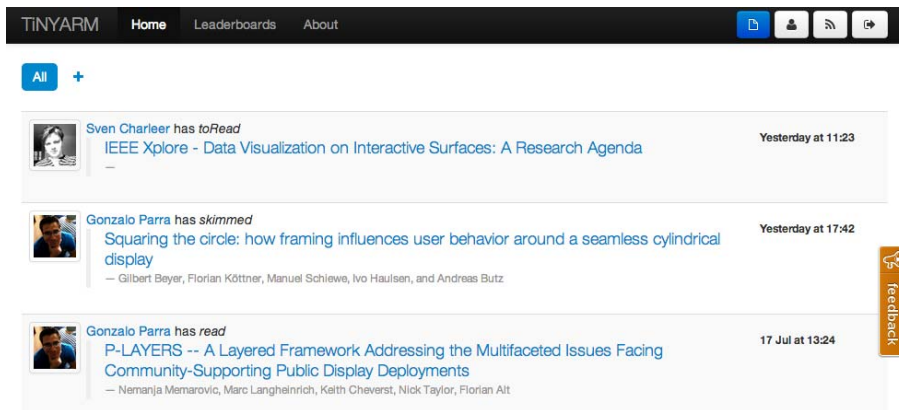


Bibliography

Van den Berg, H., Betti, A., Speckmann, B., Verbeek, K., Huijgen, P., Kramer, J., Barendregt, P., *Mapping Philosophy*. A version of this paper will be presented at the GlaMMMap Kickoff Meeting 2013, May 3, VU Amsterdam, The Netherlands.

7.3 Awareness tools

7.3.1 TiNYARM



description of functionality

web application to make researchers aware of what their peers are reading

URL

<http://atinyarm.appspot.com/>

contact

gonzalo.parra@cs.kuleuven.be

licence

Free

programming language

Java (GAE)

current use

HCI group KUL

deployment

<http://atinyarm.appspot.com/>

bibliography

Parra, G., Klerkx, J., Duval, E.: TiNYARM: This is Not Yet Another Reference Manager. 3rd International Workshop on Motivational and Affective Aspects part of the European Conference on Technology Enhanced Learning (ECTEL) vol: 957 pages:1-4.

Parra, G., Klerkx, J., Duval, E.: What should I read next? Awareness of relevant publications through a Community of Practice. In CHI EA '13. ACM, New York, NY, USA, 2375-2376.

Parra, G., Klerkx, J., Duval, E.: TiNYARM: Awareness of Research Papers in a Community of Practice. In Proceedings i-KNOW '13). (Accepted).

7.3.2 More!



*description of
functionality*

A social discovery tool for researchers

URL

<https://sites.google.com/site/kulmoreapp/>

contact

gonzalo.parra@cs.kuleuven.be

licence

Free

*programming
language*

Java, PHP

*current use
deployment*

Currently not used

<http://ariadne.cs.kuleuven.be/more/index.php?id=gonzalo-parra>

bibliography

Parra Chico, G., Klerkx, J., Duval, E. (2011). More!: mobile interaction with linked data. In Diaz, P. (Ed.), Hussein, T. (Ed.), Lohmann, S. (Ed.), Ziegler, J. (Ed.), International Workshop on Data-Centric Interactions on the Web in conjunction with the 13th IFIP TC13 Conference on Human-Computer-Interaction (INTERACT 2011). Lisbon, Portugal, 6 September 2011 (pp. 37-47) CEUR-WS.

Parra Chico, G., Duval, E. (2010). Filling the gaps to know More! about a researcher. Proceedings of the 2nd International Workshop on Research 2.0. At the 5th European Conference on Technology Enhanced Learning: Sustaining TEL. ECTEL10. Barcelona, Spain, 28 September 2010 (pp. 18-22) CEUR-WS.

Parra Chico, G., Duval, E. (2010). More! a social discovery tool for researchers. Proceedings of EdMedia10: World Conference on Educational Multimedia, Hypermedia and Telecommunications. EdMedia10. Toronto, Canada, 29 June 2010 (pp. 561-569) AACE.



7.4 Annotation tools

7.4.1 Annotateit

Annotate anything

Now you can annotate anything, anywhere on the web, by [signing up](#) and installing our bookmarklet!

- Annotate any webpage with the bookmarklet
- Add the Annotator to your own site in just two lines of code

[Sign Up Now »](#)

Follow @theannotator 140 followers Watch 587

SHUTTLEWORTH FOUNDED

description of functionality

The Annotator is an open-source JavaScript library and tool that can be added to any webpage to make it annotatable. Annotations can have comments, tags, users and more. Furthermore, the Annotator can be easily extended with new features.

URL

<http://annotateit.org/>

contact

hello@aroncarroll.com
rufus.pollock@okfn.org
nick@whiteink.com

licence

GNU GPLv3 or MIT

deployment

<http://annotateit.org/>

7.4.2 TEXTUS

TEXTUS

About | Blog | Code | Participate | Co

TEXTUS: an open source platform for collaborating around collections of texts

textus
collaborating around texts

What is TEXTUS?

In a nutshell it is an **open source** platform for working with collections of texts. It enables students, researchers and teachers to **share** and **collaborate** around texts using a simple and intuitive interface.

What does TEXTUS do?

TEXTUS currently enables users to:

SIGN UP TO THE DISCUSSION LIST

Discussion sign up

SIGN UP TO THE DEVELOPERS' LIST

Developer sign up

description of functionality

In a nutshell TEXTUS is an open source platform for working with collections of texts. It enables students, researchers and teachers to share and collaborate around texts by annotating them.



URL <http://textusproject.org/>
contact rufus.pollock@okfn.org
licence MIT
current use Open Shakespeare group (<http://openliterature.net/shakespeare/>)
deployment <http://beta.openphilosophy.org>

7.4.3 Pundit

The screenshot shows the Pundit web interface. At the top, there is a navigation bar with 'homepage' and 'semilib demo pages' (adding semantic to digital libraries). The main content area displays a page titled 'convivio/trattato primo capitolo i'. The text on the page is annotated with a yellow highlight and a context menu is open, showing options like 'Annotate text fragment', 'Connect this text to...', 'Add Comment or Tags', 'Show all annotations on this item', 'Zoom on this item', and 'Add to My Items'. On the left, there is a sidebar with 'authors' (Dante Alighieri, Ludovico Ariosto) and 'works' (Dante: Convivio/trattato 1, 2, 3, 4).

description of functionality Pundit enables users to create structured data annotating web pages or images, collect annotations and share with others to create collaborative structured knowledge. Furthermore, Pundit already has a built in entity extraction feature which annotates the texts using knowledge bases like DBpedia or Freebase

URL <http://thepund.it/>
contact pundit@netseven.it
current use Wittgenstein group in Norway (via DM2E project)
deployment http://release-bot.thepund.it/latest/examples/authors_index.html

bibliography M.Grassi, C. Morbidoni and M. Nucci. "A Collaborative Video Annotation System Based on Semantic Web Technologies". In press: Cognitive Computation. Springer-Verlag, Berlin.
M. Grassi, C. Morbidoni, M. Nucci, S. Fonda, G. Ledda. "Pundit: Semantically Structured Annotations for Web Contents and Digital Libraries". Annett Mitschick, Fernando Loizides, Livia Predoiu, Andreas Nürnberger, Seamus Ross (eds.): Semantic Digital Archives 2012. Proceedings of the Second International Workshop on Semantic Digital Archives (SDA 2012), Paphos, Cyprus, September 27, 2012, CEUR-WS.org/Vol-912, urn:nbn:de:0074-912-6.
M. Nucci, M. Grassi, C. Morbidoni, F. Piazza. "Enriching Digital Libraries Contents with SemLib Semantic Annotation System". Proceedings of the



Digital Humanities 2012 Conference, Hamburg, Germany, 16th-20th July, 2012.

C. Morbidoni, M. Grassi, M. Nucci. "Introducing SemLib Project: Semantic Web Tools for Digital Libraries", International Workshop on "Semantic Digital Archives - sustainable long-term curation perspectives of Cultural Heritage" held as part of the 15th International Conference on Theory and Practice of Digital Libraries (TPDL). 29.09.2011 in Berlin.

7.4.4 OpenAnnotation



Open Annotation Collaboration

Welcome | Phase III | Open Annotation Data Model | Community & Resources | Phase I - [ended] | Phase II - [ended]

Main Menu

Open Annotation

- Welcome
- Phase III
- 1.0 Data Model [NEW]
- Data Model Rollouts [NEW]
- OA Community Examples
- About Open Annotation
- The Collaboration

Welcome

We are pleased to announce the publication of the 1.0 release of the Open Annotation Data Model & Ontology. This work is the product of the W3C Open Annotation Community Group jointly founded by the Annotation Ontology and the Open Annotation Collaboration. The OA Community Group will be hosting three public rollout events, U.S. West Coast, U.S. East Coast, and in the U.K. this Spring and early Summer. Implementers, developers, and information managers who attend one of these meetings will learn about the OA Data Model & Ontology firsthand from OA Community implementers and see existing annotation services that have been built using the OA model.

Recent Releases

Open Annotation 1.0



A collaboration tool that aims:

- To facilitate the emergence of a Web and Resource-centric interoperable annotation environment that allows leveraging annotations across the boundaries of annotation clients, annotation servers, and content collections. To this end, interoperability specifications will be devised.
- To demonstrate through implementations an interoperable annotation environment enabled by the interoperability specifications in settings characterized by a variety of annotation client/server environments, content collections, and scholarly use cases.
- To seed widespread adoption by deploying robust, production-quality applications conformant with the interoperable annotation environment in ubiquitous and specialized services, tools, and content used by scholars -- e.g.: Zotero, AXE, LORE, Co-Annotea, Pliny; JSTOR, AustLit, MONK.

description of functionality

URL

<http://www.openannotation.org/>



7.4.5 DocumentCloud



description of functionality

DocumentCloud is a tool for annotating documents and sharing those annotations on the web. DocumentCloud runs every document you upload through OpenCalais and extracts entities (people, places and organizations) mentioned in it.

URL

<https://www.documentcloud.org/>

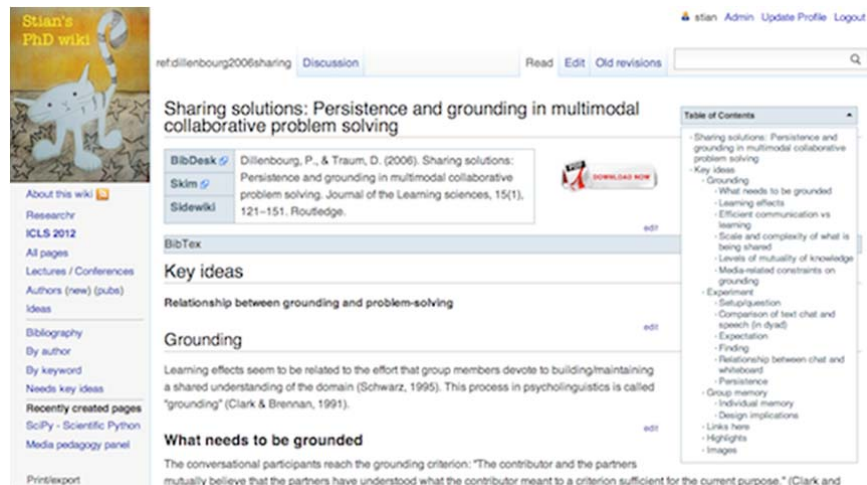
contact

support@documentcloud.org

deployment

<https://www.documentcloud.org/>

7.4.6 Researchr



description of functionality

Academic information management workflow, which is at the same time individual and happens on the local computer, but also set up to be easily shared with others. It's really a whole framework with a bunch of applications (the key ones being BibDesk, Skim, DokuWiki, Chrome).

URL

<http://reganmian.net/wiki/researchr:start>

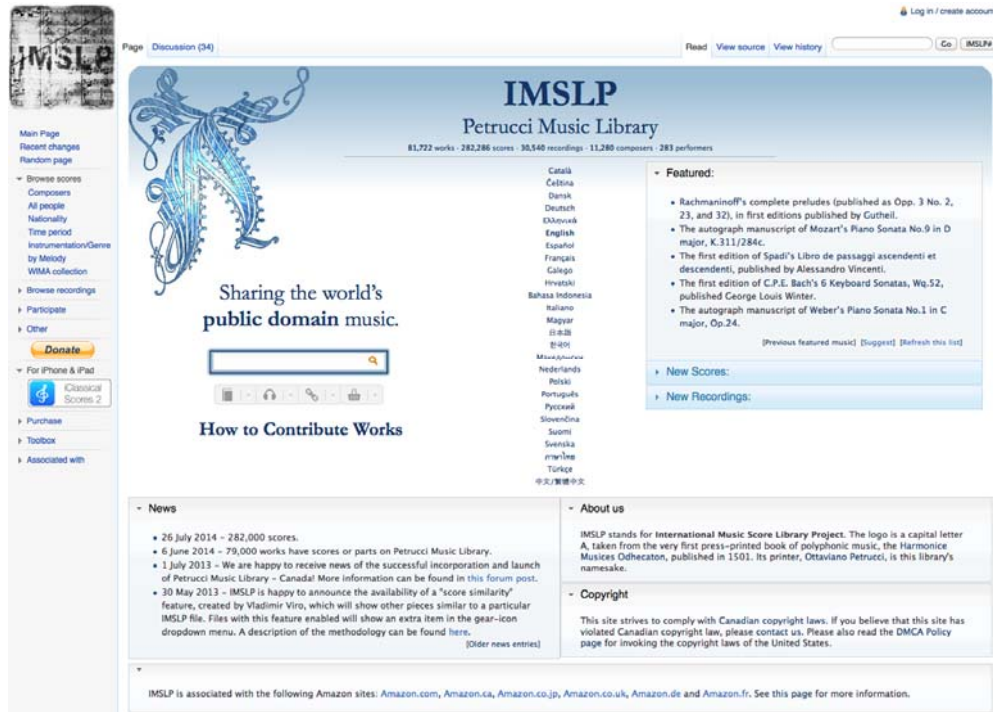


contact shaklev@gmail.com
licence Free
deployment <http://reganmian.net/wiki/researchr:start>

8 Appendix: Additional tools for musicologists

8.1 Search tools

8.1.1 IMSLP Petrucci Music



*description of
functionality*

The Petrucci Music Library (after publisher Ottaviano Petrucci) is a virtual library of public domain music scores. Since its launch on February 16, 2006, over 258.000 scores and 23.000 recordings for over 73.000 works by 7.000 composers have been uploaded (December 2013).

URL

<http://imslp.org/wiki/>

contact

<http://imslpforums.org/> | http://imslp.org/wiki/IMSLP:IRC_Access

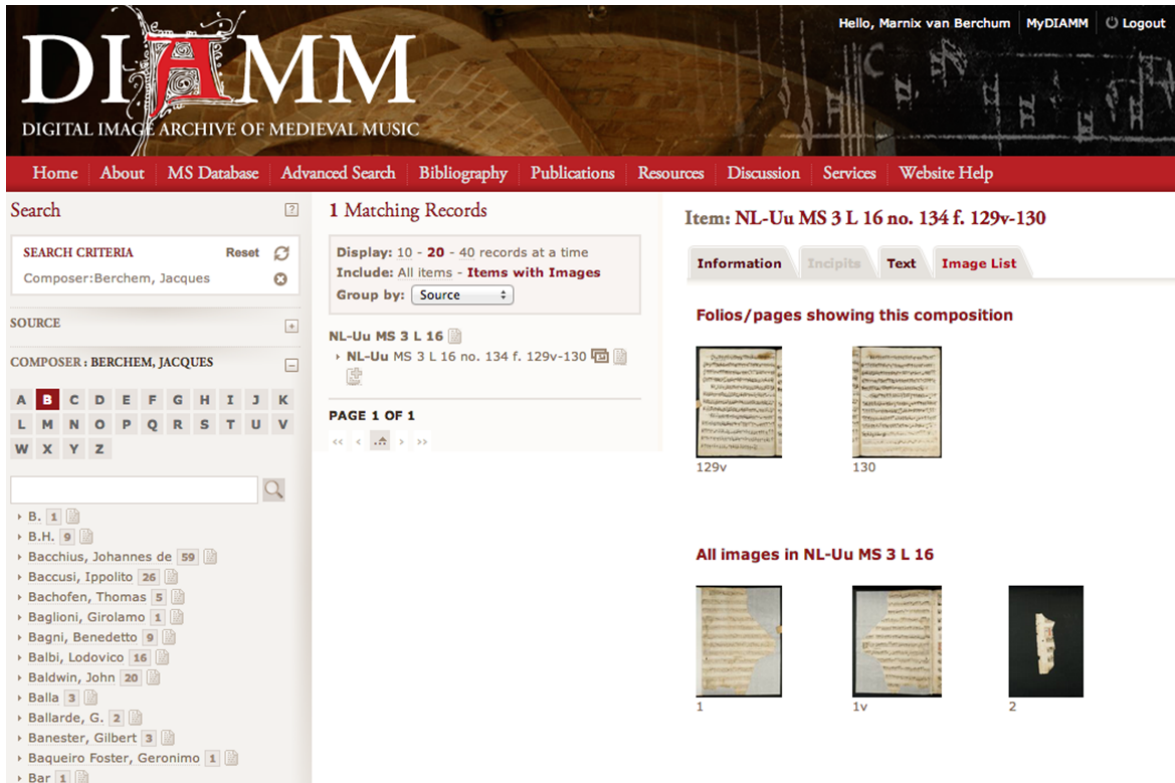
licence

(CC BY-SA 4.0)

deployment

<http://imslp.org>

8.1.2 Digital Image Archive of Medieval Music (DIAMM)



The screenshot shows the DIAMM (Digital Image Archive of Medieval Music) website interface. At the top, there is a navigation bar with links: Home, About, MS Database, Advanced Search, Bibliography, Publications, Resources, Discussion, Services, and Website Help. The main content area is divided into several sections:

- Search:** A search box with 'SEARCH CRITERIA' and 'Reset' buttons. Below it, 'Composer: Berchem, Jacques' is entered. A 'SOURCE' section is also visible.
- 1 Matching Records:** A section showing search results. It includes 'Display: 10 - 20 - 40 records at a time', 'Include: All items - Items with Images', and 'Group by: Source'. Below this, a specific record is shown: 'NL-Uu MS 3 L 16' and 'NL-Uu MS 3 L 16 no. 134 f. 129v-130'. It also indicates 'PAGE 1 OF 1'.
- Item: NL-Uu MS 3 L 16 no. 134 f. 129v-130:** A detailed view of the selected item. It has tabs for 'Information', 'Incipits', 'Text', and 'Image List'. Under 'Image List', there is a section 'Folios/pages showing this composition' with two thumbnail images labeled '129v' and '130'. Below that, 'All images in NL-Uu MS 3 L 16' shows three more thumbnail images labeled '1', '1v', and '2'.

description of functionality

DIAMM (the Digital Image Archive of Medieval Music) is a resource for the study of medieval manuscripts. The project presents images and metadata for thousands of manuscripts. DIAMM also provides a home for scholarly resources and editions, undertakes digital restoration of damaged manuscripts and documents, publishes high-quality facsimiles, and offers expertise through consulting.

URL

<http://www.diamm.ac.uk/> <http://imsip.org>

contact

<http://www.diamm.ac.uk/about/contact-diamm/>

license

<http://www.diamm.ac.uk/about/copyright/>

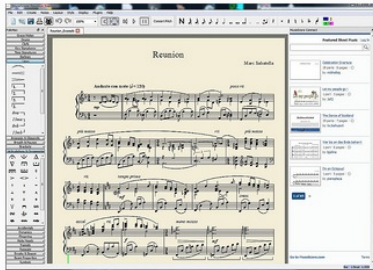
deployment

<http://www.diamm.ac.uk/> <https://www.documentcloud.org/>

8.2 Creation of music sheets

8.2.1 MuseScore

Create, play and print beautiful sheet music



[Take a tour](#)



[Other systems and platforms](#)

[What's new in MuseScore 1.3](#)

[Donate](#)

[News](#)

[Students Announced for Google](#)

MuseScore Software

- [Download](#)
- [Tour](#)
- [Handbook](#)
- [Tutorials](#)
- [Forums](#)
- [Services](#)
- [Development](#)
- [Donate](#)
- [FAQ](#)
- [How to](#)
- [Links](#)
- [Plugins](#)

Getting Started

[Part 1: Score Setup](#)

*description of
functionality*

MuseScore is a free and open source music notation software for Windows, Mac and Linux. It is easy to use and makes beautiful sheet music. MuseScore is often praised as a cost effective alternative for Sibelius and Finale.

URL

<https://musescore.com/>

contact

<https://musescore.com/contact>

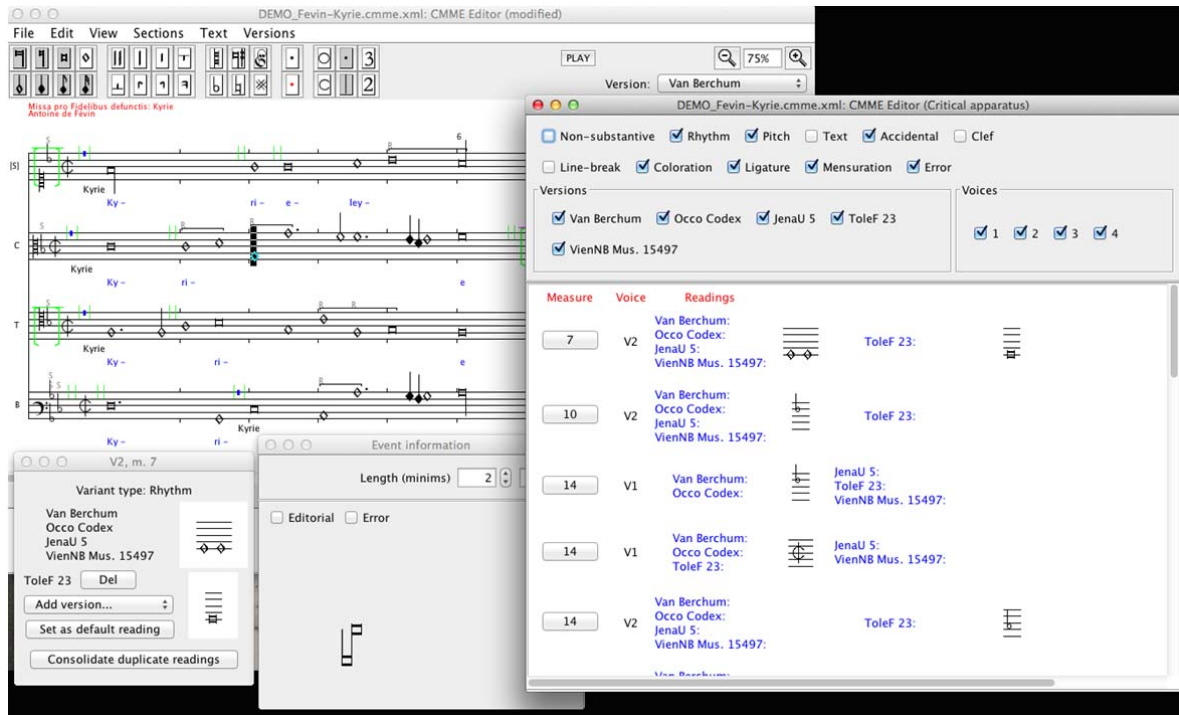
licence

Copyright ©

deployment

<https://github.com/musescore/MuseScore>

8.2.2 CMME



description of functionality

CMME offers software for creating, editing and viewing encoded score files of mensural notated music (roughly from the 15th and 16th century). The software is written in Java, and platform independent. The underlying encoding scheme is CMME-XML, a variant of the MusicXML format. On its website, CMME offers a corpus of free scores with additional metadata.

URL

<http://www.cmme.org>

contact

<http://www.cmme.org/contact>

license

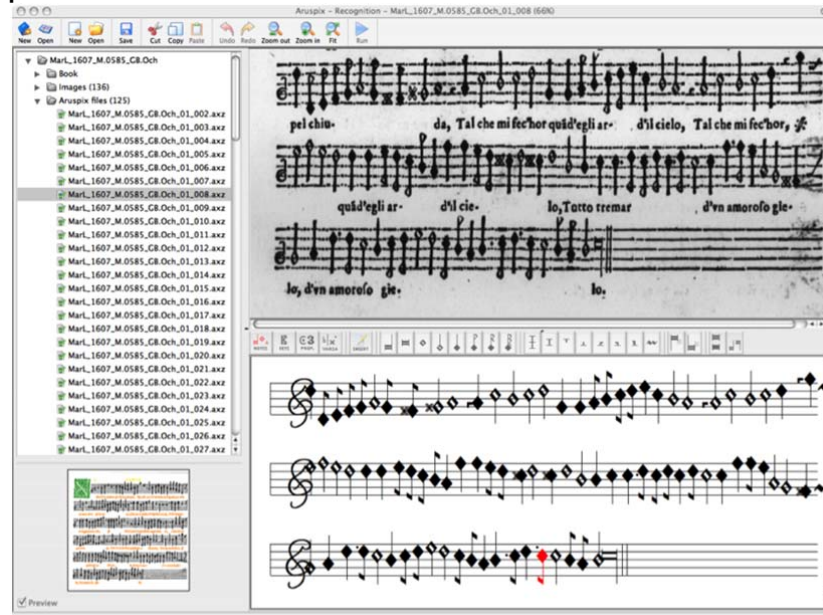
GNU General Public License

deployment

<https://github.com/tdumitrescu/cmme-editor>
<https://www.documentcloud.org/>

8.3 Optical Music Recognition

8.3.1 Aruspix



description of functionality

Aruspix is a software application for the optical recognition, the superimposition and the collation of early music prints (16th century). The optical recognition part of Aruspix uses an innovative technique in OMR based on Hidden Markov Models to identify and classify the particular musical characters found in early music prints, after an initial pre-treatment phase performed on scanned images of microfilm and facsimiles. The integrated music editor allows corrections of recognition errors. One original aspect of Aruspix is that it learns and optimizes itself dynamically as soon as a page is processed and corrected by the user.

URL

<http://www.aruspix.net/index.html>

contact

lxpugin@gmail.com | laurent@music.mcgill.ca

license

N/A

deployment

N/A <https://www.documentcloud.org/>



8.4 Classification tools

8.4.1 WEKA



Machine Learning Group at the University of Waikato

Project **Software** Book Publications People Related

Weka 3: Data Mining Software in Java

Weka is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset or called from your own Java code. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes.

Found only on the islands of New Zealand, the Weka is a flightless bird with an inquisitive nature. The name is pronounced like **this**, and the bird sounds like **this**.

Weka is open source software issued under the **GNU General Public License**.

Yes, it is possible to apply Weka to **big data**!

Data Mining with Weka is a 5 week MOOC, which was held first in late 2013. Check out the **MOOC site** for video lectures and details on how to enrol into this course and a new, advanced Weka course.

Getting started

- Requirements
- Download

Further information

- Citing Weka
- Datasets

Developers

- Development
- History

*description of
functionality*

Weka is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset or called from Java code. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well suited for developing new machine learning schemes.

URL

<http://www.cs.waikato.ac.nz/ml/weka/>

contact

<http://list.waikato.ac.nz/mailman/listinfo/wekalist>

license

GNU General Public License

deployment

<http://weka.wikispaces.com/wiki/changes>
<https://www.documentcloud.org/>



8.5 Music Analysis

8.5.1 Music21

The screenshot shows the Music21 documentation page. At the top, there are navigation links: 'music21', 'About Music21', 'previous', 'next', 'modules', and 'index'. The main heading is 'Examples and Demonstrations'. Below it, a paragraph states: 'The following examples provide a few samples of some of the possibilities available when working with music21.' The next section is 'Creating a Reduction and Labeling Intervals'. A paragraph explains: 'This example, after parsing a polyphonic work stored as an Opus and creating a Score, presents and labels the intervals of the resultant chords of all distinct harmonies.' Below this is a code block with Python code for parsing a score and creating a reduction. A note below the code says: '(Note that this example shows some octaves retained...this has been fixed in newer releases) TODO: Fix example'. Below the code is a musical score snippet. On the right side, there is a sidebar with a 'Table Of Contents' and a 'Table Of Contents' section. The sidebar includes links for 'Previous topic', 'Next topic', 'Applications and Extensions of music21', 'Examples and Demonstrations', 'About Music21', and 'Music21 Concepts Overview'.

*description of
functionality*

Music21 is a Python-based toolkit for computer-aided musicology. Applications of this toolkit include computational musicology, music information, musical example extraction and generation, music notation editing and scripting, and a wide variety of approaches to composition, both algorithmic and directly specified.

URL

<http://web.mit.edu/music21/doc/about/index.html>

contact

<http://web.mit.edu/music21/doc/about/about.html>

license

GNU Public License

deployment

<http://web.mit.edu/music21/doc/developerReference/index.html>



8.5.2 The Humdrum Toolkit

🏠 humdrum-tools

[Humdrum Installation Guide](#)
[The Humdrum User Guide](#)
[Humdrum Extras](#)
[The Humdrum Cookbook](#)
Humdrum Command Manual
[Command: assemble](#)
[Command: barks](#)
[Command: cbr](#)
[Command: census](#)
[Command: cents](#)
[Command: cleave](#)
[Command: cocho](#)
[Command: context](#)
[Command: correl](#)
[Command: deg](#)
[Command: degree](#)
[Command: diss](#)
[Command: ditto](#)
[Command: dur](#)
[Command: ekern](#)

Docs » [Humdrum Command Manual](#) » [Command: iv](#)

Command: iv

COMMAND

`iv` - determine interval-class vectors for successive vertical sonorities for Humdrum inputs

SYNOPSIS

```
`` iv [inputfile...] [> outputfile.iv]``
```

DESCRIPTION

The `iv` command is used to determine the *interval-class vector* for any of five set-theory related inputs: pitch (`**semit`s), pitch-class (`**pc`), normal form (`**nf`), prime form (`**pf`), or Fortean set name (`**pcset`). An interval-class vector is a six-element numerical list that indicates the abundance of various interval-classes (from 1 semitone to 6 semitones) for some pitch-class set. See [REFERENCES](#) below.

When provided with `**semit`s or `**pc` inputs, `iv` treats each input record as a set of pitches. Unisons and other pitch-class duplications have no effect on the output. Rests within a set of pitches are ignored; where an input record consists solely of one or more rests, a null-token is output.

description of functionality

Humdrum is a set of command-line tools that facilitates musical analysis, as well as a generalized syntax for representing sequential streams of data. As a set of command-line tools, it is programming language agnostic. Many have employed Humdrum tools in larger scripts that use PERL, Ruby, Python, Bash, LISP, and C++.

URL

<http://www.humdrum.org/>

contact

<https://groups.google.com/forum/#!forum/starstarhug>

license

N/A

deployment

<http://web.mit.edu/music21/doc/developerReference/index.html>
<https://www.documentcloud.org/>