



Digitised Manuscripts to Europeana

Final Public Report

1st February 2012 – 31st January 2015



The project was co-funded by the European Union, through the **ICT PSP** programme http://ec.europa.eu/ict_psp

DM2E was coordinated by Humboldt-Universität zu Berlin



Contents

1 Summary	3
2 Project Objectives	4
3 Consortium	8
4 The Project's Work and Results	10
4.1 Content provided to Europeana	10
4.2 DM2E Model	14
4.3 Interoperability Infrastructure	15
4.4 Scholarly Domain Model	17
4.5 Scholarly Research Platform.....	18
4.6 Digital Humanities Experiments	20
5 Dissemination	22
6 Sustainability	30

1 Summary

This final public report marks the end of the *Digitised Manuscripts to Europeana* Project (DM2E) which was funded by the Europeana Commission under the CIP ICT PSP Programme. The project started on 1 February 2012 and ran for three years.

The Linked Data Web has huge potential for bringing textual scholarship to new levels. Scholars working in the humanities need the tools that will help them recognise this future-oriented potential. DM2E responded to this challenge by developing a suite of Linked Data tools that enables libraries, archives and research institutions to provide digital content to *Europeana* – Europe's digital library. Since different providers make their data available in different formats, the developed toolset is able to convert metadata from a diverse range of source formats into the DM2E model, an application profile of the Europeana Data Model (EDM). The developed software additionally allows the contextualisation and linking of this cultural heritage data sets.

The technical development is based on digitised content provided by some of Europe's leading cultural heritage institutions such as Staatsbibliothek zu Berlin, Berlin-Brandenburgische Akademie der Wissenschaften, the Austrian National Library, the Wittgenstein Archive at University of Bergen, the National Library of Israel plus a number of other sources. In line with the Europeana Data Exchange Agreement, all contributing institutions have agreed to make their metadata available under the Creative Commons Public Domain Dedication license (CC-0), which will make it easier to reuse in diverse contexts.

Finally, DM2E provided the aggregated data in a format enabling scholars, students and the wider public to interact with cultural artefacts and their associated metadata in innovative ways. A "Scholarly Domain and Discourse Model" was created in the project (which draws on John Unsworth's "Scholarly Primitives" and resulting work) in order to describe basic functions of humanities scholars work and has been applied to the *Scholarly Research Platform* (Pundit) as an ontology.

2 Project Objectives

The main focus of the project was on the development of Linked Data based solutions and tools that enabled libraries and archives to provide digital content to Europeana – Europe’s cultural heritage portal and virtual source repository. The second focus was on the development of a data model for the domain of manuscripts. Based on the requirements of its data providers, the project developed the DM2E model¹. This new data model is a specialisation of the *Europeana Data Model* (EDM)².

The developed software allows the conversion of diverse metadata formats to the DM2E and Europeana Data Models and the contextualisation and linking of vast cultural heritage datasets.

The project ingested 20.08 million manuscript pages from libraries, archives and research institutions into its Linked Data platform and also into Europeana. The resources came from national and state libraries as well as research institutes in Austria, Germany, Israel, Norway, Italy and Hungary.

During the project it was possible to recruit six new providers to deliver their content to the project and to Europeana. This made it possible to exceed the amount initially stated in the Description of Work by about 708,000 pages of relevant humanities content.

The Digital Humanities (DH) community is a community of “power users”, which worked with the DM2E project in order to validate the results of the projects work. The project created a “Scholarly Domain Model” in order to describe basic functions of humanities scholars work. This model was partly applied to the *Scholarly Research Platform Pundit*³. The platform provides a set of instruments for scholars to work with the DM2E data by making intellectual connections between different documents.

The DM2E project was comprised of five work packages (WP), whereas WP5 was dedicated to the project management. The interaction of WP1–WP3 is shown in Figure 1. The objectives of the work packages in detail were:

WP1: Content

The objectives of WP1 were to:

- Collect metadata formats and relational backend structures and requirements towards the EDM and DM2E mapping.
- Provide WP2 and WP3 with precise information about metadata and APIs available for accessing digitised objects contributed by each content provider.
- Implement the requirements into concrete workflows by using the tools provided by WP2, which were evaluated for their suitability.
- Test the content translation interface provided by WP2.
- Organise together with WP2 the Europeana ingestion of the metadata integrated in the DM2E platform.

¹ DM2E documentation: <http://dm2e.eu/outputs> [23.03.2015].

² EDM documentation: <http://pro.europeana.eu/share-your-data/data-guidelines/edm-documentation> [30.03.2015].

³ Pundit: <http://labs.europeana.eu/apps/Pundit> [24.03.2015].

- Set up a test scenario on semantic enrichment of digital material and on rich feature extraction from digital objects for the scholarly research platform of WP3 in close cooperation with the humanities scholars.
- Organise agreements regarding Europeana's Public Domain policy.

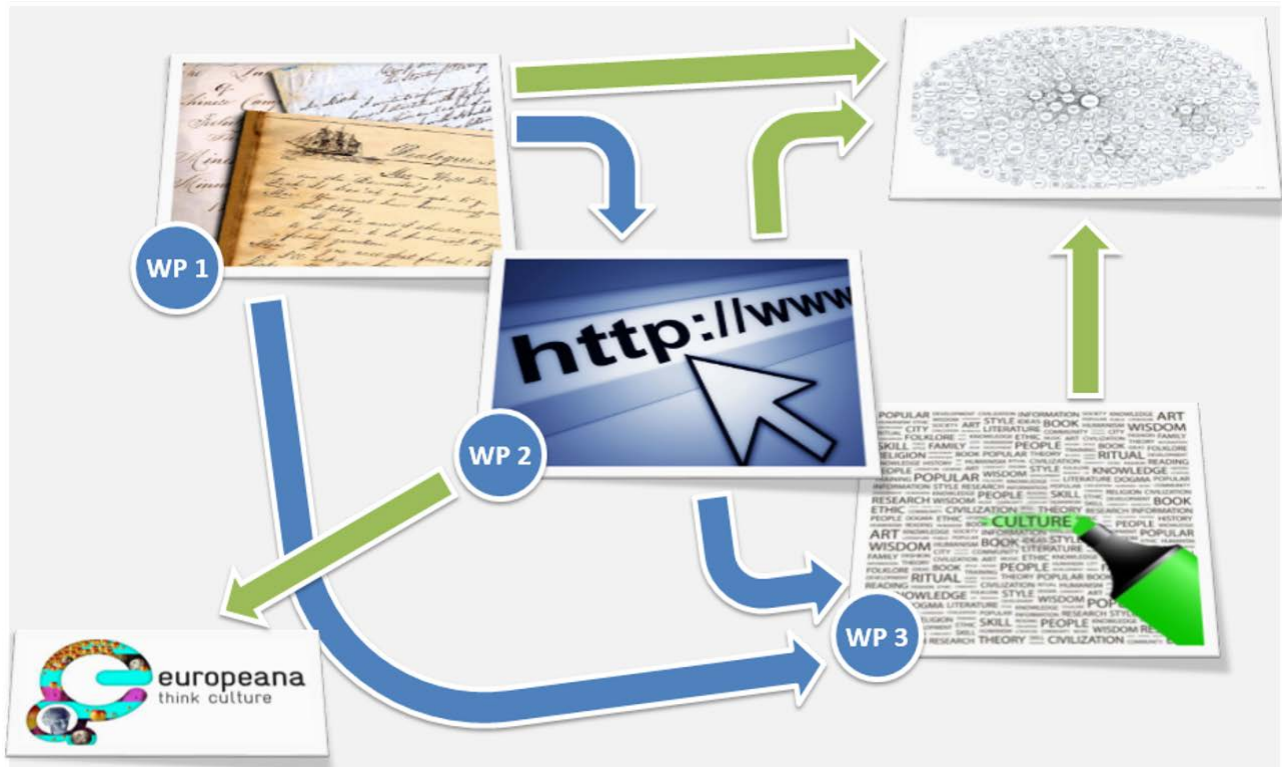


Figure 1: The DM2E architecture

WP2: Interoperability Infrastructure

The objectives of WP2 were to:

- Provide the interoperability infrastructure for translating content from current source formats into the EDM. The work was built on existing tools, adjusted to the DM2E requirements and chained into a scalable translation workflow. The workflow included the following three steps:
 - RDFization of existing content and metadata
 - Mapping into the EDM
 - Contextualisation and interlinking at the levels of legacy data.

The WP furthermore

- evaluated the tools regarding the development of the infrastructure for RDF conversion,
- aggregated the metadata and explored the content of the extracted metadata,
- developed together with WP1 the DM2E model and
- provided a graphical user interface which allowed the project partners to configure and schedule the translation of their content into the EDM.

The whole interoperability infrastructure is made available under an open source license and can thus be used by other content providers.

WP3: Digital Humanities Requirements and Related Engineering

The objectives of WP3 were to:

- Develop a prototyping platform named “Scholarly Research Platform” (Pundit) for sampling and annotating scholarly web resources.
- Lower the barrier to access for developers and advanced users in the Digital Humanities to reusing and recombining EDM data with other external data sources and user generated content by providing appropriate open source tools, annotation facilities and APIs.
- Demonstrate the added value of enhanced usability of Europeana by developing EDM support and demonstrating usefulness of co-presence of digital objects for reuse of content in the DH and thus creating an incentive for content custodians to release content for open and free development by expert communities of practice such as the DH.
- Encourage learning and community uptake and by providing input, in terms of tools, documentation and curated content collections for WP4 contest activity, the *Open Humanities Award*.
- Identify what are the “functional primitives” of the Digital Humanities and to create a “Scholarly Domain and Discourse Model” which draws on John Unsworth’s “Scholarly Primitives” in order to describe basic functions of humanities scholars work. This model was applied to the Scholarly Research Platform as an ontology.
- To conduct experiments in order to investigate how interpretative approaches of humanists can be operationalised in the particular context of Linked Data and Pundit and its components. For this purpose, humanists were confronted in real-life working contexts with the formal and explicit approach of Linked Data and semantic annotation.

The Digital Humanities community, a community of “power users”, worked with the DM2E project in order to validate the results of the projects work. An advisory board monitored this work. It consisted of the following members:

- Dr. Tobias Blanke (King's College, London)
- Sally Chambers (DARIAH-D)
- Prof. Dr. Gerhard Lauer (Göttingen Center for the Digital Humanities)
- Dr. Laurent Romary (French Institute for Research in Computer Science and Automation, INRIA)
- Prof. Dr. Felix Sasaki (DFKI and head of the German W3C-Office)
- Prof. Dr. Susan Schreibman (Trinity College Dublin)
- Dr. Claire Warwick (University College, London)
- Dr. Alois Pichler (Wittgenstein Archives, University of Bergen)
- Prof. Dr. Jürgen Renn (Max-Planck-Institut für Wissenschaftsgeschichte)

The meetings of the advisory board were chaired by Prof. Dr. Stefan Gradmann (Humboldt-Universität zu Berlin), Prof. Vivien Petras, PhD (Humboldt-Universität zu Berlin) and Dirk Wintergrün (Max-Planck-Institut für Wissenschaftsgeschichte).

WP4: Community Building and Dissemination

The objectives of WP4 were to:

- Disseminate the results of the project.
- Build and strengthen the community around Europeana.
- Support and strengthen the network of open metadata evangelists (via the OpenGLAM network) who helped to raise awareness of legal and technical best practises in a variety of different domains.
- Organise a series of events and update the community documentation on DM2E tools and workflows via the DM2E wiki.
- Organise a contest⁴ in order to encourage and highlight innovative work building on the technology developed as part of the project and for partnership working between developers and non-technical researchers.
- Facilitate conversations between technologies and researchers within the Europeana community.

⁴ The Open Humanities Awards: <http://openhumanitiesawards.org> [24.03.2015].

3 Consortium

The DM2E consortium consisted of experienced and highly competent partners in the field of digitised manuscripts both from a scholarly and from a technical perspective (such as University of Bergen or Net7), from the Linked Data community (such as with Kai Eckert who is conducting research internationally in this area) and from the Europeana context (such as Humboldt-Universität zu Berlin with Stefan Gradmann and Vivien Petras as experts in the Europeana community or the Austrian National Library with Max Kaiser as a leading figure in several Europeana related projects). The consortium brought together a good mix of cultural heritage institutions, academic partners and two SMEs specialised in the field (ExLibris and Net7).



Figure 2: The DM2E consortium in Pisa, 12 December 2014

Beneficiaries

Humboldt-Universität zu Berlin, Berlin School of Library and Information Science, Germany

<https://www.ibi.hu-berlin.de>

European Association for Jewish Culture, France

<http://www.jewishcultureineurope.org>

ExLibris, Germany

<http://www.exlibrisgroup.com>

Max-Planck-Institut für Wissenschaftsgeschichte, Germany

<https://www.mpiwg-berlin.mpg.de>

Net7 S.r.l., Italy

<http://www.netseven.it/en>

National Technical University of Athens, Greece

http://www.ntua.gr/index_en.html

Open Knowledge Foundation, United Kingdom

<https://okfn.org>

Österreichische Nationalbibliothek, Austria

<http://www.onb.ac.at>

Staatsbibliothek zu Berlin, Germany

<http://staatsbibliothek-berlin.de/en>

University of Bergen, Wittgenstein Archives, Norway

<http://wab.uib.no>

Universität Mannheim, Germany

<http://dws.informatik.uni-mannheim.de/en>

Associated Partners

University Library Johann Christian Senckenberg Frankfurt, Germany

http://www.ub.uni-frankfurt.de/judaica/judaica_en.html

Bulgarian Academy of Sciences, Bulgaria

<http://www.bas.bg>

Georg Eckert Institute for Textbook research, Germany

<http://www.gei.de/en/home.html>

EUROCORR, Italy

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

Petőfi Literary Museum (PIM), Hungary

<https://www.pim.hu>

University Archives & Special Collections Department of Brandeis University,
United States

<http://lts.brandeis.edu/research/archives-speccoll>

American Jewish Joint Distribution Committee

<http://archives.jdc.org>

Center for Jewish History, United States

<http://www.cjh.org>

4 The Project's Work and Results

4.1 Content provided to Europeana

One objective of the DM2E project was to enable new content providers to contribute their content to Europeana. During the project term it was possible to gather six new partners who provided their content. This new partners were

- University Library Johann Christian Senckenberg Frankfurt, Germany
- The Bulgarian Academy of Sciences, Bulgaria
- The Georg Eckert Institute for Textbook research, Germany
- EUROCORN, Italy
- Petőfi Literary Museum (PIM), Hungary
- University Archives & Special Collections Department of Brandeis University, United States.

All associated partners made a significant investment by contributing their content to the DM2E project without being funded in the project. The partners have been supported through the project in the creation of their mappings and in the ingestion of the new data into the DM2E platform and Europeana.

All content provided to Europeana can be found on the Europeana portal <http://www.europeana.eu/portal>. In the following the content is described for each provider:

The **Austrian National Library** (ONB) provided heterogeneous content from the Google Books Public Private Partnership ABO (Austrian Books Online)⁵ as well as digitised manuscripts (codices). The ABO content includes mainly books, but also newspapers and already digitised manuscripts in different languages. The content is in German, Latin, French, Italian, English, Czech and other various languages. During the Austrian Books Online project, the Austrian National Library digitises and makes available online its complete holdings of historical books from the 16th to the 19th century. The manuscripts (codices) are in various languages such as Latin, German, Greek, Italian, Arabic, Turkish or French.

Kalliope, the Union Catalogue for Personal Paper and Autograph Collections⁶ in Germany, is hosted by the **Berlin State Library** (SBB). Metadata from 491 Cultural Heritage Institutions have been aggregated to date. About 20,000 collections and 1.64 million autographs covering a time span from the early 16th to the late 20th century can be retrieved online. The SBB provided the personal papers (Bequest) of the German writer Gerhart Hauptmann and the Bequest of the German poet and botanist Adalbert von Chamisso. While the material from Hauptmann is mainly in German, about 30% of the content from Chamisso includes foreign-language letters and texts. The SBB additionally offered collections with high relevance to the digital humanities community: the Publisher Archives of Gebauer & Schwetschke and Western Manuscripts. The Publisher Archives range from the 18th to the 19th century and include letters and manuscripts e.g. concerning Johann Georg Walch, 1693-1775 (Walch Edition of Luther's works), Johann Chr. Gatterer, 1727-1799, (Allgemeine Historische Bibliothek), etc. Furthermore, SBB provided from its manuscript division another 45,000 images of 164 western manuscripts ranging from the 9th to the 16th century.

⁵ Austrian Books Online: <http://www.onb.ac.at/ev/about/austrianbooksonline.htm> [30.03.2015].

⁶ Kalliope: <http://kalliope.staatsbibliothek-berlin.de> [30.03.2015].

The Wittgenstein Archives at the **University of Bergen** (WAB) contributed the Wittgenstein Nachlass⁷ (bequest) kept at Trinity College Cambridge (TCC), Wren Library, Cambridge, in facsimile and text editions as well as metadata.

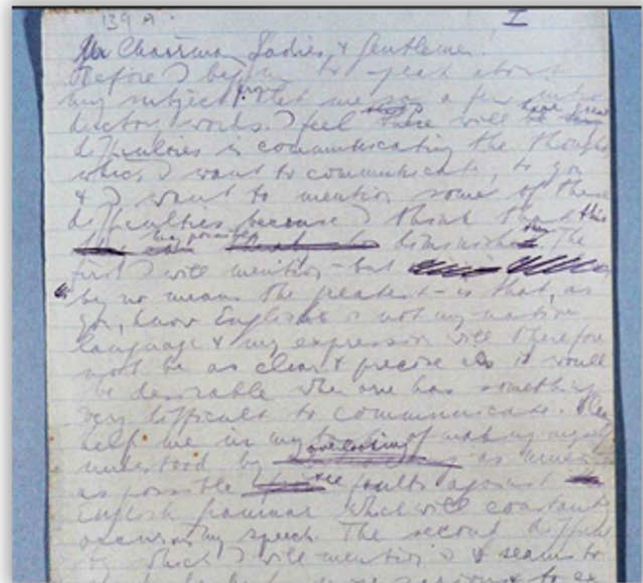
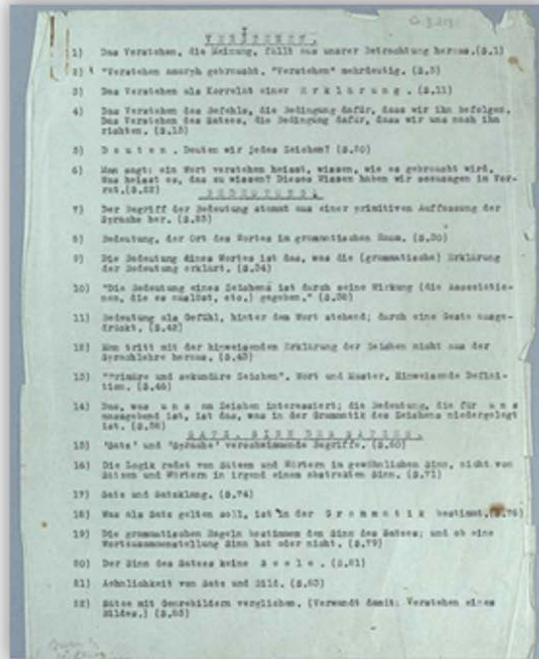


Figure 3: Manuscripts by Ludwig Wittgenstein: "The Big Typescript" and "Lecture on Ethics"

The bequest includes:

- Ts-201a1 and Ts-201a2 (1913-14) from the "Notes on Logic" corpus: This contains the first expression of Wittgenstein's view on philosophy and specific material on logic.
- Ms-139a and Ts-207 (1929) from the "Lecture on Ethics" corpus: The "Lecture on Ethics" is the only "popular" lecture which Wittgenstein ever held. It is a masterpiece expressing his early views on tensions between meaningful language and nonsensical language in which we expose ourselves when speaking about ethics.
- Ms-114, Ms-115 (first part), Ms-148, Ms-149, Ms-150, Ms-153a, Ms-153b, Ms-154, Ms-155, Ms-156a, Ts-212 and Ts-213 (1931-34) from the "Big Typescript" corpus: The Big Typescript is the great "summa" of Wittgenstein's latest thoughts elaborated from 1929 to 1933. It addresses most philosophical subjects Wittgenstein has ever been interested in.
- Ms-115 (second part), Ms-140 (p. 39v), Ms-141, Ms-152 and Ts-310 (1934-36) from the "Brown Book" corpus: The "Brown Book" was first lectured to students in Cambridge and is thus laid out very pedagogically. One aspect is the introduction of "language games" which shed light on the complexity of our language. The Book was revised several times. The last revision led to the Philosophical Investigations.

⁷ The Wittgenstein Archives at the University of Bergen: <http://wab.uib.no> [30.03.2015].

In terms of philosophical development, the bequest included four high points in Wittgenstein's philosophical development from 1913 to 1936; in terms of philosophical themes, they encompass all themes addressed by Wittgenstein incl. philosophy of language, logics, mathematics, psychology, ethics, metaphilosophy, etc.; in terms of type of Nachlass material they include all kinds of manuscripts written by Wittgenstein: first drafts, lecture notes, notebooks, copybooks, typescript cuttings, and elaborated materials such as typescripts, and materials prepared in cooperation with others. In terms of language, they contain both German and English materials.

Humboldt-Universität zu Berlin (Humboldt University Berlin, UBER) supplied the "Polytechnisches Journal". In 1820, the German chemist and industrialist J. G. Dingler first started publishing the »Polytechnisches Journal«. The journal was published over a period of 111 years and has become an important and European wide source for the history of knowledge, culture, and technology — at least in Germany it is unparalleled. The project was a cooperation between the Institute for Cultural Studies (Institut für Kulturwissenschaft) at the Humboldt-Universität zu Berlin, the Sächsische Landesbibliothek — Staats- und Universitätsbibliothek Dresden (SLUB – Saxon State and University Library Dresden) and an experienced service provider for digitisation, Editura GmbH. Access to Dingler-Online is free of charge.⁸

Besides digitisation of the images also the OCRed text is encoded according to the Text Encoding Initiative Guidelines TEI-P5. On the one hand, this enables to grasp the editors' true spirit of literally collecting polytechnical works; on the other hand, a very comfortable access to the articles is provided.

The **Max-Planck-Institut für Wissenschaftsgeschichte** (Max Planck Institute for the History of Science, MPIWG) provided Islamic scientific manuscripts (ISMI)⁹ from the State Library for the project. ISMI is an international collaborative endeavour that is working to make available a vast array of information online about the exact sciences in the premodern Islamic world. The MPIWG also offered the collection ECHO which comprises the metadata of the institute's rare book collection which consists of approximately 2.500 volumes of which most are also available in digital form. Additionally, digitised full text is present for about 100 volumes. MPIWG furthermore provided the autographs of Thomas Harriot¹⁰ to the project.

The **European Association for Jewish Culture** (EAJC) served as aggregator for the content of the American Jewish Joint Distribution Committee (JDC), Center for Jewish History (CJH) and the National Library of Israel (NLI).

The American Jewish Joint Distribution Committee shared the earliest text collection in their archives, the records of the American Jewish **Joint Distribution Committee** of the years 1914–1918¹¹. The material includes chronologies of events, correspondence, cables, reports, minutes and transcripts of meetings and conferences that document the origins, beginning in 1914, of systematic financial aid from American Jewish organisations to imperilled Jews abroad.

⁸ Polytechnisches Journal Online: <http://www.polytechnischesjournal.de/en> [24.03.2015].

⁹ The Islamic Scientific Manuscripts Initiative: <http://ismi.mpiwg-berlin.mpg.de> [24.03.2015].

¹⁰ Echo Cultural Heritage Online: http://echo.mpiwg-berlin.mpg.de/content/scientific_revolution/harriot [24.03.2015].

¹¹ Archives of the American Jewish Joint Distribution Committee: <http://archives.jdc.org> [30.03.2015].

The **Center of Jewish History** (CJH)¹² is one of the foremost Jewish research and cultural institutions in the world, having served over one million people in more than 100 countries. It is home to five partner organisations — American Jewish Historical Society, American Sephardi Federation, Leo Baeck Institute, Yeshiva University Museum and YIVO Institute for Jewish Research. The Leo Baeck Institute (LBI) and YIVO both hold diverse collections on European Judaica and CJH contributed metadata for digitised objects of these collections to DM2E via EJAC.

The **National Library of Israel** (NLI) provided metadata for items from the Judaica collections via EAJC as well for items from the general collections directly to DM2E. NLI is an associated partner of Judaica Europeana (JE) and provided content from different selections from its Rosetta repository via EAJC. The books, manuscripts and archival material is in Hebrew, Yiddish and Arabic.

The **University Archives & Special Collections Department of Brandeis University** (BRANDEIS) offered data on selected collections to DM2E via Judaica Europeana with the support of EAJC.

The Deutsches Textarchiv (German Text Archive, DTA)¹³ of the **Berlin-Brandenburgische Akademie der Wissenschaften** (BBAW) provided digitised and transcribed 1st edition publications from the 17th to the end of the 19th century online. In close collaboration with SBB, BBAW brought in these data into the interoperability infrastructure of WP2. When contextualised and transformed into EDM, SBB made these data available for ingest by Europeana.

University Library Johann Christian Senckenberg Frankfurt (UBFFM) provided DM2E with its collection of medieval manuscripts back to the 9th century¹⁴. The medieval, oriental and modern manuscripts represent a significant contribution of annotatable online manuscript content for the scholarly community. Furthermore UBFFM offered the Bequest of Max Horkheimer¹⁵.

The **Bulgarian Academy of Sciences** (BAS) provided DM2E with the full electronic corpus and digitised images from the 10th century Codex Suprasliensis, listed in the UNESCO's Memory of the World Register since 2007¹⁶.

The **Georg Eckert Institute for Textbook research** (GEI) contributed metadata and digital objects from its digital collection GEI-Digital. The collection contains scanned images in full text.

EUROCORR is an ERC AdG project which aims at publishing the **European Correspondence to Jacob Burckhardt** (1842-1897). The project's main output is the platform Burckhardtsource.org¹⁷, which hosts the manuscripts facsimile as well as XML TEI P5 transcriptions and semantic annotations on the letters. EUROCORR provided metadata and facsimiles from its digital collection to the project.

The **Petőfi Literary Museum** (PIM) contributed articles from the Hungarian avant-garde Magazine "A Tett", edited by Lajos Kassák to the project¹⁸.

¹² Center for Jewish History: <http://www.cjh.org> [30.03.2015].

¹³ Deutsches Textarchiv: <http://www.deutschestextarchiv.de> [30.03.2015].

¹⁴ Medieval Manuscripts at UBFFM: <http://sammlungen.ub.uni-frankfurt.de/msma?lang=en> [24.03.2015].

¹⁵ Bequest Max Horkheimer: <http://sammlungen.ub.uni-frankfurt.de/horkheimer>

¹⁶ Codex Suprasliensis: <http://suprasliensis.obdurodon.org> [24.03.2015].

¹⁷ The European Correspondence to Jacob Burckhardt: <http://www.burckhardtsource.org> [24.03.2015].

¹⁸ Kassák Museum: http://www.kassakmuzeum.hu/en/index.php?p=lajos_kassak [24.03.2015].

Related documents: [D1.1 - Requirements Report](#) and [D1.2 - Final Integration Report](#)

4.2 DM2E Model

The DM2E model, developed within the DM2E project, is a specialisation of the Europeana Data Model (EDM)¹⁹ for the domain of manuscripts²⁰. The DM2E ontology is machine-readable and described with full provenance metadata.

The ontology has been developed regarding the requirements of the projects' data providers and is fully based on the Linked Data paradigm. It thus allows the easy distribution of the ontology beyond the project's end via different ontology repositories. The final version of the model is distributed twofold in order to achieve higher visibility and sustainability of the model. Apart from DM2E's own ontology publishing service²¹, an installation of the Web-based RDF Schema vocabulary publishing system Neologism, the model is also published on external ontology libraries like: (<http://lov.okfn.org/>, <http://vocab.deri.ie/>, <http://metadataregistry.org/>, <http://datahub.io/>, <https://onki.fi/>). As long as the DM2E server will be maintained by UBER's Computer- and Media Service, the ontology will be available for supporting content negotiation.

The EDM has been developed within the Europeana v1.0 project as an RDF-based data model for describing rich metadata records for Europeana. It can handle huge metadata record collections represented by heterogeneous metadata standards that have to be accessible via the same platform. The EDM covers Cultural Heritage Objects (CHOs) that are collected and delivered to Europeana by diverse cultural heritage institutions. The model is very generic to cover Europe's rich and diverse CHOs but it can be specialised for domain-specific descriptions like it was the case in DM2E.

An important contribution of DM2E was the fundamental work toward EDM application profiles²², where DM2E provided an application-specific, richer view on the metadata. Questions to be answered are how to express and represent the relationship between EDM and its application profiles like DM2E in the data model, how to express and represent constraints on the data within the data model and how to ensure validity of the data at all levels: EDM, DM2E and vocabularies reused in EDM like Dublin Core. To avoid isolated solutions, DM2E initiated a task group at the Dublin Core Metadata Initiative²³ to address these questions together with Europeana and other stakeholders in the cultural heritage community. Starting point of the task group's work was the comparison of elements and validation rules in EDM, the DM2E model and the EDM application profile of the German digital library (DDB). The validation rules will be formalised and categorised. Application profile use cases from other projects and institutions are also considered leading to a formulation of best-practices for the creation of RDF application profiles in general.

The group members have conducted a case study for the validation of data described with the DM2E model, which was published and presented during a workshop titled "RDF

¹⁹ EDM documentation: <http://pro.europeana.eu/edm-documentation> [30.04.2015].

²⁰ Our definition of manuscript is very broad and includes not only handwritten manuscripts but all kind of text documents. Thus, the model also includes classes for physical objects like journals and even letters.

²¹ <http://onto.dm2e.eu/> [30.03.2015].

²² Application profiles like the DM2E model can specialise or refine vocabularies for a specific application, e.g. by mixing and matching certain ontology elements or by providing narrower element definitions.

²³ "RDF Application Profiles" Task Group: http://wiki.dublincore.org/index.php/RDF_Application_Profiles [30.03.2015].

Application Profiles in Cultural Heritage”²⁴ at the SWIB conference in December 2014. The RDF-AP task group is chaired by Antoine Isaac (Europeana) and Karen Coyle (kcoyle.net). Evelyn Dröge (UBER, DM2E) co-chairs the editorial board of the group. Kai Eckert (UMA), who is also chair of the DCMI Technical Board, as well as four other people from WP2 contribute to the task group, where DM2E is a flagship case study. The research activity on a global solution for RDF data validation initiated by the Dublin Core task group on RDF Application Profiles (RDF-AP) within the Dublin Core Metadata Initiative will be further continued.

The project presented an analysis of the metadata mappings from its different providers to the DM2E model. It described the differences between individual data providers and their respective metadata mapping cultures and provided explanations on how the providers map the metadata from different institutions, different domains and different metadata formats and supported it by visualizations. The analysis of the mappings served the project to evaluate the DM2E model and provided strategic insight for improving both mapping processes and the model itself. The results were presented at the Dublin Core Conference 2014 and published.²⁵

Related documents: [D1.1 - Requirements Report](#) and [DM2E Model V 1.2 Specification](#)

4.3 Interoperability Infrastructure

All metadata in DM2E stems from various cultural heritage institutions across Europe and is maintained and described in various formats, among others MARC 21, MAB2, TEI, and METS-MODS. These formats provide enough degrees of freedom to reflect specific requirements of the providers. Changes in the original metadata and adaptations of the transformation process lead to new versions of the published data, for which the provenance has to be stored and provided.

The interoperability infrastructure consists of the DM2E model describing the digitised manuscripts, a workflow ontology describing the transformation of the metadata, as well as the metadata provenance, a workflow execution engine with a graphical user interface (OmNom), the MINT mapping tool to support the graphical generation of mappings, a Silk-based contextualisation service for the generation of links, a Linked Data API that exposes the generated RDF data to the tools developed in WP3, an OAI-PMH service to provide the transformed data as EDM data to Europeana, and finally a search-and-browse interface as main access point for the users.

The central element is the DM2E GUI. It subsumes and integrates four different GUIs: OmNom, MINT, Silk, and the User Management. The user accesses the system via the OmNom Workflow Management System, logs into the system and can then create, edit, and use various workflows to ingest data to the DM2E data store. A typical workflow consists of a transformation based on a mapping created in MINT and the generation of links to external sources based on linkage rules created in the SILK Workbench. All GUIs are interlinked and provide single-sign-on, i.e., once a user logs in to one GUI, she or he is automatically logged in to other GUIs as well; the user experiences an integrated system. The single-sign-on is realized via a central authentication server providing an additional GUI for the user

²⁴ Semantic Web in Libraries 2014 Conference: <http://swib.org/swib14/programme.php> [30.03.2015].

²⁵ The International Conference on Dublin Core and Metadata Applications 2014: <http://dcevents.dublincore.org/IntConf/dc-2014> [24.03.2015].

management (registration of new users, editing of accounts, retrieval of forgotten passwords etc.).

Figure 4 shows an overview of the interoperability infrastructure.

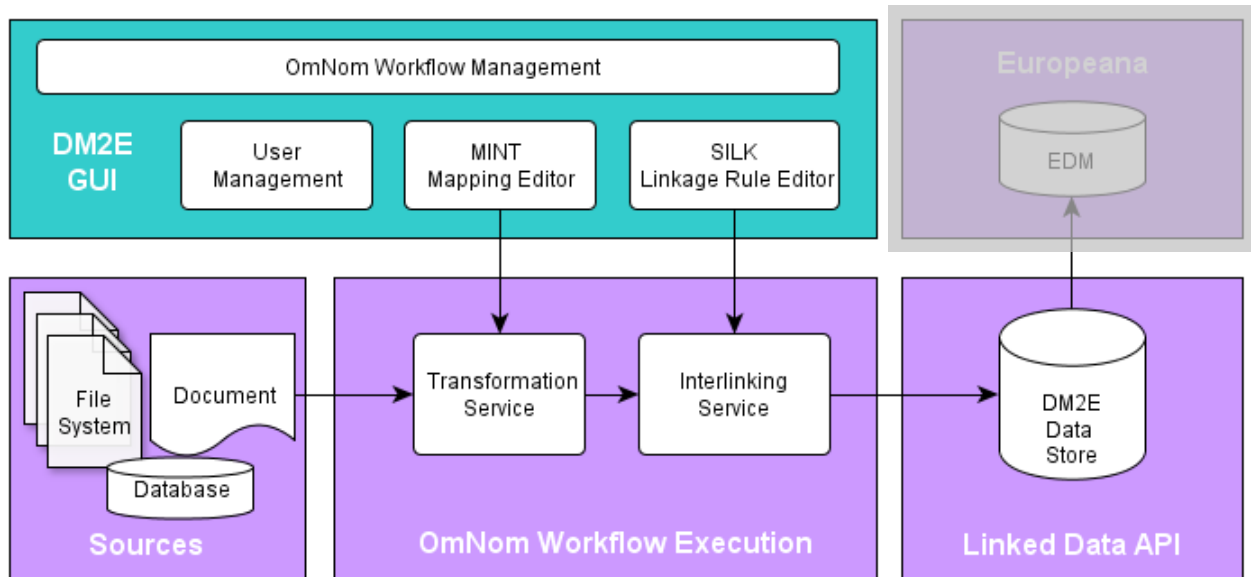


Figure 4: Overview of the interoperability infrastructure

An important difference to the original workflow is the provision of the data via the DM2E data store. The data is not directly ingested to Europeana for a simple reason: in DM2E, cultural heritage objects (CHOs) can be further granularised. Every single page of a manuscript, a chapter, an article in a journal, every transcription of a page or parts of a page like paragraphs or even individual sentences can be represented as single CHOs. The level of granularity and the identification of the single CHOs are determined by the providers during the mapping process. This granularity usually varies from provider to provider and from dataset to dataset but is generally much more fine-grained than the granularity of the CHOs that are to be ingested into Europeana, which is usually the level of whole manuscripts. Notable exceptions from this rule occur, when individual pages or images are deemed noteworthy enough to be presented in Europeana as a CHO of its own, with an individual page in the Europeana portal. This means that access to the DM2E data is provided via a dedicated Linked Data API backed by the DM2E data store. A subset of this data will subsequently be transformed from the DM2E model to the Europeana Data Model (EDM) and then ingested into Europeana. This transformation is straight forward as the DM2E model is designed as a full specialisation of EDM.

With the development of the interoperability infrastructure the DM2E project fulfilled two primary objectives:

- The transformation of various metadata and content formats describing and representing digital cultural heritage objects (CHOs) in the realm of digitised manuscripts from as many providers as possible into the EDM.
- The stable provision of the data as Linked Data and the creation of tools and services to reuse the data in the Digital Humanities, i.e., to support the so-called "Scholarly

Primitives". The basis is the possibility to annotate the data, to link the data, and to share the results as new data.

Related documents: [D2.1 - Initial Version of the Interoperability Infrastructure](#), [D2.2 - Intermediate Version of the Interoperability Infrastructure](#) and [D2.3 - Final Version of the Interoperability Infrastructure](#)

4.4 Scholarly Domain Model

The SDM has been devised based on the assumption that understanding what John Unsworth (2000) had originally proposed in terms of Scholarly Primitives is central to any such approach at modelling the digital scholarly domain. Unsworth's Primitives are understood as "basic functions common to scholarly activity across disciplines, over time, and independent of theoretical orientation"²⁶. Like other models, the SDM takes up the notion of Primitives and develops them further. Based on analysing and observing the practices of digital scholarship, we are endeavouring to acquire a better understanding of the requirements for instructing the development of sustainable infrastructures that enable scholars to harness the potential of digital technology and hence to develop appropriate digital methodologies. This requires to proceed beyond the establishment of static models to the iterative and continuous activity of "modelling"²⁷. For this reason, the SDM is conceived as an explicit but not definite set of the constituents of the domain of digital scholarship in the humanities. Similar to Manfred Thaller in his talk "Praising Imperfection"²⁸, we believe that modelling is the goal, not the model.

In this regard, *Linked Data* standards²⁹ such as the Resource Description Framework (RDF), Resource Description Framework Schema (RDFS), and Web Ontology Language (OWL) constitute a well suited means for the development of the SDM, because they allow the process of modelling to be iterative and continuous since the graph of semantic statements created is extensible. As we will see, this is also an instance of a still uncommon and emerging way to think of Linked Data as an art with epistemological implications for the practice of modelling the domain of digital scholarship in the humanities³⁰.

One of the main activities of the DM2E project has been working on further developing a digital humanities collaboration environment which is built around the Scholarly Research Platform. Pundit along with additional modules enables scholars to work with digitised manuscripts in the Linked Open Data (LOD)³¹ Web. The development of this collaborative

²⁶ Unsworth, J. (2000). Scholarly Primitives: What methods do humanities researchers have in common, and how might our tools reflect this? <http://people.brandeis.edu/~unsworth/Kings.5-00/primitives.html> [30.03.2015].

²⁷ McCarty, W. (2004). Modeling: A Study in Words and Meanings. In Schreibman, S., Siemens, R. and Unsworth, J. (eds), *A companion to digital humanities*. Malden, MA: Blackwell. <http://www.digitalhumanities.org/companion> [30.03.2015]. McCarty, W. (2005). Humanities computing. Houndmills: Palgrave Macmillan. Beynon, M.; Russ, S. and McCarty, W. (2006). Human Computing: Modelling with Meaning, *Literary and Linguistic Computing*, 21 (2), pp. 141–157.

²⁸ Thaller, M. (2013). Praising Imperfection: Why editions do not have to be finished. Lecture held at Culture & Technology - The European Summer School in Digital Humanities, Leipzig, July 2013. http://www.culingtec.uni-leipzig.de/ESU_C_T/node/292 [30.03.2015].

²⁹ Cf. for the following standards http://www.w3.org/standards/techs/rdf#w3c_all and http://www.w3.org/standards/techs/owl#w3c_all.

³⁰ Cf. Oldman, D., Doerr, M. and Gradmann, S. (n.d.). ZEN and the Art of Linked Data. New Strategies for a Semantic Web of Humanist Knowledge. To be published in Schreibman, S., Siemens, R. and Unsworth, J. (eds), *A new Companion to Digital Humanities*. Oxford: Blackwell [preprint].

³¹ W3C "Linked Data": <http://www.w3.org/standards/semanticweb/data> [30.03.2015].

research environment and the modelling process of the SDM have partly informed each other. In particular the project focused on how better models of the way in which students and scholars conduct research can be used to support the development of tools that enable users to interact with collections of texts and metadata - including transcription, translation, annotation, and curation.

The research gap identified and addressed within the DM2E project is the lack of a model which stresses the importance of creating a bridge connecting the analogue and digital scholarly practices and, most importantly, stresses the recursive relationship between these scholarly practices and the models and applications reflecting on them. This kind of research falls within what is typically called “Digital Humanities” and which is understood here as a community of practices, regardless of their particular materiality. In order to be able to discuss the “Digital Humanities” in a way that goes beyond simply discussing infrastructure so that the aforementioned challenge can be overcome, it has to be started from a “modelling process” that allows for the systematic and theoretically grounded building of bridges between practices of humanist research approaches in both the analogue and digital world.³²

The results of DM2E are intended to contribute to the emerging digital, networked and distributed environments, well beyond traditional working paradigms in the scholarly culture of the humanities. The SDM plays a pivotal role in this respect as a framework for better understanding scholarly research practices and the ways digital working modes might evolve in the future.

The work on the SDM has continuously been monitored and counselled by the Digital Humanities Advisory Board (DHAB) where DM2E has brought together scholars of the digital humanities in Europe.³³

Related document: [D3.4 - Research Report on DH Scholarly Primitives](#)

4.5 Scholarly Research Platform

Within the DM2E project, a Scholarly Research Platform (Pundit)³⁴ was developed and constantly improved. Pundit is an open source semantic annotation tool developed by Net7 allowing users to create structured data annotating the Web. Pundit not only enables users to annotate Web pages in various ways (comments, citations, bookmarks, etc.) but also to convert these annotations into semantically structured data that can be later integrated to the “Semantic Web”.

The Scholarly Research Platform has a two-fold target:

- To provide support for common scholarly activities such as commenting, annotating, collecting, and sharing.
- To provide a framework for developers to build on, by reusing data and annotations, deploying the tools for specific communities and by providing specialised visualisations/explorations of data.

The Scholarly Research Platform is an *ecosystem* of software components. They interact with each other – mostly using web APIs – and provide building blocks for implementing

³² Cf. McCarty (2005) as well as Beynon et al. (2006) which delineates our theoretical background of the modelling process.

³³ Digital Humanities Advisory Board: <http://dm2e.eu/dhab/> [30.03.2015].

³⁴ Pundit: <http://www.thepundit.it> [30.03.2015]. Pundit is originating from the SemLib project.

end-user workflows. The most important parts of the workflows (also known as primitives) are annotation and (semantic) augmentation. By annotating and augmenting the content, scholars can collect items of interest and use them to add meaning, or context, to the content itself.

The system supports deep linking of text and images to the Linked Data Web and to controlled vocabularies and taxonomies. It allows the creation of semantically rich annotations using typed relations, for example connecting a portion of a writing to another one asserting that they disagree with each other. The knowledge collaboratively created by users, which maintains provenance and attribution to annotators, forms a big semantic network, which the system manages and exposes in RDF format.

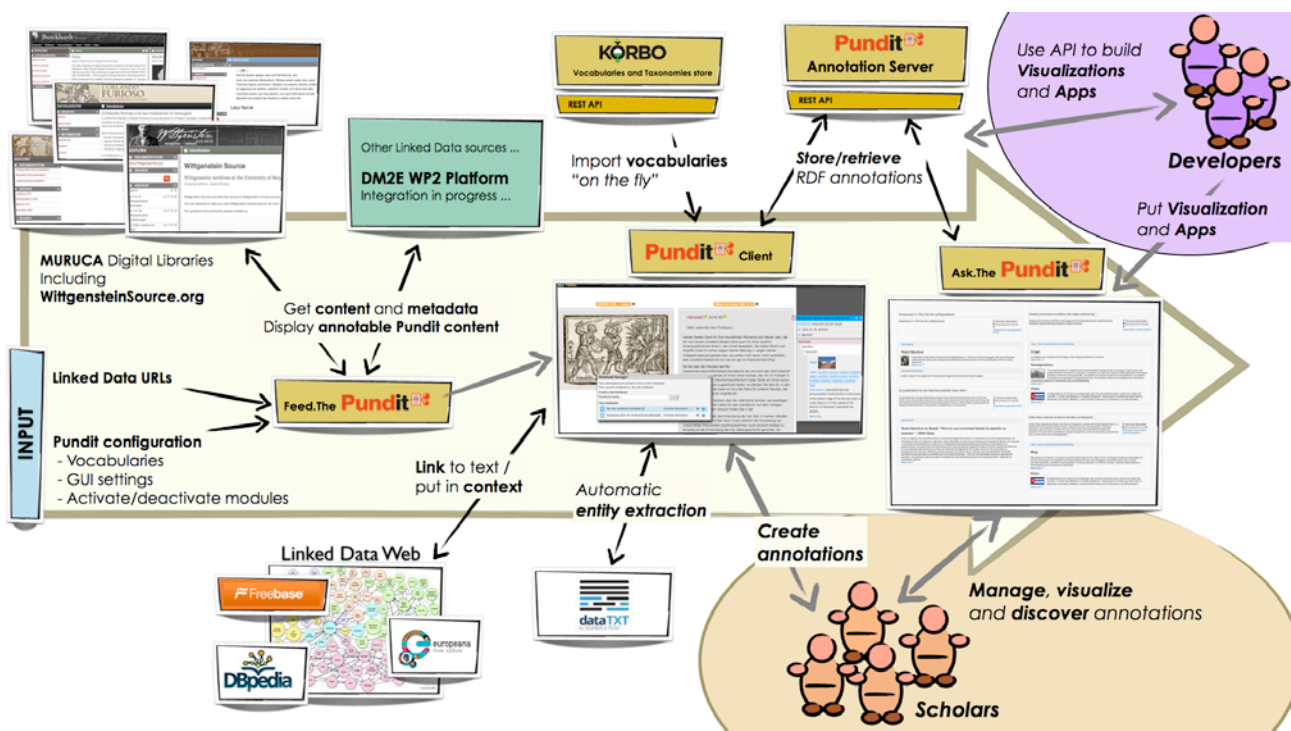


Figure 5: Scholarly Research Platform "Pundit": The software ecosystem overview

The core components of the infrastructure are (see Figure 5):

- *Pundit* as a configurable annotation tool:
 - *Pundit Server* to store annotations as RDF and provide REST API to data.
 - *Pundit Client* provides a highly configurable, web based user interface for annotating online content.
- *Ask* for collecting/managing/searching public and private annotations as well as to access vertical applications, which are, for example, visualisations of data for a specific domain;
- *Feed* to expose the *Pundit client* as a REST web service and connect it to external web applications

These four components cooperate to create end-user workflows where knowledge is created via annotations, shared and accessed via Ask, and consumed by additional applications to meet specific needs.

The framework to represent the annotations is based on the Open Annotation model. The ability to express semantic relationships between resources relies on the use of ontologies and vocabularies (such as Freebase.com and Dbpedia.org) that can be configured according to the needs of the communities using Pundit. The re-use of resources already available on the web provides solutions to multilinguality or disambiguation related challenges. In Scholarly Research Platform each annotation is described by data structured in RDF which connects the annotated object (e.g. a manuscript, a video) to other objects or entities available as Linked Data. Semantic relationships based on URIs are created between the different entities identified in a given web page. Pundit stores annotations on an RDF based server which exposes APIs to consume annotations along with their structured content.

Pundit is also a collaborative research platform which allows users to team up during the annotation creation process, to share their results and also re-use annotations created by other users. It therefore participates in the dialogue between scholars and in the creation of new knowledge (see Figure 5).

Related documents: [D3.1 - Initial Specification Report](#), [D3.2 - Prototyping Platform Implemented](#) and [D3.3 - E-learning Courses published](#)

4.6 Digital Humanities Experiments

The research interest of the experiments was to investigate how interpretative approaches of humanists can be operationalised in the particular context of Linked Data and Pundit and its components. For this purpose, humanists were confronted in real-life working contexts with the formal and explicit approach of Linked Data and semantic annotation.

The experiments particularly focused on the Scholarly Activities *annotating* and *visualising* both of which are seen as being pivotal to most humanists research activities. The aim was to investigate how these two activities materialise in different real-life use cases focusing on interpretative approaches of humanists which have no prior knowledge of Linked Data or semantic annotation tools such as Pundit.

The principal topical and temporal horizon for the experiments was Historical Sciences and Contemporary History (19th/20th century)³⁵. The three distinct use cases chosen for the experiment belong to the historical-archival domain. The first use case which has been created in cooperation with the Berlin-Brandenburgische Akademie der Wissenschaften (BBAW) and the Fachhochschule Potsdam (FHP) focused on the editorial and archival sciences. The use case created with the Georg-Eckert-Institut (GEI) focused on educational history. The use case devised with the historical seminar of the Humboldt-Universität zu Berlin (HUB) focused on visual history and the didactics of history, i.e. using tools such as Pundit for teaching history to students.

The experiments investigated the following principal research questions:

- How can genuine research questions and interests be operationalised within the context of Linked Data and the particular context of Pundit?
 - How do non-experts deal with Linked Data concepts and approaches?
- How are the Scholarly Activities annotation, visualisation, and interpretative modelling reflected as Scholarly Operations?

³⁵ The main reason for choosing this topical and temporal orientation was that the two organisers of the experiments are trained historians which was expected to facilitate and support the experiments.

- Which “statements” do humanists consider necessary in their particular use case?
- Which potential do they see in Linked Data and Pundit for applications in the humanities?

In sum, data has been collected by open **interviews**, i.e. open discussions with the participants during every phase of the experiments, **observation** during the workshops including the research data created and collected in the notebooks, and via a common **questionnaire** at the end of each experiment.

The main outcome of the experiments is the recognition of a basic tripartite research and reasoning process revolving around the Scholarly Activities of Interpretative Modelling and Annotating in the context of Linked Data based VREs. It involves first creating or choosing an annotation vocabulary, applying the vocabulary to the source material, and exploring the created annotations in order to create new hypotheses. This tripartite process can be conceptualised as an expression of research processes on the level of the Scholarly Operations and described using terminology from the SDM. Encouraging humanists to work with Linked Data requires taking a step towards translating the objects of study and methods of humanist research in Linked Data paradigm and taking a step away from concentrating on what can be automated: The tripartite process can be understood as the principle answer to the question which kinds of reasoning humanists want to see enable in Linked Data based VREs: The process enables them to apply their own reasoning practices to a certain extent and also provides a framework for further systematic investigation into the question how interpretative approaches of humanists might translate and be applied in Linked Data annotation environments.

Related document: [D3.4 - Research Report on DH Scholarly Primitives](#)

5 Dissemination

One of the primary targets of the community and dissemination work undertaken by the DM2E project is the European cultural heritage sector. As it was the project's goal to develop a workflow for cultural heritage institutions to integrate their metadata into Europeana and the Linked Data Web as well as providing applications for scholarly use of that material it was important to raise awareness about the tools being developed, so that feedback was provided to the project's developers and to guarantee uptake of the final toolset. As the focus of the project is on digitised manuscripts dissemination efforts were mainly concentrated on libraries and archives as opposed to galleries and museums (although these were not excluded).

The Open Knowledge Foundation was assigned within WP4 to disseminate the results of the project. OKFN organised a series of events and updated the community documentation on DM2E tools and workflows via the DM2E wiki. They built and strengthened the community around Europeana and facilitated conversations between technologies and researchers within the Europeana community. Furthermore the network of open metadata evangelists (via the OpenGLAM network) were supported and strengthened, who helped to raise awareness of legal and technical best practises in a variety of different domains.

Related document: [D4.4 - Dissemination and Engagement Plan](#)

The Network of Open Metadata Evangelists – “Open GLAM”

The community of open metadata evangelists which has been established by DM2E was specifically branded as the “OpenGLAM” community to allow for it to continue to exist beyond the lifetime of the project and to give it sufficiently broad appeal³⁶. The OpenGLAM community is established in such a way that it can be led by volunteers without the need for central funding. The DM2E project has invested significant resources in establishing the community structures that will allow the network of open data evangelists to expand beyond the lifetime of the DM2E project.

The Open Knowledge Foundation is committed to supporting this community and area of activity as part of its global network of thematic Working Groups. It will continue to host OpenGLAM.org and all other web based community infrastructure such as the OpenGLAM discussion and mailing lists.

Related document: [D4.4 - Dissemination and Engagement Plan](#)

Contest Award

One of the key components of the dissemination work provided by WP4 was to run a contest. Under the title ‘Open Humanities Awards’, two competition rounds took place between 2013-2014 with the goal of rewarding, encouraging and highlighting innovative work building on the technology developed as part of work packages 1, 2 and 3 as well as encouraging partnerships working between developers and non-technical researchers.

A dedicated website was set up: <http://openhumanitiesawards.org>.

³⁶ OpenGLAM: <http://openglam.org> [30.03.2015].

The first round of the Open Humanities Awards began in year two of the project and focused on supporting open source innovation based on open humanities data. This was in part a practical decision given that the DM2E tools were not yet in a state of readiness to run a compelling competition with. A total of over 50 applications were received and a prestigious judging panel from the Digital Humanities sphere selected two winning projects:

- Dr Bernhard Haslhofer (University of Vienna) for the project 'Maphub'
- Dr Robyn Adams (Centre for Editing Lives and Letters, University College London) for the project 'Joined Up Early Modern Diplomacy'

The next phase of the Open Humanities Awards was launched in year three. This second round consisted of two tracks: an Open track, following on the success of the first round, inviting all submissions using open data or open content, and a dedicated DM2E track, focused on projects building on the research being done in the DM2E project. In this round, a total of 21 applications was received (2 applications to the DM2E track, and 19 to the Open track). The following winners were selected by the high-profile judging panel of Digital Humanities specialists:

Open track:

- Dr. Rainer Simon (AIT Austrian Institute of Technology), Leif Isaksen & Pau de Soto Cañameres (University of Southampton) and Elton Barker (The Open University) for the project 'SEA CHANGE'
- Dr.-Ing. Michael Piotrowski (Leibniz Institute of European History (IEG)) for the project 'Early Modern European Peace Treaties Online'

DM2E track:

- Dr. Maximilian Hadersbeck (Center for Information and Language Processing (CIS), University of Munich (LMU)) for the project finderApp WITTFind.

Related document: [D4.6 – Contest Award Fund](#)

Publications

All publications are accessible online via <http://dm2e.eu/outputs/#Publications>.

Olenksy, M., Stiller, J., & Dröge, E. (2012). Poisonous India or the Importance of a Semantic and Multilingual Enrichment Strategy. In Juan Manuel Dodero, Manuel Palomo-Duarte, Pythagoras Karampiperis (Eds.), *Metadata and Semantics Research. Proceedings of the 6th Research Conference, MTSR 2012, Cádiz, Spain, November 28-30, 2012* (pp. 252-263). Springer: Berlin Heidelberg.

Kai Eckert (2013) Provenance and Annotations for Linked Data. *Proceedings of the International Conference on Dublin Core and Metadata Applications 2013*.

Evelyn Dröge, Julia Iwanowa, Violeta Trkulja, Steffen Hennische, Stefan Gradmann (2013) Wege zur Integration von Ontologien am Beispiel einer Spezifizierung des Europeana Data Model. *Proceedings des 13. Internationalen Symposiums für Informationswissenschaften (ISI 2013)*.

Falch, R.J., Erbacher, C. & Pichler, A. (2013). Some observations on developments towards the Semantic Web for Wittgenstein scholarship. In: *Mind, Language and Action*.

Contributions of the Austrian Ludwig Wittgenstein Society. Edited by Danièle Moyal-Sharrock, Annalisa Coliva and Volker Munz, pp. 119-121.

Stefan Gradmann, Julia Iwanowa, Evelyn Dröge, Steffen Hennicke, Violeta Trkulja, Marlies Olensky, Christian Stein, Alexander Struck, Konstantin Baierer (2013) Modellierung und Ontologien im Wissensmanagement. Information Wissenschaft & Praxis 2013.

Marco Grassi, Christian Morbidoni and Michele Nucci (2013) Pundit: augmenting web contents with semantics. Literary and Linguistic Computing, Vol. 28, No. 4, 2013. Oxford University Press.

Mácha, J., Falch R.J. & Pichler, A. (2013). Overlapping and Competing Ontologies. In: DH-CASE '13 Proceedings of the 1st International Workshop on Collaborative Annotations in Shared Environment: metadata, vocabularies and techniques in the Digital Humanities.

Pichler, A. & Zollner-Weber, A. (2013). Sharing and debating Wittgenstein by using an ontology. In: Lit Linguist Computing, 28(4): 700-707.

Schreibman, S., Gradmann S., Hennicke S., Blanke, T., Chambers, S., et. al. (2013). Beyond Infrastructure – Modelling Scholarly Research and Collaboration. In: Digital Humanities 2013: Conference Abstracts, University of Nebraska–Lincoln, USA, 16-19 July 2013. pp. 386-389.

Eckert, K., D. Ritze, K. Baierer, and C. Bizer. (2014). RESTful open workflows for data provenance and reuse. In: Proceedings of the companion publication of the 23rd international conference on World wide web companion (WWW Companion '14). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 259-260. DOI=10.1145/2567948.2577347

Hennicke, S., Dröge, E., Trkulja, V., & Iwanowa, J. (2014). From ESE to EDM and Beyond: How Europeana Provides Access to its Cultural Heritage Objects. In: M. Ockenfeld (Ed.): Informationsqualität und Wissensgenerierung. Proceedings der 3. DGI-Konferenz, 66. Jahrestagung der DGI, pp. 129-140. Frankfurt am Main: DGI, ISBN 978-3-925474-73-6.

Winer, D. (EAJC), Judaica Europeana: an Infrastructure for Aggregating Jewish Content (2014). In: Judaica Librarianship. Journal of the Association of Jewish Libraries (AJL), Vol. 18 ISSN: 2330-2976.

Dröge, E., Iwanowa, J. and Hennicke, S. (2014). A specialisation of the Europeana Data Model for the representation of manuscripts: The DM2E model, In: Libraries in the Digital Age (LIDA) Proceedings, Volume 13, 2014.

Baierer, K., Dröge, E., Petras, V. & Trkulja, V. (2014). Linked Data Mapping Cultures: An Evaluation of Metadata Usage and Distribution in a Linked Data Environment. In: Proceedings of the International Conference on Dublin Core and Metadata Applications DC-2014, 8-11 October 2014, Austin Texas, USA, pp. 1-11.

Dröge, E., Hennicke, S., Iwanowa, J., Olensky M., Rühle, S. & Trkulja, V. (2015). In: E. Euler, P. Klimpel (Eds.): Von ESE zu EDM und darüber hinaus: Wie Europeana Zugang zu Objekten des kulturellen Erbes ermöglicht. In: Der Vergangenheit eine Zukunft. Kulturelles Erbe in der digitalen Welt (pp.98-123). Berlin: iRights.Media.

Presentations

1st Year

Conference	Date	Speaker	Link to presentation
Medieval Cultures on the Web, Italy	7th-9th March 2012	Stefan Gradmann and Christian Morbidonni	NA
Data Modelling in the Humanities, Brown University, US	14th-16th March 2012	Stefan Gradmann	http://www.youtube.com/watch?v=gLf24W1eBy8&feature=youtu.be
Digital Humanities, Luxembourg	20th-21th March 2012	Stefan Gradmann	NA
WWW-Conference: Digital Humanities Panel, Semantics and the Web, Lyon, France	20th April 2012	Stefan Gradmann	NA
Conference of the Labor Manuscripts Section, France	30th May 2012	Stefan Gradmann	NA
Judaica Europeana Session, Israel	11th June 2012	Stefan Gradmann, Dov Winer and Ido Ivri	NA
Digital Humanities: Digital Diversity, Germany	16th July 2012	Marco Grassi	NA
COST Training School Meeting	28th September 2012	Stefan Gradmann and Christian Morbidonni	NA
Semantic Media Web Workshop, Germany	22nd October 2012	Kai Eckert	http://www.semantic-media-web.de/
Leipzig eHumanities Seminar, Germany	7th November 2012	Stefan Gradmann	http://www.slideshare.net/DM2E/beyond-infrastructure-stefan-gradmann-leipzig-digital-humanities-seminar-7th-november-2012
Minerva Israel	13th November 2012	Sam Leon and Christian Morbidoni	http://www.slideshare.net/DM2E/beyond-infrastructure-stefan-gradmann-leipzig-digital-humanities-seminar-7th-november-2012
Semantic Web in Libraries, Cologne, Germany	26th - 28th November 2012	Kai Eckert	http://swib.org/swib12/slides/Eckert_SWIB12_137.zip
Conference of the European Society for Textual Scholarship, Netherlands	24th November 2012	Simone Fonda	http://ests.huygens.knaw.nl
Everything is on the Move: The Mamluk Empire as a Node in (Trans-) Regional Networks, Germany	December 2012	Kilian Schmidtner	http://www.slideshare.net/DM2E/kilian-schmidtner-klaus-thoden-the-missing-links-defining-the-mamluk-empire-as-a-node-and-what-the-network-knows-about-it
Open Digital Humanities: use and reuse of digital data in the Humanities, Netherlands	17th January 2013	Joris Pekel	http://www.slideshare.net/DM2E/joris-pekkel-okfndm2e-and-the-digital-humanities

Conference	Date	Speaker	Link to presentation
Berlin Colloquium of Library Science, Germany	29th January 2013	S. Gradmann, S. Hennicke, J. Iwanowa, E. Dröge	http://www.slideshare.net/DM2E/edm-und-dm2e-anmerkungen-zum-heuristischen-potential-des-europeana-data-model

2nd Year

Conference	Date	Speaker	Link to presentation
5th BID Congress	11/03/2013	Kai Eckert, Steffen Hennicke, Julia Iwanowa	http://www.bid-kongress-leipzig.de/t3/index.php?id=26
Wittgenstein's Philosophical Investigations: Re-Evaluating a Project	21/03/2013	Alois Pichler	http://wittgensteinpi.squarespace.com/
13th International Symposium of Information Science	21/03/2013	Evelyn Dröge, Julia Iwanowa, Violeta Trkulja, Steffen Hennicke, Stefan Gradmann	http://de.slideshare.net/DM2E/towards-integrating-ontologies-an-edmbased-approach
KIM Workshop	26/03/2013	Evelyn Dröge	http://de.slideshare.net/evelyndr/spezialisierung-des-edm-durch-dm2e
Easy Tools for Difficult Texts	19/04/2013	Stefan Gradmann	http://easytools.huygens.knaw.nl/?page_id=18
Annual convention of the Austrian Library Network	15/05/2013	Doron Goldfarb	http://de.slideshare.net/DM2E/obv-linked-opendatafinal
Pundit Presentation at DARIAH-DE Expert Workshop	17/06/2013	Romeo Zitarosa, Simone Fonda	NA
CERN Workshop on Innovations in Scholarly Communication (OAI8)	19/06/2013	Stefan Gradmann, Alessio Piccioli	https://indico.cern.ch/event/211600/session/1/contribution/24
LODLAM Summit	19/06/2013	Simone Fonda	NA
APEX-Conference	26/06/2013	Steffen Hennicke	http://de.slideshare.net/DM2E/130627-apex
36th International Wittgenstein Symposium 2013	17/07/2013	Alois Pichler	http://www.alws.at/index.php/review/view/review_nachlese_zum_36_wittgenstein_symposium_2013/
Digital Humanities Conference 2013 in Nebraska	19/07/2013	Stefan Gradmann, Susan Schreibman	http://dh2013.unl.edu/schedule-and-events/program/
10th Congress of European Jewish Studies Association in Paris	20/07/2013	Christian Morbidonni	NA
16th World Congress of Jewish Studies in Jerusalem	28/07/2013	Dov Winer	http://www.jewish-studies.org/?cmd=world_congress
International Conference on Dublin Core and Metadata Applications	02/09/2013	Kai Eckert	http://dcevents.dublincore.org/IntConf/dc-2013/paper/view/154
DH Case: Collaborative Annotations in shared environments (Demo tools)	10/09/2013	Christian Morbidonni, Simone Fonda	http://www.cs.unibo.it/dh-case/pdf/Morbidoni.pdf
DH Case: Collaborative Annotations in shared	10/09/2013	Alois Pichler	http://www.cs.unibo.it/dh-case/pdf/Macha.pdf

Conference	Date	Speaker	Link to presentation
environments (Data models and digital editions)			
TPDL Conference	22/09/2013	Valentine Charles, Antoine Isaac, Vassilis Tzouvaras, Steffen Henniecke	http://tpdl2013.upatras.gr/tut-edm.php
Herbsttagung der Fachgruppe Dokumentation	15/10/2013	Evelyn Dröge	http://de.slideshare.net/DM2E/herbsttagung-der-fachgruppe-dokumentation-evelyn-drge
Meeting of National Europeana Partners	08/11/2013	Evelyn Dröge	http://de.slideshare.net/DM2E/national-aggr-droege
Semantic Web in Libraries (SWIB)	25/11/2013	Steffen Henniecke, Evelyn Dröge, Julia Iwanowa, Kai Eckert, Violeta Trkulja	http://swib.org/swib13/

3rd Year

Conference	Date	Speaker	URL
ExLibris Semantic Web Special Interest Group meeting	18/03/2014	Konstantin Baierer	http://de.slideshare.net/DM2E/dm2e-interopability-platform-omnompubby-konstantin-baierer
Deutsches Kulturerbe auf dem Weg in die Europeana II	20/03/2014	Vivien Petras	http://www.armubi.de/tagung2014/downloads/DM2E.pdf
WWW2014 - 23rd International World Wide Web conference	07/04/2014	Kai Eckert; Dominique Ritze; Konstantin Baierer; Christian Bizer	dl.acm.org/citation.cfm?doid=2567948.2577347
EOD conference 2014: Sustaining the networked future: use and reuse of digital content	11/04/2014	Lieke Ploeger	http://de.slideshare.net/DM2E/openglam-presentation-at-eod-conference-11-april-2014
KIM Workshop 2014	15/04/2014	Evelyn Dröge	http://de.slideshare.net/DM2E/das-dm2e-modell-evelyn-drge-kim-workshop-der-dini-kim-ag-mannheim-15042014
11th EDRENE: Educational Repositories Network Conference	06/05/2014	Dov Winer	http://de.slideshare.net/dovw/20140506-edrene-athenswiner
EVA Conference on Electronic Imaging and the Visual Arts	07/05/2014	Dov Winer	http://www.judaica-europeana.eu/docs/DM2E_Judaica_Winer-EVA_Florence_2014.pdf
3. DGI-Konferenz 2014	09/05/2014	Steffen Henniecke, Evelyn Dröge, Violeta Trkulja, Julia Iwanowa	http://de.slideshare.net/DM2E/von-ese-zu-edm-und-darber-hinaus-wie-europeana-zugang-zu-objekten-des-kulturellen-erbes-ermglicht
"Digital Wittgenstein Scholarship" Sommerschule	05/06/2014	Alois Pichler, Rune Falch, Øyvind Gjesdal, Heinz W. Krüger, Deirdre Smith	http://bit.ly/1hQKV06

Conference	Date	Speaker	URL
Libraries in the Digital Age (LIDA)	18/06/2014	Evelyn Dröge, Julia Iwanowa, Steffen Henniscke	http://bit.ly/1qTdwCs
10th Congress of the European Association for Jewish Studies	21/07/2014	Lena Stanley-Clamp	http://www.judaica-europeana.eu/Downloads/JE-DM2E_PptParis21July2014.pdf
IFLA 2014 Satellite Meeting: Linked Data in Libraries: Let's make it happen!	14/08/2014	Marko Knepper, Valentine Charles	http://dm2e.okblogfarm.org/files/MakinglibrarylinkeddatausingtheEuropeanaDataModel.pdf
"Paratext in Digital Culture: Is Paratext Becoming the Story?"	28/08/2014	Alois Pichler	http://elmcip.net/node/9646
International Conference on Dublin Core and Metadata Applications	09/10/2014	Konstantin Baierer, Evelyn Dröge	http://de.slideshare.net/DM2E/dc-2014-baiererdroege
EVA/Minerva Jerusalem International Conference on Digitisation of the Culture	10/11/2014	Kristin Dill	http://dev.thepund.it/tutorials/evaminerva2014/ http://de.slideshare.net/evaminerva/1-kristin-dillpunditdm2e

Events

- [Open Data in Cultural Heritage workshop](#), 20 April 2012, Berlin, Germany
- [OpenGLAM legal workshop – part 1](#) and [part 2](#), 27 April and 29 May 2012, Paris, France
- [Open Data in Cultural Heritage workshop](#), 13 June 2012, London, UK
- [Open Culture Hackday](#) (OKFest), 18 September 2012, Helsinki, Finland
- [Open Data in Cultural Heritage workshop](#) (OKFest), 22 September 2012, Helsinki, Finland
- [Open Humanities Hack](#), 21-22 November 2012, London, UK
- [Pundit hackday](#), 3-4 May 2013, Pisa, Italy
- [Pundit workshop](#), 23 May 2013, The Hague, The Netherlands
- [The Web as Literature](#), 10 June 2013, London, UK
- [Open Data in Cultural Heritage workshop](#) (OKCon), 16 September 2013, Geneva, Switzerland
- [Wittgenstein Incubator Pundit training](#), 6 December 2013, Bergen, Norway
- [Pundit UI/UX event](#), 2 April 2014, Berlin, Germany
- [Open Data in Cultural Heritage workshop](#), 15 July 2014, Berlin, Germany
- [Pundit community session](#) at the DARIAH-EU 4th General VCC meeting, 18 September 2014, Rome, Italy
- [Putting Linked Library Data to Work: the DM2E Showcase](#), 18 November, Vienna, Austria
- [Open Humanities Hack](#), 28 November, London, UK
- Workshop "[RDF Application Profiles in Cultural Heritage](#)" during [Semantic Web in Libraries \(SWIB\) conference](#), 1-3 December, Bonn, Germany

- [Enabling humanities research in the Linked Open Web – DM2E final event](#), 11 December, Navacchio, Italy

Related document: [D4.2 - Workshop Programme](#)

Deliverables

- [Deliverable 1.1 – WP1 – Requirements Report](#)
- [Deliverable 1.2 – WP1 – Final Integration Report](#)
- [Deliverable 2.1 – WP2 – Initial Version of the Interoperability Infrastructure](#)
- [Deliverable 2.2 – WP2 – Intermediate Version of the Interoperability Infrastructure](#)
- [Deliverable 2.3 – WP2 – Final Version of the Interoperability Infrastructure](#)
- [Deliverable 3.1 – WP3 – Initial Specification Report](#)
- [Deliverable 3.2 – WP3 – Prototyping Platform Implemented](#)
- [Deliverable 3.3 – WP3 – E-learning Courses published](#)
- [Deliverable 3.4 – WP3 – Research Report on DH Scholarly Primitives](#)
- [Deliverable 4.2 – WP4 – Workshop Programme](#)
- [Deliverable 4.3 – WP4 – Documentation Package](#)
- [Deliverable 4.4 – WP4 – Dissemination and Engagement Plan](#)
- [Deliverable 4.5 – WP4 – Digital Humanities Advisory Board Convention](#)
- [Deliverable 4.6 – WP4 – Contest Award Fund](#)
- [Deliverable 5.10 – WP5 – Results Transfer Plan](#)

6 Sustainability

The original plan for the sustainability of project results for DM2E was (from the DoW, Task 5.6):

“In order to guarantee reuse of the project results and thus their sustainability, HUB, the Europeana Foundation and TEL will work together on a sustainability plan. The overall objective is not so much to create a new organization at the end of the project but to make sure that data, technical building blocks and platform infrastructure created in DM2E will be transferred to Europeana/The European Library to be hosted and reused. Input from the leaders of WPs 2, 3 and 4 is required for this plan. The outcome of this task is twofold: a results transfer plan will need to be created roughly 6 months before the end of the project and the actual transfer of results needs to take place at the end of the project.”

In an early stage it became clear that Europeana will not be able to follow the plan as laid out in Description of Work. There are no technical or other barriers from the DM2E side to transfer data and infrastructure to Europeana, most of which resides on the DM2E server, it is rather a problem of Europeana not being equipped to process the project outcomes.

The D5.10 “Results Transfer Plan” presents the solutions and proposals for the sustainability of the DM2E project results, worked out together with all work package leads and the Europeana Foundation. The Results Transfer Plan is structured in that way, that minimal requirements are discussed in order to ensure long-term preservation but also maximum options are depicted, to propose solutions for the future. The Results Transfer Plan

- introduces the achievements of the created Linked Data infrastructure within the DM2E project
- describes the conducted contacts of the project to potential cooperation partners from the Europeana cluster and representatives from other institutions in order to sustain the project results
- describes the archiving, access and integration of DM2E object metadata with Europeana
- explains where and how the developed infrastructure will be sustained in order to be available in the future
- points out where the DM2E model, a specialisation of the Europeana Data Model (EDM) will be accessible and explains the initiated task group on RDF Application Profiles at the Dublin Core Metadata
- describes where deliverables, the DM2E Wiki, software, data model and learning videos will be documented
- and explains the established network around the DM2E.

Regarding the sustainability of the DM2E data, WP2 prepared a proposal for representing DM2E content in the highest possible granularity in EDM and in Europeana. This innovative proposal (included as appendix A in D5.10) allows the scholarly research platform “Pundit” to be linked from Europeana.

Related document: [D5.10 Results Transfer Plan](#)

The DM2E Wiki

The DM2E Wiki on <http://wiki.dm2e.eu> hosts all core English documentation in order to enable community input and annotation. It also links to all non-English DM2E documentation

hosted on the DM2E Bitbucket. The Wiki has been designed by the OKFN so as to create separate areas for different kinds of DM2E documentation such as OmNom Documentation and Pundit Documentation. This makes the Wiki easier to use and maintain and allows such a large resource to remain approachable for a broad community of users.

WP2 and WP3 (UBER, Net7, UMA) supplied the underlying technical documentation which was used by OKFN to create the DM2E Wiki which hosts a full set of technical and user resources supplemented by a series of screencasts and introductory videos on DM2E tool functionality and the value of the project for digital scholarship. Other documentation was added such as the “Introduction to Open Cultural Data”, Pundit user tutorials and a section of the DM2E Wiki focussed on the metadata formats and mappings to the EDM used by project beneficiaries. Translations of the documentation were completed by OKFN, Net7, UBER, UIB, and MPIWG.

Related document: [D4.3 - Documentation Package](#)

Outcomes

All outcomes of the DM2E project are available on the DM2E website <http://dm2e.eu>. The Europeana Pro website hosts the DM2E deliverables and documents on the DM2E data model <http://pro.europeana.eu/get-involved/projects/project-list/dm2e>. The ontology and all source code developed in DM2E is available at <https://github.com/DM2E>.

The Digitised Manuscript to Europa Project was coordinated by

Prof. Vivien Petras, PhD
Berlin School of Library and Information Science
Humboldt-Universität zu Berlin
Unter den Linden 6, 10099 Berlin, Germany
Tel: +49 30 2093-4325
Fax: +49 30 2093-4335
E-mail: vivien.petras@ibi.hu-berlin.de

Grant Agreement No: ICT-PSP-297274

Project acronym: DM2E

Project website: <http://dm2e.eu>