

ARCHIVES
PORTALEUROPE
Foundation



Stichting Archives Portal Europe Foundation



Cracking digital archival research and metadata: Archives Portal Europe

DARIAH Virtual Annual Event, 18 November 2020

Speakers:

- DARIAH-EU Working Group “Sustainable publishing of (meta)data” – Johan Van der Eycken, johan.vandereycken@arch.be
- Archives Portal Europe – Kerstin Arnold, kerstin.arnold@archivesportaleurope.net
- PARTHENOS project – Klaus Illmayer, klaus.illmayer@oeaw.ac.at
- SSHOC project – Frank Fischer, frank.fischer@dariah.eu; Laure Barbot, laure.barbot@dariah.eu

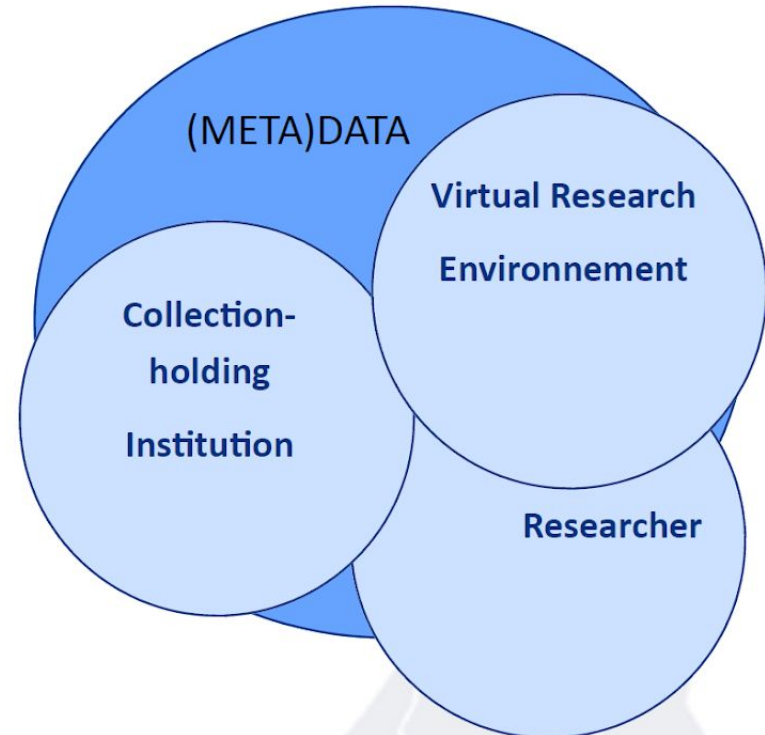
Working Group Sustainable publishing of (meta)data

Johan Van der Eycken
Brussels, 18 November 2020

Opportunities

- New technologies offer for researchers
- There has never been so much data and metadata available
- Need to maximise the use of these new opportunities

to bring together international expertise from different international networks



Focus on cooperation

- **Knowledge exchange**
 - We all face the same difficulties seen from our own research background and perspective.
 - We are all working on thinking about solutions for these same problems
- **Cooperation versus competition**
 - Resources and staff are a scarce commodity and are distributed between different RI and instances = competition
 - Awareness of the usefulness of collaboration and of sharing knowledge and know-how, without losing sight of the individual interest or the interest of one's own institution = collaboration becomes political

Field of collaboration

Metadata, a path to standardisation.

Specific challenges regarding sustainable archiving and availability of digital information, such as the use and implementation of international standards.

Metadata and Society.

Focussed on new opportunities and challenges arising in a digital society, e.g., the implementation of the General Data protection Regulation (GDPR) for cultural heritage (EU) and eIDAS.

Metadata communication and interoperability.

The Communication between archival institutions and researchers, both at the level of technological components and at the level of researchers' expectations and options

Activities

- Bring together and communicate about, knowledge, hands-on expertise and tools concerning sustainable digital publishing of metadata (e.g. 2016 - <https://hal.inria.fr/hal-01281442>)
- Offer tools for self analysis (e.g. 2015-2017 - Maddlain - <https://www.kbr.be/en/projects/maddlain>)
- Integrate experiences from research infrastructures ([EHRI](#), Europeana, Archives Portal Europe, [CESSDA](#) and others) and collection holding institutions (e.g. 2017-2022 - [Social Sciences and Humanities Data Archive](#))
- Keep the dialogue between archival institutions and research infrastructures alive (e.g. 2017-2018 - [Trust and Understanding, The value of metadata in a digitally joined-up world](#))
- Explore options to overcome challenges in reaching the goal of producing standards-compliant, publishable data and metadata (e.g. [PROMISE \(PReserving Online Multiple Information: towards a Belgian StratEgy\)](#))

WORKING GROUP:

<https://www.dariah.eu/activities/working-groups/sustainable-publishing-of-metadata/>

CONTACT:

JOHAN VAN DER EYCKEN - JOHAN.VANDEREYCKEN@ARCH.BE

ERIC DE RUYTER - EDR@IISG.NL

DORIEN STYVEN - DORIEN.STYVEN@KAZERNEDOSSIN.BE

TOM GHELDOF - TOM.GHELDOF@KULEUVEN.BE



**ARCHIVES
PORTALEUROPE
Foundation**

Stichting Archives Portal Europe Foundation

Kerstin Arnold,
Technical Coordinator

Twitter: @kerstarno

Background and general introduction

Network

34

Countries

incl. overseas territories

24+

Languages

all official EU languages, plus Georgian, Hebrew, Icelandic

5

Alphabets

ქართული, Ελληνικά, Српски

Types of institutions

National archives
City archives
Private archives
Museum archives
Church archives
University archives
Business archives
etc....

7,000+

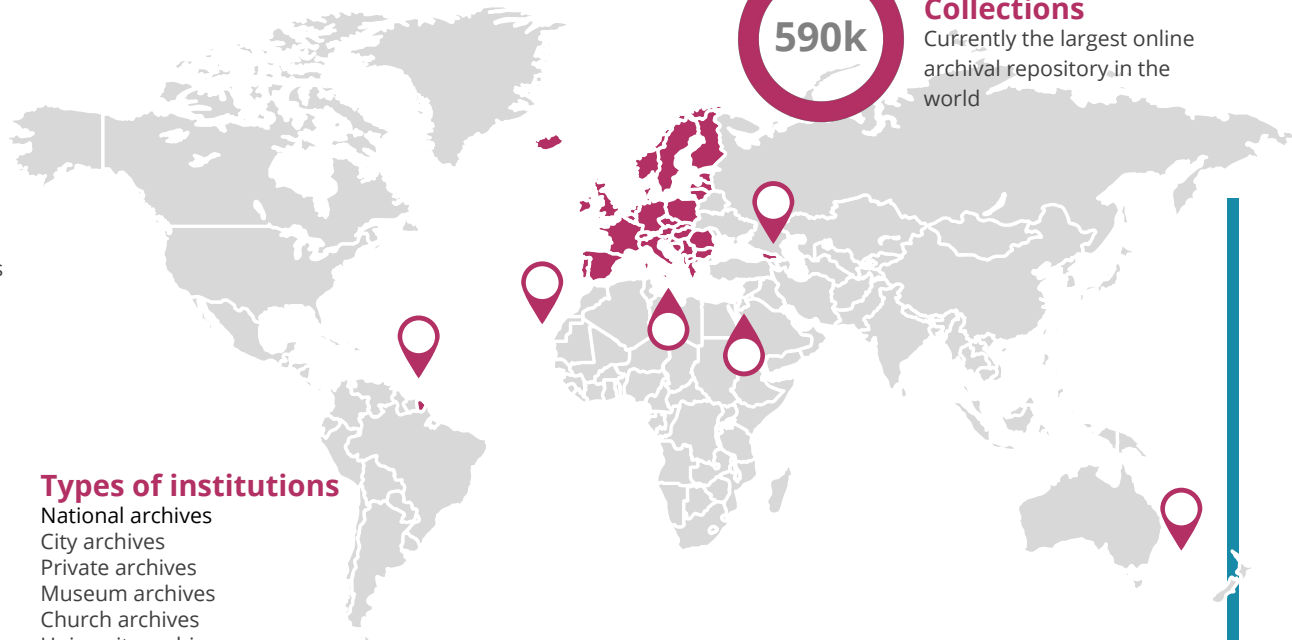
Institutions

1,100 actively contribute with content

590k

Collections

Currently the largest online archival repository in the world

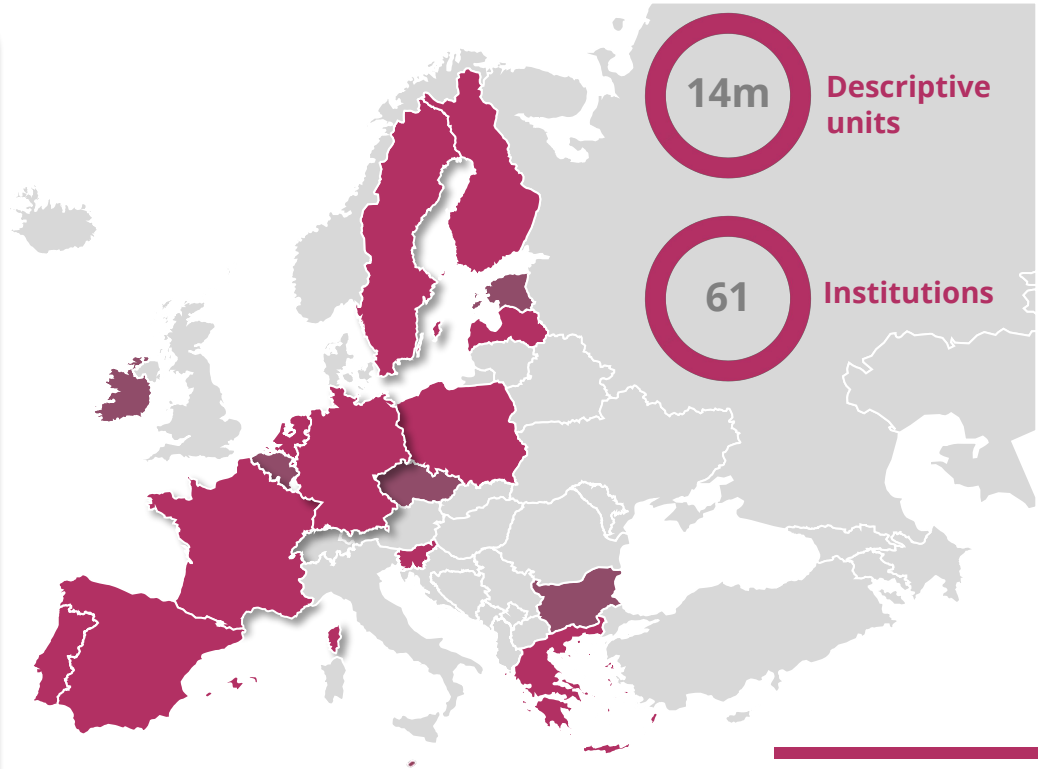


Archives Portal Europe network (APEnet, 2009-2012)



APEnet

- Display information about archival collections, their creators, the institutions;
- Searchable through a single engine;
- Link to the individual digital objects where applicable
- Aggregator for Europeana

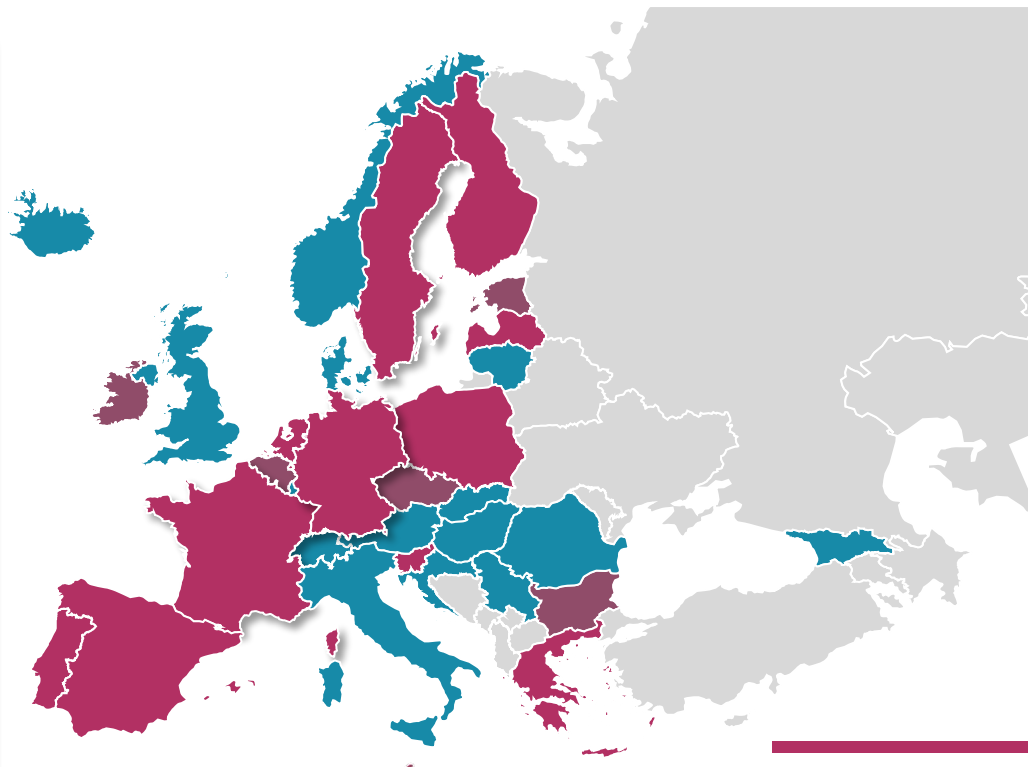


Archives Portal Europe network of eXcellence (2012-2015)

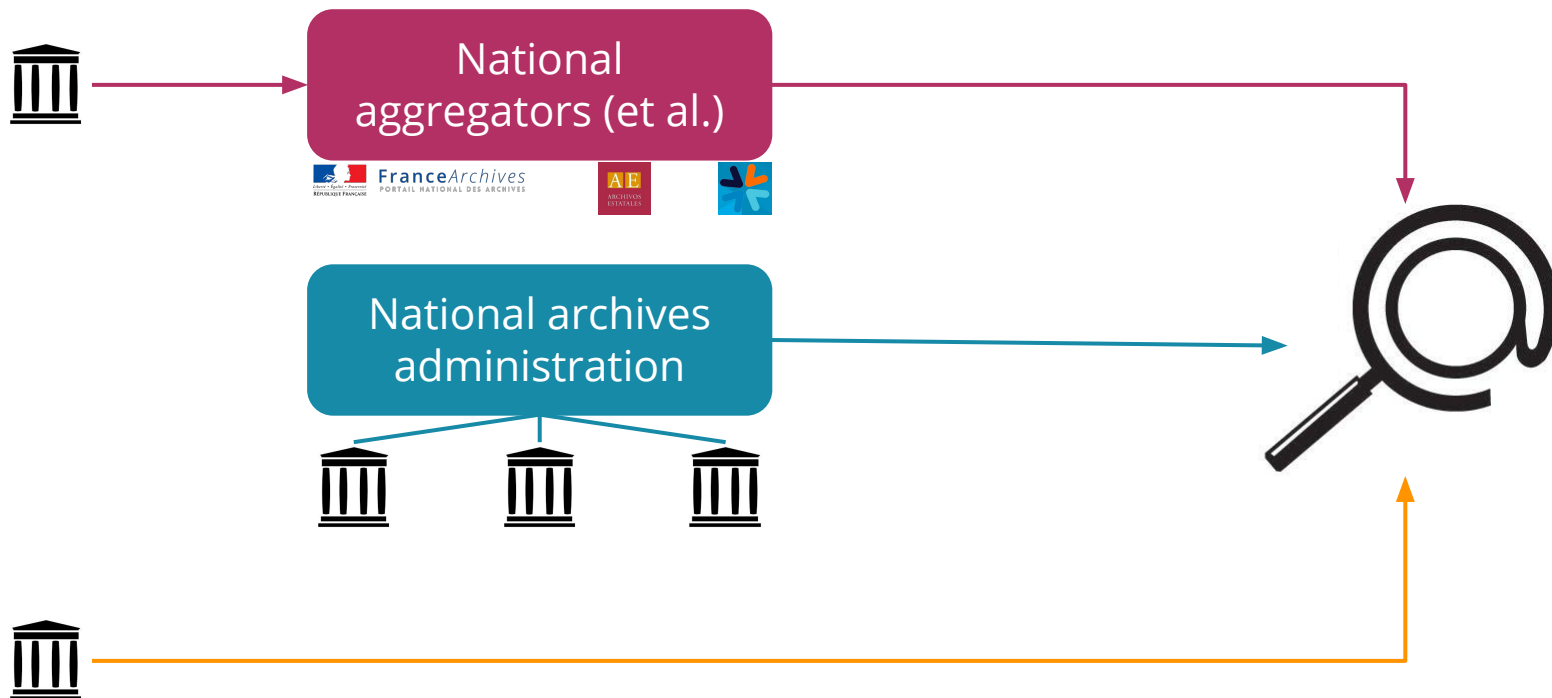


APEx

- Strengthen the portal
- Open up to other countries
- Open up to other archives



Ways to deliver content to the portal



Information about an institution (old design)

- Directory
 - ▶ Albania
 - ▶ Austria
 - ▶ Belgium
 - ▶ Bulgaria
 - ▶ Croatia
 - Hrvatski državni arhiv**
 - Državni arhiv u Bjelovaru
 - Državni arhiv u Dubrovniku
 - Državni arhiv u Gospiću
 - Državni arhiv u Karlovcu
 - Državni arhiv za Međimurje
 - Državni arhiv u Osijeku
 - Državni arhiv u Pazinu
 - Državni arhiv u Rijeci
 - Državni arhiv u Sisku
 - Državni arhiv u Slavonskom Brodu
 - Državni arhiv u Splitu
 - Državni arhiv u Šibeniku
 - Državni arhiv u Varaždinu
 - Državni arhiv u Virovitici
 - Državni arhiv u Vukovaru
 - Državni arhiv u Vukovaru - Odjel Vinkovci
 - Državni arhiv u Zadru
 - Državni arhiv u Zagrebu
 - ▶ Czech Republic
 - ▶ Denmark
 - ▶ Estonia
 - ▶ Finland
 - ▶ France
 - ▶ Georgia
 - ▶ Germany

Hrvatski državni arhiv

Print 

CONTACT DETAILS (SEE MORE)

Visitor's address:	Marulićev trg 21, 10000 Zagreb
Country:	Hrvatska
E-mail address:	Send an e-mail
Webpage:	Go to the website
Telephone:	+385 1 4801999

ACCESS & SERVICE INFORMATION (SEE MORE)

Opening hours:	Mon 8:15 - 15:45 Tue 8:15 - 15:45 Wen 8:15 - 15:45 Thu 8:15 - 17:45 Fri 8:15 - 13:45
Access conditions:	Accessible to the public
Searchroom:	160 places

ARCHIVES & HOLDINGS DESCRIPTION (SEE MORE)

Extent of holdings:	28700 m
---------------------	---------

OTHER INFORMATION

Type of archive:	National archives
------------------	-------------------

ARCHIVAL MATERIAL

[List of Finding aids](#)

2015

Search results (old design)

Select types/dates ⓘ

Select the element in which you want to search:

Select the type of document:

Dates of creation (yyyy/mm/dd), from: to:

Exact date search

Search ⓘ

View result(s) in context

Match at least one word

Only material with digital objects

[Clear search](#)

Search results:

List view **Context view** ⓘ

[Save search](#)

Search results: 1 - 10 / 222 Results per page: 10 ▼

<< < 1 2 3 4 5 6 7 > >>

Selected criteria:

COUNTRY:

- Slovenia (136)
- Germany (33)
- Netherlands (21)
- Switzerland (19)
- United Kingdom (6)
- Austria (3)
- France (3)
- Croatia (1)

ARCHIVAL INSTITUTION:

DIGITAL OBJECT TYPE:

- Unspecified (173)
- Image (30)
- Text (19)

DATE TYPES:

- Full date (166)
- Only descriptive date (40)
- No date specified (16)

START TIMESPAN:

END TIMESPAN:

Sorting your search:

Relevance | Date | Title | Reference code | Finding aid no

Nasa Snaga (Zagreb) (Mirko Kus-Nikolajev). 1924.

1 Brief Nasa Snaga (Zagreb) Mirko Kus-Nikolajev

Contains digital objects

Title: Karl Kautsky Papers
Reference code: <http://hdl.handle.net/10622/ARCH00712 - 637>
Netherlands - Internationaal Instituut voor Sociale Geschiedenis

[View only results from this finding aid](#)

Nasa Snaga (Zagreb) (Mirko Kus-Nikolajev). 1924.

1 Brief Nasa Snaga (Zagreb) Mirko Kus-Nikolajev

Contains digital objects

Title: Karl Kautsky Papers
Reference code: <https://hdl.handle.net/10622/ARCH00712 - 637>
Netherlands - Internationaal Instituut voor Sociale Geschiedenis

SI AS 1899 Slovenski izseljenski odbor Zagreb

Slovenski izseljenski odbor Zagreb (1941-1945) Link to original

Contains digital objects

Title: SI AS 1899 Slovenski izseljenski odbor Zagreb
Reference code: SI AS 1899

<http://hdl.handle.net/10622/ARCH00712 - 637>

Nasa Snaga (Zagreb) (Mirko Kus-Nikolajev). 1924.


1924.

Extent

1 Brief

Keywords

Personal names: Mirko Kus-Nikolajev
Corporate names: Nasa Snaga (Zagreb)



[View finding aid](#)

Collection details (old design)



Sign In

English

HOME SEARCH DIRECTORY FEATURED DOCUMENTS API

Search in all content Search

Netherlands > [Internationaal Instituut voor Sociale Geschiedenis](#)

- Karl Kautsky Papers
 - More before...
 - KAUTSKY-FAMILIEARCHIV
 - KORRESPONDENZ (Nr. 1-1811)
 - Briefe von/an KARL KAUTSKY (Nr. 1-809)
 - More before...
 - Briefe an Karl Kautsky (Nr. 70-809)
 - More before...
 - Nasa Snaga (Zagreb) (Mirko Kus-Nikolajev). 1924.
 - More after...
 - More after...
 - More after...
 - More after...
 - More after...



Bookmark this



<http://hdl.handle.net/10622/ARCH00712 - 637>

Nasa Snaga (Zagreb) (Mirko Kus-Nikolajev). 1924.

1924.

Extent

1 Brief

Digital objects



Digital object on provider's website

The screenshot displays a digital object viewer interface. At the top, there is a navigation bar with a back button, a page indicator 'Image 2 of 2', a 'Go' button, and a forward button. On the right side of the top bar are icons for a grid and settings. The main content area is titled 'Nasa Snaga (Zagreb) (Mirko Kus-Nikolajev). 1924.' and features a large central image of a handwritten document. The document is dated 'Juli 1924' and is addressed to 'Herrn Karl Kautsky, Berlin'. The text is in German and discusses political matters. To the left of the main image is a 'CONTENTS' sidebar with two thumbnail images. Below the main image is a dark box with the text 'Collection International Institute of Social History (Amsterdam)'. On the right side, a 'MORE INFORMATION' sidebar contains the following details:

- About the Item**
- Dates: 1924.
- Physical description: 1 Brief
- Part of: Karl Kautsky Papers
- Inventory number: 637
- Refer to this record: <https://hdl.handle.net/10622/ARCH00712.637>
- Refer to this Item: <https://hdl.handle.net/10622/ARCH00712.637>
- Use restrictions for this item: Please consult the IISH copyright statement before using this item

At the bottom left of the viewer are download and share icons, and at the bottom right are full-screen and zoom icons.

Research and metadata standards

A multilingual and multicultural environment

34

Countries

incl. overseas territories

24+

Languages

all official EU languages, plus
Georgian, Hebrew, Icelandic

5

Alphabets

ქართული, Ελληνικά, Српски

Full text search (new design)

The screenshot displays a search interface with the following elements:

- Three filter buttons: "Search in archives (0)", "Search in names", and "Search in Institutions".
- A search input field with a dropdown menu set to "Search in all countries".
- The search query: "Nap?!* Ναπολέων ნαპολεον Наполеон".
- A search button (magnifying glass icon).
- Advanced search options: "Show advanced search options" (dropdown), "Show only materials with digital objects" (checkbox), and "Search each term separately" (checkbox).

- Combining search terms in different languages and scripts
- Using wildcards for more inclusive coverage

Three types of searches (new design)



English ▾

→ Create account

👤 Login

Advanced search

Find an institution

Explore

About us

Tools

Blog

Help

🏠 > Search in archives

Advanced search

Search in archives

When looking for archival materials such as manuscripts, administrative records, maps etc.



📄 Search in archives

👤 Search in names

🏢 Search in institutions

Search in all countries ▾

Type search term...



Three main metadata standards

Encoded Archival Description

- XML format
- Archival material
- apeEAD profile
- Linked to apeMETS for digital objects

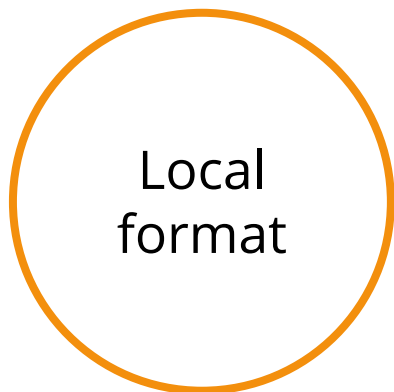
Encoded Archival Context

- XML format
- Records creators
- apeEAC-CPF profile
- Potential extension to other entities

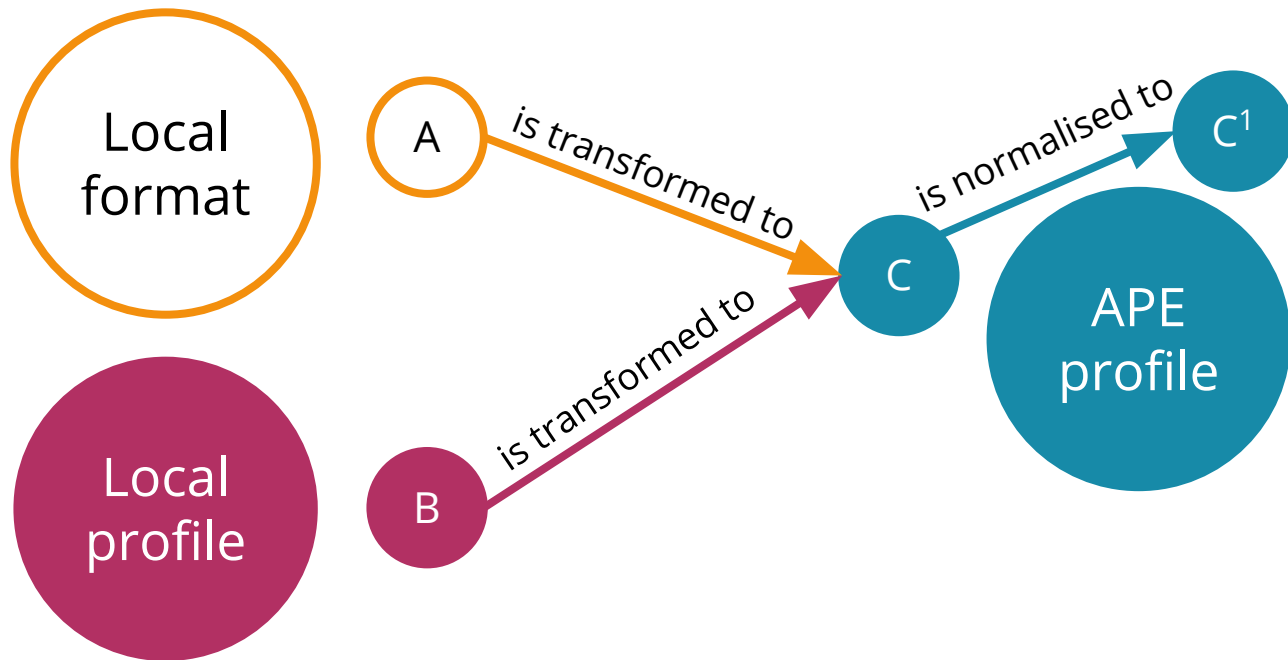
Encoded Archival Guide

- XML format
- Institutions with archival holdings
- EAG2012 standard

Different ways to say the same thing



Transformation and normalisation



Advanced search options (new design)

Search in all countries ▼ Type search term...

Hide advanced search options ^ Show only materials with digital objects ? Search each term separately ?

These advanced options will focus on results from "Search in archives".

? Search using ▼ ? Document type ▼ ? Date search to Exact date search

View in context ?

Filters for search results (new design)

Search filters:

[?](#) Countries ▼

[?](#) Archival institutions ▼

[?](#) Finding aids ▼

[?](#) Document type ▲

- Finding aid
- Holdings guide
- Source guide

[?](#) Levels ▲

- Other descriptions
- Fonds description

[?](#) Date type ▼

[?](#) Start Timespan ▲

Specify exact year

YYYY 📅

Or a time period

Select Century ▼

Select Decade

Select Year

[?](#) End Timespan ▼

[?](#) Search using ▼

Based on...

- Administrative aspects
 - Country and institution
- Type and structure of archival collections
 - Finding aid, holdings guide, source guide
 - Level of description
- Specific parts of the archival description
 - Title, identifier, scope and content note
 - Dates
 - Digital objects

Managed in the portal's back-end

Country level

[Home](#)

Country Manager for Croatia

Manage archival landscape

Manage content

- Manage institution managers
 - Download institution statistics
-

Institution level

[Home](#) - *Hrvatski državni arhiv*

Content provider information

Content information

Preferences for uploaded files

Create automatic harvesting function

Upload content

Create EAC-CPF

Content manager

Managed in the metadata as provided

```
<archdesc level="fonds" type="inventory">
  <did>
    <unitid type="call number">ABC</unitid>
    <unittitle>Collection of photographs</unittitle>
  </did>
  <dsc>
    <C>
      <did>
        <unitid type="call number">1</unitid>
        <unittitle>Photograph of the Croation State Archives in Zagreb</unittitle>
        <unitdate>18 November 1990</unitdate>
        <dao xmlns:xlink="http://www.w3.org/1999/xlink" xlink:href="http://www.arhiv.hr/portals/0/Kvadratici/Arhiv.jpg"
          xlink:title="Outside of the Croatian State Archives in Zagreb"/>
      </did>
      <scopecontent>
        <p>The picture shows the building of the Croation State Archives in Zagreb.</p>
      </scopecontent>
    </C>
  </dsc>
</archdesc>
```

Normalised during conversion

<unitdate>18 November 1990</unitdate>

becomes

<unitdate normal="1990-11-18">
18 November 1990</unitdate>

Additions based on provider's input

Default type for <dao> items	<input checked="" type="checkbox"/> Take from file (<dao@xlink:role>) if existing
Default rights statement for digital objects	<input type="checkbox"/> TEXT <input checked="" type="checkbox"/> IMAGE <input type="checkbox"/> SOUND <input type="checkbox"/> VIDEO <input type="checkbox"/> 3D <input type="checkbox"/> Unspecified
Description	<input type="text" value="Creative Commons CC0 Public Domain Dedication"/>
Rights holder	<input type="text"/>
Default rights statement for EAD data	<input type="text" value="Creative Commons CC0 Public Domain Dedication"/>
Description	<input type="text"/>
Rights holder	<input type="text"/>

Searching by topics (new design)

Highlights

Topics

How about looking at our content from a different angle? We have selected numerous topics that allow you to dive into the archives material on a thematic or genre-related approach.



Social history

Documents that relate to the everyday life of the population, living and working conditions, welfare; collections related to historians of social history

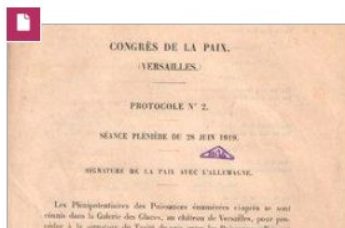
100,000+ results



Industry

Documents related to companies, industrial production, the second sector at large

100,000+ results



Genealogy

Documents related to the study of families, family histories, and lineage

100,000+ results



Source guides and subject headings

Source guides

- Special document type
- Combining collections from one or several institutions relating to the same topic

Subject headings

- Part of the archival descriptions
- To certain extent based on national vocabularies
- Potential for Linked Data

Topic management by the providers

Select topic*:

Colonialism ▼

Select source guide:

--- Please select one --- ▼

Specify a keyword for <controlaccess><subject>:
(Split multiple keywords by | character.)

Topic extensions by registered users (new design)

 **Bookmark**

 **Make a suggestion**


 **Contact institution**

 **Share**

 **Download / print**

Make a suggestion

Make a suggestion to further the use of this page for yourself and - potentially - other users in future. You can assign this page to one of our topics, suggest a translation, connect this page to an external resource, e.g. a Wikipedia page, or add to it in other ways, e.g. by uploading a related image. We aim at reviewing your suggestion as soon as possible and will get back to you accordingly.

 **What would you like to do?**

Assign to topic

Suggest translation

Connect to another resource

Other

Your full name*

Full name

Your email address*

Email address

Please provide details on your suggestion*

Normal  **B** *I* U    

Outlook: users forming a network

Welcome Archives!

We need a couple of things from you before we can proceed...

01. APE T&C acceptance

02. Settings & preferences

You can change this at any time in your account area

What would you like to be kept up to date on?

Newsletter

We will email you once a month, keeping you up to date on our latest blog articles, events, conferences, grants available. You can opt out at any time.

Researcher materials

We will email you with suggestions of new topics or functionalities related to your research, networking opportunities and current events.

Updates to my saved content

We will email you when a change has been made to your bookmarks or new content has been added matching your saved searches.

Outlook: users forming a network

The screenshot shows the Archives Portal Europe user interface. At the top, there is a navigation bar with the logo on the left, a language dropdown set to 'English', and a 'My account' dropdown. Below this is a dark red navigation menu with links for 'Advanced search', 'Find an institution', 'Explore', 'About us', 'Tools', 'Blog', and 'Help'. The main content area has a breadcrumb trail: 'Home > Account Dashboard > Saved searches, bookmarks and collections'. The title is 'Saved searches, bookmarks and collections'. Below the title is a paragraph of placeholder text: 'Nunc ex libero, suscipit at lectus et, congue tincidunt diam. Integer sodales quam in mollis interdum. Ut quis luctus magna. Cras semper nisi eget lorem tristique suscipit. Nullam vel bibendum sem.' There are three tabs: 'Saved searches', 'Bookmarks', and 'My collections', with 'My collections' being the active tab. Below the tabs is a search bar with the placeholder 'Find collections...', a '+ Create collection' button, and a 'Sort by: Collection name' dropdown. A table displays a single collection entry with columns for ID, Description, Searches, Bookmarks, and Saved on. The table row shows: (3004) Collection name displays here, Aliquam vulputate justo ipsum, eu pretium ipsum mollis eget. Aliquam diam erat, ornare quis magna eget, convallis dictum augue., 3, 2, and DD-MM-YYYY. At the bottom of the table row are links for 'Collection contents', 'Copy', 'Edit collection', and 'Delete collection'.

ARCHIVES PORTAL EUROPE

English My account

Advanced search Find an institution Explore About us Tools Blog Help

Account Dashboard > Saved searches, bookmarks and collections

Saved searches, bookmarks and collections

Nunc ex libero, suscipit at lectus et, congue tincidunt diam. Integer sodales quam in mollis interdum. Ut quis luctus magna. Cras semper nisi eget lorem tristique suscipit. Nullam vel bibendum sem.

Saved searches Bookmarks My collections

Find collections... + Create collection Sort by: Collection name

(ID) Collection name	Description	Searches	Bookmarks	Saved on
(3004) Collection name displays here	Aliquam vulputate justo ipsum, eu pretium ipsum mollis eget. Aliquam diam erat, ornare quis magna eget, convallis dictum augue.	3	2	DD-MM-YYYY

Collection contents Copy Edit collection Delete collection

PARTHENOS and the requirement of structured archive data of high quality

Klaus Illmayer (Austrian Centre for Digital Humanities and Cultural
Heritage at the Austrian Academy of Sciences - ACDH-CH/OEAW)

klaus.illmayer@oeaw.ac.at

<https://orcid.org/0000-0001-7253-996X>

Topics

- PARTHENOS
 - Short description
 - Implementation of archive material
 - Experiences
- Archives
 - Issues regarding structured data as expected by harvesters like PARTHENOS
 - Description vs. data
 - Records in Contexts (RiC) as game changer?
- Cracking the archive
 - Experiences with RiC
 - Entity driven recording
 - Archives Portal Europe as harvester compared to PARTHENOS and SSHOC

PARTHENOS: introduction



PARTHENOS

- Was a Horizon 2020 funded project (2015-2019):
<https://www.parthenos-project.eu/>
- Is an **acronym** that stands for “Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies”
- Strong **focus** on a technical solution for **harvesting data** from **many sources** to gain **cross-disciplinary insights**: Virtual Research Environment ([VRE](#))
- Other outcomes: [training suite](#), [standardization survival kit](#), [hub](#), [guidelines and policies](#), [deliverables](#)
- Important **goal**: apply a **linked data model** to the **harvesting mechanism** and **prepare a VRE** to allow researchers **querying** for **complex research questions** (training & awareness for standards, policies, technical framework)

PARTHENOS: aim



PARTHENOS

- **Semantic rich data** from many sources can help in **discovering new research insights, cross-disciplinary exchange** and **reviewing** of thesis/claims/opinions
- PARTHENOS **collects data from humanities and social sciences disciplines** e.g. linguistic studies, history studies, archaeology, cultural heritage
- **Relies on already existing data collections** by providers like CLARIN, DARIAH, ARIADNE
- VRE runs natively on a **triple store** and data is modelled in **CIDOC CRM**
- Development of a **tool** that supports **mapping** from any source

PARTHENOS: archive material



PARTHENOS

- Two main **modes of interest**:
 - **Combining data** from different spheres e.g. archaeology, literature and theatre studies - when thinking about how a theatre play from ancient Greece was performed (= theatre studies) it is helpful to have resources that gives you background information on the architecture of theatre buildings (= archaeology/heritage science) and interpretations of the play (= literature studies)
 - **Discovering new connections** e.g. history studies and meteorology/climate studies - transcription of a notebook of a monk in the medieval times (= history studies) may include observations on climate events (= meteorology/climate studies)
- Both modes differ in **granularity of details** that is necessary: **first may work with simple** keywords and a string search in title whereas the **second** needs **background details** on the **content** of an archive holding
- PARTHENOS VRE as a **framework to provide both modes to researchers**

PARTHENOS: discovery



PARTHENOS

Discovery Service NLP Hub Text Analytics Services Visual Media Service Documentation Members

Parthenos Discovery

SPARQL Quick Links - Login ?



<https://parthenos.acdh.oeaw.ac.at/Start>

PARTHENOS Discovery

Your gateway to the PARTHENOS Entities Dataspace. Metadata from all partner Research Infrastructures is aggregated here under a common semantic model, allowing for unprecedented interdisciplinary view on these heterogeneous datasets.



Explore/Search

PARTHENOS Discovery application offers multiple options to explore the dataset. You can search by keyword, by semantic relations, by sources or via the underlying model.

by keyword

Keyword search with faceting over selected properties

by semantic relations

Advanced semantic search using query builder

by sources

Dataset divided by the contribution from individual sources

by model

Explore the data based on the underlying ontology the [PARTHENOS Entities Model](#)

Additionally you can dive into the dataset through one of the main types: actors, datasets, services, software, projects and places. Either through keyword search or by exploring the entities in the model.

(The numbers indicate how many instances of given type (including subclasses) are there, distinguishing between explicitly created instances and all instances.)

Highlights



Main Entities

Browse through the manually crafted dataset of top-level entities like project partners, services and datasets, for which descriptions did not previously exist.



Showcases

This is a hand curated list of entities-detail views and other specialized pages to highlight specific aspects of the application or the dataset.



Images

Browse through records featuring an image representation.



Places

Browse through geo-referenced records. (Only available for records with geo-coordinates).

Screenshot of the PARTHENOS Discovery (part of PARTHENOS VRE) start page

PARTHENOS: experiences



PARTHENOS

- **Semantic rich data** sets are **not easy to gather**
- Even more **complicated** to **model and map** it in the destination data model
- CIDOC CRM allows a **lot of expression** but becomes very fast **very complicated** - it is usually also necessary to expand the basic model (in this case introducing the PARTHENOS Entity Model)
- The innovative **approach** of PARTHENOS (the vision was very optimistic but nevertheless an important forward-looking proof of concept) relies strongly on the availability of a **lot of data** with a **high quality** and an extensive detailed and structured **“rich description”** (source data model and fields are crucial)
- **Archive material important** but was not much harvested due to less archive data collections in consortium & **missing linked data** approaches by archives


Archives: structured data


- A **shift** is necessary **from descriptive data to structured data** if we like to take the **next step** from digital search to **digital analysis and research**
- This is true for **all kind of data**: (meta)data on archive material, research data, transcriptions and even publications
- **FAIR data principles as a guidance**: important are persistent identifiers (**PIDs**) to find the **data points for connection** between data collections and **machine-readable data**: clear distinction of fields and entities (highlighting by structure entities like persons, places, events)
- Currently **establishing harvesting** platforms is already a **tough job** (finding **common grounds** not easy: different schemas, software, mapping tools, recording traditions, challenges for infrastructure and staff) - PARTHENOS shows future paths to be taken to **expand the digital research potential**

Archives: PARTHENOS - information in title

Parthenos Discovery SPARQL Quick Links Login

[Home](#) / [Search](#)

 Search
Full-text search with faceting.

 Help

holding

Literal Resource Class

Hide Filter

type

about

source

Search literals...

ARIADNE (179)

European Holocaust Research Infrastructure (35)

LRE MAP (1)

found 35 matches

Subject	Type
Drawing by Alexander Bogen of a female partisan standing and holding a rifle	Linguistic Object, Collection
Fort Polk Historical Holding U.S. Army medical blanket collection	Linguistic Object, Collection
Drawing by Alexander Bogen of a bearded partisan, seated and holding a rifle	Linguistic Object, Collection
Original Acquisition Event of American soldiers holding Nazi flag	Acquisition
Collection Event for American soldiers holding Nazi flag	Production
Collection Event for Drawing by Alexander Bogen of a female partisan standing and holding a rifle	Production
Time span of Collection Event for American soldiers holding Nazi flag	Time-Span
Original Acquisition Event of Photograph of the Rejers holding their infant daughter	Acquisition
Collection Event for Drawing by Alexander Bogen of a bearded partisan, seated and holding a rifle	Production
Moskovits Archive: Court cases against Dierig Holding Company Archivo Moskovits: Prozess gegen Firma Dierig Holding	Linguistic Object, Collection

« 1 2 3 4 »

Screenshot of the PARTHENOS Discovery (part of PARTHENOS VRE) search results page

Archives: PARTHENOS - information in keywords

Drawing by Alexander Bogen of a bearded partisan, seated and holding a rifle Feedback

URI

<http://parthenos.d4science.org/handle/EHRI/PORTAL/Thing/88F1AD94-759E-30D9-AEBF-C75D210ADC9D>

[View in RDF visualizer](#)

Type

Linguistic Object, Collection

Source

European Holocaust Research Infrastructure (http://parthenos.d4science.org/handle/api_____::parthenos___:ehri::units)

Minimal metadata

[Source \(Raw triples\)](#)

[Viewer](#)

[Knowledge Graph](#)

Type

Object, Art

is About

[Bogen, Alexander., World War, 1939-1945--Underground movements--Lithuania., Guerillas--Belarus., World War, 1939-1945--Jewish resistance--Belarus., Holocaust, Jewish \(1939-1945\), in art., World War, 1939-1945--Underground movements--Belarus., Jewish ghettos--Lithuania--Vilnius., Guerillas--Lithuania., World War, 1939-1945--Jewish resistance--Lithuania.](#)

Archives: Description vs. data

- In PARTHENOS most **information gained** from **archive material** is in the **title** and in the **description** - sometimes you also have **keywords**
- What PARTHENOS **missed** is the **hierarchical structure** in the archives that also gives implicit information (but also usually due to the title/description of the holding/fonds e.g. the description of “scope and content”)
- Both **title and description** as well as keywords is **not structured enough**:
 - There is the need to **explicit highlight by structure the important entities** that are mentioned or somehow related to an archive material e.g. persons, places, concepts
 - There is also the need to **say something about the relation** how an entity is connected to an archive material e.g. archive material *is a **photograph** showing **person X** who is *photographed by **person Y** at **place A** on **time B** as part of the **event Z****
 - The bold entities and the italic relations all **need to be in a structured data format** e.g. linked data formats like RDF in a structure like the one defined by the **ontology** CIDOC CRM and using common agreed **vocabularies**

Archives: structured information?



Sign In

English

HOME SEARCH DIRECTORY FEATURED DOCUMENTS API

Search in all content Search

Netherlands > Internationaal Instituut voor Sociale Geschiedenis

Labour and Socialist International Archives

- More before...
- SEKRETARIAT
 - More before...
 - Kontakte mit angeschlossenen Parteien und Informationen betr. die verschiedenen Ländern
 - Deutschland
 - More before...
 - Österreich -
 - More before...
 - Sonstige Korrespondenz
 - Korrespondenz mit Österreicher bis Febr. 1934 -
 - More before...
 - Illmayer, F. Bregenz. 1926. 1 Bl.**
 - More after...
 - More after...
 - More after...
 - More after...
 - More after...
 - More after...

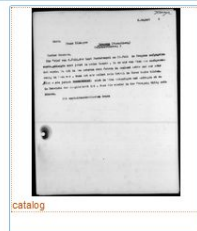
Bookmark this Print

<http://hdl.handle.net/10622/ARCH01368> - 3866

Illmayer, F. Bregenz. 1926. 1 Bl.

1926.

Digital objects



cat@log

Send feedback to the archive

Screenshot of the Archives Portal Europe search detail result page

Archives: structured information?

CONTENTS << Illmayer, F. Bregenz. 1926. 1 Bl.

3866

Image 2 of 2 Go

Illmayer

6. August 6

Herrn Franz Illmayer Bregenz (Vorarlberg)
Weissenröthweg 1

Wertes Genosse,

Ihr Brief vom 9. Juli, der laut Poststempel am 22. Juli in Bregenz aufgegeben wurde, gelangte erst jetzt in meine Hände, da er mir von Wien aus nachgeschickt wurde. Da ich in den letzten drei Jahren im Ausland lebte und nur sehr wenig in Wien war, kann ich mir selbst kein Urteil in Ihrer Sache bilden. Ich würde jedoch ~~inzwischen~~ mich in Wien erkundigen und anfragen ob es in Bereiche der Möglichkeit ist, dass Sie wieder in der Bewegung tätig sein können.

Mit sozialdemokratischem Gruss

MORE INFORMATION

About this item

Dates

1926.

Part of

Labour and Socialist International Archives

Inventory number

3866

Refer to this record

<https://hdl.handle.net/10622/ARCH01368.3866>

Refer to this item

<https://hdl.handle.net/10622/ARCH01368.3866>

Use restrictions for this item

Please consult the IISH copyright statement before using this item

{IISG}

Page

Refer to this image

<https://hdl.handle.net/10622/E05BEFD8-F38D-4DC9-A545-279CD794A83D>

Collection International Institute of Social History (Amsterdam)

Screenshot of an item from an Archives Portal Europe search detail result page at the source provider

Archives: structured information?



Navigation bar with a background image of a historical map and document. It includes a "Sign In" button in the top right corner, a language dropdown menu set to "English", and a search bar with the text "Search in all content" and a "Search" button. The navigation menu contains links for HOME, SEARCH, DIRECTORY, FEATURED DOCUMENTS, and API.

Switzerland > Universitätsbibliothek Basel

Enzeldokumente Universitätsbibliothek Basel

- More before...
- Brief an Christoph Bernoulli-Paravicini / von Alexander von Humboldt
- More after...

Content provider
Basel UB

Extent
1 Bogen

Dimensions
4'

General remarks
Bemerkung
Die Unterschrift ist herausgeschnitten (vgl. dazu die Notiz über dem Textbeginn von C. Chr. Bernoulli, 1890).

Source of acquisition
Akzession
Kauf. Herkunft: Erasmushaus. Datum: 11.06.1996. Akz.-Nr.: 0285919. Eigentümer: Öffentliche Bibliothek der Universität Basel

Conditions governing reproduction
Benutzungsbeschränkung
Es gelten die generellen Benutzungsregeln für den Sonderlesesaal

Keywords
Personen
Personal names:Humboldt, [Alexander](#) <<von>> (1769-1859) (Autor) Bernoulli, Christoph (1782-1863) (Adressat)
Körperschaften
Corporate names: Erasmushaus (Firma) (Früherer Eigentümer)
Orte
Geographic names:
Paris

[Send feedback to the archive](#)

Screenshot of the Archives Portal Europe search detail result page

Archives: RiC

- Necessary **linked data oriented format** is in general on the rise but still in an **early stage** of implementation
- For archives there is a promising development: ICA's new standard **Records in Contexts (RiC)** => <https://www.ica.org/en/records-in-contexts-ontology>
- RiC is oriented to **linked data expressions** and can be compared to an ontology like CIDOC CRM therefore allows **easy mappings** to ontologies
- It is downward compatible to ICA ISAD(G) standard but allows **new additional and detailed structuring of information**
- Thus it enables archives and researchers **to enrich already recorded** archive holdings and to **start new archive recordings** with a rich **relation and entity related data model**

Cracking the archive: experiences with RiC

- RiC allows very **detailed recording of data on archive material** without loss of archive structure rules and benefits of descriptive oriented recording
- For the **full effect** it is necessary to **record details of archive** material which means **more effort** when doing the recording: typing something in a description field is way more faster and easier for humans than **setting relations to entities** where ideally **identifiers should be applied**
- Also it implies **more intensive research** when it comes to details
- Is also true for **research results and data** where the descriptive approach (publications) has a strong tradition and advantages but will also need to **orient more** on the **structured way of data representation** (like TEI for digital editions)
- Advantages of such a switch to structured data will **show up in the future**

Cracking the archive: experiences with RiC

Archiv TFM [Startseite](#) [Impressum](#) [Datenschutzerklärung](#) [Anmelden](#)

Das digitale Archiv TFM ist im Aufbau. Daten werden beständig eingetragen, Lücken in der digitalen Abbildung des Bestands sind deswegen möglich. Zudem können sich Inhalte ändern. Bei Fragen und für Hinweise freuen wir uns über eine Nachricht an: archiv.tfm@univie.ac.at

Suche nach Titel von Einträgen

Abendroth, Irene

Identifizier: <https://archiv-tfm.univie.ac.at/record-set/8537> (Last change: 05.09.2019 - 14:54)

Record Set

Shared Properties

Local Identifier: 414.P0002
General Note
Autograph Irene Abendroth

Record Set Properties

Type: collection
Classification: Aufführungsdokumentation
Physical or Logical Extent
1 graphic record
1 other record

HIC Properties

Stempel Zentralinstitut: Nein

Record Set Relation

is member of
[Foto- und Graphiksammlung](#)

Agent Relations

has subject
[Abendroth, Irene](#)

Concept/Thing relations

has subject
[Die lustigen Weiber von Windsor \[Oper\] / Otto Nicolai](#)

Archiv TFM [Startseite](#) [Impressum](#) [Datenschutzerklärung](#) [Anmelden](#)

Das digitale Archiv TFM ist im Aufbau. Daten werden beständig eingetragen, Lücken in der digitalen Abbildung des Bestands sind deswegen möglich. Zudem können sich Inhalte ändern. Bei Fragen und für Hinweise freuen wir uns über eine Nachricht an: archiv.tfm@univie.ac.at

Suche nach Titel von Einträgen

Abendroth, Irene

Identifizier: <https://archiv-tfm.univie.ac.at/agent/8535> (Last change: 05.09.2019 - 14:50)

Agent 

Shared Properties

Global Persistent Identifier: <http://d-nb.info/gnd/116002506>

[View Record Set](#)

Agent Properties

Type: [person](#)

[has subject](#)
[is subject of](#)
[Abendroth, Irene](#)



Screenshot of the website of the theatre historical collection and archive of the Department of Theatre, Film, and Media studies at Vienna University - combining two detail pages

Cracking the archive: entity driven recording

- The **more entities and relations** are **identified** and bound to PIDs the more **complex and innovative research** will be possible
- Identifying entities is **time consuming** as it involves reading the content
- I'm sceptical that this task can be solved with **machine-learning**: support yes, but in the end it needs **humans to gain high-quality (meta)data**
- At first it looks like as the burden lies **at the source provider** like archives
- But it is unrealistic that the **necessary resources** are allocated there to create high quality and detail data (also archives should be neutral agencies)
- Instead **investment in solutions** where researchers and research projects can **enrich data from archives** (e.g. crowdsourcing) and **delivers back** this contextual information into platforms like APE, PARTHENOS - [EHRI](#) as a good example for such an approach

Cracking the archive: APE and PARTHENOS

PARTHENOS Discovery Lab

Here are a few more advanced options to explore and experiment with the dataset.

Screenshot of the
PARTHENOS
Discovery start page
(showing one part)



SPARQL

User interface to run any SPARQL query on the underlying dataset.



Statistics

A suite of aggregation queries offering an overview over various aspects of the dataspace.



Data Notebooks

Connect to the dataset via a Jupyter Notebook and do your own processing.



Curation

Collection of custom pages to run experiments on or to showcase data quality aspects of the dataset.

- Having **enriched structured data based on FAIR data principles** is a main **requirement** for **pushing digital research** and digital work with archive materials to the next level
- It allows to do **complex queries** across many domains (disciplines, archives, methods, research questions) by using methods like SPARQL queries, data notebooks, innovative analysis and supporting the two PARTHENOS modes

Social Sciences & Humanities Open Cloud



Frank Fischer

Higher School of Economics (Moscow) & DARIAH-EU

Laure Barbot

DARIAH-EU

Project:



SSHOC

social sciences & humanities open cloud



Horizon 2020
European Union Funding
for Research & Innovation

Type of action & funding:
Research and Innovation action
(INFRAEOSC-04-2018)

Partner institutions: 48

SSH ESFRI Landmarks and Projects
& international SSH data infrastructures
230+ people involved

Project budget:

€ 14,455,594.08

Duration: 40 months
(January 2019 – 30 April 2022)

Project website:
www.SSHOpenCloud.eu



Why Social Sciences and Humanities Open Cloud?

Objectives:

1. **To build the SSH Cloud:** creating the social sciences and humanities (**SSH**) part of European Open Science Cloud (**EOSC**)
2. **To maximise data re-use:** maximising **re-use** through **Open Science** and **FAIR** principles (standards, common catalogue, access control, semantic techniques, training)
3. **To interconnect RIs:** interconnecting existing and new infrastructures (clustered cloud infrastructure)
4. **To set up governance:** establishing appropriate **governance model** for SSH-EOSC

Being realised through 9 Work Packages:

WP1 Project Management and Administration

WP2 Communication, Dissemination and Impact

WP3 Lifting Technologies and Services into the SSH Cloud

WP4 Innovations in Data Production

WP5 Innovations in Data Access

WP6 Fostering Communities, Empowering Users & Building Expertise

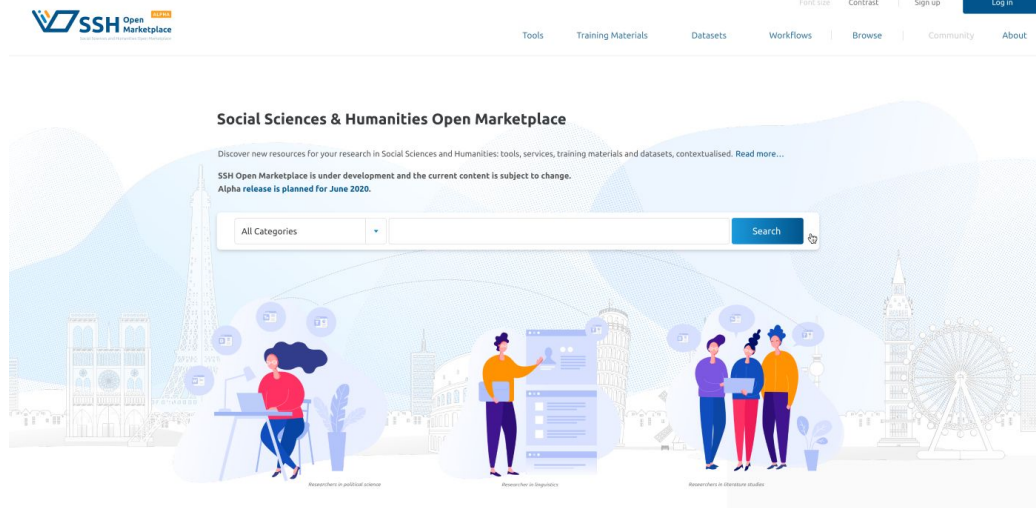
WP7 Creating the SSH Open Marketplace

WP8 Governance, Sustainability, Quality Assurance

WP9 Data Communities

SSH Open Marketplace

- Discovery portal for SSH resources
 - Tools & services
 - Training materials
 - Workflows
 - Datasets
 - Publications
- 3 guiding principles
 - Contextualisation
 - Curation
 - Community

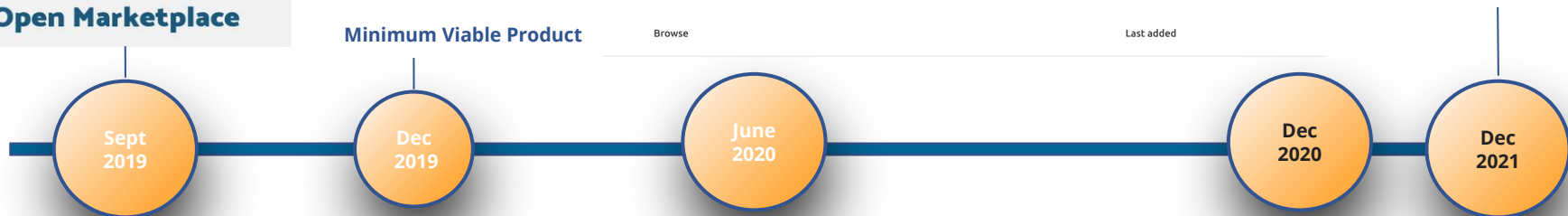
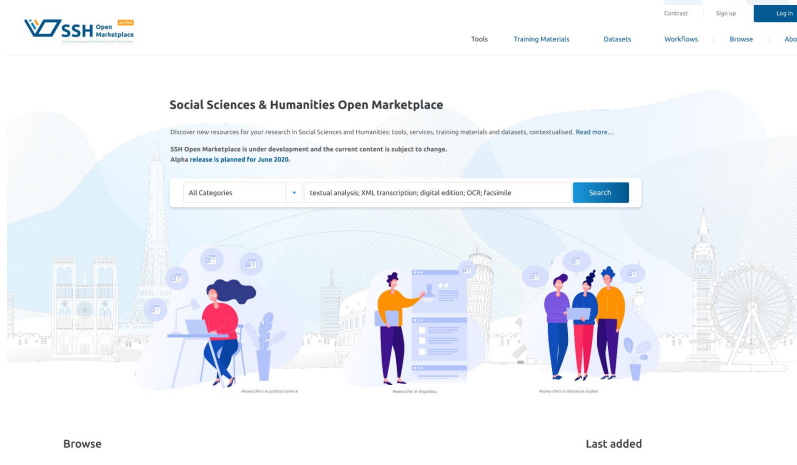


Creation of the SSH Open Marketplace – Timeline

DOI: [10.5281/zenodo.3547649](https://doi.org/10.5281/zenodo.3547649)



**D7.1 System Specification
SSH Open Marketplace**



Minimum Viable Product

Browse

Last added

Final release

**MS42 – Alpha release –
Marketplace available for
testers and internal project
review**

Beta (first public) release

The SSH Open Marketplace: discovery portal and metadata

- Metadata aggregation
 - Identification and prioritisation of sources
 - Beta release: TAPoR, SSK, Programming Historian, DH conference papers, CLARIN Switchboard, EOSC catalogue
- Mappings between data models
 - Flexible and evolutive data model - dynamic properties
- Enrichment and Curation

Referencing semantic artifacts in the SSH Open Marketplace?

- Forthcoming SSHOC conversion Hub as part of the MP (T3.5)
- Semantic artifacts as items in the MP:
 - controlled vocabularies
 - dictionaries and thesauri
 - metadata schemas
 - ontologies

Cf. D7.6 “Resources for Marketplace content description” to be released at the end of the year

Metadata discussions within SSHOC

- WP3 - Lifting Technologies and Services into the SSH Cloud - CLARIN
 - [D3.1 Report on SSHOC \(meta\)data interoperability problems](#)
 - [Inventory of metadata sets, standards and taxonomies used in SSH](#)

ARCHIVES
PORTALEUROPE



Thank you very much for your attention
and participation!

www.archivesportaleurope.net
info@archivesportaleurope.net