



# The Europeana Use Case

## Multilingual Terminology Mapping at Europeana

Vivien Petras

Berlin School of Library and Information Science

18 April 2013

Linked Heritage Seminar on Multilingualism and Terminology

# Contents

- Europeana: Multilingual Collections & Users
- EDM and the Semantic Data Layer
- Multilingual Terminology Alignment EuropeanaConnect
- Mapping in Europeana-related Projects
- Automatic Multilingual Enrichment in the Europeana portal
- Preview: New Enrichment Ideas



# Europeana

Home My Europeana

Choose a language ▾



Search ▾

Search


Help

## New exhibition : Leaving Europe

Why did 30 million Europeans emigrate to America in the 19th/20th centuries? This exhibition of rarely seen pictures tells their story.

Open Exhibition



From the blog 



### Meet Europeana: Michelle van Duijn

Today, we introduce you to Michelle van Duijn, Europeana Web Communications Assistant. Find out what she does all day and why she loves Alice in Wonderland... Michelle is 26 years old and has a bachelor degree in Economics (majoring in ...



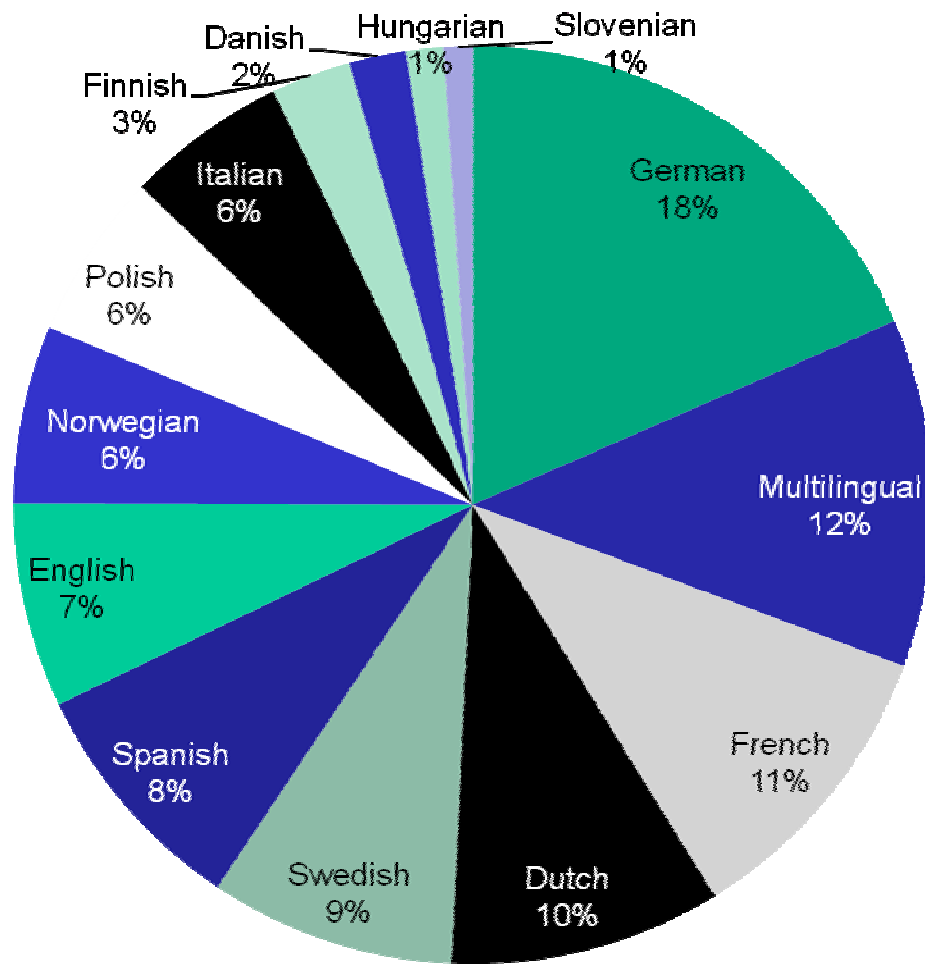
### Faithful Companions

Maybe you wouldn't say it, but donkeys have played a prominent role in the human world for a long time. Not only have they been helping with work on farms and transport for at least 5,000 years, they play a part in ...

- 26.7 million objects
- 15.2 million images
- 10.6 million texts
- 450,000 sound files
- 170,000 video files

- > 2,200 institutions
- > 30 countries

# Europeana Multilingual Collections

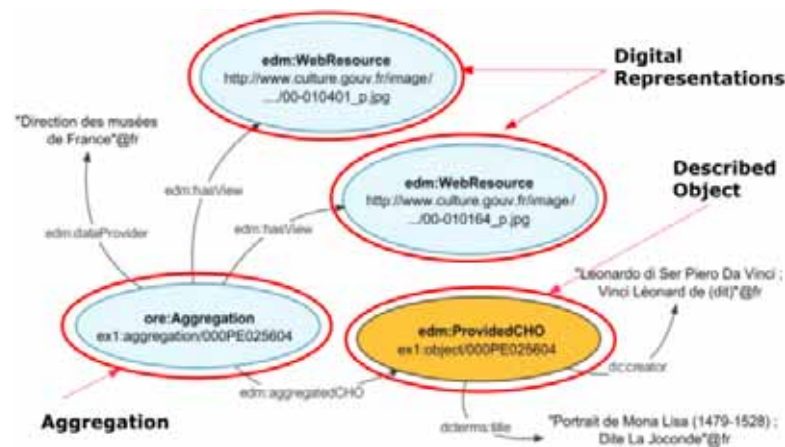


→ Many Europeana objects are language-independent (e.g. images), but the meta-data is multilingual.

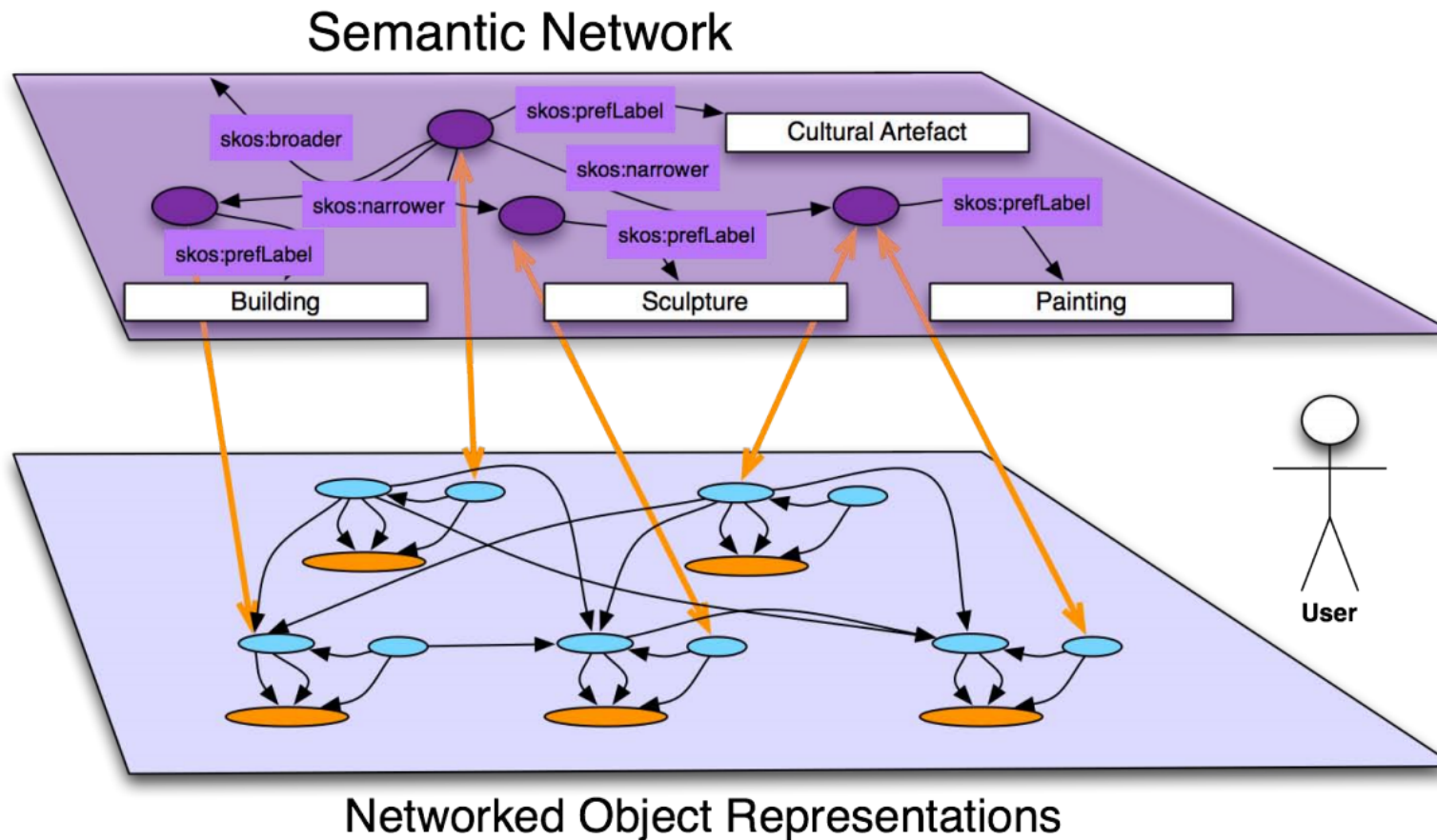


# Europeana Data Model (EDM)

- linked data model
- representation of cultural heritage objects from libraries, archives and museums
- unites several standards and vocabularies
- is as generic as possible
- can be specialised for different domains
- Allows alignment to vocabularies (KOS)  
→ semantic data layer

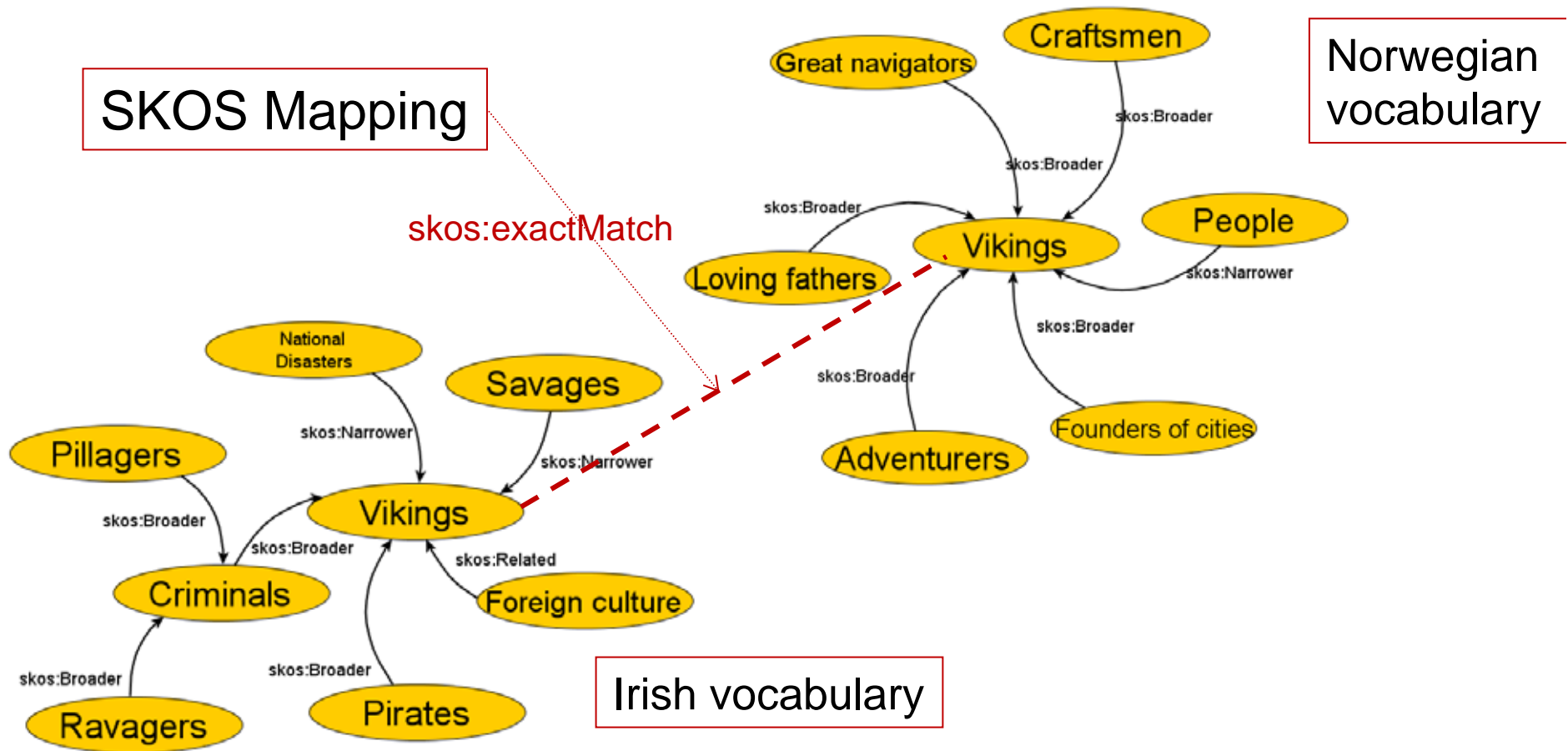


# Europeana Semantic Data Layer



Doerr, M.; Gradmann, S.; Hennicke, S.; Isaac, A.; Van de Sompel, H. (2010).  
The Europeana Data Model (EDM). 76th IFLA General Conference and  
Assembly 10-15 August 2010, Gothenburg, Sweden.

# Semantic Data Layer Alignment Example



Cousins, Jill (2010). *Europeana Overview*.  
Europeana Open Cultures Conference, 14-15 October Amsterdam

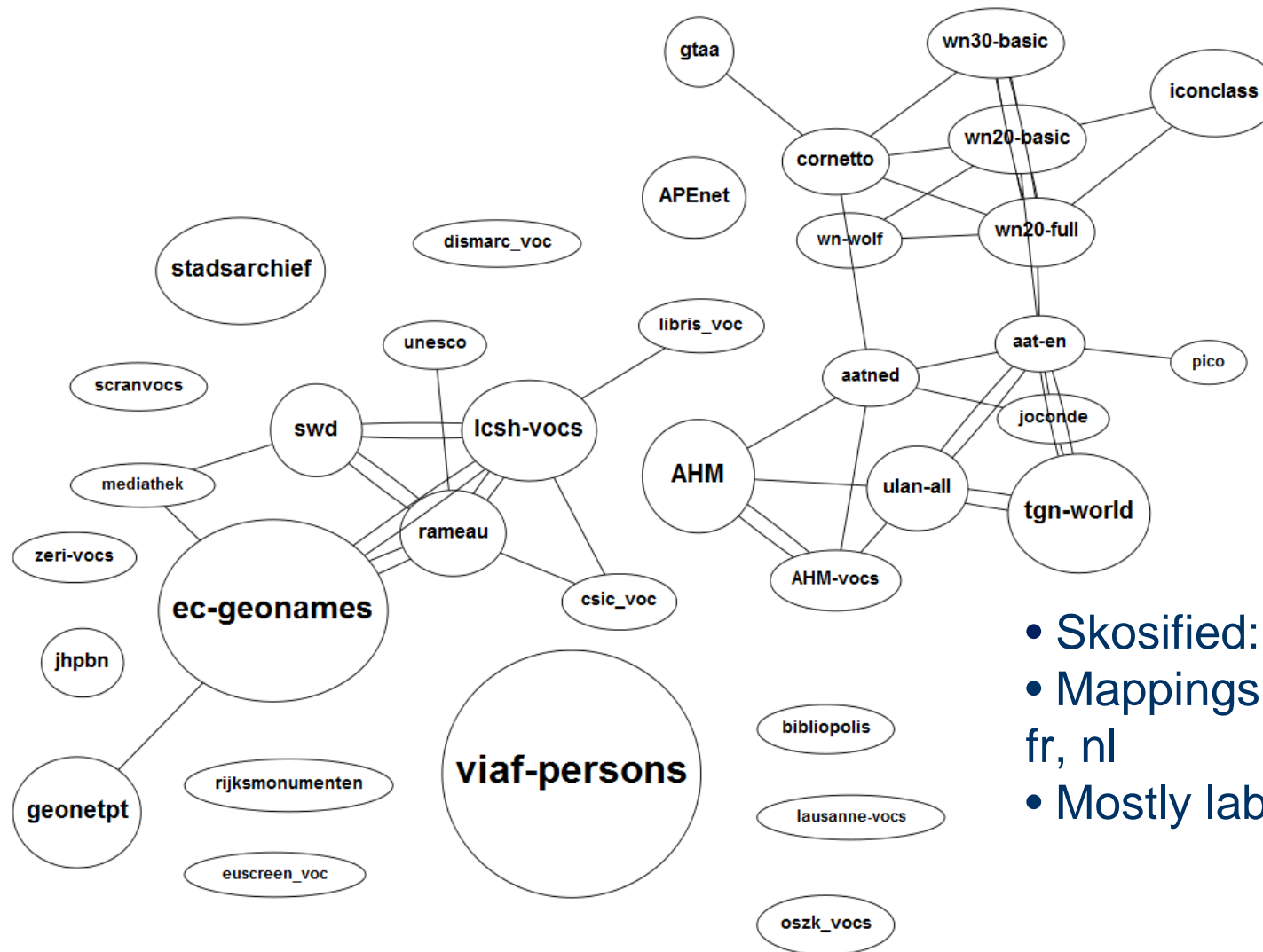
# Multilingual Terminology Alignment: EuropeanaConnect

- Alignment to **pivot vocabularies** (e.g. UDC, DDC, VIAF, TGN, Geonames, Wordnets, dbPedia)
- skos:exactMatch
- Methodology:
  - Conversion to SKOS/RDF
  - Different alignment methods (Lexical matching, Structure-based matching, Instance-based matching)
  - Disambiguation of matching candidates
  - Combining alignments

→ AMsterdam ALignment GenerAtion Metatool  
([semanticweb.cs.vu.nl/amalgame/](http://semanticweb.cs.vu.nl/amalgame/))



# Semantic Alignments of Vocabularies



- Skosified: en, fr, de, nl, hu
- Mappings (>500,000): en, fr, nl
- Mostly label matches

# The European Library and the MACS Initiative

- MACS: Multilingual Access to Subjects
- Initiative to map LCSH – Rameau – SWD subject headings

**macs**  
multilingual access to subjects

**Search Results**  
*Lp-super/super/Work*

Readers  
Annotators  
Search Links  
List Terms  
Help  
Logout

Editors  
Admins  
Add New Link  
To Do  
To Know  
Tag Mgt

Supervisors  
Maintenance

Searched for \* in SWD.  
Search retrieved 61567 links in 3.30 seconds.

1/6157 Next Page

SWD	RAMEAU	LCSH	Domains	MACS Link	Dups
100Cr6	--null--	--null--	600	MACS0132433	
360-Grad-Feedback	Cadres (personnel) -- Évaluation	360-degree feedback (Rating of employees)	650	MACS0128696	
360-Grad-Feedback	Personnel -- Évaluation	360-degree feedback (Rating of employees)	330 350	MACS0133464	
A7V Sturmpanzerwagen	A7V (char d'assaut)*	A7V (Tank)	355 600	MACS0134758	
Aachener Frieden <1748>	Traité d'Aix-la-Chapelle (1748)	--	940	MACS0058040	
Aachener Printe	Pain d'épice AND Aix-la-Chapelle (Allemagne)	Gingerbread AND Aachen (Germany)*	390 641	MACS0134979	
Aale	Anguillidés	Anguillidae	590	MACS0119013	
Aalenium	Aalénien	--	550 560	MACS0076568	
Aalquappe	Lota	Lota	590	MACS0016185	
ab	--null--	--null--	400	MACS0130006	

1/6157 Next Page

# The European Library and the MACS Initiative

- TEL: automatic mapping to MACS (subjects), VIAF (persons), Geonames (places)

Home → Authority search for "voyages"

Everything

Full text

Remote Search

Authority

REFINE

Authority Type: **Subject**

Authority Type

Subject (103)

[Voyages interplanétaires](#) (Subject)

Voyages interplanétaires (French)

Interplanetary voyages (English)

[Voyages -- Cartes](#) (Subject)

Voyages -- Cartes (French)

[Reise](#) (Subject)

Reise (German)

Voyages (French)

Voyages and travels (English)

[Imaginäre Reise](#) (Subject)

Imaginäre Reise (German)

Voyages imaginaires (French)

Voyages, Imaginary (English)

Phantasiereise (German)

[Étudiants -- Voyages](#) (Subject)

Étudiants -- Voyages (French)

Students -- Travel (English)

Schulfahrt (German)

Élèves -- Voyages (French)

Zoom

# Europeana 1914-18

- UGC related to WWI
- Keywords translated & searched in 8 languages
- Next: keyword alignment to LCSH

The screenshot shows the Europeana 1914-18 search results page for the keyword 'Trench Life'. The page features a navigation menu with 'Home', 'Add', 'Explore', and 'An Unlikely Frier'. A search bar at the top right allows for language selection. The main content area displays a grid of search results, each with a thumbnail image and a brief description. The results include items such as 'Isenkrammeren, der lavede genstande af patronhylstre i skyttegravene', 'Thomas Clarkson - Bravery In The Trenches, 1916', 'World War 1 - Picture of soldiers in trench (Somme)', 'Feldpostkarte - "Eine Villa unter der Erde"', 'Fred - War diary of Frederick Charles Rawlings', 'My grandfather in Royal Naval Division 1915-1919 - Able Seaman's Diary', 'Max Jacoby on the Eastern Front - A Jewish doctor in the German Army', 'My Great Uncle Albert who I never knew - Letters between brother and sister', 'Christmas Truce Of 1914-Grandad Williams Story - My Great Granddad William Wood and My Granddad Thomas Wood', 'Andrew Neill - his life story', 'Poem Passed Down - Poem about going to war and coming back injured', 'From Geordie Land To No Mans Land - Diaries turned into a book', 'My poetry on WW1 - Poem "The 14-18 war"', 'Andrej Majnik na Ruski in Soki fronti ter v ruskem ujetnistvu - Andrej Majnik on Russian and Isonzo front', and 'Feldpostgruße von Johann Enzelmann - Field postcards of Johann Enzelmann'. The 'Keywords' section on the left sidebar is circled in red, and the 'Trench Life' keyword is also circled in red.

Matches for: Trench Life ( 749 stories )

Select a language

Home Add Explore An Unlikely Frier

Back to search results

## World War 1

Translate to ...  
Powered by Microsoft® Translator

**Story**  
My grandfather is one of the men in the picture. He was a brave brav serve his country. We miss him every single day.

**Contributor**  
Rebecca

**Summary description of items**  
Picture of soldiers in a trench. Royal Irish Rifles ration party Somme.

**The story begins ...**  
1916 – 1916

**Language**  
English

**Keywords**  
Trench Life

**Theatres**  
Western Front

Isenkrammeren, der lavede genstande af patronhylstre i skyttegravene  
The hardware dealer who made things from gun cartridges

Thomas Clarkson - Bravery In The Trenches, 1916  
March 13th 1916 at Fleurbaix

World War 1  
Picture of soldiers in trench (Somme)

Feldpostkarte - "Eine Villa unter der Erde"  
Field Postcard - "A villa in the ground"

Fred  
War diary of Frederick Charles Rawlings

My grandfather in Royal Naval Division 1915-1919  
Able Seaman's Diary

Max Jacoby on the Eastern Front  
A Jewish doctor in the German Army

My Great Uncle Albert who I never knew  
Letters between brother and sister

Christmas Truce Of 1914-Grandad Williams Story  
My Great Granddad William Wood and My Granddad Thomas Wood

Andrew Neill - his life story

Poem Passed Down  
Poem about going to war and coming back injured

From Geordie Land To No Mans Land  
Diaries turned into a book

My poetry on WW1  
Poem "The 14-18 war"

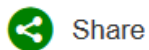
Andrej Majnik na Ruski in Soki fronti ter v ruskem ujetnistvu  
Andrej Majnik on Russian and Isonzo front

Feldpostgruße von Johann Enzelmann  
Field postcards of Johann Enzelmann

# Automatic Multilingual Enrichment in Europeana



View item at  
Stadtgeschichtliches  
Museum Leipzig [↗](#)



Share



Cite on Wikipedia



Translate details

## Prosit Neu Jahr 1891. Familie Drescher Leipzig

Search also for:

Title

[Prosit Neu Jahr 1891. Familie Drescher](#)

Auto-generated tags ▾

What ▾

**Concept Term:** <http://www.eionet.europa.eu/gemet/concept/10851>

**Concept Label:** [infrastructură a comunității] (ro); [kamu tesisleri] (tr); [fysisk fellesfasilitet] (no); [közösségi létesítmény] (hu); [kopienai piederošie resursi] (lv); [bendruomenės infrastruktūra] (lt); [gemeinschaftseinrichtung] (de); [community facility, 社区设施] (def); [julkinen palvelu] (fi); [équipement collectif] (fr); [sammhällets facilitet] (sv); [общественно средство] (bg); [javna zgradba/oprema/storitev] (sl); [verejné zariadenie] (sk); [fællesskabsfacilitet] (da); [komunitatearen ekipamendu; europako batasunaren ekipamendu] (eu); [attrezzatura collettiva] (it); [tagħmir komunitarju/ kollettiv] (mt); [مرفق مجتمعي] (ar); [zařízení veřejné] (cs); [δημόσια εγκατάσταση] (el); [budynek użyteczności publicznej] (pl); [recursos colectivos] (pt); [community facility] (en); [общественное сооружение] (ru); [üldkasutatav rajatis] (et); [equipamiento colectivo] (es); [gemeenschappelijke voorziening] (nl)

**Concept Term:** <http://www.eionet.europa.eu/gemet/concept/1410>

**Concept Label:** [cerkev] (sl); [kostol] (sk); [biserică] (ro); [kirke] (da); [chiesa] (it); [kirke] (no); [templom] (hu); [bažnyčia] (lt); [kostel] (cs); [kirche] (de); [εκκλησία/(ιερός) ναός] (el); [church, 教堂] (def); [kirkko] (fi); [budynek sakralny] (pl); [igrejas] (pt); [църква] (bg); [église] (fr); [kyrka] (sv); [church] (en); [церковь] (ru); [kirik] (et); [iglesia] (es); [kerk] (nl)

**Concept Broader Label:** <http://www.eionet.europa.eu/gemet/concept/10851>

When ▾

**Period Term:** [http://semium.org/time/18xx\\_4\\_quarter](http://semium.org/time/18xx_4_quarter)

**Period Label:** [4e quart 19e siècle] (fr); [4 quarter of the 19th century] (en); [4-я четверть 19-го века] (ru)

**Period Term:** <http://semium.org/time/1890>

**Period Label:** [1890] (def)



# Automatic Multilingual Enrichment in Europeana

Vocabulary	Tag type	Enriched metadata fields
GEMET Thesaurus	Concept	dc:subject
		dc:type
		dcterms:alternative
DBpedia	Agent	dc:contributor
		dc:creator
Semium Time Ontology	Period	dc:date
		dc:coverage
		dcterms:temporal
GeoNames	Place	dc:coverage
		dcterms:spatial



# Poisonous India...



Search

Search

Help

[Return to search results](#)

[< Previous](#) [Next >](#)



View item at  
[Musées lausannois](#)

Share

Cite on Wikipedia

Translate details

Select language

Powered by Microsoft Translator

**Creator:** Huber, Claude

**Date:** 1967; <http://semium.org/time/1967>

**Format:** photographie

**Subject:** Inde; <http://www.eionet.europa.eu/gemet/concept/6358>

**Identifier:** P.2.E.19.301.001; 141728

**Rights:** Musée historique de Lausanne

**Provider:** Musées lausannois

**Providing country:** Switzerland

[Auto-generated tags](#)

[What](#)

**Concept Term:** <http://www.eionet.europa.eu/gemet/concept/2233>

**Concept Label:** [cauză a bolii] (ro); [hastalık nedeni] (tr); [sykdomsårsak] (no); [betegség oka] (hu); [slimības cēlonis] (lv); [ligos priežastis] (lt); [krankheitsursache] (de); [disease cause, 致病原因] (def); [taudin syy] (fi); [étiologie] (fr); [sjukdomsorsak] (sv); [причина за болестта] (bg); [povzročitelj bolezni] (sl); [příčina ochorenia] (sk); [gaixotasunaren jatorri; gaixotasunaren kausa] (eu); [sygdomsårsag] (da); [causa di malattia] (it); [kawża ta' mard] (mt); [سبب المرض] (ar); [příčina onemocnění] (cs); [αίτιο (αιτιολογία) νόσου] (el); [causas de doença] (pt); [przyczyna choroby] (pl); [disease cause] (en); [причина болезни] (ru); [haiguse põhjus] (et); [causa de enfermedad] (es); [ziekteoorzaak] (nl)

**Concept Term:** <http://www.eionet.europa.eu/gemet/concept/6358>

**Concept Label:** [strup] (sl); [jed] (sk); [gift] (da); [otravã] (ro); [veleno] (it); [gift] (no); [méreg] (hu); [nuodai] (lt); [jed] (cs); [gift] (de); [δηλητήριο] (el); [poison, 毒藥] (def); [myrkky] (fi); [venenos] (pt); [trucizna] (pl); [отрова] (bg); [poison] (fr); [gift] (sv); [poison] (en); [яд] (ru); [mürk] (et); [venenos] (es); [vergift] (nl)

**Concept Broader Label:** <http://www.eionet.europa.eu/gemet/concept/2233>

**Search also for:**

**Who**

[Huber, Claude \(207\)](#)

**What**

[Inde \(2093\)](#)

**Provider**

[Musées lausannois \(16416\)](#)

# Automatic Multilingual Enrichment Challenges

- Metadata quality & sparsity
- Vocabulary ambiguity
  - domain  
*GEMET print* → (German) *Druck* → *pressure*
  - language  
*electrical Power* → (German) *Strom* → (Czech) *strom* → *tree*
  - context  
*Córdoba* = *Spain* | *Argentina*




*Olensky, M., Stiller, J., Dröge, E. (2012). Poisonous India or the Importance of a Semantic and Multilingual Enrichment Strategy. In: Proc. of MTSR 2012: Metadata and Semantics Research Conference, Nov. 2012, Cádiz, Spain.*


# Preview: New Multilingual Enrichment Plans

→ transition to linked data-based Europeana Data Model (EDM) in production system

- links to contextual vocabularies from providers
- enrich during ingestion



[View](#)



[View item at](#)

**Buccin trombone. Nominal pitch: B?.**

**Title:**

**Description:** Technical description: Brass; ligature fitting on bell section at joint; stockings on main slides. Bell with one coil, angled to face forwards. Repair History: Main slide possibly not original (tenon of slide section of joint is tapered, bell section joint for cylindrical tenon).

**Date:** Circa 1840

**Type:** <http://www.mimo-db.eu/InstrumentsKeywords/4378>; <http://www.mimo-db.eu/HornbostelAndSachs/356>

**Identifier:** #UEDIN:214

**Data provider:** University of Edinburgh

**Provider:** MIMO - Musical Instrument Museums Online

**Providing country:** FR

[Auto-generated tags](#)

[What](#)

**Concept Term:** <http://www.mimo-db.eu/InstrumentsKeywords/4378>

**Concept Label:** [Buccin] (de); [Buccin] (); [Buccin-trombone] (it); [423.22 Labrosones with slides, Trumpets] (def); [Buccin] (sv); [Buccin] (fr); [Buccin] (en); [Buccin] (nl)

**Search also for:**

**Title**  
[Buccin trombone. Nominal pitch: B?.](#)  
(1)

**Provider**  
[University of Edinburgh](#) (3981)  
[MIMO - Musical Instrument Museums Online](#) (41553)

# Preview: Multilingual Enrichment Ideas

- Improved heuristics of enrichment and stricter normalization
- Metadata annotation through user input (social tagging, classification)
- Geoparser and Gazetteer for creation of geodata based on place names
- Open ontology for named periods to use in enrichments
- Extended enrichment of Agents and Concepts based on DBpedia