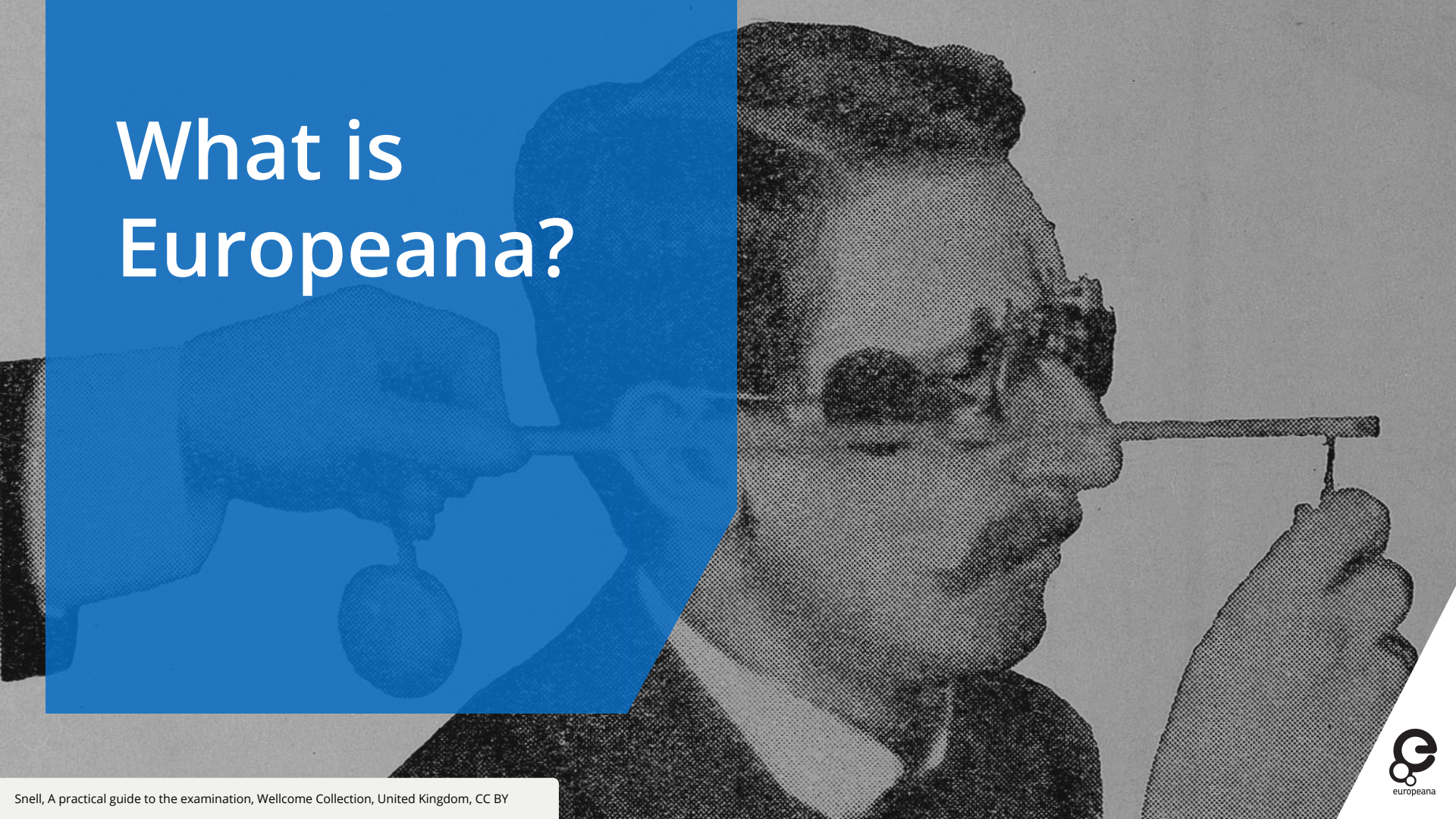


Building a metadata space for cultural heritage at Europeana, using best practices for modeling, enriching and sharing data

Antoine Isaac

with slides from Hugo Manguinhas, Juliane Stiller, Mónica Marrero

What is Europeana?



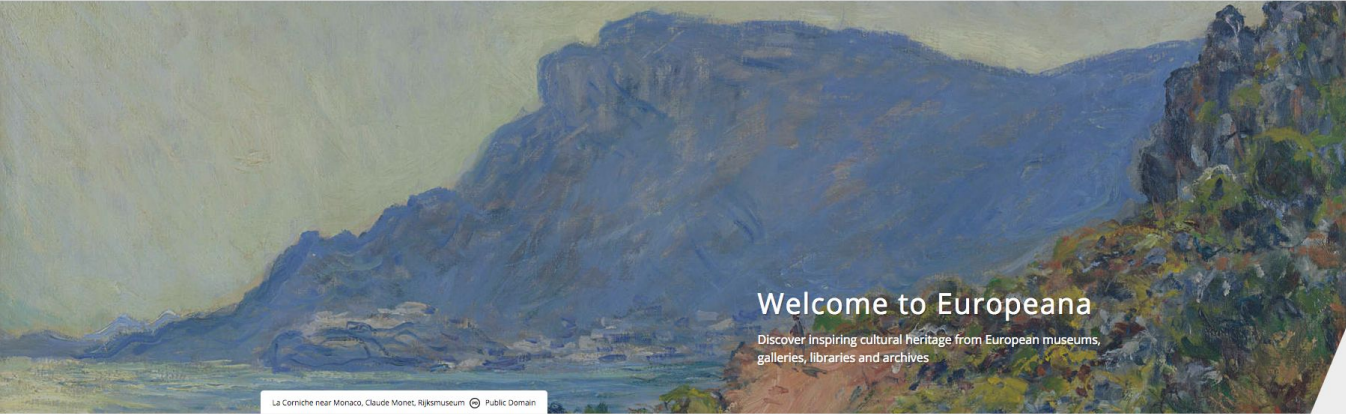
Europeana.eu

- Publishes 52M digitized objects
- From 4,000 libraries, archives and museums
- From 44 countries

Search millions of items

COLLECTIONS TEACHERS ABOUT US

[You're viewing the new Europeana experience. Go to the original Europeana.](#)



Welcome to Europeana

Discover inspiring cultural heritage from European museums, galleries, libraries and archives

La Corniche near Monaco, Claude Monet, Rijksmuseum Public Domain

What's new



Discovering Europe

Discover beautiful and historical



BLOG

Czar Peter House in...



BLOG

Abraham Tuschinski's...



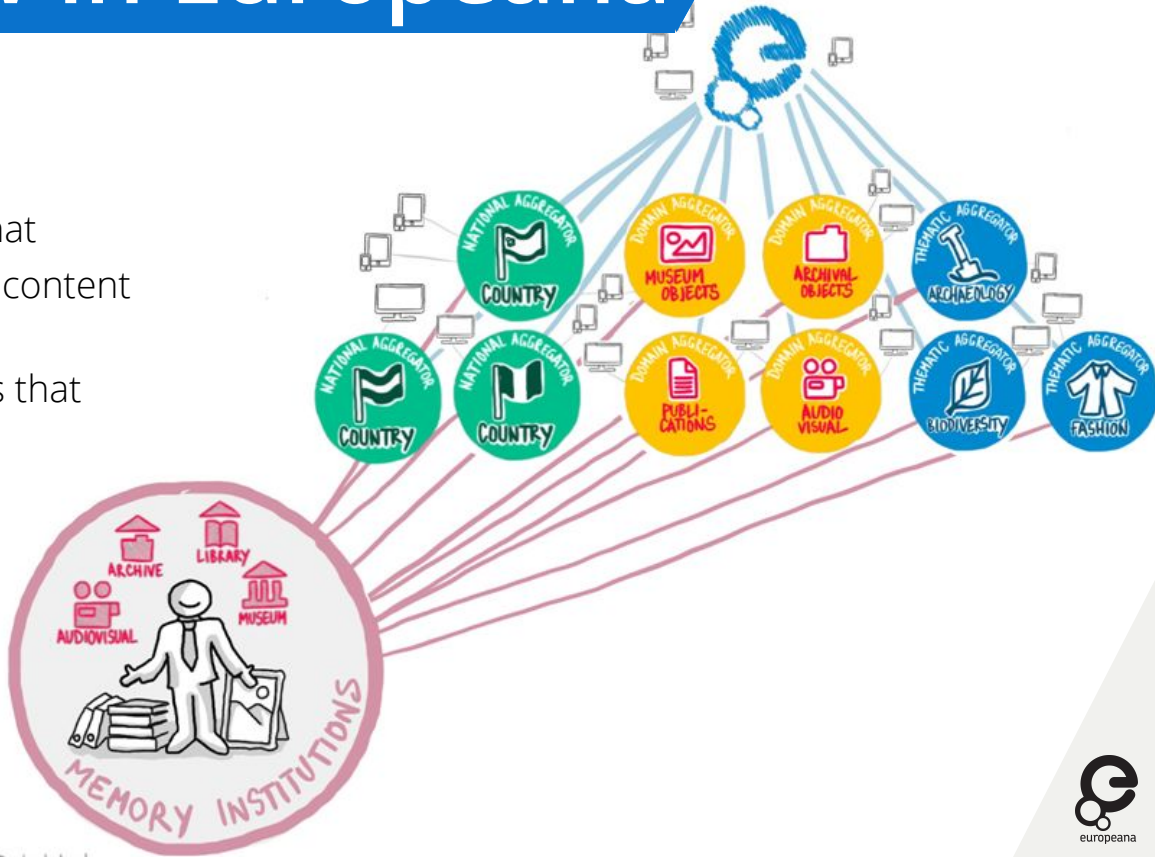
BLOG

Studying abroad

(Meta)data flow in Europeana

Data providers: cultural institutions that provide metadata and links to digitized content

Aggregators: organizations or projects that gather data from a specific country or domain (music, fashion, archaeology...)



A diverse set of data(sets)

We must harvest and process metadata that is very heterogeneous

Many different domains, themes and types of object

Books, newspapers, letters, paintings, maps, photographs, music, fashion, sculpture, 3D objects...

Different data models at source

MARC, LIDO, DC, METS and many more

Plus controlled vocabularies (thesauri, name authority lists...)

Many languages

Datasets officially use 38 languages but many more appear in metadata fields

How to get more homogeneous, richer & multilingual data?



Data modeling for richer, more interoperable metadata

Like many aggregators, we ask providers to use one model

- the [Europeana Data Model](#) (EDM)

But we cannot do whatever we like: we do not operate in isolation!

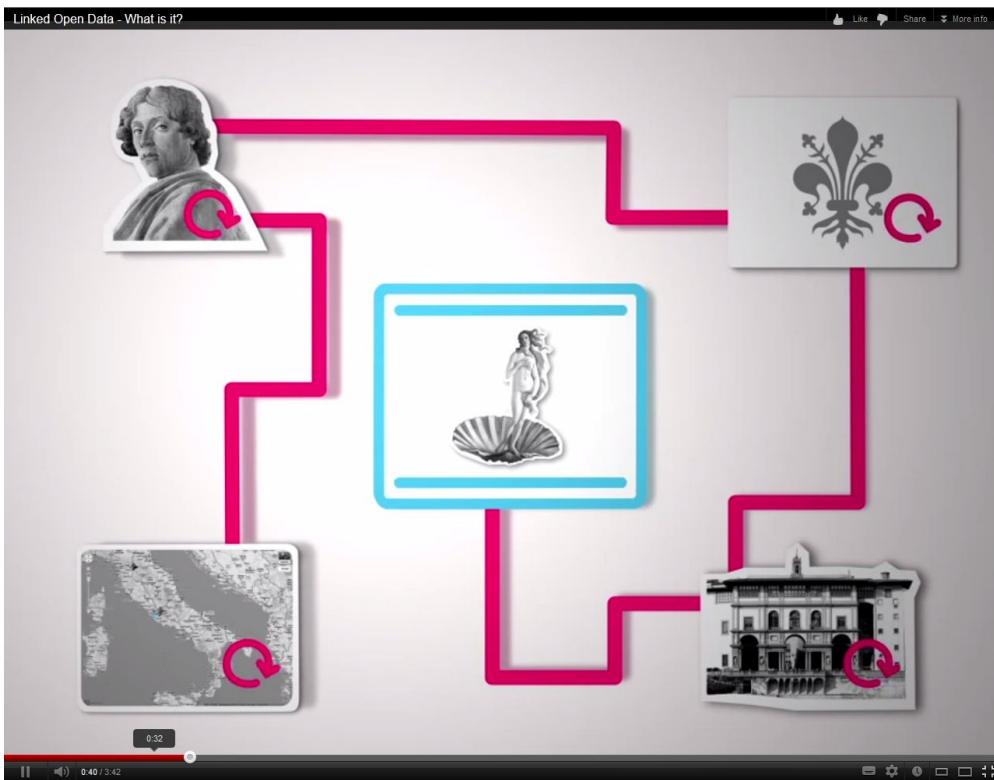
Our approach must be

- easy and rewarding for our partners
- based on community-agreed best practices

Following Linked Open Data Best Practices

Linked Open Data - What is it?

<http://vimeo.com/36752317>



Data modeling for richer, more interoperable metadata

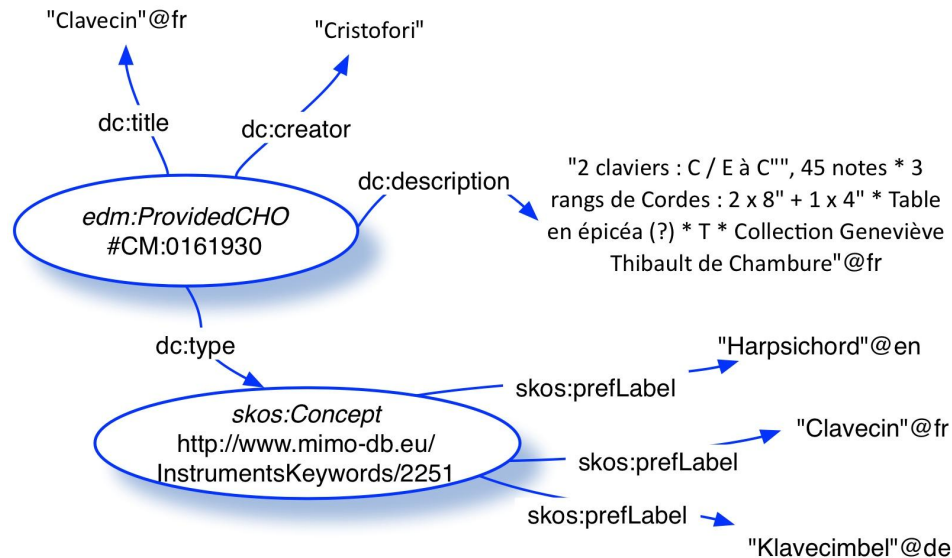


Clavecin, Bartolomeo Cristofori

Cité de la Musique,

MIMO - Musical Instruments Museums Online | CC BY-NC-SA

https://www.europeana.eu/en/item/09102/CM_0161930



Europeana Data Model example

<https://pro.europeana.eu/post/the-europeana-data-model-a-living-model-5-years-on>

Massive re-use of vocabularies



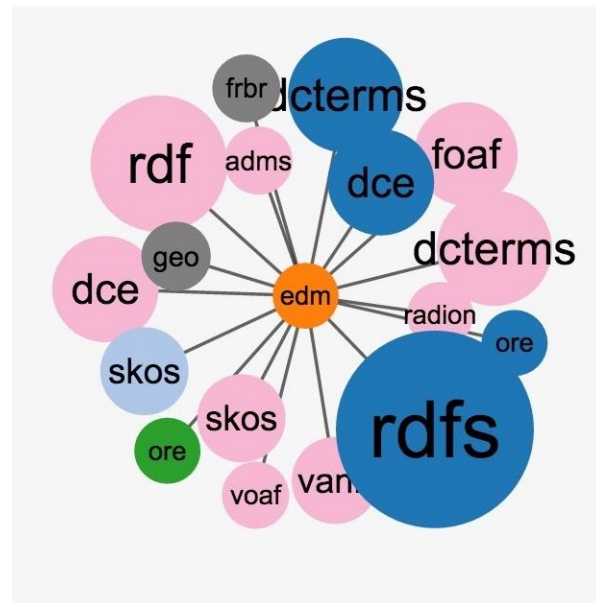
OAI-ORE



FOAF

Plus (sometimes only for one property!)

- Web Annotation
- RDA
- WGS84
- EBUcore
- ccRel
- ODRL
- DOAP
- SVCS
- DCAT
- ADMS
- ...



EDM in Linked Open Vocabularies (LOV)

<https://lov.linkeddata.es/dataset/lov/vocabs/edm>

Enriching metadata

EDM gives a base for (linking to) multilingual metadata
data as resources with web URIs, not only character strings

We encourage data providers to contribute their own links/data to local or external vocabularies

Linked Open Data vocabularies

currently recognized by Europeana in providers' metadata:

The Getty - Art & Architecture Thesaurus (AAT)	http://vocab.getty.edu/aat/
The Getty - Union List of Artist Names (ULAN)	http://vocab.getty.edu/ulan/
Getty Thesaurus of Geographic Names (TGN)	http://vocab.getty.edu/tgn/
Virtual International Authority File (VIAF)	http://viaf.org/viaf/
Geonames	https://sws.geonames.org/
Iconclass	http://iconclass.org/
Gemeinsame Normdatei (GND)	http://d-nb.info/gnd
Israel Museum Jerusalem Concepts	http://www.imj.org.il/imagine/thesaurus/objects/
Library of Congress Subject Headings (LCSH)	http://id.loc.gov/authorities/subjects/
Europeana Sounds Genres	http://data.europeana.eu/concept/soundgenres/
UDC	http://udcdata.info/rdf/
UNESCO Thesaurus	http://vocabularies.unesco.org/thesaurus/
YSO - General Finnish ontology	https://finto.fi/yso/en/
Wikidata	http://www.wikidata.org/entity/
Fashion Thesaurus	http://thesaurus.europeanafashion.eu/thesaurus/
MIMO Concepts	http://www.mimo-db.eu/

Enriching metadata

EDM gives a base for (linking to) multilingual metadata
data as resources with web URIs, not only character strings

We encourage data providers to contribute their own links/data to local or external vocabularies

In parallel, we apply automatic enrichment to link object metadata to reference datasets for places, persons, concepts

Contextual entities

We are building an "Entity Collection":

- Centralized point of reference and access to data about contextual entities: places, agents (persons and organizations), concepts...
- Caching and curating data from the wider Linked Open Data cloud
- A sort of Europeana "knowledge graph"

Data currently in the Entity Collection

216,302 Places

- a subset of **Geonames**, corresponding to places which are part of European countries and of some specific feature classes

165,005 Agents

- a subset of **DBpedia** corresponding to most of the instances of dbp:Artist with some exceptions, and integrated from 49 DBpedia language editions

1,572 Concepts

- a subset of **DBpedia** corresponding to a selection concepts matching the needs from Europeana Collections (e.g., WWI battles).
- Europeana Sounds music genres (obtained from **Wikidata**)
- **Photo Consortium's photography vocabulary**

Entity data example

DBpedia resource for Mozart

```
{
  "@context": http://data.europeana.eu/entity/context.json,
  id: http://data.europeana.eu/agent/base/146951,
  type: "Agent",
  prefLabel: {
    ar: "فولفغانغ أماديوس موتسارت",
    az: "Volfqanq Amadej Mosart",
    be: "Вольфганг Амадэй Моцарт",
    bg: "Волфганг Амадеус Моцарт",
    bs: "Wolfgang Amadeus Mozart",
    ca: "Wolfgang Amadeus Mozart",
    cs: "Wolfgang Amadeus Mozart",
    cy: "Wolfgang Amadeus Mozart",
    da: "Wolfgang Amadeus Mozart",
    de: "Wolfgang Amadeus Mozart",
    el: "Βόλφγκανγκ Αμαντίους Μότσαρτ",
    en: "Wolfgang Amadeus Mozart",
    es: "Wolfgang Amadeus Mozart",
    et: "Wolfgang Amadeus Mozart",
    eu: "Wolfgang Amadeus Mozart",
    fi: "Wolfgang Amadeus Mozart",
    fr: "Wolfgang Amadeus Mozart",
    ga: "Wolfgang Amadeus Mozart",
    gd: "Wolfgang Amadeus Mozart",
    gl: "Wolfgang Amadeus Mozart",
    he: "וולפגאנג מוצרט",
    hi: "वोल्फगांग आमडेयुस मोसार्त",
    hr: "Wolfgang Amadeus Mozart",
    hu: "Wolfgang Amadeus Mozart",
    hy: "Վոլֆգանգ Ամադեուս Մոցարտ",
    is: "Wolfgang Amadeus Mozart",
    it: "Wolfgang Amadeus Mozart",
    ja: "ヴォルフガング・アマデウス・モーツァルト",
    ka: "ვოლფგანგ ამადეუს მოცარტი",
    ko: "볼프강 아마데우스 모차르트",
    lt: "Volfqangas Amadejus Mocartas",
    lv: "Volfqangs Amadejs Mocarts",
    mk: "Волфганг Амадеус Моцарт",
    nl: "Wolfgang Amadeus Mozart",
    no: "Wolfgang Amadeus Mozart",
    pl: "Wolfgang Amadeusz Mozart",
    pt: "Wolfgang Amadeus Mozart",
    ro: "Wolfgang Amadeus Mozart",
    ru: "Моцарт, Вольфганг Амадеевич",
    sk: "Wolfgang Amadeus Mozart",

```

Preferred labels

```
    sl: "Wolfgang Amadeus Mozart",
    sq: "Wolfgang Amadeus Mozart",
    sr: "Волфганг Амадеус Моцарт",
    sv: "Wolfgang Amadeus Mozart",
    tr: "Wolfgang Amadeus Mozart",
    uk: "Вольфганг Амадей Моцарт",
    yi: "וואלפגאנג אמאדעוס מאצארט",
    zh: "沃尔夫冈·阿马德乌斯·莫扎特"
  },
  altLabel: {
    en: "Mozart, Johann Chrysostom Wolfgang Amadeus",
    "en.en": "Mozart, Wolfgang Amadeus"
  },
  biographicalInformation: [...],
  dateOfBirth: [
    "1756-01-27"
  ],
  dateOfDeath: [
    "1791-12-05"
  ],
  end: [
    "1791-12-05"
  ],
  identifier: [
    "32197206"
  ],
  placeOfBirth: [
    {
      "id": http://dbpedia.org/resource/Salzburg
    },
    {
      "id": http://dbpedia.org/resource/Austria
    }
  ],
  placeOfDeath: [
    {
      "id": http://dbpedia.org/resource/Salzburg
    },
    {
      "id": http://dbpedia.org/resource/Vienna
    }
  ],
  sameAs: [
    http://sw.dbpedia.org/resource/Wolfgang\_Amadeus\_Mozart,
    http://ro.dbpedia.org/resource/Wolfgang\_Amadeus\_Mozart,
    http://bg.dbpedia.org/resource/Волфганг\_Амадеус\_Моцарт,
    http://nds.dbpedia.org/resource/Wolfgang\_Amadeus\_Mozart,
    http://pa.dbpedia.org/resource/ਵੋਲਫਗਾਂਗ\_ਅਮਾਦੇਅਸ\_ਮੋਜ਼ਾਰਟ,
    http://pc.dbpedia.org/resource/Wolfgang\_Amadeus\_Mozart,

```

Internal links

External (coreference)
links
(e.g. Wikidata)

Normalizing metadata – Languages

An example of structured data under the hood...

We receive many values in our language metadata field

`fr, français, french...`

We normalize it to ISO code

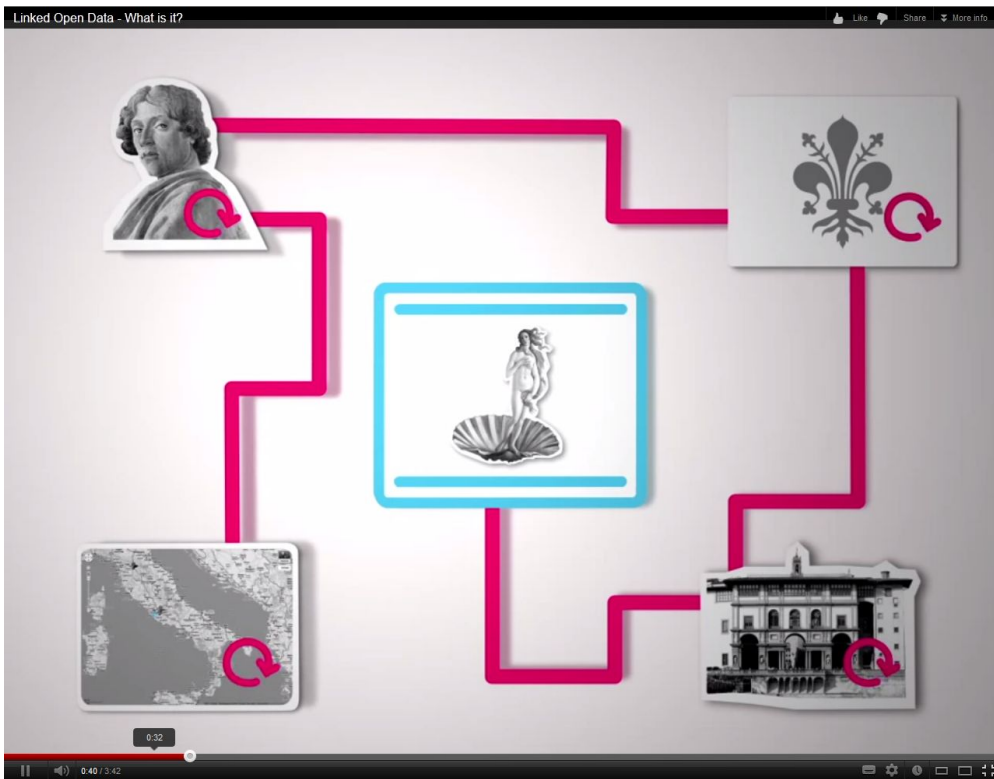
`fr`

Behind the scene, we use EU's [Language Name authority list](#)



Following community best practices and making it known

Linked Open Data Best Practices



Linked Open Data - What is it?

<http://vimeo.com/36752317>

Data on the Web Best Practices

W3C Working Group (2014-2016)

Best practices that were especially relevant for us:

- BP 15: Reuse vocabularies, preferably standardized ones
- BP 16: Choose the right formalization level

<https://www.w3.org/2013/dwbp/>





The acrobatics of metadata

DISTANCE 8 MÈTRES

A community sport

EDM development has involved experts from libraries, archives, museums and academics



*Europeana Assembly General Meeting,
Rijksmuseum, Amsterdam, 2015*

EUROPEANATECH COMMUNITY

Who we are

EuropeanaTech is the community of experts, developers, and researchers from the R&D sector within the greater Europeana Network Association.

What we do

We coordinate research and development endeavors that seek to improve the standing of European digital cultural heritage and facilitate knowledge across the sector at large.

Why join?

- Network and collaborate with multidisciplinary technical professionals from around the world in Europeana's largest digital cultural heritage community

<https://pro.europeana.eu/europeana-tech>

Lessons worth sharing (again)

All this requires a lot of researching and talking

Having best practices is useful, yes!

But access to resources (vocabularies) is also crucial

Flexibility is required:

- We sometimes changed definitions to avoid semantic overcommitment
- Re-using is easier when one has a cool-head approach to formal semantics

ENDORSE.

16-19 MARCH 2021

THE EUROPEAN DATA CONFERENCE ON REFERENCE DATA AND SEMANTICS

Thank you!



europeana



europeana.eu



@EuropeanaEU



Co-financed by the Connecting Europe
Facility of the European Union