



Insects and Fruit
Jan van Kessel
1660 - 1665, Rijksmuseum
Netherlands, Public Domain

Europeanana

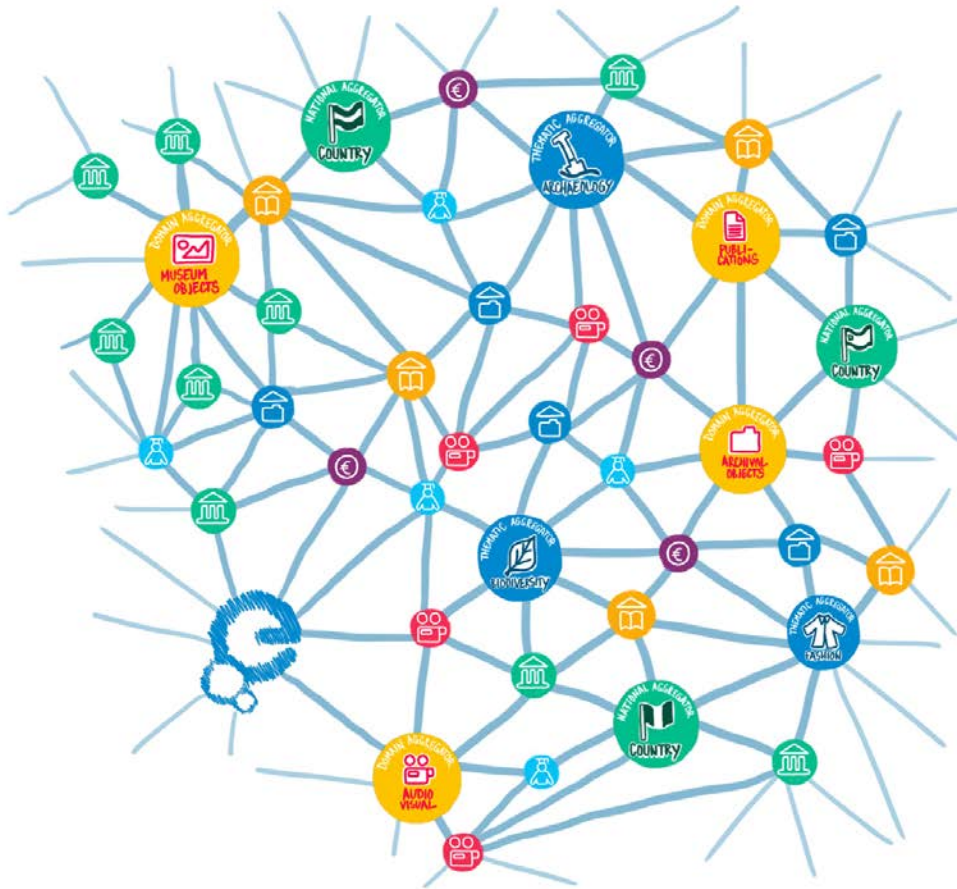
Entitifying Europeanana: Building an ecosystem of networked references for Cultural Objects

Hugo Manguinhas, Valentine Charles, Antoine Isaac, Tim Hill
Europeanana Foundation

J.V. KESSEL

What is Europeana?

The Platform for Europe's Digital Cultural Heritage



We aggregate metadata:

- *From all EU countries*
- *~3,500 galleries, libraries, archives and museums*
- *More than 53M objects*
- *In about 50 languages*
- *Huge amount of references to places, agents, concepts, time*

[Europeana aggregation infrastructure](#)

Europeana | CC BY-SA

DENKSCHETS.NL

Europeana Linked Data Strategy

Our efforts and lines of work

- Europeana Data Model (EDM) offers a base for linking data
- We apply automatic enrichment to link source data to reference data
- We encourage data providers to contribute their own vocabularies so that we can benefit from data links made at data providers' level
- We encourage alignment activities between domain vocabularies

Significant progress have been made, most of it presented in past SWIB!

Europeana Linked Data Strategy

A strategy for Entities

As a cornerstone for our strategy we are building an "Entity Collection"

- A service that acts as a centralized point of reference and access to data about contextual entities
- Caching and curating data from the wider Linked Open Data cloud
- A sort of Europeana "knowledge graph"

Europeana Linked Data Strategy

Motivation

- *Improve user experience*
 - *Support better ways of searching and navigating through the collections, eliminating ambiguity and clarifying the meaning of descriptions*
 - *Adapt better to the language of the user*
- *by improving the interlinking of data*
 - *Brings more context to the objects*
 - *Alleviates polysemy issues*
 - *Expands language coverage*
 - *Contributes to build a web of data ('knowledge graph') that third parties can use to improve their users' experience*

The Entity Collection

Use Cases

Enrichment of Provider's Data

- A controlled vocabulary to help identify named references to entities

Crowdsourcing

- Objects can be annotated with references to entities
- A controlled vocabulary for client applications

Entity Collection

Europeana Collections Portal

- *Findability*: users can look for entities, not only records (Entity-Based Search)
- *Understandability*: Entity Pages group and present all assertions about an entity
- *Exploration*: Navigation along relationships becomes possible

Republication for Re-use

- Entities can be republished as an open source to the community

The Entity Collection

What can it enable?

Food and Drink

- + Agriculture 25266
- + Beverages 14856
- + Cuisine 26796
- + Eating behaviors 27721
- + Food and drink by country 10858
- + Food and drink preparation 28386
- + Food and drink terminology 10581
- + Food and the environment 1027
- + Food culture 27553
- + Food decorations 2166
- + Food festivals 6

Entity based facets

- + 10861432 1568
- + Africa 3776
- + Americas 92
- + Antarctica 4
- + Arabian Peninsula 33
- + Asia 2074
- + Atlantic Ocean 45
- + Australia (continent) 3
- + Bering Sea 1
- + British Empire 12
- + Community of Latin American and Caribbean



Semantic auto-completion

Search Leo

- Leonardo da Vinci
- Leonard Cohen
- Leo Tolstoy

Semantic and Metadata annotations

Create new annotation

Brass Instrument

My Items are only available to logged users. Please log in and use this section to store and use items.

Technical description: Brass; ligature fitting on bell section at joint; stockings on main slides. Bell with one coil, angled to face forwards. Repair

Entity Pages

Leonardo da Vinci

Painter

Leonardo di ser Piero da Vinci was an Italian Renaissance polymath: painter, sculptor, architect, musician, mathematician, engineer, inventor, anatomist, geologist, cartographer, botanist, and writer. [Wikipedia](#)

Born: April 15, 1452, Vinci, Italy

Died: May 2, 1519, Amboise, France

Full name: Leonardo di ser Piero da Vinci

Period: High Renaissance

Parents: Piero Fruosino di Antonio da Vinci, Caterina da Vinci

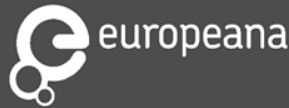
Artwork

- Mona Lisa 1517
- Virgin of the Rocks 1486
- Lady with an Ermine 1490
- Ginevra de' Benci 1478
- The Baptism of Christ 1475

Food & Drink Project

Pundit Annotation Client

Google Knowledge Card



The Entity Collection

How do we choose our target vocabularies?

As defined in the recent [Europeana Tech Task Force on enrichment and evaluation](#) ([presented](#) last year), we consider the following criteria when selecting a vocabulary:

- *Properly documented and supported by a community*
- *Technically available on the web according to the Linked Data best practices and recipes*
- *Available under an open licence*
- *Multilingual*
- *Abide to a minimal ontological commitment principle*
- *Apply the best practices and standards for the representation, structure and description of vocabularies*
- *Well-connected internally and externally to other vocabularies (preferably spine vocabularies)*

The Entity Collection

Which target vocabularies are we using?

For historical reasons, the target vocabularies correspond to the ones being used for [Semantic Enrichment](#) (as of November 2016):

214,307
resources

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

165,008
resources

- *Agents*
*a subset of **DBpedia** corresponding to most of the instances of `dbp:Artist` with some exceptions, and integrated from [49 DBpedia language editions](#).*

274
resources

- *Concepts*
*a subset of **DBpedia** corresponding to a handful of concepts matching the needs from **Europeana Collections**.*

2,566
resources

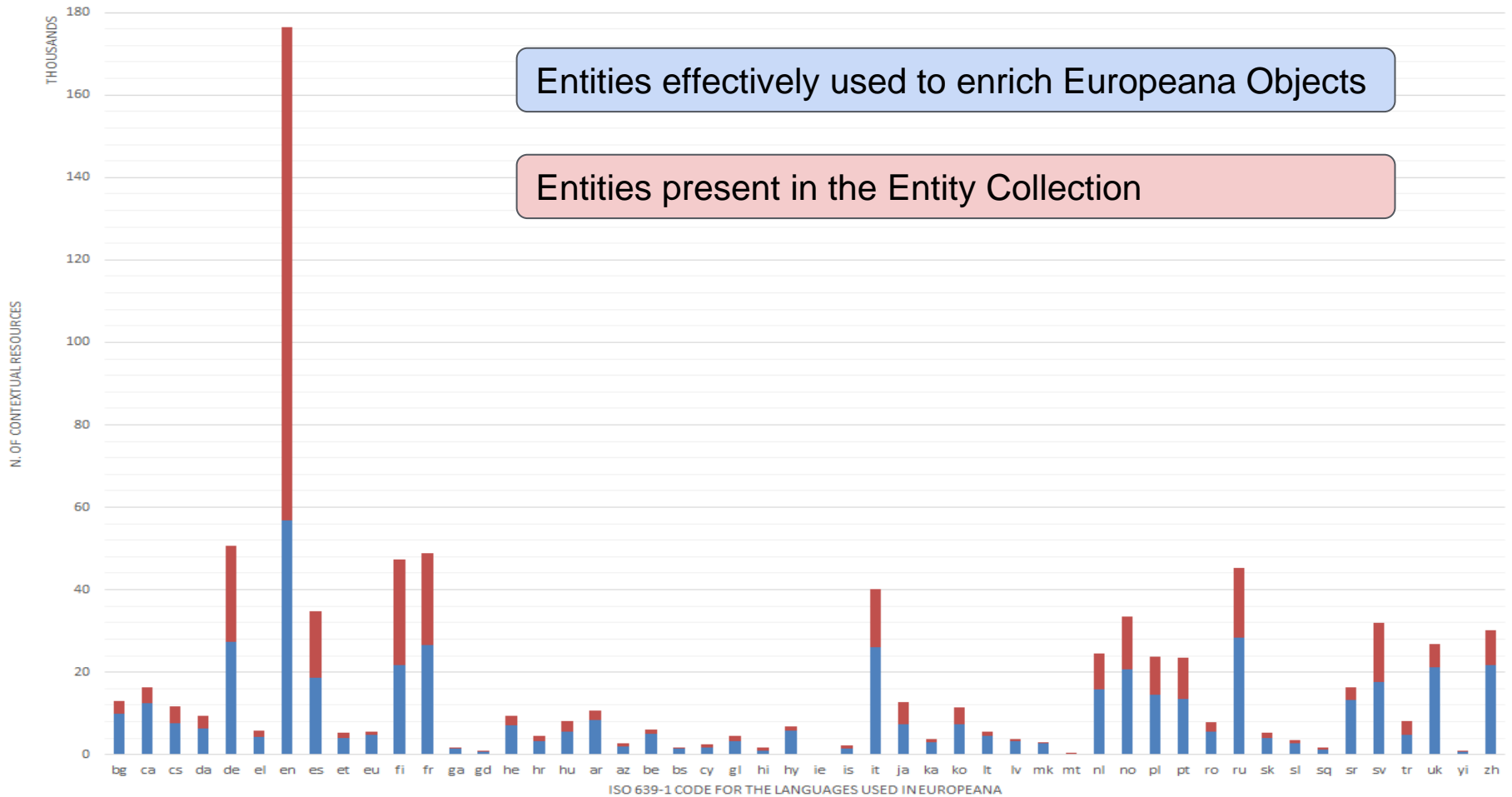
- *Time Spans*
*The chronological periods from **Semium Time**.*

The Entity Collection

Contribution to multilingual coverage

Europeana contextual resources with labels in a given language

■ Entities effectively used to semantically enrich Europeana objects ■ Entities present in the Entity Collection



Entities effectively used to enrich Europeana Objects

Entities present in the Entity Collection

The Entity Collection

Are these target vocabularies enough?

- *Not enough coreferencing information to other vocabularies*
 - *particularly to the ones we receive from data providers (e.g. musical instruments, MIMO)*
- *Labels and values are not always accurate and normalized*
 - *need for better reference data (e.g. VIAF)*
- *Missing relevant information*
 - *e.g. roles and professions*
- *Need to expand coverage to other types of entities*
 - *namely Works and Events*

The Entity Collection

Challenges

Investigate and design strategies for:

- *Integrating new vocabularies that can further improve*
 - *entity descriptions and multilingual coverage (e.g. VIAF)*
 - *linking between entities (e.g. Wikidata)*
- *Integrating alignments, in particular:*
 - *links between local/domain vocabularies to pivot vocabularies*
- *Supporting manual curation of existing and new entities*
- *Keeping up-to-date the information collected from external sources*

The Entity Collection

Our roadmap for the next years

- ✓ › [Mint Europeana URIs for Entities](#) and update internal references
- ✓ › *Make entity services and data available via an API*
 - *Make use of the API in the Collections Portal*
 - *Implement support for new vocabularies and entity types*

The Entity Collection

Alpha release of our new Entity API

Entity Collection - REST API

This Swagger API console provides an overview of and interface to the Entity Collection REST API. For more help and information, head to our comprehensive [online documentation](#).

[Terms of use](#)

Entity retrieval

Show/Hide | List Operations | Expand Operations

GET /entity/{type}/{namespace}/{identifier} Retrieve a known entity

GET /entity/{type}/{namespace}/{identifier}.jsonld Retrieve a known entity

Discovery API Search Controller

Show/Hide | List Operations | Expand Operations

GET /entity/suggest Suggest entities for the given text query. Suported values for type: Agent, Place, Concept, Timespan, All

GET /entity/suggest.jsonld Suggest entities for the given text query. Suported values for type: Agent, Place, Concept, Timespan, All

[BASE URL: / , API VERSION: 0.0.1]

VALID {...}

More methods will come, for:

Creation, Update and Delete; URI resolution to Europeana Entities

The Entity Collection

DBpedia resource for "Mozart" in our data

```

{
  "@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 Amadey 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 Αμαντιέους 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: "בֶּרְנַרְדְּטְ מוֹצָרְט",
    hi: "वोल्फगांक आमडेयुस मोसार्त",
    hr: "Wolfgang Amadeus Mozart",
    hu: "Wolfgang Amadeus Mozart",
    hy: "Վոլֆգանգ Ամադեյուս 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: "Volfgangas Amadėjus Mocartas",
    lv: "Volfgangs 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 Амадей",
    sk: "Wolfgang Amadeus Mozart",
    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: "בֶּרְנַרְדְּטְ מוֹצָרְט",
    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/Paris
    },
    {
      "@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://oc.dbpedia.org/resource/Wolfgang Amadeus Mozart,
  ]
}

```

Preferred labels for 48 languages

Inter-linking information... still need to switch references to link to Europeana Entities

Coreference links to 6 other datasets (e.g. Freebase, Wikidata)

The Entity Collection

Entity API - suggest method

/entity/suggest.json?text=neo&lang=en&rows=6

The screenshot shows a search interface with a search bar containing 'neo'. Below the search bar, there is a list of search results. Each result includes a title, a date of birth (if applicable), a date of death (if applicable), and a type icon. The results are:

Entity	Type
Neolithic	CONCEPT
Neoclassicism	CONCEPT
Neofit Bozvelievo	PLACES
Neo Rauch	PEOPLE
Neon Park	PEOPLE
Neofit Rilski	PLACES

```
{
  "@context": [
    https://www.w3.org/ns/ldp.jsonld,
    http://www.europeana.eu/schemas/context/entity.jsonld,
  ],
  "@language": "en"
},
  "contains": [
    {
      id: http://data.europeana.eu/concept/base/211,
      prefLabel: "Neolithic",
      type: "Concept"
    },
    {
      id: http://data.europeana.eu/concept/base/214,
      prefLabel: "Neoclassicism",
      type: "Concept"
    },
    {
      id: http://data.europeana.eu/place/base/133388,
      isPartOf: [
        {
          id: http://data.europeana.eu/place/base/131184,
          prefLabel: "Obshtina Momchilgrad"
        },
        {
          id: http://data.europeana.eu/place/base/76,
          prefLabel: "Republic of Bulgaria"
        }
      ],
      prefLabel: "Neofit Bozvelievo",
      type: "Place"
    },
    {
      dateOfBirth: "1960-04-18",
      id: http://data.europeana.eu/agent/base/73775,
      prefLabel: "Neo Rauch",
      type: "Agent"
    },
    {
      dateOfBirth: "1940-12-28",
      dateOfDeath: "1993-09-01",
      id: http://data.europeana.eu/agent/base/72702,
      prefLabel: "Neon Park",
      type: "Agent"
    },
    {
      id: http://data.europeana.eu/place/base/132103,
      isPartOf: [
        {
          id: http://data.europeana.eu/place/base/131003,
          prefLabel: "Obshtina Vetrino"
        },
        {
          id: http://data.europeana.eu/place/base/76,
          prefLabel: "Republic of Bulgaria"
        }
      ],
      prefLabel: "Neofit Rilski",
      type: "Place"
    }
  ],
  totalItems: 6,
  type: "BasicContainer"
}
```


Conclusion

- *A Strategy for Entities is a “must” for Europeana*
- *There is no “one fits all” vocabulary*
- *We have a long way to go...*
...but we are making progress

Flowers
Anonymous
1700-1799, Rijksmuseum
Netherlands, Public Domain

Thank you!



europaena

#AllezCulture

hugo.manguinhas@europeana.eu