



europæana  
food and drink

**Grant Agreement 621023**

## ***Europeana Food and Drink***

# **Semantic Demonstrator Extended**

<b>Deliverable number</b>	<i>D3.20d</i>
<b>Dissemination level</b>	<i>CO</i>
<b>Delivery date</b>	<i>20 July 2016</i>
<b>Status</b>	<i>Final</i>
<b>Author(s)</b>	<i>Vladimir Alexiev (ONTO) Andrey Tagarev (ONTO) Laura Tolosi (ONTO)</i>



This project is funded by the European Commission under the  
*ICT Policy Support Programme* part of the  
*Competitiveness and Innovation Framework Programme*.

## Abstract

This document describes the additional development on the EFD Semantic Demonstrator performed after the official D3.20 deliverable (M22). It describes work performed between 31 October 2015 and 20 July 2016 (M31), the achieved results, the created data and enrichments, and the extended application functionality. It is an addition to the D3.20 deliverable, and therefore should be read in addition to it.

## Revision History

Rev	Date	Author	Org	Description
v0.1	10/7/2016	Vladimir Alexiev,	ONTO	Initial version
V0.2	15/7/2016	Andrey Tagarev	ONTO	Add semantic enrichment service
v0.3	20/7/2015	Laura Tolosi	ONTO	Add comparison by language

**Statement of originality:**

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

## Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	Additional Work	5
1.2	Abbreviations	5
<b>2</b>	<b>Additional Collections and Objects</b>	<b>6</b>
2.1	Examples From the Collections	7
<b>3</b>	<b>Semantic Enrichment in French</b>	<b>14</b>
3.1	Language Selection	14
3.2	French Enrichment	14
3.3	Enrichment Evaluation	14
<b>4</b>	<b>Extend the FD Classification</b>	<b>17</b>
4.1	Comparison of FD Trees Across Languages	17
4.2	French FD Categories	18
4.3	FD Classification as SKOS	19
<b>5</b>	<b>Geographic Mapping</b>	<b>20</b>
5.1	Hierarchical Place Processing	20
5.2	Coordinate Processing	20
5.3	Geographic Mapping in the Semapp	23
<b>6</b>	<b>Semantic Enrichment Web Service</b>	<b>24</b>
6.1	Service Input	24
6.2	Output in JSON-LD	24
6.3	Output in Turtle	25
<b>7</b>	<b>Other Sem App Enhancements</b>	<b>26</b>
<b>8</b>	<b>Dissemination</b>	<b>27</b>
8.1	Engaging Bulgarian Community	27
8.2	Sem App Access Statistics	27
8.3	Further Participation in Working Groups	32
<b>9</b>	<b>Sustainability</b>	<b>33</b>
9.1	Food and Drink Classification	33
9.2	Semantic Enrichment Service	33
9.3	Semantic Demonstrator	34
<b>10</b>	<b>References</b>	<b>35</b>

## Figures

Figure 1	BG-ONTO: Bulgarian Recipes	8
Figure 2	IT-ALI: Old Photos	8
Figure 3	UK-HM: Ethnographic Collection	9
Figure 4	BE-MRAC: Ethnographic Collection, Old Photos	9
Figure 5	BE-CAG: Mostly Modern Food-related Artefacts	10
Figure 6	UK-HP: London Pubs	10
Figure 7	IE-LGMA: Irish Plants and Foods	11
Figure 8	UK-TOP: Mix of Old Photos and Modern Food & Drink Photos	11
Figure 9	UK-CT: Victoria & Albert Food-related Artefacts	12
Figure 10	LT-VUFC: Old Lithuanian Recipes	12
Figure 11	UK-WAM: Various Wolverhampton Artefacts	13

Figure 12 Raw English FD Categories per Level.....	17
Figure 13 Comparing 4 Languages to English.....	17
Figure 14 Comparing English to 4 Languages.....	18
Figure 15 EFD Semapp Geographic Clusters .....	20
Figure 16 EFD Map Showing Individual Object .....	21
Figure 17 Zooming Into Jittered Coordinates.....	22
Figure 18 Zoomed Into Individual Object.....	22
Figure 19 “View on Map” Button .....	23
Figure 20 EFD Data Documentation.....	26
Figure 21 EFD Ontology Documentation .....	26
Figure 22 Monthly History of Visits .....	28
Figure 23 Country Distribution of Visitors, 2015 .....	28
Figure 24 Country Distribution of Visitors, 2016 .....	29

## Tables

Table 1 Semapp Collections as of Oct 2015.....	6
Table 2 Semapp Collections as of Jul 2016.....	6
Table 3 Enrichment Evaluation Results .....	15
Table 4 French FD: Manual to Automatic Comparison.....	15
Table 5 French Places: Manual to Automatic Comparison .....	16
Table 6 Number of Visitors and Visits.....	27
Table 7 City Distribution of Visitors, 2015 .....	29
Table 8 City Distribution of Visitors, 2016 .....	30

## 1 Introduction

This document describes the additional development on the EFD Semantic Demonstrator (EFD semapp) performed after the official D3.20 deliverable (M22). It describes work performed between 31 October 2015 and 20 July 2016 (M31), the achieved results, the created data and enrichments, and the extended application functionality. It is an addition to the D3.20 deliverable, and therefore should be read in addition to it. Note: periodic progress reports were also submitted: D3.20a (at M18) and D3.20b (at M21).

### 1.1 Additional Work

The following additional work was performed and is described in detail:

- Add extra collections and additional objects in existing collections
- Extend semantic enrichment to French (in addition to English)
- Extend the FD Classification to French, and further elaboration of the FD classification through bottom-up augmentation
- Add a geographic map in addition to the hierarchical browsing by Place
- Establish a semantic enrichment web service to suggest automatic enrichments (provide semantically enriched content) that is used by the Crowdsourcing Enrichment application (developed by D3.5 Technical Demonstrator and T5.2 Community/crowdsourcing platform)
- Further participation in Europeana and DBpedia working groups
- Ensure sustainability of the Classification and Sem app

Note: the task to discover additional Europeana CHOs related to FD was undertaken by NTUA. We helped them using the FD Classification for this task, and with some technical problems related to crawling CHOs from the Europeana portal.

### 1.2 Abbreviations

Abbrev	Description
API	Application Programming Interface
CH	Cultural Heritage
EDM	Europeana Data Model
EFD	Europeana Food and Drink
FD	Food and Drink
JSON	JavaScript Object Notation
RDF	Resource Description Framework, the semantic data format
SPARQL	SPARQL Protocol and RDF Query Language, the semantic query language
UI	User Interface
URL	Uniform Resource Locator
UTF-8	The most commonly used Unicode Transformation Format

## 2 Additional Collections and Objects

The initial release of the semapp (Oct 2015) included only English-language objects from the following collections:

**Table 1 Semapp Collections as of Oct 2015**

Collection	Obj
BG-ONTO	9071
IT-Alinari	498
UK-Horniman	4352
UK-Wolverhampton	439
UK-TopFoto	1814
<b>Total</b>	<b>16174</b>

Our first extension task was to add extra collections and additional objects in existing collections. Since all collections had completed their data collection and conversion to EDM, in most cases it was easy for us to get the data from MINT and apply automatic English enrichment.

**Table 2 Semapp Collections as of Jul 2016**

Provider	Name	Lang	Objects	FD	Places	FD/Obj	PI/Obj
BE-CAG	Centrum Agrarische Geschiedenis	EN	999	899	540	0.90	0.54
BE-MRAC	Musée royal de l'Afrique centrale	FR	7500	14959	16843	1.99	2.24
BG-ONTO	Ontotext & Bulgariana	BG	9071	119142	9071	13.13	1.00
IE-LGMA	Local Government Management Agency	EN	2000	4910	961	2.46	0.48
IT-ALI	Fratelli Alinari	EN	498	88	526	0.18	1.06
LT-VUFC	Vilnius University Faculty of Communication	LT,EN	1007	3984	321	3.96	0.32
UK-CT	Collections Trust & Victoria and Albert Museum	EN	6502	15946	4141	2.45	0.64
UK-HM	Horniman Museum and Gardens	EN	4351	19980	10834	4.59	2.49
UK-HP	HistoryPin (Shift) & National Brewery Heritage Trust	EN	3416	10513	6017	3.08	1.76
UK-TOP	TopFoto & thePictureKitchen	EN	7838	39698	6709	5.06	0.86
UK-WAM	Wolverhampton Arts and Museums	EN	503	347	101	0.69	0.20
	<b>Total</b>		<b>43685</b>	<b>230466</b>	<b>56064</b>	<b>5.28</b>	<b>1.28</b>

The increase is from 16k to 43k objects, or 2.7x.

Specific notes on some collections:

- BE-MRAC is an important ethnographic collection described in FR. To enrich it, we extended the enrichment service with French, which involved significant work (see next section).

- For BG-ONTO we developed an enrichment based on regular expression (regex) processing. We gathered the 328 most popular words/phrases and matched them to en.dbpedia manually, e.g.

агне\w*	- Lamb and mutton
айвар	- Ajvar
айр\wn	- Ayran
алкохол\w*	- Alcoholic beverage
ашуре	- Ashure
баклава	- Baklava

- (w\* means “any number of word-chars, so e.g. the first line matches агне, агнешко, агнешка, etc).
- This produced a large number of specific enrichments. Each object is also associated with one place (Bulgaria).
- LT-VUFC use a bilingual thesaurus (LT-EN), so we leveraged the EN keywords to apply the EN enrichment
- UK-CT is a V&A Museum collection. We obtained the metadata from Europeana. Since that dataset uses the “EDM External” schema (unlike MINT that uses “EDM Internal”), we had to modify the enrichment pipeline and the semapp to cater to this slightly different data model.
- UK-HP is a new collection, the images of London pubs used in the EFD Book publication
- UK-TOP has added objects from an external partner provider, thePictureKitchen.
- Wolverhampton is actually 2 providers: Arts and Museums and Archives and Local Studies

## 2.1 Examples From the Collections

The following shots show some examples from the different collections. Unfortunately some images are missing from Europeana Cloud, e.g.

- 31% from BE-MRAC
- All thumbnails of BE-CAG. Therefore we use image URLs pointing to NTUA MINT, e.g.
  - [http://foodanddrink.image.ntua.gr/image/CAG/00009803\\_1.JPG](http://foodanddrink.image.ntua.gr/image/CAG/00009803_1.JPG) instead of
  - [https://cloud.europeana.eu/api/data-providers/CAG/records/00009803\\_1.JPG/representations/presentation/CAG/00009803\\_1.JPG](https://cloud.europeana.eu/api/data-providers/CAG/records/00009803_1.JPG/representations/presentation/CAG/00009803_1.JPG)



**Картофени кюфтета на фурна**

Варим картофите, белим ги и ги намачкваме в голям дълбок съд. Добавяме яйцето, предварително задушените



**Градинска мусака**

Задушаваме ориза в олиото заедно с нарязания на ситно лук. Наливаме 1-та чаша прясно мляко, объркваме сместа и я прехвърляме я



**Картофена супа с месо**

Месото се нарязва на неголеми порционни парчета и се вари за около 40 минути, посолва се. Картофите се измиват, обелват се и се



**Салата с маринувани домати**

В съд на котлона кипнете оцета, зехтина, със захарта, солта и пасирания чесън. Полейте нарязаните малки



**Телешки гувеч по манастирски**

Обезкостеното, измито и почистено месо се нарязва на парчета и се задушва в сгорещеното олио и малко



**Баница с кайма и праз**

В 3-4 с.л. олио се запържва ситно нарязаният праз. Добавя се каймата и след като и тя се задуши - черен пипер и сол на вкус. Намазва се



**Пилешка чорба по селски**

Порционирайте пилешкото и сложете в тенджера заедно с 3 л студена вода, малко сол. Добавете дреболиите,



**Салата Айсберг плюс**

Зеленчуците се измиват и отцеждат. Салатата се накъсва на парчета, а лукът, краставицата, доматиите и яйцата се нарязват на

**Figure 1 BG-ONTO: Bulgarian Recipes**



**View of Stia**

WARNING: Permission must be required for non editorial use. Please contact Alinari Archives;



**Water vender's kiosk in Naples**

CREATIVE USE: allowed;



**A lemon from the Alinari Garden in Florence**

CREATIVE USE: allowed;



**Pea plant, Cascine Park, Florence**

WARNING: Permission must be required for non editorial use. Please contact Alinari Archives;



**A lovely prickly pear cactus in Taormina**

WARNING: Permission must be required for non editorial use.



**View of the Gate and part of the city wall of Jesi. A farmer working in a garden in is in the foreground.**



**Farmhouse in San Manno in the environs of Perugia**

WARNING: Permission must be required for non editorial use.

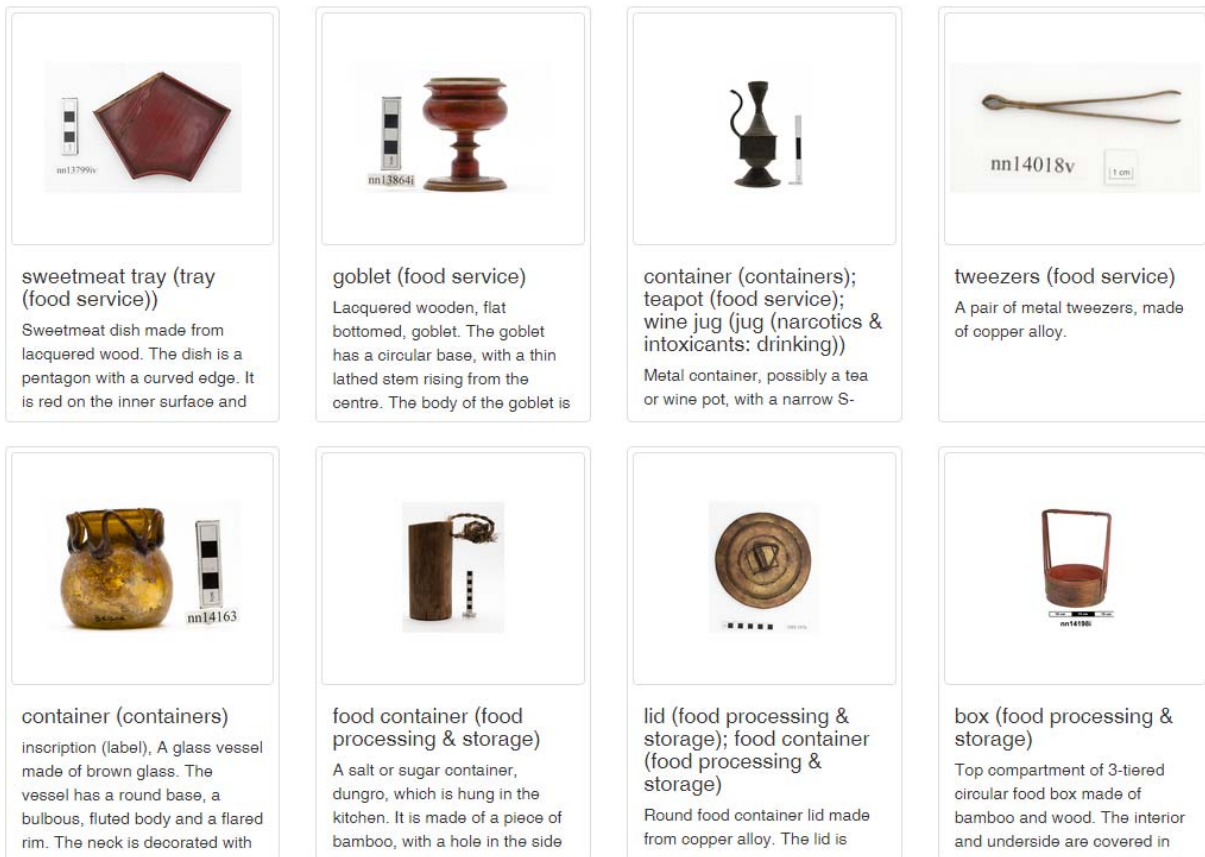


**Pair of lovers eating breakfast: work preserved in the Gemaldegalerie of Dresden.**

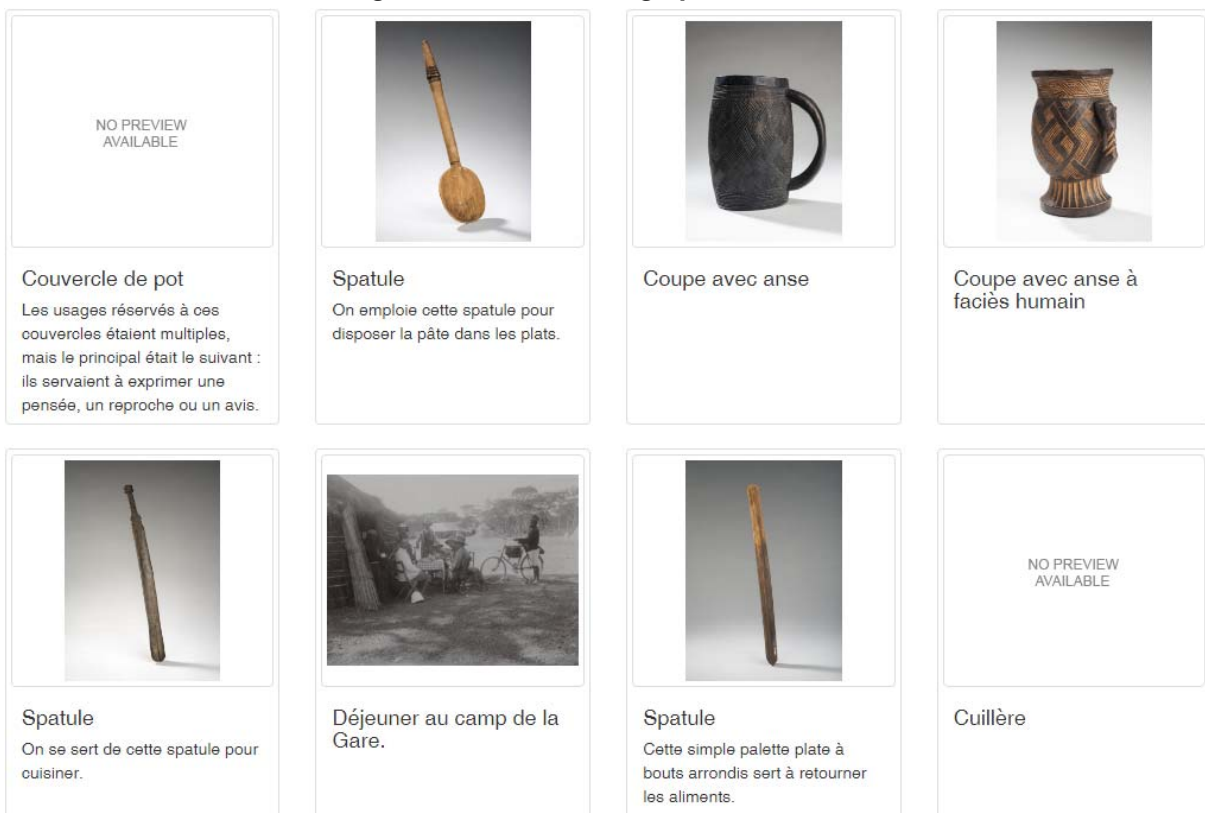
**Figure 2 IT-ALI: Old Photos**



### D3.20d Semantic Demonstrator Extension



**Figure 3 UK-HM: Ethnographic Collection**



**Figure 4 BE-MRAC: Ethnographic Collection, Old Photos**









D3.20d Semantic Demonstrator Extension

<p><b>Recipient to pour beer</b> Measuring jug made of pewter, used to pour half a liter of beer. The pitcher is marked for this exact quantity with hallmarks, placed on the edge.</p>	<p><b>Verpakking Le Pastourel Dubbel Creme</b> Dit is een verpakking van Le Pastourel Dubbel Creme. Ze bevat 45 tot 52 procent</p>	<p><b>Thuis diepvriezen</b> Vanaf de jaren '50 verschijnt de huisdiepvries eerst op het platteland. De stedelijke gebruikers volgen pas in de jaren '70.</p>	<p><b>Vruchtensap Libby's</b> Reclame voor (en consumptie van) groentensappen is veel zeldzamer dan pakweg die voor soep in blik of zakjes. Deze advertentie komt uit het</p>
<p><b>Recipient to pour beer</b> Measuring jug with ear and spout made from stanniferous faience and pewter lid, used to pour beer. On the front of the jug is a star shaped cartouche</p>	<p><b>Boterlepel</b> Metalen voorwerp met plastic lemmet om harde boter te schrapen.</p>	<p><b>Voorflap boek Boomteelt</b> Auteurs als Van Hulle bouwen voort op de vergaarde kennis van de "pomologen" uit de achttiende en negentiende</p>	<p><b>Handle to operate a beer tap</b> Faience handle for an English tap system. The handle was placed over a copper bar and fixed with a screwcap. The</p>









**Figure 5 BE-CAG: Mostly Modern Food-related Artefacts**

<p><b>Portsmouth Arms, Pentonville</b> 139/143 Pentonville Road, London This photo was taken by the Charrington Brewery during an architectural survey of their</p>	<p><b>Crown, Billericay</b> High Street, Billericay This photo was taken by the Charrington Brewery during an architectural survey of their pubs. These photographs, now part of the</p>	<p><b>Rabbits, West Thurrock</b> West Thurrock This photo was taken by the Charrington Brewery during an architectural survey of their pubs. These photographs, now part of the</p>	<p><b>Brewer's Arms, Burgess Hill</b> Burgess Hill This photo was taken by the Charrington Brewery during an architectural survey of their pubs. These</p>
<p><b>All Hallows Club</b> All Hallows on Sea This photo was taken by the Charrington Brewery during an architectural survey of their pubs. These photographs, now part of the</p>	<p><b>Gibraltar</b> Gibraltar, Gibraltar Walk, Bethnal Green This photo was taken by the Charrington Brewery during an architectural survey of their pubs. These</p>	<p><b>Man of Kent, Sydenham</b> 173 Sydenham Road, SE26 This photo was taken by the Charrington Brewery during an architectural survey of their pubs. These photographs, now</p>	<p><b>King's Arms, Fulham</b> 425 New Kings Road/High Street, Fulham, SW6 4RN This photo was taken by the Charrington Brewery during an architectural survey of their</p>

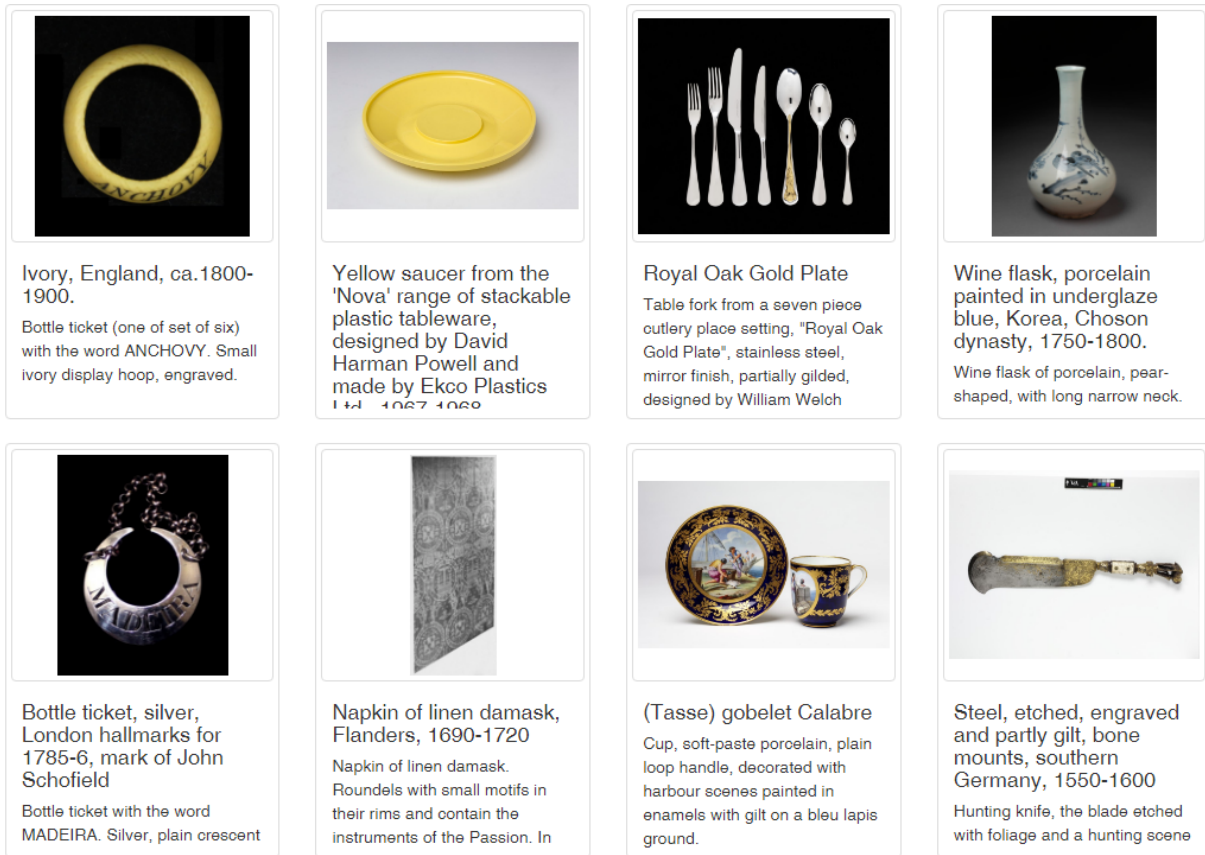
**Figure 6 UK-HP: London Pubs**

 <p>stream and plants</p>	 <p>Wild Mushrooms There are an abundance of wild mushrooms to be found throughout the forests of Ireland.</p>	 <p>beetroot in colourful variety Wild beets are native on Irish and European shores and come in several different colour variations.</p>	 <p>Lucky Bag Pie Ingredients The Lucky Bag Pie is a very versatile dinner because it uses whatever is available as ingredients. Irish mums love</p>
 <p>halloween treats Children (and adults) enjoy the many kinds of halloween treats available.</p>	 <p>Fallen apples<sup>5</sup> Kingfishers spear and slice the soft flesh of the fallen apple, an important part of many animals' diet.</p>	 <p>Cliffs of Moher Cake The Cliffs of Moher are an Irish landmark. This cake is a fusion of a natural Irish Heritage Site and Irish cooking.</p>	 <p>Cabbage Leafy Cabbage is a traditional Irish vegetable that can be grown for the majority of the year.</p>

**Figure 7 IE-LGMA: Irish Plants and Foods**

 <p>Detail of country hedge with blackberries and view across to Ashdown Forest, Sussex, England credit: Marie-Louise Avery / thePictureKitchen / TopFoto</p>	 <p>Mobile ice cream car Mobile ice cream car</p>	 <p>Glass half full of lager in sunshine at seaside cafe credit: Marie-Louise Avery / thePictureKitchen / TopFoto Glass half full of lager in</p>	 <p>Quadragesimal 1521 © TopFoto Quadragesimal 1521 © TopFoto</p>
 <p>RITA FROGGARTT BARMAID/ ; 17 MAY 1962 RITA FROGGARTT BARMAID/ ; 17 MAY 1962</p>	 <p>Detail of country hedge with blackberries and view across to Ashdown Forest, Sussex, England credit: Marie-Louise Avery / thePictureKitchen / TopFoto</p>	 <p>White lilies in tall vase in pale interior with blurred girl on sofa credit: Marie-Louise Avery / thePictureKitchen / TopFoto</p>	 <p>Libyan Market - man drinking Libyan Market - man drinking</p>

**Figure 8 UK-TOP: Mix of Old Photos and Modern Food & Drink Photos**



Ivory, England, ca.1800-1900.  
Bottle ticket (one of set of six) with the word ANCHOVY. Small ivory display hoop, engraved.

Yellow saucer from the 'Nova' range of stackable plastic tableware, designed by David Harman Powell and made by Ekco Plastics Ltd. 1967-1968

Royal Oak Gold Plate  
Table fork from a seven piece cutlery place setting, "Royal Oak Gold Plate", stainless steel, mirror finish, partially gilded, designed by William Welch

Wine flask, porcelain painted in underglaze blue, Korea, Choson dynasty, 1750-1800.  
Wine flask of porcelain, pear-shaped, with long narrow neck.

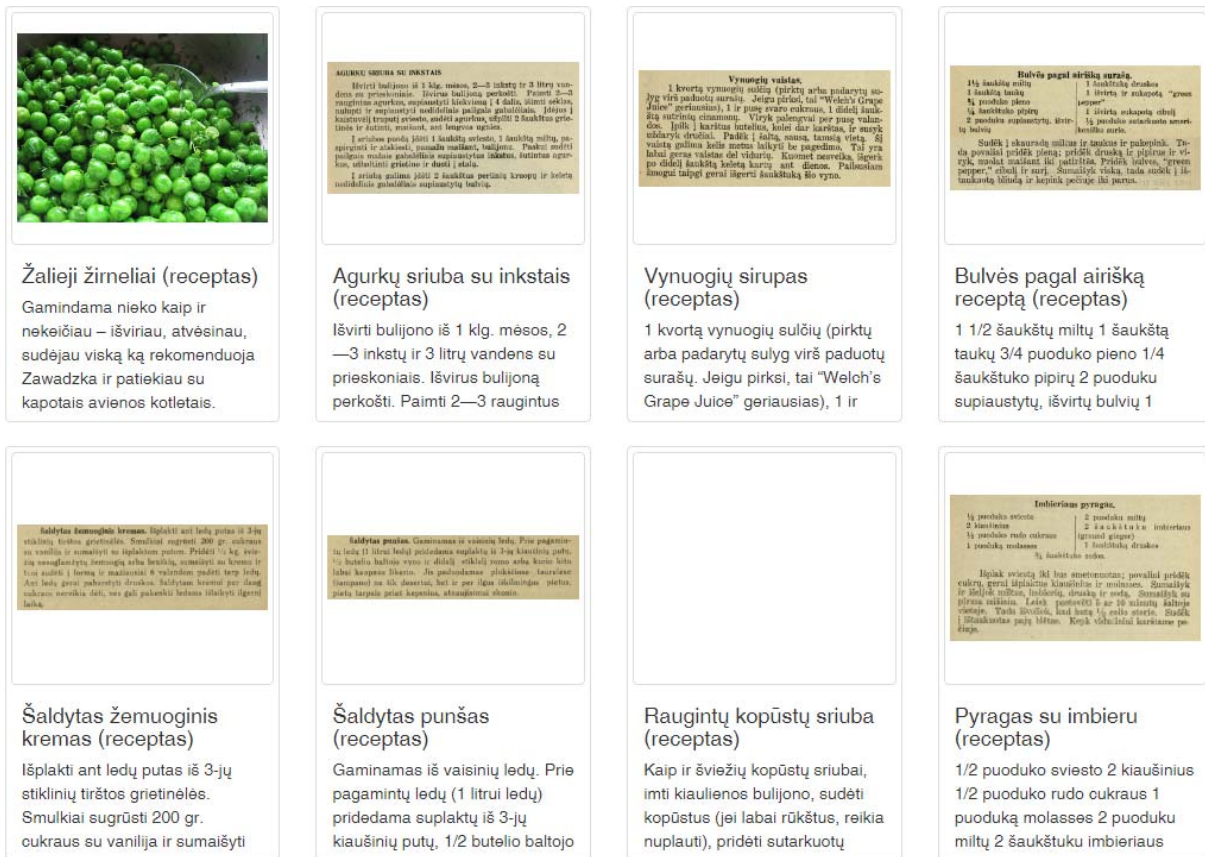
Bottle ticket, silver, London hallmarks for 1785-6, mark of John Schofield  
Bottle ticket with the word MADEIRA. Silver, plain crescent

Napkin of linen damask, Flanders, 1690-1720  
Napkin of linen damask. Roundels with small motifs in their rims and contain the instruments of the Passion. In

(Tasse) gobelet Calabre  
Cup, soft-paste porcelain, plain loop handle, decorated with harbour scenes painted in enamels with gilt on a bleu lapis ground.

Steel, etched, engraved and partly gilt, bone mounts, southern Germany, 1550-1600  
Hunting knife, the blade etched with foliage and a hunting scene

Figure 9 UK-CT: Victoria & Albert Food-related Artefacts



Žalieji žirneliai (receptas)  
Gamindama nieko kaip ir nekeičiau – išviriau, atvėsinau, sudėjau viską ką rekomenduoja Zawadzka ir patiekiau su kapotais avienos kotletais.

Agurkų sriuba su inkstais (receptas)  
Išvirti buljono iš 1 kg. mėsos, 2 —3 inkstų ir 3 litrų vandens su prieskoniais. Išvirus buljoną perkošti. Paimti 2—3 raugintus

Vynuogių sirupas (receptas)  
1 kvortą vynuogių sulčių (pirktų arba padarytų sulyg virš paduotų surašų. Jeigu pirksi, tai "Welch's Grape Juice" geriausias), 1 ir

Bulvės pagal arišką receptą (receptas)  
1 1/2 šaukštų miltų 1 šaukštą taukų 3/4 puoduko pieno 1/4 šaukštuko pipirų 2 puoduku supiautytų, išvirtų bulvių 1

Šaldytas žemuoginis kremas (receptas)  
Išplakti ant ledų putas iš 3-jų stiklinių tirėtos grietinėlės. Smulkiai sugrūsti 200 gr. cukraus su vanilija ir sumaišyti








Šaldytas punšas (receptas)  
Gaminamas iš vaisinių ledų. Prie pagamintų ledų (1 litru ledų) pridėdama suplaktų iš 3-jų kiaušinių putų, 1/2 butelio baltojo

Raugintų kopūstų sriuba (receptas)  
Kaip ir šviežių kopūstų sriubai, imti kiaulienos buljono, sudėti kopūstus (jei labai rūkštus, reikia nuplauti), pridėti sutarkuotų

Pyragas su imbieru (receptas)  
1/2 puoduko sviesto 2 kiaušinius 1/2 puoduko rudo cukraus 1 puoduką molasės 2 puoduku miltų 2 šaukštuku imbieriaus

Figure 10 LT-VUFC: Old Lithuanian Recipes

D3.20d Semantic Demonstrator Extension

			
<p><b>Plate</b> A bread plate entitled 'Old English' by Barratts of Staffordshire.</p>	<p><b>Magic lantern slide</b> Magic lantern slide depicting a work by William Hogarth.</p>	<p><b>Dish</b> An Indian earthenware dish which is partially perforated on the base with concave rim.</p>	<p><b>Snuff Box</b> Decorated circular snuff-box, painted with a portrait of a Georgian gentleman taking punch. Reads 'Punch to a T.' with gilt interior.</p>
		<p>NO PREVIEW AVAILABLE</p>	
<p><b>Bowl</b> Earthenware bowl with basic lined pattern around sides.</p>	<p><b>Tankard</b> German gilt tankard and cover repousse or beaten chased work, Embossed with classical figures, mask, fruit and scrolls. The knob of the top formed by a</p>	<p><b>Vessel, Oinochoe</b> An oinochoe vessel which is undecorated with a bulbous body, flaring lip and round mouth.</p>	<p><b>Vessel, Lekythos</b> LEKYTHOS, black glazed ware decorated in the Attic Red Figure technique.</p>

**Figure 11 UK-WAM: Various Wolverhampton Artefacts**

### 3 Semantic Enrichment in French

#### 3.1 Language Selection

We conducted a poll amongst EFD partners on which second language to add to semantic enrichment (in addition to English). The criteria for evaluation were:

- Number and quality of EFD collections in that language
- Number of CHOs in that language in Europeana
- Size of the respective Wikipedia
- Number of Wikipedia categories starting from the FD root in the respective language
- How many of these categories have sameAs correspondences to English (see sec 4.1)
- Density of category-article categorizations
- ONTO NLP experience with the language

We considered the following languages (ordered approximately by preference) EN, BG, FR, CAT, ES, PL, HU, DE, LT, NL, GR, IT.

At the end, French was selected for enrichment. Two EFD collections feature FR content:

- BE-CAG: 1000 objects, whose text however is also available in EN
- BE-MRAC: 7500 objects in an important ethnographic collection, which we were able to enrich

#### 3.2 French Enrichment

Developing a French enrichment pipeline was a major undertaking for ONTO. We developed it:

- Based on our English pipeline, adapting all language-specific components to French
- Deployed publically available French language resources
- Based on data from EN DBpedia, FR DBpedia and Wikidata. Leveraged the EN-FR Inter-Language Links between articles
- Leveraged Machine Learning models about word/phrase collocation for disambiguation

The initial version showed very low recall (0.14) and modest precision (0.46), which was due to the high ambiguity present in French relative to the FD domain. For example, “Coupe” (an often appearing artefact) has at least 19 meanings in <https://fr.wikipedia.org/wiki/Coupe>. It was hard to disambiguate this to [https://fr.wikipedia.org/wiki/Coupe\\_\(récipient\)](https://fr.wikipedia.org/wiki/Coupe_(récipient)), because many BE-MRAC texts are quite short, and we didn’t have enough collocation Gold Standard data.

Therefore, we added a rule-based approach, which increased both recall (0.93) and precision (0.67).

#### 3.3 Enrichment Evaluation

Table 2 shows the number of FD and Place enrichments per collection, and averages per object. However, it is also important to evaluate the accuracy (Precision) and completeness (Recall) of these enrichments. We performed a detailed evaluation based on manual enrichment of a subset (gold standard):

- 20 objects from each English collection (total 180)
- 100 objects from the French collection (BE-MRAC), developed together with the MRAC museum

**Table 3 Enrichment Evaluation Results**

	TP	FP	FN	Precision	Recall	F1
<b>EN FD</b>	355	46	303	0.885	0.540	0.670
<b>EN Places</b>	102	25	55	0.803	0.650	0.718
<b>FR FD (Initial)</b>	19	22	116	0.463	0.141	0.216
<b>FR FD (With Rules)</b>	127	62	9	0.672	0.934	0.782
<b>FR Places</b>	19	36	114	0.345	0.143	0.202

TP: true positives, FP: false positives, FN: false negatives. F1: harmonic mean of Precision and Recall.

The above are “raw” precision and recall numbers. Often there are very small differences between the manual and automatic annotation, so the automatic annotation should also be counted as correct. Some examples (we provide the original French text and English translations. We don't show concepts that are common between Manual and Automatic enrichment (counted as TP above).

**Table 4 French FD: Manual to Automatic Comparison**

Text	Manual	Automatic	Comment
Pêcheurs à la ligne Fishers “on a line”	Pêche_à_la_mouche Fly-fishing	Pêche_(halieutique) Fishing	Auto is more accurate
faire de l'huile alimentaire à base de noix de palme preparation of edible oil based on palm nuts	Huile_de_palme Palm oil	Alimentation; Huile_alimentaire; Noix_de_coco Food; Edible oil; Coconut	Coconut is wrong, the oil palm (Elaeis) is different from the Coconut palm (Cocos). The other 2 are correct
Coupe à boire. Coupe polie. Drinking cup. Polished? Cup.		Boisson Drink	(Coupe is common) Man has skipped a concept. Auto is more accurate
Coupe avec anse. Cup with handle		Poignée Handle	(Coupe is common) Man has skipped a concept. Auto is more accurate
Couvercle de pot. Lid for a pot		Pot_(récipient)	(Couvercle is common) Man has skipped a concept. Auto is more accurate
récipient que l'on verse la première gorgée d'eau au nouveau-né Container from which a first sip of water is given to a newborn	Eau Water	Eau_potable Drinking water	Auto is more specific

Regarding places: many MRAC places are small/obscure places in Congo and other places in Africa, that are not comprehensively described in fr.wiki, therefore we cannot recognize them. It would be possible to do this with additional effort, by bringing in Wikidata and doing additional curation.

**Table 5 French Places: Manual to Automatic Comparison**

Text	Manual	Automatic	Comment
Kasai; Luangue (Region between Kasai river and Luangue river)	<a href="#">Lwange (cours d'eau)</a>		(Kasaï_(rivière) is common). Auto didn't recognize Luangue because there's only a note "Loange (ou Luangue en portugais)" in the page, but no redirect (explicit alias). Wikidata has 3 entries for a Luangue river (need to be checked and merged)
Congo; Kasai	République_démocratique_du_Congo; Kasaï-Oriental_(province_historique); Kasaï-Occidental	Congo_(fleuve); Kasaï_(rivière)	Congo should have been recognized as the country. Kasai is highly ambiguous, even MRAC don't know to which of two possible provinces their data refers
Central Africa	Afrique_centrale	Afrique	Auto is less specific
Congo; Uele	République_démocratique_du_Congo	Uele_(rivière)	Congo should have been recognized as the country. Man decided that Uele does not refer to the river, but that's better than nothing (Uele river is indeed in Congo)

Overall an F-score of about 0.7 across all collections is fairly good.

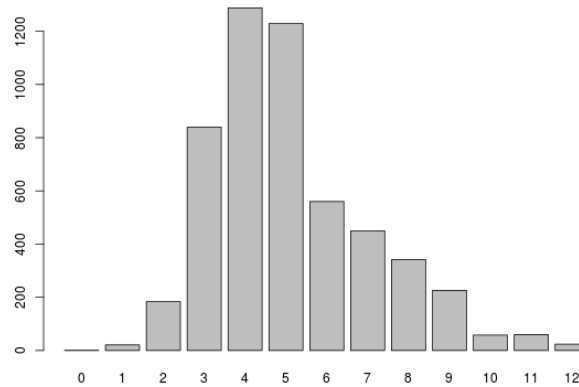


## 4 Extend the FD Classification

As preparation for French enrichment, we had to extend the FD Classification with categories from fr.wiki.

### 4.1 Comparison of FD Trees Across Languages

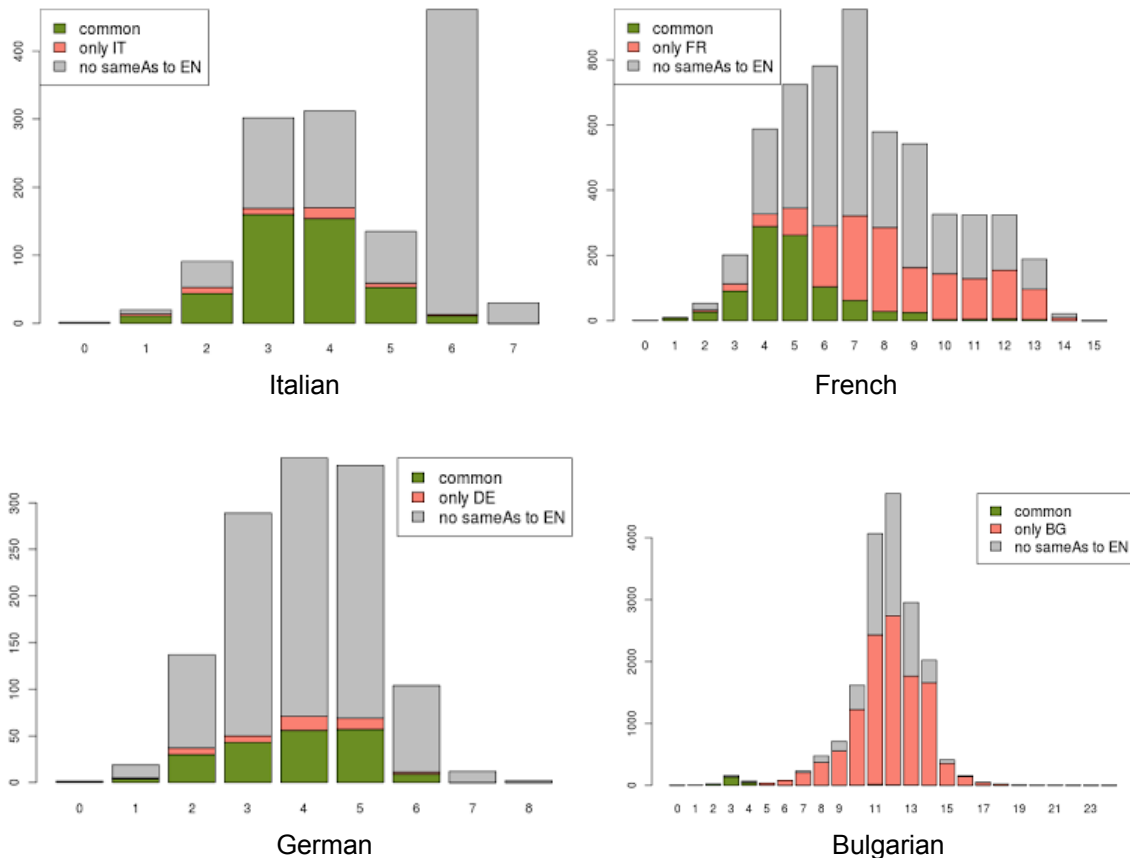
One of the criteria for selecting a second language for enrichments was the number of sameAs (Inter-Language Links) to the English FD categories. Starting from the FD root, the number of English categories per level is as follows:



**Figure 12 Raw English FD Categories per Level**

This is called “raw” number since it shows the categories before pruning. The FD Classification is built by pruning irrelevant branches, and reducing the path to root (thus reducing “semantic drift”).

We compared this distribution to FD categories in 4 candidate languages



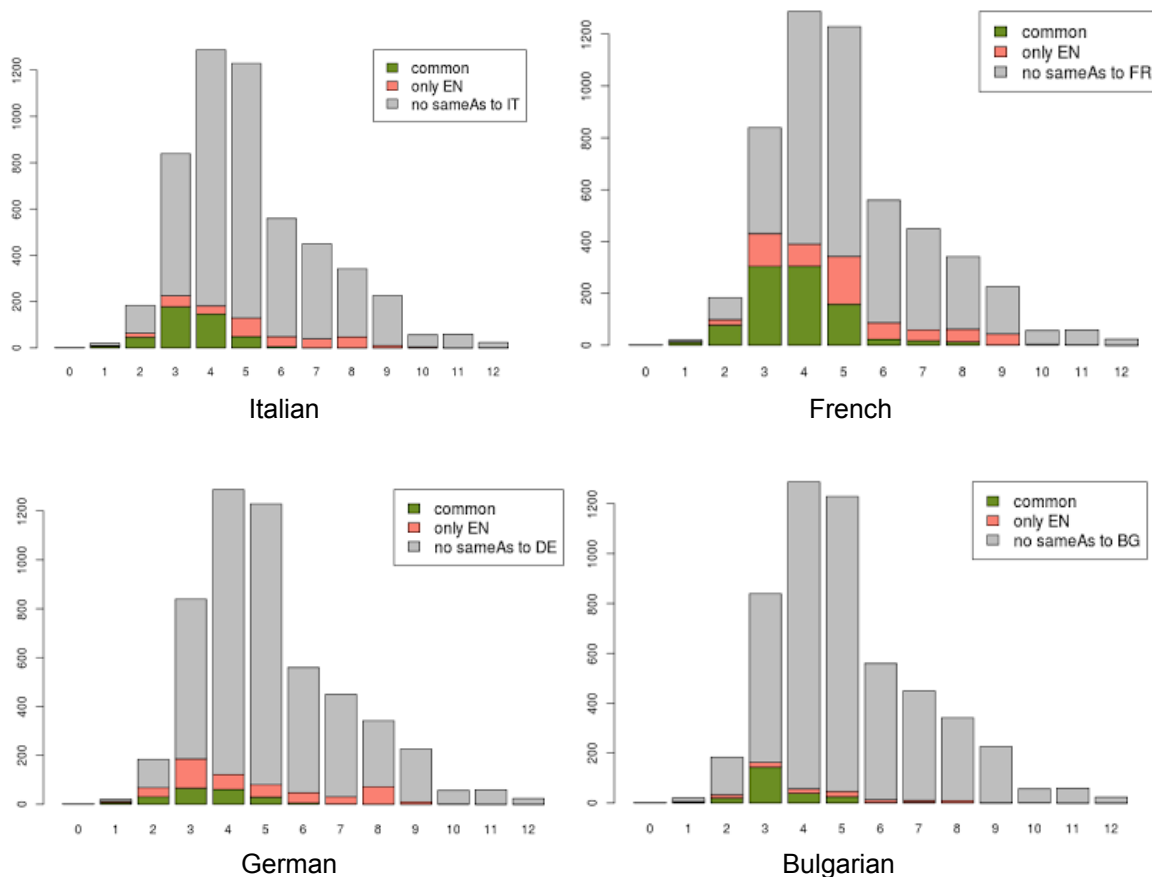
**Figure 13 Comparing 4 Languages to English**

French shows the best numbers, e.g. it has some 300 FD categories at level 4 that correspond to English. Bulgarian shows the worst numbers, with only 150 corresponding categories up to level 5.

Legend:

- **Green:** XX FD categories that have *sameAs* to English FD
- **Red:** XX FD categories that have *sameAs* to English, but outside FD
- **Grey:** XX FD categories that don't have *sameAs* to English

We also analyzed this in the opposite direction, from the perspective of the English FD tree. You can see that English (the biggest Wikipedia) has a significantly larger number of categories.



**Figure 14 Comparing English to 4 Languages**

## 4.2 French FD Categories

We constructed the French FD tree as follows:

- Loaded the fr.dbpedia, including owl:sameAs correspondences to en.dbpedia, the so-called Inter-Language Links
- Found all FR categories that have EN FD correspondence and added them to the tree
- Found “nearby” categories: if a category has both FD parent and FD child, we added it to the tree even if it doesn't have an EN counterpart

We further elaborated the FD classification through bottom-up augmentation using FD articles discovered during the enrichment evaluation (see sec. 3.3).

### 4.3 FD Classification as SKOS

We make a FD gazetteer from all Wikipedia articles classified by the FD Categories. Consider for example the article Kashkaval (a type of yellow cheese). We use its title and all aliases (redirects), as well as its categories, e.g.:

```
# label and types
dbr:Kashkaval
  a yago:Cheese107850329, yago:DairyProduct107843775,
    yago:Food100021265, yago:Food107555863, yago:Foodstuff107566340;
  rdfs:label "Kashkaval"@en.

# aliases (redirects)
dbr:Кашкавал    rdfs:label "Кашкавал";    dbo:wikiPageRedirects dbr:Kashkaval.
dbr:Caşcaval    rdfs:label "Caşcaval";    dbo:wikiPageRedirects dbr:Kashkaval.
dbr:Kashawan    rdfs:label "Kashawan";    dbo:wikiPageRedirects dbr:Kashkaval.
dbr:Kashkawan   rdfs:label "Kashkawan";   dbo:wikiPageRedirects dbr:Kashkaval.
dbr:Kashkawane  rdfs:label "Kashkawane";   dbo:wikiPageRedirects dbr:Kashkaval.
dbr:Kashqawan   rdfs:label "Kashqawan";   dbo:wikiPageRedirects dbr:Kashkaval.
dbr:Kaskaval    rdfs:label "Kaskaval";    dbo:wikiPageRedirects dbr:Kashkaval.
dbr:Kaşkaval    rdfs:label "Kaşkaval";    dbo:wikiPageRedirects dbr:Kashkaval.
dbr:Kaškaval    rdfs:label "Kaškaval";    dbo:wikiPageRedirects dbr:Kashkaval.

# categories
dbr:Kashkaval dct:subject
  dbc:Bulgarian_cheeses, dbc:Romanian_cheeses, dbc:Serbian_cuisine,
  dbc:Turkish_cheeses, dbc:Cow's-milk_cheeses, dbc:Sheep's-milk_cheeses,
  dbc:Republic_of_Macedonia_cheeses, dbc:Serbian_cheeses.

# broader categories
dbc:Bulgarian_cheeses a skos:Concept;
  rdfs:label "Bulgarian cheeses"@en;
  skos:broader dbc:Bulgarian_cuisine, dbc:Cheeses_by_country.
dbc:Cheeses_by_country a skos:Concept;
  rdfs:label "Cheeses by country"@en;
  skos:broader dbc:Cheeses, dbc:Cuisine_by_nationality.
  # also dbc:Categories_by_country but that's not in the F&D tree
```

The categories are expressed in SKOS: they have type `skos:Concept` and use `skos:broader`. But the articles are not `skos:Concept`, since they can be any specific type (e.g. `yago:Cheese107850329` as above, `dbo:Food`, `dbo:Person`, etc).

## 5 Geographic Mapping

Based on the Place enrichments and the Geonames place hierarchy, we added a Geographic Map, in addition to the hierarchical browsing by Place. It complements the existing lightbox (thumbnail grid). This involved the following subtasks.

### 5.1 Hierarchical Place Processing

- Eliminate **superfluous ancestor** places. E.g. if a CHO is tagged with Rome and Italy, we remove the parent place Italy, else the same CHO will appear with two different markers on the map
- Complement with **ancestors with coordinates**: If a CHO is marked with "Fleet Street" and neither GeoNames nor DBpedia have coordinates about it, we need to add its most specific ancestor that has coordinates (in this case, "City of London" and not "London", which is a greater area)

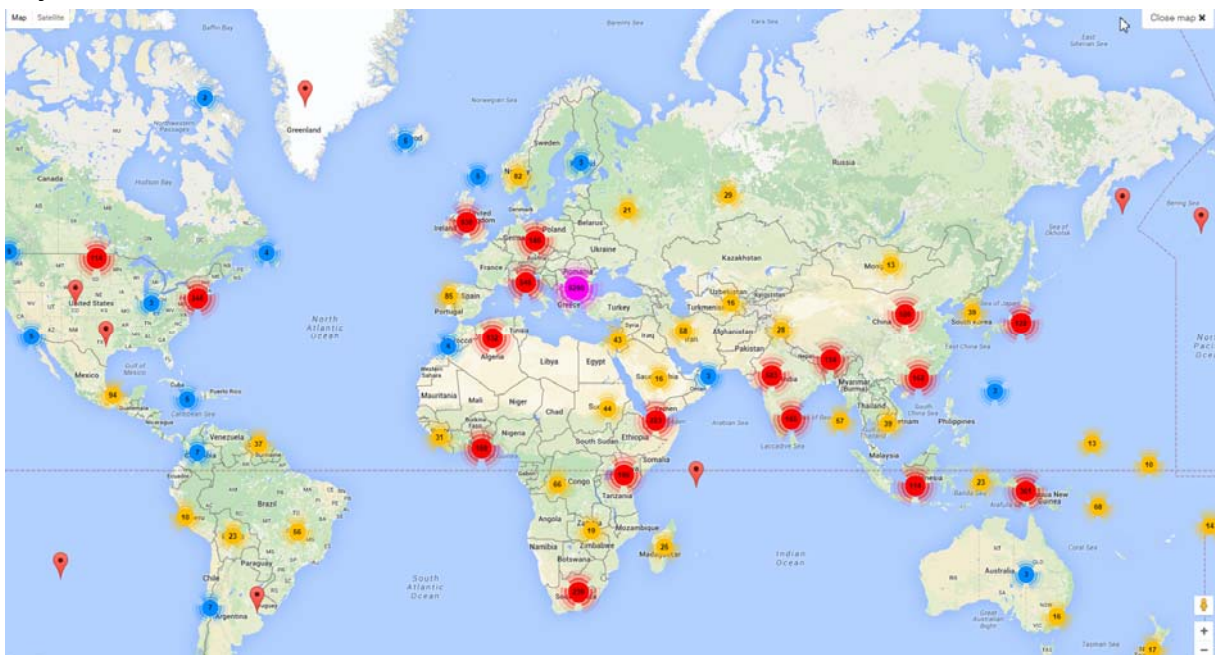
### 5.2 Coordinate Processing

- **Average coordinate** values.

DBpedia and Geonames places have slightly different coordinates for the same place. We averaged the coordinates of the same place, to ensure one marker per place.

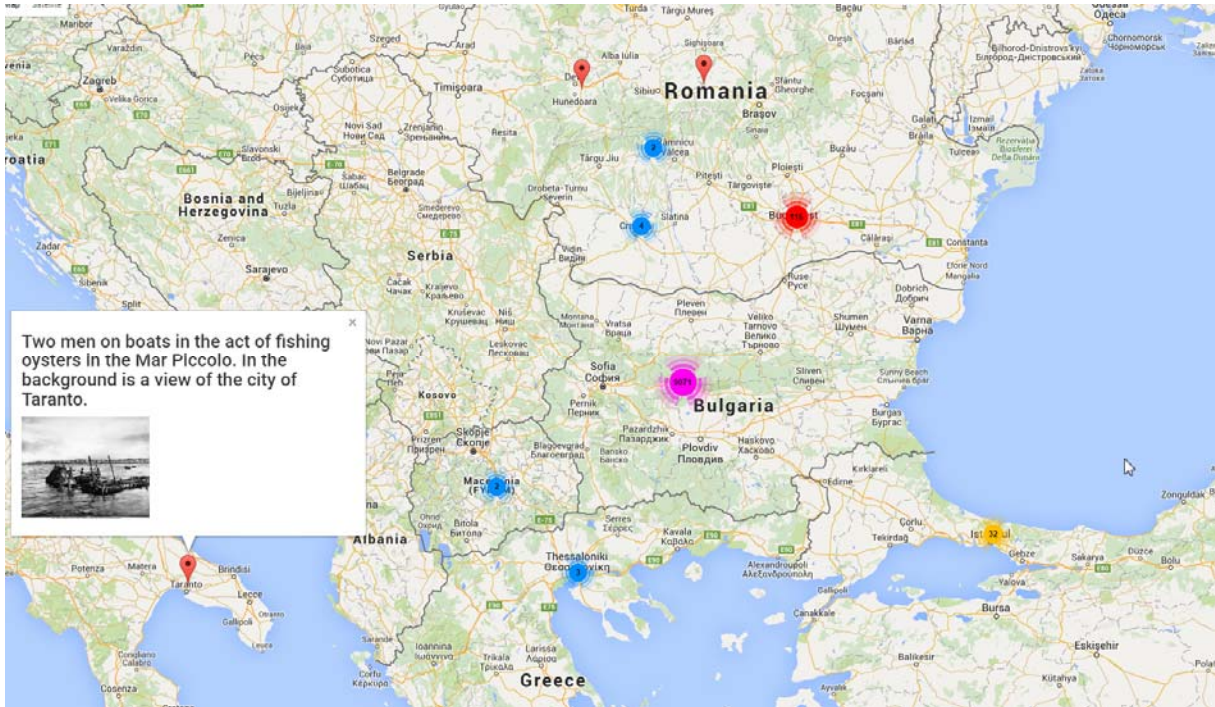
- **Jitter coordinates**.

We use the “marker cluster” library. It can display many thousands of places by using clusters of objects that are close to each other, with the number of markers. When you zoom in, the cluster is broken down into smaller clusters, down to individual objects that are shown as markers.



**Figure 15 EFD Semapp Geographic Clusters**

Then you can click on a marker to see the object info; and click once more to see the full object record.



**Figure 16 EFD Map Showing Individual Object**

But if several objects reference the same place, you cannot “break the cluster” to get to individual objects.

Consider the 9k BG-ONTO objects: they all refer to the same place, and are shown somewhere in the middle of Bulgaria (near the [Tsarichina natural reserve](#)). To allow the user to zoom-in to individual objects, we have introduced jitter (randomness) in the coordinates associated with every object (see Figure 17 and Figure 18). We want to shift the coordinates by up to 10km:

- 10km of latitude equals 0.090 degrees everywhere on Earth
- 10km of longitude equals 0.122 degrees in Bulgaria (along parallel 42), less closer to the equator, and more closer to the poles. We need the “cosine” function to compute longitude-dependent jitter range, but SPARQL doesn’t include trigonometric functions. Therefore we use constant jitter (0.122 degrees), which is good enough.

We introduced jitter with the following SPARQL query:

```

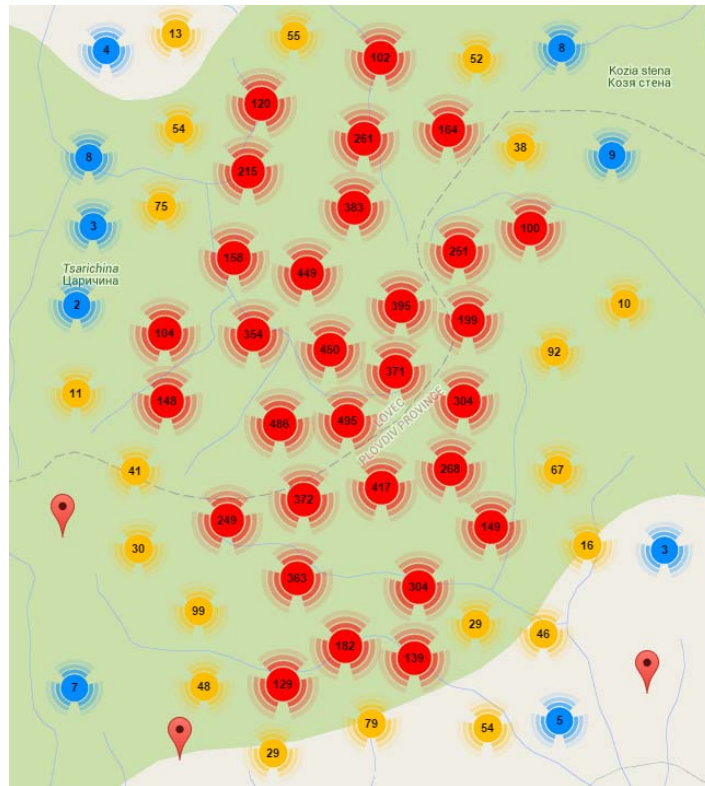
construct {
  ?cho dct:spatial
    [wgs:lat ?rand_lat; wgs:long ?rand_long; rdfs:label ?name]
} where {
  {select ?cho ?place ?name
    (average(?rand_lat) as ?rand_lat) (average(?rand_long) as ?rand_long) {
    ?cho a edm:ProvidedCHO; dct:spatial ?place.
    ?place a dbo:Place; wgs:lat ?lat; wgs:long ?long; rdfs:label ?name.
    bind(?lat+rand()*0.090 as ?rand_lat)
    bind(?long+rand()*0.122 as ?rand_long)
    } group by ?cho ?place ?name}
}

```

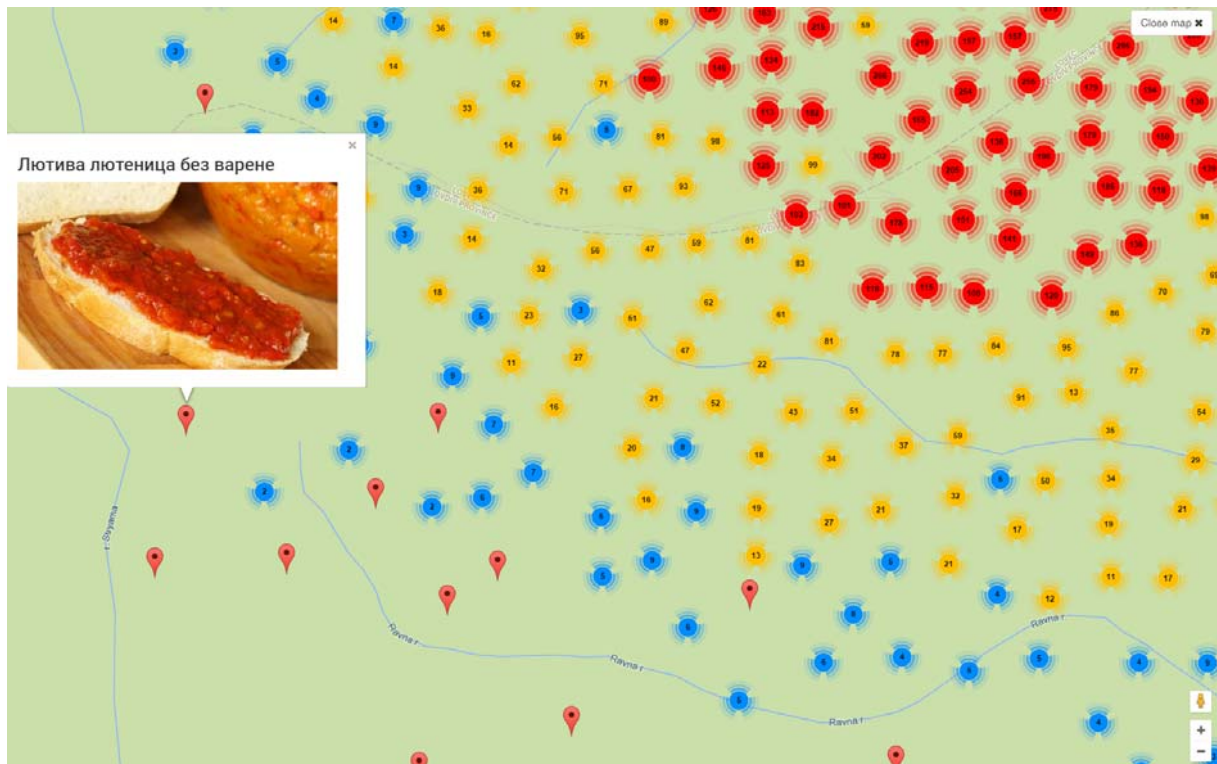
This records the jittered places as blank nodes without rdf:type, which allows us to find them for the map (SPARQL query below), but skip them when displaying place enrichments.

## D3.20d Semantic Demonstrator Extension

```
select ?cho ?title ?lat ?long ?place_name {  
  ?cho dct:spatial ?place.  
  filter not exists {?place a ?type} # bit of a dirty hack, but so what  
  ?place wgs:lat ?lat; wgs:long ?long; rdfs:label ?place_name.  
}
```



**Figure 17 Zooming Into Jittered Coordinates**



**Figure 18 Zoomed Down To Individual Object**

### 5.3 Geographic Mapping in the Sem App

We have added a View on Map button to the sem app.

The screenshot shows the 'Europeana Food & Drink' application. At the top left is a logo with a green fork, knife, and wine glass. The title 'Europeana Food & Drink' is centered. To the right are the 'ontotext' and 'made with europeana' logos. Below the title is a search bar and a 'View on map' button circled in red. The main content area is divided into a left sidebar with filters and a main grid of results. The filters include 'Food and Drink' (with sub-categories like Agriculture, Beverages, Cuisine, etc.) and 'Places'. The results grid shows four items, each with an image and a title: 'View of Stia', 'A lemon from the Alinari Garden in Florence', 'A lovely prickly pear cactus in Taormina', and 'Farmhouse in San Manno in the environs of Perugia'. Each item has a warning or creative use notice below the title.

**Figure 19 “View on Map” Button**

It shows the result set (objects selected by the current query) on a geo map, using clustering as described above. The user can click on a marker to see the object image and description, and click further to see the detailed object record.

An important consideration is the total response size, which for the initial view includes all 43k objects. In order to avoid making the user wait for a long time, we adopted these strategies:

- Asynchronous loading of data and dynamic clustering as more data comes in
- Deployed a cache to speed up responses (see sec 7).
- Limited the info per object to the bare minimum (coordinates and URL). When the user clicks, an extra request is made to fetch the info for the popup.

## 6 Semantic Enrichment Web Service

We established a semantic enrichment web service that provides semantically enriched content. It suggests automatic enrichments to crowd-source curators and is used by the Crowdsourcing Enrichment application (developed by D3.5 Technical Demonstrator and T5.2 Community/crowdsourcing platform). We tackled the important issues of availability, performance and monitoring to establish this service. The same service could be used to provide semantically enriched content to other application creators as well.

The service is deployed at <http://efd.ontotext.com/enrichments/extract> and

### 6.1 Service Input

The service is used as follows. Consider the following example:

```
curl -X "@example.txt" -H content-type:text/plain
http://efd.ontotext.com/enrichment/extract?uri=http://example.org/objects/0001 >
example.jsonld
```

where example.txt has the following text:

- A piece of strawberry sponge cake on a white plate with a small blue and white spotted mug of black coffee credit: Marie-Louise Avery / thePictureKitchen / TopFoto. baking; strawberries; cakes; europeana food and drink; eufd; sugar; food; GEN; sweet; cooking; cookery; teatime; cream; slice; whipped; icing; break  
The picture was taken in London or maybe Seattle.

The **uri** parameter is the object URL against which enrichments are emitted. The service discovers FD topics and Places, and does the special place processing described in sec 5.1. Gor now it handles EN enrichment (but returns concept labels in English and French).

### 6.2 Output in JSON-LD

The service output is in JSON-LD, which is both convenient for web apps (JSON is easy to handle) and is a valid RDF representation. For the above example it is:

```
[
  { "@id" : "http://dbpedia.org/resource/Cake",
    "rdfs:label" : [ {"@value" : "Cake"} ]
  },
  ...
  { "@id" : "http://dbpedia.org/resource/London",
    "efd:name" : [ {"@value" : "London"} ],
    "wgs:lat" : [ {"@type" : "http://www.w3.org/2001/XMLSchema#double",
                  "@value" : "51.507877"} ],
    "wgs:long" : [ {"@type" : "http://www.w3.org/2001/XMLSchema#double",
                   "@value" : "-0.12662"} ]
  },
  ...
  { "@id" : "http://example.org/objects/0001",
    "dct:spatial" : [ {"@id" : "http://dbpedia.org/resource/London"},
                     {"@id" : "http://dbpedia.org/resource/Seattle"} ],
    "dct:subject" : [ {"@id" : "http://dbpedia.org/resource/Cake"},
                      {"@id" : "http://dbpedia.org/resource/Cooking"},
                      {"@id" : "http://dbpedia.org/resource/Coffee"},
                      {"@id" : "http://dbpedia.org/resource/Sponge_cake"},
                      {"@id" : "http://dbpedia.org/resource/Strawberry"},
                      {"@id" : "http://dbpedia.org/resource/Sugar"},
                      {"@id" : "http://dbpedia.org/resource/Food"} ]
  }
]
```



### 6.3 Output in Turtle

The same output converted in Turtle is easier to understand:

```
dbr:Cake rdfs:label "Cake" .
dbr:Coffee rdfs:label "Coffee" .
dbr:Cooking rdfs:label "Cooking" .
dbr:Food rdfs:label "Food" .
dbr:Sponge_cake rdfs:label "Sponge cake" .
dbr:Strawberry rdfs:label "Strawberry" .
dbr:Sugar rdfs:label "Sugar" .

dbr:London rdfs:label "London" ;
  wgs:lat "51.50788"^^xsd:double ;
  wgs:long "-0.12662"^^xsd:double .

dbr:Seattle rdfs:label "Seattle" ;
  wgs:lat "47.606213"^^xsd:double ;
  wgs:long "-122.33207"^^xsd:double .

<http://example.org/objects/1000>
  dct:spatial dbr:London , dbr:Seattle ;
  dct:subject dbr:Cake , dbr:Cooking , dbr:Coffee , dbr:Sponge_cake ,
    dbr:Strawberry , dbr:Sugar , dbr:Food .
```

It first emits the labels and coordinates of all concepts and places found by enrichment, and then emits links from the given CHO URI to those concepts and places (dct:subject and dct:spatial respectively).

## 7 Other Sem App Enhancements

- Deployed a SQUID proxy to improve application response time (from minutes to seconds)
- Wrote a detailed description of <http://efd.ontotext.com/data/>

# EFD Semantic Data

## Table of Contents

- [1. Intro](#)
- [2. Ontology](#)
- [3. FD Classification](#)
- [4. Place Data](#)
- [5. EFD Collections](#)
- [6. Semantic Enrichments](#)

## 1 Intro

This page allows you to download all data used by the [EFD semantic application](#).

Most files are provided in Turtle 1.1. Some of them use SPARQL-style prefixes, eg

```
prefix dct: <http://purl.org/dc/terms/>
# instead of
@prefix dct: <http://purl.org/dc/terms/>.
```

**Figure 20 EFD Data Documentation**

- Created an EFD ontology and documented it with Parrot

The screenshot displays the documentation for the 'Europeana Food and Drink ontology (Ontology)'. It includes a 'Definition' section with the URI 'http://data.foodanddrinkeurope.eu/ontology' and a 'Preferred Namespace URI' of 'http://data.foodanddrinkeurope.eu/ontology#'. A 'Properties' section lists various ontology properties such as 'ancestor place', 'EFD ancestor', 'EFD child', 'EFD descendant article count', 'EFD evidence score 1', 'EFD evidence score 2', 'EFD local article count', 'EFD local evidence count', 'EFD subject', 'in topic tree', and 'Not FD relevant'. The bottom part of the image shows the documentation for the 'ancestor place (Property)', defined as 'From Geonames place to itself and all its ancestor places: Kleene closure on:parentFeature\*'. A navigation bar at the bottom allows switching between Ontologies, Properties, Individuals, and Expand/Collapse all, along with 'Go to top' and 'Back to form' buttons.

**Figure 21 EFD Ontology Documentation**

## 8 Dissemination

Additional publications in 2016 related to the EFD topics are listed in the last section.

### 8.1 Engaging Bulgarian Community

We advertised the sem app at the Bulgarian Google group [cultural-heritage-digitalization](#):

- [Announcement](#) of the BG-ONTO collection and the semapp
- [Invitation](#) to Bulgarian FD-related museums to contribute content

We also advertised at the <http://bulgariana.eu/> website:

- [Description](#) of the Bulgarian Recipes collection
- [Announcement](#) of the collection and the sem app

### 8.2 Sem App Access Statistics

We have logged all traffic to the semapp web application, and analyzed them using AWstats (Advanced Web Statistics).

**Table 6 Number of Visitors and Visits**

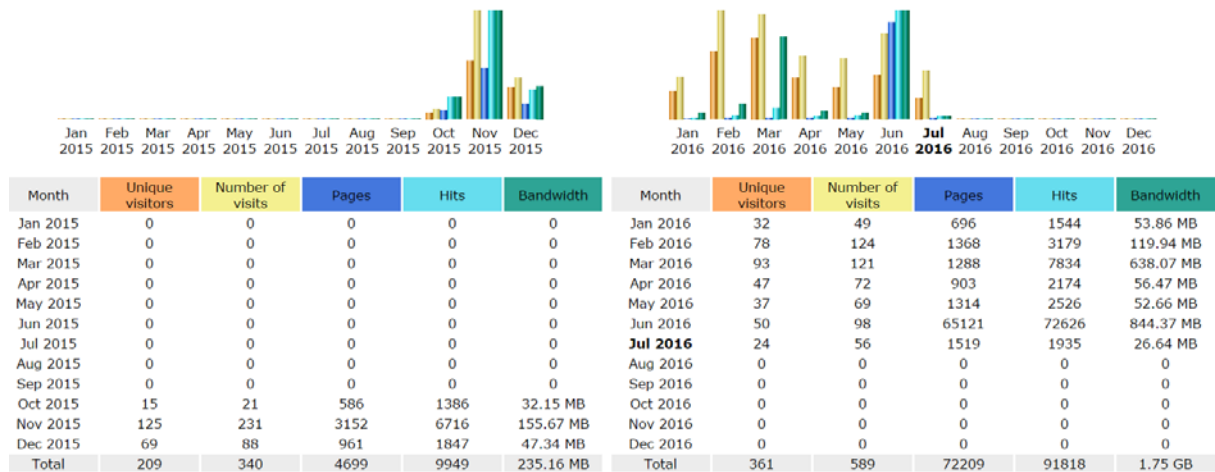
Reported	Year 2015				
First visit	29 Oct 2015 - 06:52				
Last visit	28 Dec 2015 - 22:53				
	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Viewed traffic *	<b>&lt;= 209</b>	<b>340</b> (1.62 visits/visitor)	<b>4699</b> (13.82 Pages/Visit)	<b>9949</b> (29.26 Hits/Visit)	<b>235.16 MB</b> (708.24 KB/Visit)
Not viewed traffic *			<b>2394</b>	<b>2758</b>	<b>24.62 MB</b>

Reported	Year 2016				
First visit	01 Jan 2016 - 09:49				
Last visit	19 Jul 2016 - 20:40				
	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Viewed traffic *	<b>&lt;= 361</b> Exact value not available in 'Year' view	<b>589</b> (1.63 visits/visitor)	<b>72209</b> (122.59 Pages/Visit)	<b>91818</b> (155.88 Hits/Visit)	<b>1.75 GB</b> (3115.45 KB/Visit)
Not viewed traffic *			<b>73018</b>	<b>75091</b>	<b>497.84 MB</b>

- \* Not viewed traffic includes traffic generated by robots, worms, or replies with special HTTP status codes.
- Unique visitors: Exact value not available in 'Year' view

We have 209 unique visitors in 2015 (2 months) and 361 unique visitors in 2016 (6.5 months). The number of visits is 340 and 589 respectively. Given that we have not disseminated the semapp extensively, that is not bad.

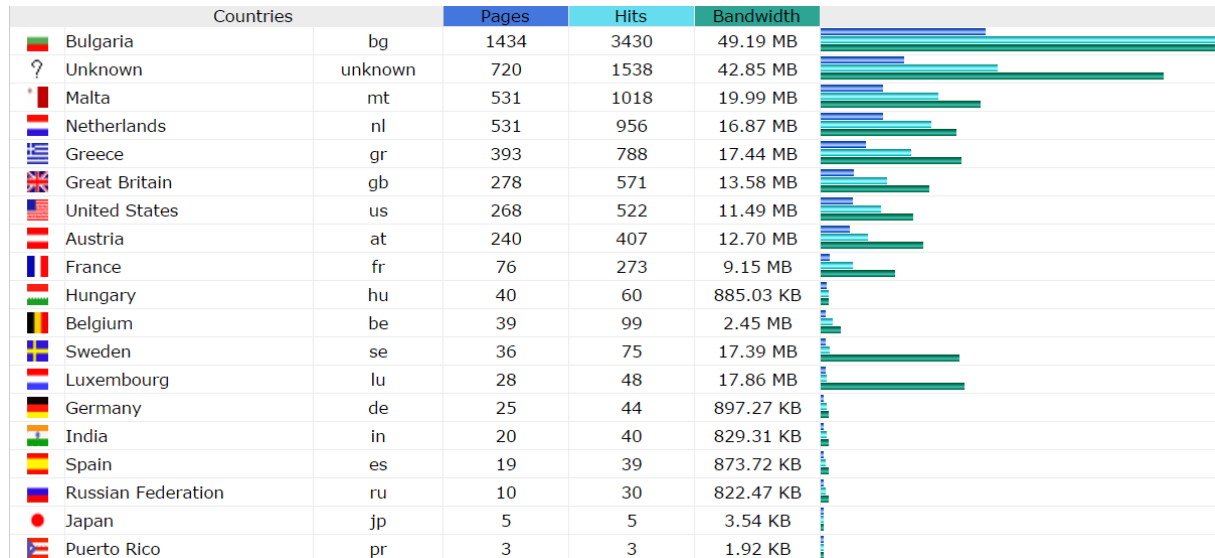
### D3.20d Semantic Demonstrator Extension



**Figure 22 Monthly History of Visits**

The monthly history shows an initial peak of interest (125 visitors in Nov), which decreases later. Now that we have the extended semapp, we plan to disseminate it to increase traffic.

The geographic distribution of visitors is quite wide, although most are from Bulgaria.



**Figure 23 Country Distribution of Visitors, 2015**

Interestingly, in 2016 we have a wider distribution, and the visits are dominated by Hungary not Bulgaria.

## D3.20d Semantic Demonstrator Extension

Countries		Pages	Hits	Bandwidth	
	Hungary	hu	33653	34045	23.33 MB
	Bulgaria	bg	30672	35044	340.01 MB
	Romania	ro	1715	1850	15.23 MB
	Slovak Republic	sk	1321	1420	4.88 MB
	Spain	es	935	1250	43.41 MB
	Greece	gr	646	2142	123.07 MB
	Unknown	unknown	596	5892	560.38 MB
	United States	us	496	942	40.44 MB
	France	fr	495	987	28.02 MB
	Netherlands	nl	297	504	19.21 MB
	Great Britain	gb	237	614	25.07 MB
	Belgium	be	206	445	12.07 MB
	Germany	de	195	4941	493.70 MB
	Lithuania	lt	193	348	9.68 MB
	Croatia	hr	152	508	29.79 MB
	Cyprus	cy	89	146	1.65 MB
	Italy	it	75	178	5.47 MB
	Austria	at	73	149	3.21 MB
	Latvia	lv	51	51	412.58 KB
	India	in	27	115	4.63 MB
	Poland	pl	23	57	1.38 MB
	China	cn	20	40	801.72 KB
	Canada	ca	18	57	2.36 MB
	Russian Federation	ru	11	11	59.81 KB
	Norway	no	5	74	3.74 MB
	Ukraine	ua	3	3	24.27 KB
	Luxembourg	lu	2	2	1.56 KB
	Israel	il	1	1	655 Bytes
	Taiwan	tw	1	1	655 Bytes
	Australia	au	1	1	1.27 KB

**Figure 24 Country Distribution of Visitors, 2016**

We also have a wide distribution of visitor cities, from Putian, China to Razgrad, Bulgaria. Please note that AWstats uses GeoIP libraries and can recognize only 27-42% of the cities.

**Table 7 City Distribution of Visitors, 2015**

Countries	Regions	Cities: 37	Hits	Percent
Bulgaria	Grad Sofiya	Sofia	1103	11 %
Greece	Attiki	Athens	756	7.5 %
Netherlands	Noord-Holland	Amsterdam	586	5.8 %
Netherlands	Zuid-Holland	Den haag	300	3 %
France	Ile-de-France	Montrouge	173	1.7 %
Great Britain	Buckinghamshire	Gawcott	114	1.1 %
United States	California	Mountain view	107	1 %
United States	Massachusetts	Lynn	94	0.9 %
France	Ile-de-France	Paris	93	0.9 %
Great Britain	Cambridgeshire	Cambridge	85	0.8 %
United States	South Carolina	Duncan	85	0.8 %
Hungary	Heves	Gyöngyös	60	0.6 %
Great Britain	Essex	Chelmsford	58	0.5 %
Luxembourg	Luxembourg	Schifflange	48	0.4 %
United States	New Jersey	Woodbridge	48	0.4 %
India	Maharashtra	Mumbai	40	0.4 %
Great Britain	London, City of	London	37	0.3 %
Germany	Baden-Wurttemberg	Karlsruhe	36	0.3 %
Netherlands	Noord-Holland	Amstelveen	36	0.3 %
Belgium	Brabant	Braine-l'alleud	35	0.3 %

D3.20d Semantic Demonstrator Extension

Belgium	Oost-Vlaanderen	Sleidinge	34	0.3 %
Greece	Thessaloniki	Thessaloníki	32	0.3 %
Sweden	Varmlands Lan	Torsby	32	0.3 %
Belgium	Antwerpen	Antwerp	30	0.3 %
United States	Delaware	Wilmington	28	0.2 %
United States	New York	Staten island	28	0.2 %
United States	California	San francisco	13	0.1 %
Germany	Mecklenburg-Vorpommern	Kiez	7	0 %
United States	Ohio	Columbus	5	0 %
Japan	Osaka	Osaka	4	0 %
United States	Texas	Mcallen	3	0 %
United States	Washington	Seattle	3	0 %
United States	District of Columbia	Washington	2	0 %
United States	Indiana	Indianapolis	2	0 %
Poland	Malopolskie	Kraków	1	0 %
Germany	Brandenburg	Potsdam	1	0 %
Japan	Okayama	Tama	1	0 %
Others/Unknown			5829	58.5 %

**Table 8 City Distribution of Visitors, 2016**

Countries	Regions	Cities: 83	Hits	Percent
Bulgaria	Grad Sofiya	Sofia	13086	14.2 %
Germany	Mecklenburg-Vorpommern	Kiez	4491	4.8 %
Greece	Attiki	Athens	1094	1.1 %
Bulgaria	Razgrad	Razgrad	672	0.7 %
Croatia	Grad Zagreb	Zagreb	508	0.5 %
Bulgaria	Varna	Varna	359	0.3 %
Bulgaria	Stara Zagora	Stara zagora	330	0.3 %
France	Ile-de-France	Paris	323	0.3 %
Netherlands	Zuid-Holland	Den haag	316	0.3 %
France	Haute-Normandie	Le havre	307	0.3 %
Belgium	Brabant	Tervuren	271	0.2 %
United States	California	Mountain view	254	0.2 %
Spain	Cataluna	Barcelona	171	0.1 %
Great Britain	Buckinghamshire	Gawcott	147	0.1 %
Cyprus	Limassol	Lemesos	139	0.1 %
Bulgaria	Sliven	Sliven	133	0.1 %
Germany	Bayern	Munich	130	0.1 %
Germany	Berlin	Berlin	121	0.1 %
Italy	Emilia-Romagna	Modena	119	0.1 %
United States	Montana	Missoula	116	0.1 %
Great Britain	Wolverhampton	Wolverhampton	107	0.1 %
Great Britain	Glasgow City	Glasgow	93	0.1 %
Germany	Niedersachsen	Bramsche	91	0 %
Greece	Kozani	Ptolemais	80	0 %
Norway	Oslo	Oslo	74	0 %
Netherlands	Noord-Holland	Amsterdam	73	0 %
United States	Louisiana	New orleans	71	0 %
Lithuania	Vilniaus Apskritis	Vilnius	68	0 %
Germany	Bayern	Nürnberg	63	0 %
United States	Michigan	Ann arbor	57	0 %

### D3.20d Semantic Demonstrator Extension

United States	New York	New york	56	0 %
Italy	Toscana	Lucca	48	0 %
France	Languedoc-Roussillon	Montpellier	46	0 %
Canada	Quebec	Blainville	45	0 %
Spain	Castilla-La Mancha	Pantoja	42	0 %
Bulgaria	Burgas	Karnobat	41	0 %
France	Ile-de-France	Montrouge	41	0 %
United States	Delaware	Wilmington	41	0 %
United States	District of Columbia	Washington	39	0 %
Belgium	West-Vlaanderen	Kortrijk	38	0 %
Poland	Mazowieckie	Warsaw	36	0 %
Germany	Hessen	Marburg	36	0 %
France	Provence-Alpes-Cote d'Azur	Nice	35	0 %
Spain	Cataluna	Sant pere de ribes	34	0 %
Belgium	West-Vlaanderen	Oostende	34	0 %
Belgium	Antwerpen	Geel	34	0 %
China	Beijing	Beijing	31	0 %
United States	North Carolina	Charlotte	29	0 %
Greece	Imathia	Náousa	29	0 %
Greece	Iraklion	Iráklion	29	0 %
Greece	Thessaloniki	Thessaloníki	29	0 %
United States	California	El segundo	29	0 %
Belgium	Brussels Hoofdstedelijk Gewest	Brussel	29	0 %
Spain	Cataluna	Santa maria del camí	28	0 %
Great Britain	London, City of	London	25	0 %
Poland	Malopolskie	Kraków	20	0 %
Netherlands	Utrecht	Utrecht	18	0 %
United States	Ohio	Columbus	15	0 %
Italy	Veneto	Rovigo	11	0 %
United States	New Jersey	Woodbridge	10	0 %
Germany	Bremen	Bremen	9	0 %
Spain	Islas Baleares	Palma	9	0 %
United States	Indiana	Indianapolis	9	0 %
United States	Georgia	Alpharetta	8	0 %
Cyprus	Paphos	Páfos	7	0 %
Canada	British Columbia	Abbotsford	6	0 %
Canada	Ontario	Ottawa	6	0 %
Greece	Attiki	Glyfáda	4	0 %
China	Fujian	Putian	4	0 %
Netherlands	Zuid-Holland	Delft	4	0 %
United States	California	San francisco	3	0 %
United States	Pennsylvania	Malvern	2	0 %
United States	New York	Tonawanda	2	0 %
Slovak Republic	Bratislava	Bratislava	1	0 %
Belgium	Brussels Hoofdstedelijk Gewest	Brussels	1	0 %
Taiwan	T'ai-pei	Taipei	1	0 %
Hungary	Budapest	Budapest	1	0 %
United States	Arizona	Phoenix	1	0 %
Poland	Swietokrzyskie	Herby	1	0 %
Bulgaria	Razgrad	Sandrovo	1	0 %

India	Maharashtra	Mumbai	1	0 %
Great Britain	Bath and North East Somerset	Bath	1	0 %
Israel	Tel Aviv	Herzliya	1	0 %
Others/Unknown			66893	72.8 %

### 8.3 Further Participation in Working Groups

We continued our participation in working groups and community initiatives.

- In 2016 Vladimir Alexiev was elected to the Europeana Members Council<sup>1</sup>. The MC helps Europeana establish its working strategy, and sets the agenda for the Annual General Meeting (in 2016 it will be in Riga, Latvia).
- We also participate in the Europeana Data Quality Committee, which allows us to help with technical approaches for improving quality, and push for better quality in Europeana

We are also active in DBpedia:

- Participate in the DBpedia Ontology and Data Quality committee
- Will participate in the DBpedia Citation Challenge judging group in Sep 2016
- Active in DBpedia semi-annual meetings

Finally, we are active in Wikidata, especially the Coreferencing and Authority Control projects/ communities. By giving back to the community, this allows us to obtain better background knowledge for our semantic enrichment services.

We participated in the following meetings in 2016:

- 20160212 The Hague: DBpedia meeting. We presented on “Using DBPedia in Europeana Food and Drink” [Alexiev 2016]
- 20160222 Copenhagen: Europeana Members Council
- 20160421 The Hague: Europeana Data Quality Committee
- 20160606 Budapest: EFD closing meeting. We presented the enhancements to the sem app [Tagarev 2016]

---

<sup>1</sup> <http://pro.europeana.eu/blogpost/meet-the-members-council-vladimir-alexiev>



## 9 Sustainability

This section answers some questions regarding the sustainability of the developed artefacts after the end of the project.

### 9.1 Food and Drink Classification

- **Where does it live?**

The EFD Classification is published at <http://efd.ontotext.com/data>, which includes detailed documentation.

- **What form is it in?**

It's a tree of Wikipedia (DBpedia) categories, starting from the roots dbc:Food\_and\_drink (English) and frdbc:Alimentation (French), that are judged relevant to FD. It's in the form of Turtle files:

- [./efd-blacklist.ttl](#): 400 manual judgements that are cut-off points in the category tree
- [./efd-child-triples.ttl](#) and [./efd-child-triples-FR.ttl](#): child-parent category links. A subset of the original DBpedia skos:broader links, since only shortest links to root, and only links within the FD tree, are present.
- [./efd-intree-triples.ttl](#) and [./efd-intree-triples-FR.ttl](#): full list of FD categories
- [./efd-level-triples.ttl](#) and [./efd-level-triples-FR.ttl](#): the length of the shortest path to root for each category

This is based on the DBpedia categories for English and French. Since the DBpedia datasets are very large, it's not at the EFD Data directory. We have published only the judgments which categories are F&D relevant, and the level and parent-child info. If someone shows interest, we can also add the DBpedia triples on request. Our en.dbpedia is generated from Wikipedia dumps of 20 Oct 2015, using the open source DBpedia Extraction Framework.

- **Was it actually used in the project?**

Yes, the Classification is used by the Semantic Enrichment pipeline to find FD topics in CHOs (free-text metadata). This is used by the sem app in the FD hierarchical facet.

- **How might it be used in future?**

FD projects can use this classification directly. Other projects have shown interest in this approach, as previously explained: the Europeana for Education working group, and the Europeana Art channel.

- **Could we publish it as SKOS?**

The categories are already expressed in SKOS: they have type skos:Concept and use skos:broader. But the articles are not skos:Concept, since they can be any specific type (e.g. dbo:Food, dbo:Person, yago:Cheese107850329, or no type). See more details in sec. 4.3.

### 9.2 Semantic Enrichment Service

- **Can others use Semantic Enrichment?**

ONTO has deployed a semantic enrichment web service that is used by the Crowdsourcing Enrichment application (developed by D3.5 Technical Demonstrator and T5.2 Community/crowdsourcing platform). See sec. 6 for details.

The same could be used by others to enhance food and drink content within Europeana.

- **Are there licensing terms?**

Anyone can use the service freely. It's not subject to licensing.

- **How long will you keep it running?**

ONTO will keep this running until end-2016 (half a year after project end).

- **Could it be used to test Europeana Annotations?**

We are liaising with Europeana regarding the Europeana Annotations server developed in the context of Europeana Sounds. Later this year we should be able to submit the EFD semantic enrichments developed by us to this server. But to our knowledge, Europeana does not yet display annotations on the Europeana portal.

In the time being, Europeana can get our enrichments from

<http://efd.ontotext.com/data/enrichment/>

### 9.3 Semantic Demonstrator

- **Where does it live?**

<http://efd.ontotext.com/app> is the application

<http://efd.ontotext.com/sparql> is a SPARQL endpoint, where anyone can see and query the same data that we use in the application

<http://efd.ontotext.com/data> is the data we have used (including the EFD ontology)

- **How would anyone access and use it?**

Anyone can access the above, play with the application, query or download the data. There is no password.

- **What is the commitment to keep it accessible?**

ONTO will keep this running until end-2016 (half a year after project end).

- **What are the licensing issues?**

The sem app uses Ontotext GraphDB, which is licensed to the project. The app itself is not subject to licensing.

- **Could it be used for other projects?**

Other groups have expressed interest in the semantic classification shown by the sem app. To do it in a new domain will require assembling a domain-specific gazetteer of that domain, and some adaptation of the app.

- The app uses Elasticsearch for faceting, and the GraphDB-ElasticSearch connector.

- To use it in a production environment would require licensing Ontotext GraphDB

We have released the sem app as open source at <https://github.com/Ontotext-AD/europeana-food-and-drink>.

## 10 References

- [Alexiev 2015] Vladimir Alexiev. Europeana Food and Drink Semantic Demonstrator Delivery. Europeana Food and Drink project, Oct 2015.  
[http://vladimiralexiev.github.io/pubs/Europeana-Food-and-Drink-Semantic-Demonstrator-Delivery-\(D3.20\).pdf](http://vladimiralexiev.github.io/pubs/Europeana-Food-and-Drink-Semantic-Demonstrator-Delivery-(D3.20).pdf)
- [Alexiev 2016] Using DBPedia in Europeana Food and Drink. Presentation, DBpedia Meeting, The Hague, Netherlands, Feb 2016.  
<http://vladimiralexiev.github.io/pres/20160212-Using-DBPedia-in-Europeana-Food-and-Drink.pdf>
- [Tagarev 2015] Andrey Tagarev, Laura Tolosi, Vladimir Alexiev. Domain-specific modelling: Towards a Food and Drink Gazetteer (extended version). In *Semantic Keyword-based Search on Structured Data Sources*, LNCS 9383. Springer Verlag, Jan 2016. [http://link.springer.com/chapter/10.1007/978-3-319-27932-9\\_16](http://link.springer.com/chapter/10.1007/978-3-319-27932-9_16)
- [Tagarev 2016] Andrey Tagarev, Vladimir Alexiev, Boyan Simeonov, Laura Tolosi, and Georgi Boychev. Europeana Food and Drink Semantic Application. Europeana Food and Drink project meeting, Budapest, 6 June 2016
- [Manguinhas 2016] Hugo Manguinhas, Nuno Freire, Antoine Isaac, Juliane Stiller, Valentine Charles, Aitor Soroa, Rainer Simon, Vladimir Alexiev. Exploring comparative evaluation of semantic enrichment tools for cultural heritage metadata. In 20th International Conference on Theory and Practice of Digital Libraries (TPDL 2016), Hannover, Germany, Sep 2016.  
<http://vladimiralexiev.github.io/pubs/TPDL2016.pdf>