Grant Agreement 621023

Europeana Food and Drink

Semantic Demonstrator Extended

Deliverable number: D3.20d
Dissemination level: CO
Delivery date: 20 July 2016
Status: Final
Author(s): Vladimir Alexiev (ONTO), Andrey Tagarev (ONTO), Laura Tolosi (ONTO)

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

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

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This deliverable contains original unpublished work except where clearly indicated otherwise.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

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

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

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

<table>
<thead>
<tr>
<th>Collection</th>
<th>Obj</th>
</tr>
</thead>
<tbody>
<tr>
<td>BG-ONTO</td>
<td>9071</td>
</tr>
<tr>
<td>IT-Alinari</td>
<td>498</td>
</tr>
<tr>
<td>UK-Horniman</td>
<td>4352</td>
</tr>
<tr>
<td>UK-Wolverhampton</td>
<td>439</td>
</tr>
<tr>
<td>UK-TopFoto</td>
<td>1814</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>16174</strong></td>
</tr>
</tbody>
</table>

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**

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

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 (regexp) processing. We gathered the 328 most popular words/phrases and matched them to en.dbpedia manually, e.g.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>агне\w*</td>
<td>Lamb and mutton</td>
</tr>
<tr>
<td>айвар</td>
<td>Ajvar</td>
</tr>
<tr>
<td>айрн</td>
<td>Ayran</td>
</tr>
<tr>
<td>алкоголь\w*</td>
<td>Alcoholic beverage</td>
</tr>
<tr>
<td>ашуре</td>
<td>Ashure</td>
</tr>
<tr>
<td>баклава</td>
<td>Baklava</td>
</tr>
</tbody>
</table>

(\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)
**Figure 1 BG-ONTO: Bulgarian Recipes**

- **Картофени кофиета на фурна**
  Бяли картофи, белим ги и ги намачваме в голем дълбок съд. Добавяме вода, предварително задушени.

- **Градинска мусака**
  Задушаваме орехи в олио, добавяме на парче на ситно грунд. Намахваме 1-та чаша приносия млечно, обърваме смеса и я прекарваме я в съда за фурна. Добавяме вода, лайка, съдия се на хартия, солта и пяна от яйца. Повтаряме на съда за фурна.

- **Салата с мариновани домати**
  В съд на купона вмъкваме оцвета, лайка, със захарта, солта и пасирания чесън. Повтаряме на съда за фурна.

- **Телешки ябълка по манастирски**
  Овалното, измиват и почистете месо. Се нареза в парчета и се задушава в сгореното олио и лайка.

- **Баница с кайма и праг**
  В 3-4 с. л. олио се задушава ситно на парче, добавяме кайма и се задушава в кайма. Добавяме в парчетата и се задушава в свани на парче. Добавяме парчетата и се задушава в свани на парче.

- **Пилешка чорба по селски**
  Порционирайте пилешката и сложете в тенджера заварено с 3 л студена вода, малко сол. Добавете дребните на парче.

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

- **View of Stia**
  WARNING: Permission must be required for non editorial use. Please contact Alinari Archives.

- **Water vendor's kiosk in Naples**
  CREATIVE USE: allowed;

- **A lemon from the Alinari Garden in Florence**
  CREATIVE USE: allowed;

- **Paep 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 Gemäldegalerie of Dresden.**
<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweetmeat tray (tray (food service))</td>
<td>Sweetmeat dish made from lacquered wood. The dish is a pentagon with a curved edge. It is red on the inner surface and</td>
</tr>
<tr>
<td>Goblet (food service)</td>
<td>Lacquered wooden, flat bottomed, goblet. The goblet has a circular base, with a thin lathe trimmed stem rising from the centre. The body of the goblet is</td>
</tr>
<tr>
<td>Container (containers); teapot (food service); wine jug (jug (narcotics &amp; intoxicants: drinking))</td>
<td>Metal container, possibly a tea or wine pot, with a narrow sp</td>
</tr>
<tr>
<td>Tweezers (food service)</td>
<td>A pair of metal tweezers, made of copper alloy.</td>
</tr>
<tr>
<td>Container (containers)</td>
<td>Inscription (label): A glass vessel made of brown glass. The vessel has a round base, a bulging, fluted body and a flared rim. The neck is decorated with</td>
</tr>
<tr>
<td>Food container (food processing &amp; storage)</td>
<td>A soft or sugar container, dungho, which is hung in the kitchen. It is made of a piece of bamboo, with a hole in the side</td>
</tr>
<tr>
<td>Lid (food processing &amp; storage); food container (food processing &amp; storage)</td>
<td>Round food container lid made from copper alloy. The lid is</td>
</tr>
<tr>
<td>Box (food processing &amp; storage)</td>
<td>Top compartment of 3-tiered circular food box made of bamboo and wood. The interior and underside are covered in</td>
</tr>
</tbody>
</table>

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

**Figure 4 BE-MRAC: Ethnographic Collection, Old Photos**
Figure 5 BE-CAG: Mostly Modern Food-related Artefacts

- **Recipient to pour beer**
  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.

- **Verpakking Le Pastourel Dubbel Creme**
  Dit is een verpakking van Le Pastourel Dubbel Creme. Ze bevat 4% tot 52 procent.

- **Thuis diepvriezen**
  Vanaf de jaren ’50 verschijnt de huisdiepvries eerst op het platteland. De stedelijke gebruikers volgen pas in de jaren ’70.

- **Vruchtenas Libby’s**
  Reclame voor (en consumptie van) groentensappen is veel zeldzamer dan pakweg die voor soep in blik of zakjes. Deze advertentie komt uit het...

- **Recipient to pour beer**
  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.

- **Boterlepel**
  Metalen voorwerp met plastic lemmet om harde boter te schrapen.

- **Voorflap boek Boomstoot**
  Auteurs als Van Hulke bouwen voort op de vergaarde kennis van de “pomologen” uit de achttiende en negentiende

- **Handle to operate a beer tap**
  Faience handle for an English tap system. The handle was placed over a copper bar and fixed with a screwcap. The...

---

Figure 6 UK-HP: London Pubs

- **Portsmouth Arms, Pentonville**
  139/143 Pentonville Road, London This photo was taken by the Charrington Brewery during an architectural survey of their

- **Crown, Billericay**
  High Street, Billericay This photo was taken by the Charrington Brewery during an architectural survey of their pubs. These photographs, now part of the

- **Rabbits, West Thurrock**
  West Thurrock This photo was taken by the Charrington Brewery during an architectural survey of their pubs. These photographs, now part of the

- **Brewer’s Arms, Burgess Hill**
  Burgess Hill This photo was taken by the Charrington Brewery during an architectural survey of their pubs. These

- **All Hallows Club**
  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

- **Gibraltar**
  Gibraltar, Gibraltar Walk, Bethnal Green This photo was taken by the Charrington Brewery during an architectural survey of their pubs. These

- **Man of Kent, Sydenham**
  173 Sydenham Road, SE26 This photo was taken by the Charrington Brewery during an architectural survey of their pubs. These photographs, now

- **King’s Arms, Fulham**
  425 New King’s Road/High Street, Fulham, SW6 4RN This photo was taken by the Charrington Brewery during an architectural survey of their...
**Figure 7 IE-LGMA: Irish Plants and Foods**

- **Stream and plants**
- **Wild mushrooms**
  - There are an abundance of wild mushrooms to be found throughout the forests of Ireland.
- **Beetroot in colourful variety**
  - Wild beets are native in Irish and European shores and come in several different colour variations.
- **Lucky Bag Pie Ingredients**
  - The Lucky Bag Pie is a very versatile dinner because it uses whatever is available as ingredients. Irish mum's love
- **Halloween treats**
  - Children (and adults) enjoy the many kinds of Halloween treats available.
- **Fallen apples**
  - Kingfishers snap and slice the soft flesh of the fallen apple, an important part of many animals’ diet.
- **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.
- **Cabbage**
  - Leafy Cabbage is a traditional Irish vegetable that can be grown for the majority of the year.

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

- **Detail of country hedge with blackberries and view across to Ashdown Forest, Sussex, England**
  - Credit: Marie-Louise Avery / thePictureKitchen / TopFoto
- **Mobile ice cream car**
  - Credit: Marie-Louise Avery / thePictureKitchen / TopFoto
- **Glass half full of lager in sunshine at seaside cafe**
  - Credit: Marie-Louise Avery / thePictureKitchen / TopFoto
- **Quadragesimal 1521 © TopFoto**
- **Quadragesimal 1521 © TopFoto**
- **RITA FROGGARITT BARMAD / 17 MAY 1982**
  - Credit: Marie-Louise Avery / thePictureKitchen / TopFoto
- **Detail of country hedge with blackberries and view across to Ashdown Forest, Sussex, England**
  - Credit: Marie-Louise Avery / thePictureKitchen / TopFoto
- **White lilies in tall vase in pale interior with blurred girl on sofa credt: Marie-Louise Avery / thePictureKitchen / TopFoto**
- **Libyan Market • man drinking**
  - Credit: Libyan Market • man drinking
D3.20d Semantic Demonstrator Extension

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

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

- **Yellow saucer from the 'Novo' range of stackable plastic tableware, designed by David Harman Powell and made by Ekco Plastics.**

- **Royal Oak Gold Plate**
  Table fork from a seven piece cutlery set. **"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, plan engraver.

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

- **(Tasse) goblet Calabre**
  Cup, soft-paste porcelain, plain loop handle, decorated with harbour scenes painted in enamels with gilt on a blue laps ground.

- **Steel, etched, engraved and partly gilt, bone mounts, southern Germany, 1650-1690.**
  Hunting knife, the blade etched with foliage and a hunting scene.

---

**Figure 10 LT-VUFC: Old Lithuanian Recipes**

- **Žaliujų žirnelių (receptas)**
  Gamindama neko kaip ir nekaitinau – išvairu, užvisai, sudėjant viską ką, išorėja ir įkartoja. Įrengia ir įpakaituojasi su kepant braškas kiekiais.

- **Agurkų siruva su inkstais (receptas)**
  Išvairu buliono ir 3 kg mėsos, 2 — 3 riešų ir 3 triukų uždengdavo su prieksnio už. Išvairu bulioną perkiūrį. Pirmą 2—3 raitės

- **Vynuogių sirupas (receptas)**
  1 krontos vynuogių rūšių (pilkė arba paskirtų) arba paprastų uždengdavo 3—4 perkiūrį. Jiegių pirkštas, t. y. "Welsh’s Grape Juice" geras užsakas, 1 ir

- **Buvelės pagal ašikę receptą (receptas)**
  Į 1 1/2 puodų mažės, 1 puodą supjaustyti 1/4 puodą pipirų 2 puodu supjaustyti, išvairę bulionų 1

- **Šaltuotos žemųvoginio kremės (receptas)**
  Išplakti ant leidų putos iš 3—5 stk višnycių grietinės. Išimtų wąsą auginti 200 gr, cukrą su vaistais ir sumažinti

- **Šaltuotos punšos (receptas)**
  Gamindamas iš vaisių ir leidų. Prie pagamintų leidų (1 litro leidų) pripažįsta supjaustyti 3—5 stk vaisių putų, 1/2 butelio baltojo

- **Raugintų kopūstų siruva (receptas)**
  Kaip ir šaltuotojo kopūstų sirvūs, imti kiaušinius buliono, sudėti kopūstai (jei labai rūbštus, reikia neperkioti). Išvairių bulionų

- **Pyragas su imbriu (receptas)**
  1/2 puodą aviečio 2 kiaulienių, 1/2 puodą rodo cukraus 1 puodą vaisių riešai 2 puodų puoštų miltų 2 šaukštuvių imbranų.
<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plate</td>
<td>A bread plate entitled 'Old English' by Barratts of Staffordshire.</td>
</tr>
<tr>
<td>Magic lantern slide</td>
<td>Magic lantern slide depicting a work by William Hogarth.</td>
</tr>
<tr>
<td>Dish</td>
<td>An Indian earthenware dish which is partially restored on the base with concurve rim.</td>
</tr>
<tr>
<td>Snuff Box</td>
<td>Decorated circular snuff-box, painted with a portrait of a Georgian gentleman taking punch. Reads 'Punch to a T.' with gilt interior.</td>
</tr>
<tr>
<td>Bowl</td>
<td>Earthenware bowl with basic lined pattern around eidee.</td>
</tr>
<tr>
<td>Tankard</td>
<td>German gilt tankard and cover repousse or beaten chased work, Embossed with classical figures, mask, fruit and scroll. The knob of the top formed by a</td>
</tr>
<tr>
<td>Vessel, Oinochoe</td>
<td>An oinochoe vessel which is undecorated with a bulbous body, flaring lip and round mouth.</td>
</tr>
<tr>
<td>Vessel, Lekythos</td>
<td>LEKYTHOS, black glazed ware decorated in the Attic Red Figure technique.</td>
</tr>
</tbody>
</table>

*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 publicly 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), 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

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

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

<table>
<thead>
<tr>
<th>Text</th>
<th>Manual</th>
<th>Automatic</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pêcheurs à la ligne</td>
<td>Pêche à la mouche</td>
<td>Pêche_(halieutique) Fishing</td>
<td>Auto is more accurate</td>
</tr>
<tr>
<td>Fishers “on a line”</td>
<td>Fly-fishing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>faire de l'huile alimentaire à base de noix de palme</td>
<td>Huile de palme Palm oil</td>
<td>Alimentation; Huile_alimentaire; Noix_de_coco Food; Edible oil; Coconut</td>
<td>Coconut is wrong, the oil palm (Elaeis) is different from the Coconut palm (Cocos). The other 2 are correct</td>
</tr>
<tr>
<td>preparation of edible oil based on palm nuts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coupe à boire. Coupe polie. Drinking cup. Polished? Cup.</td>
<td>Boisson Drink</td>
<td>(Coupe is common) Man has skipped a concept. Auto is more accurate</td>
<td></td>
</tr>
<tr>
<td>Coupe avec anse. Cup with handle</td>
<td>Poignée Handle</td>
<td>(Coupe is common) Man has skipped a concept. Auto is more accurate</td>
<td></td>
</tr>
<tr>
<td>Couvercle de pot. Lid for a pot</td>
<td>Pot_(récipient)</td>
<td>(Couvercle is common) Man has skipped a concept. Auto is more accurate</td>
<td></td>
</tr>
<tr>
<td>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</td>
<td>Eau Water</td>
<td>Eau_potable Drinking water</td>
<td>Auto is more specific</td>
</tr>
</tbody>
</table>

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>
<thead>
<tr>
<th>Text</th>
<th>Manual</th>
<th>Automatic</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kasai; Luangue (Region between Kasai river and Luangue river)</td>
<td>Lwange (cours d'eau)</td>
<td></td>
<td>(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)</td>
</tr>
<tr>
<td>Congo; Kasai</td>
<td>République_démocratique_du_Congo;</td>
<td>Congo_(fleuve); Kasaï_(rivière)</td>
<td>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</td>
</tr>
<tr>
<td>Central Africa</td>
<td>Afrique_centrale</td>
<td>Afrique</td>
<td>Auto is less specific</td>
</tr>
<tr>
<td>Congo; Uele</td>
<td>République_démocratique_du_Congo</td>
<td>Uele_(rivière)</td>
<td>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)</td>
</tr>
</tbody>
</table>

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](image)

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](image)
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.

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.:

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

# aliases (redirects)
dbr:Csacaval rdfs:label "Csacaval";
   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:Kaskaval rdfs:label "Kaskaval";
   dbo:wikiPageRedirects dbr:Kashkaval.
dbr:Kaskaval rdfs:label "Kaškaval";
   dbo:wikiPageRedirects dbr:Kashkaval.
```

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](image)

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

```sparql
construct {
} where {
  {select ?cho ?place ?name
   {average(?rand_lat) as ?rand_lat} {average(?rand_long) as ?rand_long} {
     ?cho a edm:ProvidedCHO; dct:spatial ?place.
     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.
select ?cho ?title ?lat ?long ?place_name {
  filter not exists {?place a ?type} # bit of a dirty hack, but so what
}

**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.

![View on Map Button](image)

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
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:

```
<http://example.org/objects/1000>
  dct:spatial dbr:London, dbr:Seattle;
```

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)

### 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, e.g.

```
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

*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 sem app
- 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>
<thead>
<tr>
<th>Table 6 Number of Visitors and Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Reported Year 2015</td>
</tr>
<tr>
<td>First visit 29 Oct 2015 - 06:52</td>
</tr>
<tr>
<td>Last visit 28 Dec 2015 - 22:53</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Unique visitors</td>
</tr>
<tr>
<td>Viewed traffic * &lt;= 209 340</td>
</tr>
<tr>
<td>(1.62 visits/visitor) (13.82 Pages/Visit) (29.26 Hits/Visit) (708.24 KB/Visit)</td>
</tr>
<tr>
<td>Not viewed traffic *</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Reported Year 2016</td>
</tr>
<tr>
<td>First visit 01 Jan 2016 - 09:49</td>
</tr>
<tr>
<td>Last visit 19 Jul 2016 - 20:40</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Unique visitors</td>
</tr>
<tr>
<td>Viewed traffic * &lt;= 361 589</td>
</tr>
<tr>
<td>(1.63 visits/visitor) (122.59 Pages/Visit) (155.88 Hits/Visit) (3115.45 KB/Visit)</td>
</tr>
<tr>
<td>Not viewed traffic *</td>
</tr>
</tbody>
</table>

* 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.
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.

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

<table>
<thead>
<tr>
<th>Countries</th>
<th>Regions</th>
<th>Cities: 37</th>
<th>Hits</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulgaria</td>
<td>Grad Sofiya</td>
<td>Sofia</td>
<td>1103</td>
<td>11 %</td>
</tr>
<tr>
<td>Greece</td>
<td>Attiki</td>
<td>Athens</td>
<td>756</td>
<td>7.5 %</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Noord-Holland</td>
<td>Amsterdam</td>
<td>586</td>
<td>5.8 %</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Zuid-Holland</td>
<td>Den haag</td>
<td>300</td>
<td>3 %</td>
</tr>
<tr>
<td>France</td>
<td>Ile-de-France</td>
<td>Montrouge</td>
<td>173</td>
<td>1.7 %</td>
</tr>
<tr>
<td>Great Britain</td>
<td>Buckinghamshire</td>
<td>Gawcott</td>
<td>114</td>
<td>1.1 %</td>
</tr>
<tr>
<td>United States</td>
<td>California</td>
<td>Mountain view</td>
<td>107</td>
<td>1 %</td>
</tr>
<tr>
<td>United States</td>
<td>Massachusetts</td>
<td>Lynn</td>
<td>94</td>
<td>0.9 %</td>
</tr>
<tr>
<td>France</td>
<td>Ile-de-France</td>
<td>Paris</td>
<td>93</td>
<td>0.9 %</td>
</tr>
<tr>
<td>Great Britain</td>
<td>Cambridgeshire</td>
<td>Cambridge</td>
<td>85</td>
<td>0.8 %</td>
</tr>
<tr>
<td>United States</td>
<td>South Carolina</td>
<td>Duncan</td>
<td>85</td>
<td>0.8 %</td>
</tr>
<tr>
<td>Hungary</td>
<td>Heves</td>
<td>Gyöngyös</td>
<td>60</td>
<td>0.6 %</td>
</tr>
<tr>
<td>Great Britain</td>
<td>Essex</td>
<td>Chelmsford</td>
<td>58</td>
<td>0.5 %</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>Luxembourg</td>
<td>Schifflange</td>
<td>48</td>
<td>0.4 %</td>
</tr>
<tr>
<td>United States</td>
<td>New Jersey</td>
<td>Woodbridge</td>
<td>48</td>
<td>0.4 %</td>
</tr>
<tr>
<td>India</td>
<td>Maharashtra</td>
<td>Mumbai</td>
<td>40</td>
<td>0.4 %</td>
</tr>
<tr>
<td>Great Britain</td>
<td>London, City of London</td>
<td>37</td>
<td>0.3 %</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Baden-Wurttemberg</td>
<td>Karlsruhe</td>
<td>36</td>
<td>0.3 %</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Noord-Holland</td>
<td>Amstelveen</td>
<td>36</td>
<td>0.3 %</td>
</tr>
<tr>
<td>Belgium</td>
<td>Brabant</td>
<td>Braine-f’alleud</td>
<td>35</td>
<td>0.3 %</td>
</tr>
</tbody>
</table>
### Table 8 City Distribution of Visitors, 2016

<table>
<thead>
<tr>
<th>Countries</th>
<th>Regions</th>
<th>Cities: 83</th>
<th>Hits</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>Oost-Vlaanderen</td>
<td>Sleidinge</td>
<td>34</td>
<td>0.3 %</td>
</tr>
<tr>
<td>Greece</td>
<td>Thessaloniki</td>
<td>Thessaloniki</td>
<td>32</td>
<td>0.3 %</td>
</tr>
<tr>
<td>Sweden</td>
<td>Varmlands Lan</td>
<td>Torsby</td>
<td>32</td>
<td>0.3 %</td>
</tr>
<tr>
<td>Belgium</td>
<td>Antwerpen</td>
<td>Antwerp</td>
<td>30</td>
<td>0.3 %</td>
</tr>
<tr>
<td>United States</td>
<td>Delaware</td>
<td>Wilmington</td>
<td>28</td>
<td>0.2 %</td>
</tr>
<tr>
<td>United States</td>
<td>New York</td>
<td>Staten island</td>
<td>28</td>
<td>0.2 %</td>
</tr>
<tr>
<td>United States</td>
<td>California</td>
<td>San francisco</td>
<td>13</td>
<td>0.1 %</td>
</tr>
<tr>
<td>Germany</td>
<td>Mecklenburg-Vorpommern</td>
<td>Kiez</td>
<td>7</td>
<td>0 %</td>
</tr>
<tr>
<td>United States</td>
<td>Ohio</td>
<td>Columbus</td>
<td>5</td>
<td>0 %</td>
</tr>
<tr>
<td>Japan</td>
<td>Osaka</td>
<td>Osaka</td>
<td>4</td>
<td>0 %</td>
</tr>
<tr>
<td>United States</td>
<td>Texas</td>
<td>Mcallen</td>
<td>3</td>
<td>0 %</td>
</tr>
<tr>
<td>United States</td>
<td>Washington</td>
<td>Seattle</td>
<td>3</td>
<td>0 %</td>
</tr>
<tr>
<td>United States</td>
<td>District of Columbia</td>
<td>Washington</td>
<td>2</td>
<td>0 %</td>
</tr>
<tr>
<td>United States</td>
<td>Indiana</td>
<td>Indianapolis</td>
<td>2</td>
<td>0 %</td>
</tr>
<tr>
<td>Poland</td>
<td>Malopolskie</td>
<td>Kraków</td>
<td>1</td>
<td>0 %</td>
</tr>
<tr>
<td>Germany</td>
<td>Brandenburg</td>
<td>Potsdam</td>
<td>1</td>
<td>0 %</td>
</tr>
<tr>
<td>Japan</td>
<td>Okayama</td>
<td>Tama</td>
<td>1</td>
<td>0 %</td>
</tr>
<tr>
<td>Others/Unknown</td>
<td></td>
<td></td>
<td>5829</td>
<td>58.5 %</td>
</tr>
<tr>
<td>Country</td>
<td>Region</td>
<td>City</td>
<td>Percentage</td>
<td>Score</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------</td>
<td>------------</td>
<td>------------</td>
<td>-------</td>
</tr>
<tr>
<td>United States</td>
<td>New York</td>
<td>New york</td>
<td></td>
<td>56 0 %</td>
</tr>
<tr>
<td>Italy</td>
<td>Toscana</td>
<td>Lucca</td>
<td></td>
<td>48 0 %</td>
</tr>
<tr>
<td>France</td>
<td>Languedoc-Roussillon</td>
<td>Montpellier</td>
<td></td>
<td>46 0 %</td>
</tr>
<tr>
<td>Canada</td>
<td>Quebec</td>
<td>Blainville</td>
<td></td>
<td>45 0 %</td>
</tr>
<tr>
<td>Spain</td>
<td>Castilla-La Mancha</td>
<td>Pantoja</td>
<td></td>
<td>42 0 %</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>Burgas</td>
<td>Karnobat</td>
<td></td>
<td>41 0 %</td>
</tr>
<tr>
<td>France</td>
<td>Ile-de-France</td>
<td>Montrouge</td>
<td></td>
<td>41 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>Delaware</td>
<td>Wilmington</td>
<td></td>
<td>41 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>District of Columbia</td>
<td>Washington</td>
<td></td>
<td>39 0 %</td>
</tr>
<tr>
<td>Belgium</td>
<td>West-Vlaanderen</td>
<td>Kortrijk</td>
<td></td>
<td>38 0 %</td>
</tr>
<tr>
<td>Poland</td>
<td>Mazowieckie</td>
<td>Warsaw</td>
<td></td>
<td>36 0 %</td>
</tr>
<tr>
<td>Germany</td>
<td>Hessen</td>
<td>Marburg</td>
<td></td>
<td>36 0 %</td>
</tr>
<tr>
<td>France</td>
<td>Provence-Alpes-Cote d'Azur</td>
<td>Nice</td>
<td></td>
<td>35 0 %</td>
</tr>
<tr>
<td>Spain</td>
<td>Cataluna</td>
<td>Sant pere de ribes</td>
<td></td>
<td>34 0 %</td>
</tr>
<tr>
<td>Belgium</td>
<td>West-Vlaanderen</td>
<td>Oostende</td>
<td></td>
<td>34 0 %</td>
</tr>
<tr>
<td>Belgium</td>
<td>Antwerpen</td>
<td>Geel</td>
<td></td>
<td>34 0 %</td>
</tr>
<tr>
<td>China</td>
<td>Beijing</td>
<td>Beijing</td>
<td></td>
<td>31 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>North Carolina</td>
<td>Charlotte</td>
<td></td>
<td>29 0 %</td>
</tr>
<tr>
<td>Greece</td>
<td>Imathia</td>
<td>Náousa</td>
<td></td>
<td>29 0 %</td>
</tr>
<tr>
<td>Greece</td>
<td>Iraklion</td>
<td>Iráklion</td>
<td></td>
<td>29 0 %</td>
</tr>
<tr>
<td>Greece</td>
<td>Thessaloniki</td>
<td>Thessaloniki</td>
<td></td>
<td>29 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>California</td>
<td>El segundo</td>
<td></td>
<td>29 0 %</td>
</tr>
<tr>
<td>Belgium</td>
<td>Brussels Hoofdstedelijk Gewest</td>
<td>Brussel</td>
<td></td>
<td>29 0 %</td>
</tr>
<tr>
<td>Spain</td>
<td>Cataluna</td>
<td>Santa maria del camí</td>
<td></td>
<td>28 0 %</td>
</tr>
<tr>
<td>Great Britain</td>
<td>London, City of</td>
<td>London</td>
<td></td>
<td>25 0 %</td>
</tr>
<tr>
<td>Poland</td>
<td>Malopolskie</td>
<td>Kraków</td>
<td></td>
<td>20 0 %</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Utrecht</td>
<td>Utrecht</td>
<td></td>
<td>18 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>Ohio</td>
<td>Columbus</td>
<td></td>
<td>15 0 %</td>
</tr>
<tr>
<td>Italy</td>
<td>Veneto</td>
<td>Rovigo</td>
<td></td>
<td>11 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>New Jersey</td>
<td>Woodbridge</td>
<td></td>
<td>10 0 %</td>
</tr>
<tr>
<td>Germany</td>
<td>Bremen</td>
<td>Bremen</td>
<td></td>
<td>9 0 %</td>
</tr>
<tr>
<td>Spain</td>
<td>Islas Baleares</td>
<td>Palma</td>
<td></td>
<td>9 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>Indiana</td>
<td>Indianapolis</td>
<td></td>
<td>9 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>Georgia</td>
<td>Alpharetta</td>
<td></td>
<td>8 0 %</td>
</tr>
<tr>
<td>Cyprus</td>
<td>Paphos</td>
<td>Páfos</td>
<td></td>
<td>7 0 %</td>
</tr>
<tr>
<td>Canada</td>
<td>British Columbia</td>
<td>Abbotsford</td>
<td></td>
<td>6 0 %</td>
</tr>
<tr>
<td>Canada</td>
<td>Ontario</td>
<td>Ottawa</td>
<td></td>
<td>6 0 %</td>
</tr>
<tr>
<td>Greece</td>
<td>Attiki</td>
<td>Glyfáda</td>
<td></td>
<td>4 0 %</td>
</tr>
<tr>
<td>China</td>
<td>Fujian</td>
<td>Putian</td>
<td></td>
<td>4 0 %</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Zuid-Holland</td>
<td>Delft</td>
<td></td>
<td>4 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>California</td>
<td>San francisco</td>
<td></td>
<td>3 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>Pennsylvania</td>
<td>Malvern</td>
<td></td>
<td>2 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>New York</td>
<td>Tonawanda</td>
<td></td>
<td>2 0 %</td>
</tr>
<tr>
<td>Slovak Republic</td>
<td>Bratislava</td>
<td>Bratislava</td>
<td></td>
<td>1 0 %</td>
</tr>
<tr>
<td>Belgium</td>
<td>Brussels Hoofdstedelijk Gewest</td>
<td>Brussels</td>
<td></td>
<td>1 0 %</td>
</tr>
<tr>
<td>Taiwan</td>
<td>T'ai-pei</td>
<td>Taipei</td>
<td></td>
<td>1 0 %</td>
</tr>
<tr>
<td>Hungary</td>
<td>Budapest</td>
<td>Budapest</td>
<td></td>
<td>1 0 %</td>
</tr>
<tr>
<td>United States</td>
<td>Arizona</td>
<td>Phoenix</td>
<td></td>
<td>1 0 %</td>
</tr>
<tr>
<td>Poland</td>
<td>Świętokrzyskie</td>
<td>Herby</td>
<td></td>
<td>1 0 %</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>Razgrad</td>
<td>Sandrovo</td>
<td></td>
<td>1 0 %</td>
</tr>
</tbody>
</table>
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\(^1\). 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]

---

\(^1\) [http://pro.europeana.eu/blogpost/meet-the-members-council-vladimir-alexiev](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](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/](http://efd.ontotext.com/data/enrichment/)

9.3 Semantic Demonstrator

- **Where does it live?**
  [http://efd.ontotext.com/app](http://efd.ontotext.com/app) is the application
  [http://efd.ontotext.com/sparql](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](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](https://github.com/Ontotext-AD/europeana-food-and-drink).
10 References


[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