



# Using DBpedia in Europeana Food and Drink

**Vladimir Alexiev, Ontotext Corp**

DBpedia meeting, 12 Feb 2016, The Hague

---

# Content

- Europeana Food and Drink
- EFD Semantic Application

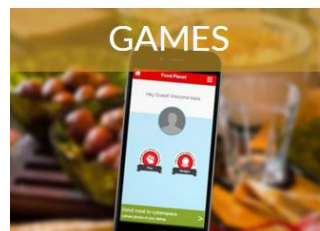
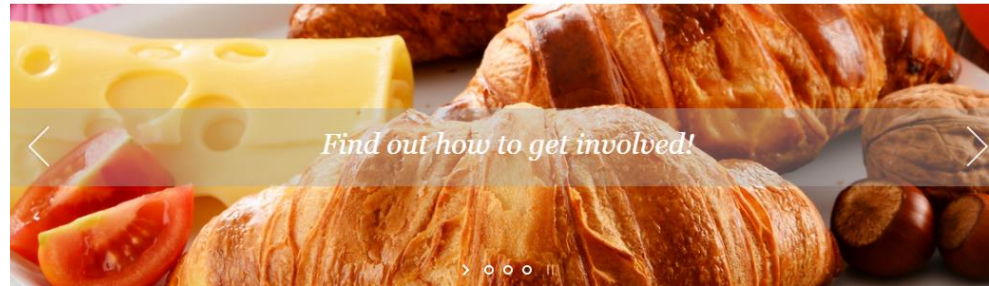
How is it done? Tech part:

- Food and Drink Classification
- Using DBpedia for Enrichment
- Using GeoNames for Parent Places

# Europeana Food and Drink



ABOUT FOOD TRAILS BOOKS PICTURE LIBRARY GAMES CAKE EXHIBITION EDUCATION *Get involved!*



- Collects FD content, makes FD creative applications
- <http://foodanddrinkeurope.eu/>

# EFD Semantic App (Demonstrator)

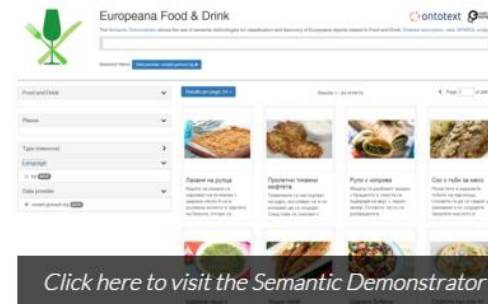


ABOUT FOOD TRAILS BOOKS PICTURE LIBRARY GAMES CAKE EXHIBITION EDUCATION *Get involved!*



The **Semantic demonstrator** demonstrates the use of semantic technologies for classification and discovery of Europeana objects related to Food and Drink; it aims to provide semantic enrichment, i.e. extract references to Food and Drink topics from free text in object metadata to strengthen the relevance of database search results on the theme of Food and Drink.

The application does this by applying a Europeana Food and Drink Classification scheme to Cultural Heritage Objects, such as those stored in **Europeana** to yield more precise groupings, relationships and categorisations for database items.



- Showcases enrichment, faceted semantic search
- [Demo](#), [Data](#), [SPARQL](#). [Homepage](#), [Description](#), [Deliverable](#)



# EFD Sem App Shots



Europeana Food & Drink

ontotext  made with europeana

The Semantic Demonstrator shows the use of semantic technologies for classification and discovery of Europeana objects related to Food and Drink. [Detailed description](#), [data](#), [SPARQL endpoint](#).

Selected filters: [Data provider: recepti.gotvach.bg](#) ✕

Food and Drink ▾

Places ▾

Type (resource) >

Language ▾

☐ bg 6419


Data provider ▾

☒ recepti.gotvach.bg 6419


Results per page: 24 ▾

Results 1 - 24 of 6419


< Page 1 of 268 >




**Лазаня на рулца**  
Корите за лазаня се нарязват на по-малки с ширина около 5 см и дължина колкото е парчето на бекона, отгоре се




**Пролетни тиквени кюфета**  
Тиквичките се настъргват на едро, посоляват се и се изчакват да се отцедят. След това се смесват с




**Руло с коприва**  
Яйцата се разбиват заедно с брашното и сместа се подправя на вкус с черен пипер. Готовото тесто се разпределя в




**Сос с гъби за месо**  
Почистете и нарежете гъбите на парченца. Сложете ги да се сварят до омекване и ги отцедете. Загрейте маслото в




**Шарена каша с**



**Чушки пане**



**Шарена бобена**



**Спагети със сос от**

- Here: Bulgarian recipes
- But this is by dataProvider, not through semantic enrichment
- Only a few BG enrichments: will add 300+ term→DBpedia mappings by regexp, as no BG NLP is planned

# EFD Sem App Shots



## Europeana Food & Drink



The Semantic Demonstrator shows the use of semantic technologies for classification and discovery of Europeana objects related to Food and Drink. [Detailed description, data, SPARQL endpoint.](#)

Selected filters: **FD: Fishing** **Place: Oceania**

Food and Drink

Agriculture 107

Beverages 3

Cuisine 178

Eating behaviors 178

Food and drink by country 1

Food and drink preparation 179

Food and drink terminology 3

Food culture 178

Food industry 36

Food politics 8

Food safety 5

Places

Africa 6

Americas 2

Asia 59

Europe 49

North America 4

Oceania 179

Southern Ocean 1


Type (resource)

Language


Results per page: 24

Results 25 - 48 of 179


Page 2 of 8



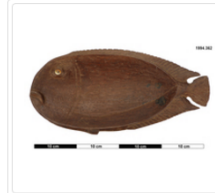
shark hook  
Shark hook made from metal with plant fibre line.



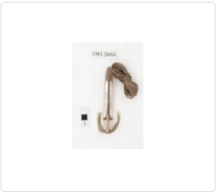
fish spear (spear (hunting, fishing & trapping))  
Ornamental fish spear made of wood and decorated with brown, black and white pigment.




fish hook  
Shark hook constructed from on elbow hafted piece of hard wood to which a barbed, potentially human, bone point has been socketed and bound, the shank



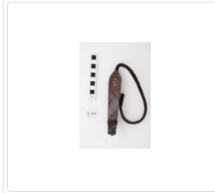
bowl (food service)  
A fish shaped bowl made of wood, which is probably a piece of tourist art. It is quite shallow with a down curved mouth and the opercula of a Turbo




fish hook  
Medium sized fish hook with line.



bowl  
Collectors' Gallery (exhibition), Frigate-Bird (Fregata)



fish hook  
Collectors' Gallery (exhibition), Fish hook made of mother of



fish hooks  
Two fish hooks.

- Fishing objects from Oceania.
- Semantic enrichment of objects using Concept Extraction (EN only)
- Like a Mini-Europeana, but adds 2 semantic hierarchical facets: FD topic and Place

# EFD Sem App Shots



## Europeana Food & Drink

The [Semantic Demonstrator](#) shows the use of semantic technologies for classification and discovery of Europeana objects related to Food and Drink. [Detailed description](#), [data](#), [SPARQL endpoint](#).

Selected filters: [Category: Beverages](#) [Place: Roman Empire](#) [Alinari](#)

Food and Drink

+ Agriculture 2

+ Beverages 4

+ Eating behaviors 1

+ Food and drink preparation 2

+ Food culture 1

+ Food industry 1

+ Foods 4

+ Gustation 1

+ Nutrition 2

Places

+ Africa 4

+ Asia 4

+ Earth 4

+ Europe 4

+ Southern Europe 4

+ Western Europe 1

Type (resource)


>

Language


>

Results per page: 24


Results 1 - 4 of 4




Milk jugs and other containers for domestic use from Pompeii and Herculaneum, now kept at the National Archaeological Museum



The house of the Dolii, named after its thirty five large terracotta oil and wine jars stored in a pit. Ostia Antica  
WARNING: Permission must



Shepherd milking a goat; relief preserved in the Roman National Museum, in Rome. The work was displayed at the Augustea Exhibit of



Ship laden with wine, relief preserved in Treviri. Work displayed at the Augustea Exhibit of the Roman World, held in Rome in 1937-

- Objects from Fratelli Alinari related to the Roman Empire and Beverages
- Semantic (conceptual) search / browse, query expansion

---

Tech part

# HOW IS IT DONE?



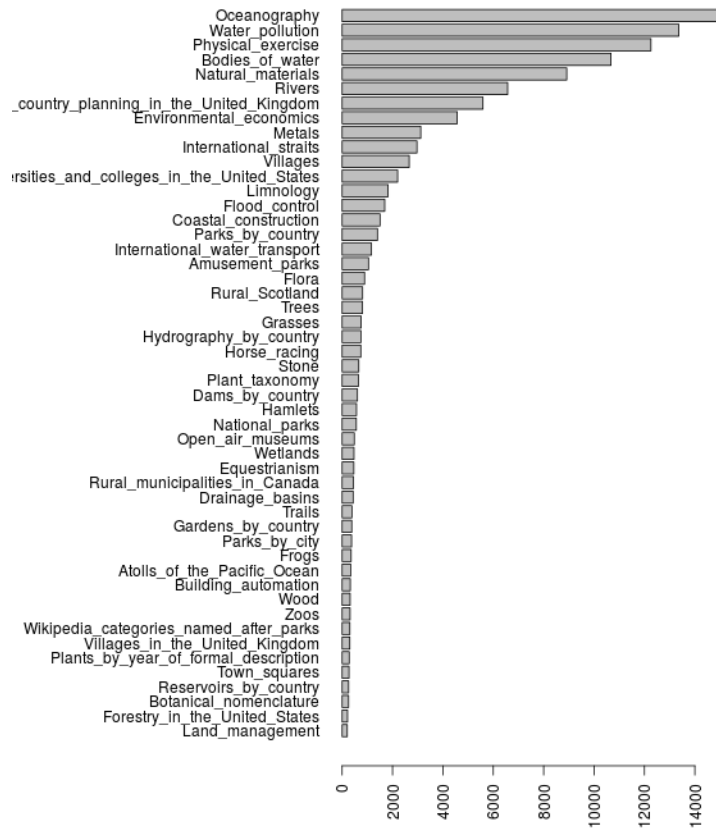
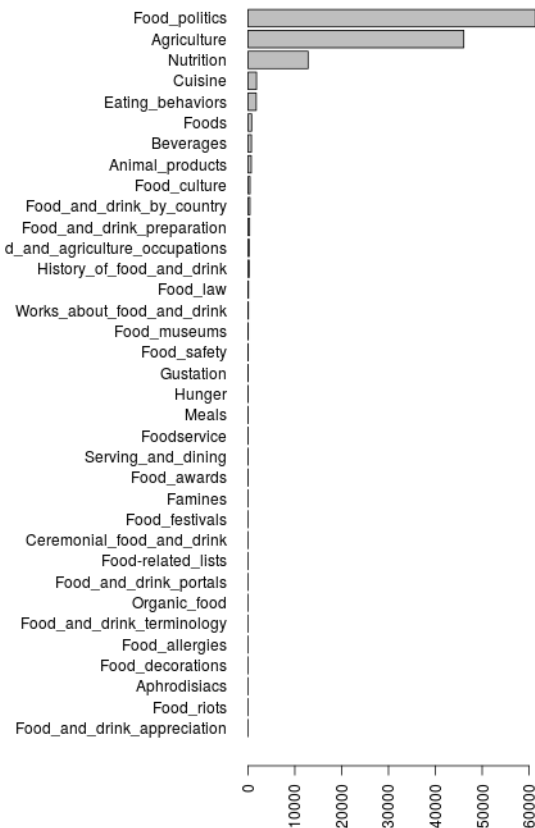
---

# Food and Drink Classification

- FD is such a broad topic: how to find classifications for it?
  - Main goal: FD gazetteer to recognize FD topics in CH objects
  - Studied a number of LOD datasets: USDA Standard Reference, FAO AgroVoc, EC EuroVoc, Getty AAT, UMBEL SuperTypes, Yago WordNet Domains, WiBi, DBTax, Wiktionary
  - Concluded we should use Wikipedia Categories
  - Use through DBpedia
  - Wikidata has expunged Categories as useless / too unorganized
  - To complement with Wikipedia lists (extend Extraction Framework), UMBEL, DBTax
- Also do Place enrichment using DBpedia/GeoNames
- EFD Classification: [Deliverable](#), [Slides](#)

# Wikipedia FD Categories

- Starting from dbc:Food\_and\_drink, you reach 887k cats, 26 levels deep, representing 80% of all categories
- Most are irrelevant to FD (left: level 2, right: level 5)



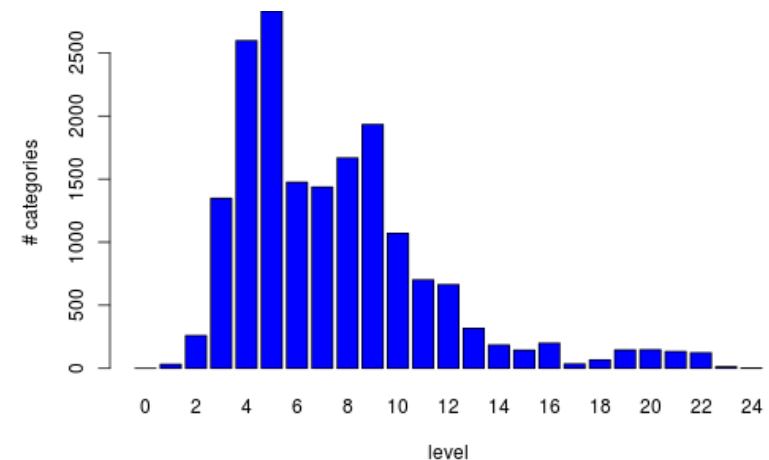
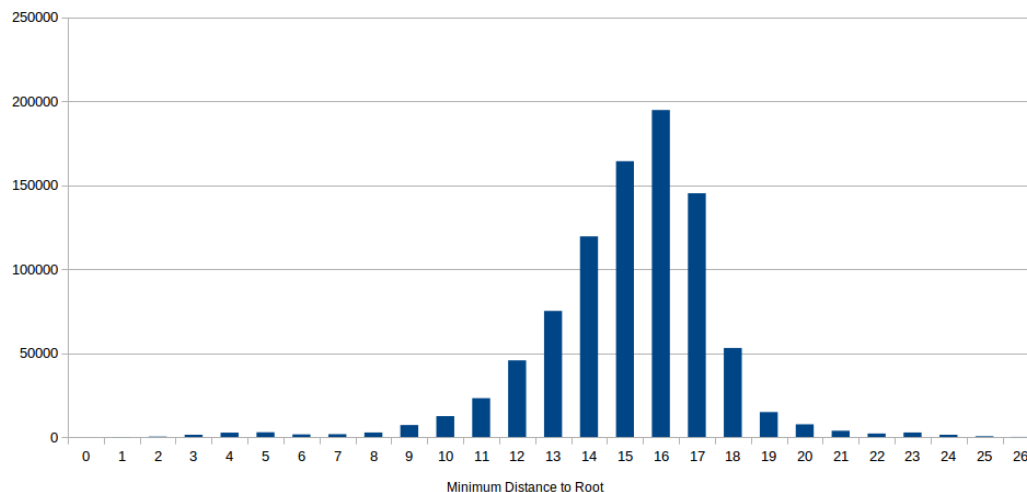
# Pruning the FD Category Tree

→ Mix of approaches:

- Statistical analysis of the category network
- Manual curation (chopping out irrelevant branches: about 300)
- Evidence-based feedback (CHOs, UMBEL, DBTax)

→ Reduced cats by 98%, level from 16 to 5 (reduce semantic drift)

- Left relevant: 17.5k cats, 221k articles, ~900k labels



→ Paper: “Domain-specific modeling: Towards a Food and Drink Gazetteer” (IKC 2015): [paper](#), [slides](#).

- Extended version (LNCS 9398): [preprint](#), [published](#)

# Using DBpedia, Editing Wikipedia

- DBpedia is a great integrated resource that we used for Categories (cat hierarchy, art→cat classification), Places, labels
- When something is missing or wrong, don't just cry! Add it to Wikipedia, e.g.
  - dbc:Bottles: added parent dbc:Drink\_containers
  - dbr:Gourd: added category dbc:Bottles
  - “Muller” (vessel for making mulled beer or wine): added as alias (Redirect) of dbr:Mulled\_wine
  - Made article dbr:Shepherd's\_crook (cut & paste)
- You learn a lot of curious facts in the process!
  - Did you know the Crook and flail are the insignia of pharaonic authority?



---

# Using GeoNames for Parent Places

- Surprisingly, DBpedia doesn't have good Parent Place info:
  - No statement that Bulgaria and France are part of Europe.
  - Have classes `yago:MemberStatesOfTheEuropeanUnion`, `yago:EuropeanCountries`, `yago:EuropeanUnionMemberEconomies`, but no relation to `dbr:Europe`
- No uniform property. E.g. `dbo:City` has **`dbo:region`** and **`dbo:country`**, while `dbo:Island` (e.g. `dbr:Andaman_Islands`) has:
  - **`dbo:archipelago`** `dbr:Andaman_and_Nicobar_Islands` (physical parent)
  - **`dbo:location`** `dbr:Bay_of_Bengal` (physical parent)
  - **`dbp:countryAdminDivisions`** `dbr:Andaman_and_Nicobar_Islands` (admin parent)
  - **`dbo:country`** `dbr:India` (admin ancestor)
  - **`dbo:capital`** `dbr:Port_Blair` (admin child)
  - **`dbo:majorIsland`** `dbr:North_Andaman_Island`, `dbr:South_Andaman_Island` (physical child, partial)
- So we decided to use GeoNames, which has uniform property `gn:parentFeature`

# Using GeoNames for Parent Places

- GeoNames → Wikipedia/DBpedia links
  - Got from geonames.org, used script from Dbpedia
  - Coverage is about 65% (Wikidata has similar coverage)
  - Added 250 links (e.g. dbr:Sloane\_Street  
gn:parentFeature  
dbr:Royal\_Borough\_of\_Kensington\_and\_Chelsea)
- About 10 fixes to GeoNames on their site (e.g. North & South America were children of “America”...a small village)
- Loaded GeoNames & DBpedia in Ontotext GraphDB, used owl:sameAs optimization to merge (smush) the corresponding nodes
- Voila! A hierarchical semantic facet
  - Faceting is done by ElasticSearch, we use the GraphDB-ElasticSearch Connector

[-] Africa	1015
[+] Ancient Rome	4
[+] Byzantine Empire	1
[+] Congo Basin	2
[+] East Africa	160
[-] Eastern Africa	315
[+] Burundi	1
[+] Eritrea	4
[+] Ethiopia	72
[+] Kenya	77
[+] Madagascar	25





**Thanks for your time!**



Project co-funded by the European Union under the ICT Policy Support Programme

