

# The French Enlightenment Novel as a Graph

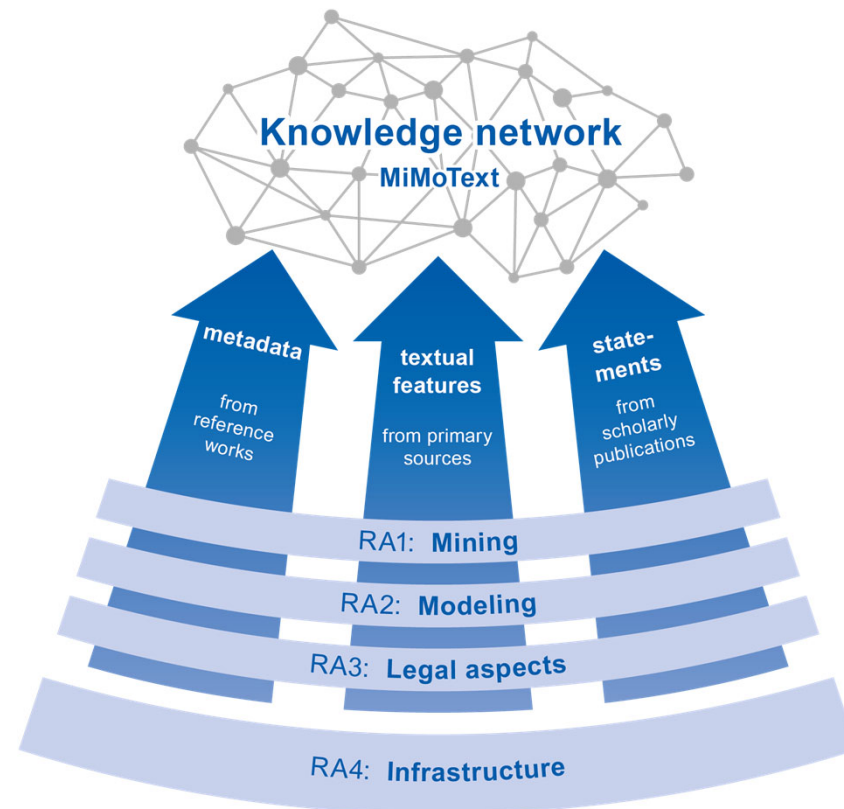
Potentials and Challenges in the Construction of a Knowledge Network

[Maria Hinzmann<sup>1</sup>, Julia Röttgermann<sup>1</sup>, Anne Klee<sup>1</sup>, Moritz Steffes<sup>1</sup>, and Christof Schöch<sup>1</sup>](#)

<sup>1</sup>Trier Center for Digital Humanities, Universität Trier



# Mining and Modeling Text



# The French Enlightenment Novel as a Graph



## Bibliographie du genre romanesque français 1751-1800

97.22 DIDEROT, Denis  
 La religieuse, par Diderot  
 Paris, Buisson, an cinquième de la République, in-8, 411p.  
 BM BN†  
 JG I 91; JP 22 vend. V  
 Gay Q Saint-Fargeau Tchermersine  
 Gay donne 'Dufart' comme éditeur.  
 Cet ouvrage a paru à la fin de 1796, date que donnent les bibliographes. Nous le classons ici en raison de la date révolutionnaire.  
 Une version remaniée et raccourcie de ce roman a été publiée dans un recueil qui a pour titre: *Amours, galanteries, intrigues, ruses et crimes des Capucins ...*, Amsterdam & Paris, 1788 (BM†, BN†). Mais la date de 1788 est fautive, et il s'agit d'une compilation du XIXe siècle.

*1re personne (mémoires); Paris, province; Marie-Suzanne Simonin, madame de Moni, la sœur Sainte-Christine, la sœur Ursule, M. Manouri; les malheurs de la victime d'une claustration forcée; thèmes philosophiques et humanitaires.*

### Autres éditions:


- Une contrefaçon anglaise de la première édition est signalée par Giles Barber, 'A note on the English piracies of *La Religieuse* and *Jacques le fataliste*, 1797', *Diderot Studies*, XVI, 1973, pp. 15 et suiv.
- Paris, Gueffier jeune, Knapen fils, an cinquième (1796 v. st.), 2t. in-18, 247, 244p. (BN†, JP 4 niv. V, Gay, Tchermersine)
- Paris, Ancelle, an V (1797), 2t. in-18, viii + 172, 170p. (Oront†)
- Paris, Taillard, 1797, in-18, 254p. (Bruxelles)
- *La religieuse, ouvrage posthume de Diderot*, Paris, chez les marchands de nouveautés, 1797, pet. in-8, 288p. (WU, Veillées I 1798 i 97-107, Tchermersine)

59.25 VOLTAIRE, François-Marie Arouet de  
 Candide ou l'Optimisme, traduit de l'allemand de Mr. le docteur Ralph  
 1759, in-12  
 BN  
 AL 1759 II 203-210; AT 1761 (1759); CorrL mars 1759  
 Bengesco Dufrenoy Gay Morize Q  
 Il paraît y avoir eu jusqu'à une vingtaine d'éditions datées de 1759. Sur la question de la véritable édition *princeps*, voir Bengesco; Morize; I.O. Wade, *Voltaire et Candide*, Princeton, 1959; B. Gagnebin, ds *Bulletin du bibliophile*, 1960, pp. 22-31; J.-D. Candaux, ds *Studies on Voltaire*, XVIII, 1961, pp. 173-178.  
 3e personne; Europe, Amérique; Candide, Cunégonde, Pangloss, Martin; voyages, aventures romanesques, désastres; thèmes philosophiques, ton satirique.

# Mining: primary texts (novels)

https://github.com/MiMoText/NER\_novels

raw_results	use subfolders for results and manually edited files	10 days ago
scripts	use raw_results/ as output folder for script	10 days ago
.gitignore	Initial commit	15 days ago
LICENSE	Initial commit	15 days ago
README.md	Img URL	21 hours ago

☰ README.md 

## NER novels

Perform Named Entity Recognition (NER) on french novels from the roman18 corpus with the help of SpaCy.

### Description

The roman18-corpus of about 100 eighteenth-century French novels in full text is analysed via SpaCy's named entity recognition. Named entity recognition (NER) is a popular information retrieval technique "to identify and segment named entities and classify or categorize them under various predefined classes" (Sarkar, 2019).

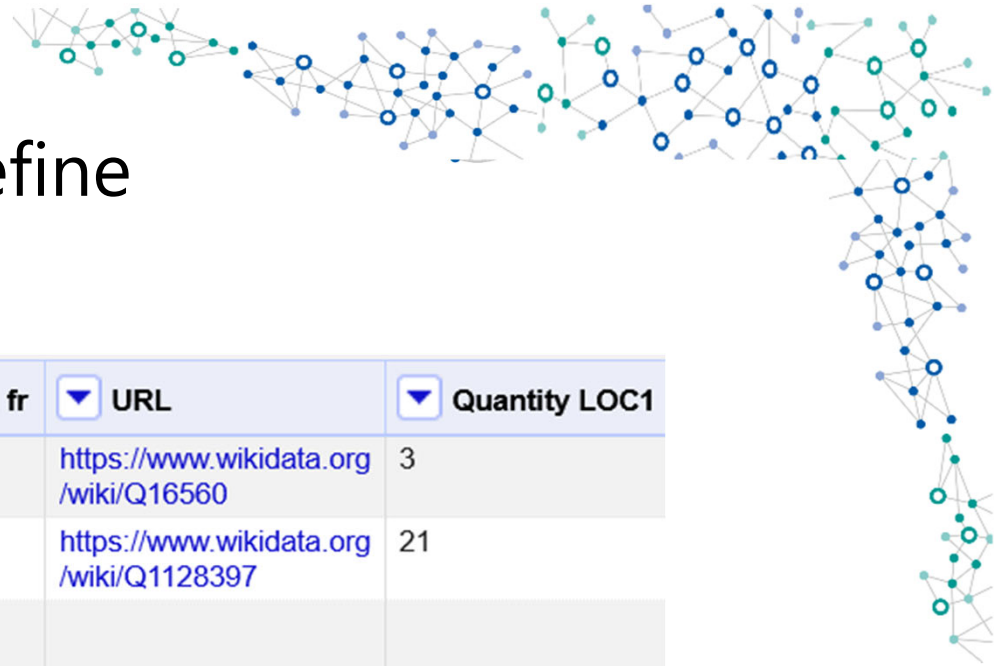
[Named entity recognition](#)

Within the French language package of SpaCy one can distinguish the following types of named entities: LOC, PER, MISC and ORG entities. The five most common "LOC" (location) entities within each novel and their numerical occurrences per text are extracted.

### Structure of this repository

The script which performs the NER and its technical documentation is under `scripts/`. The latest results of this script are stored under `raw_results/`, although this is configurable. These results still need manual corrections and supplements to be suitable for import into the [MiMoText Wikibase](#). The latest version of this is under `edited_for_import/`. In this folder there is also a file named `edits_openrefine.txt`. This has been generated by [OpenRefine](#) which is used for semi-automatic reconciliation and includes all the editing steps in this tool.

# Wikidata Reconciliation via OpenRefine



<input type="checkbox"/> All	<input type="checkbox"/> Column 1	<input type="checkbox"/> bgrf	<input type="checkbox"/> LOC1	<input type="checkbox"/> Wikidata Label fr	<input type="checkbox"/> URL	<input type="checkbox"/> Quantity LOC1
<input type="checkbox"/> <input type="checkbox"/>	1. Abbes_Voyage	58.5	Palais	palais	<a href="https://www.wikidata.org/wiki/Q16560">https://www.wikidata.org/wiki/Q16560</a>	3
<input type="checkbox"/> <input type="checkbox"/>	2. Anonym_Suzon	83.9	Couvent	couvent	<a href="https://www.wikidata.org/wiki/Q1128397">https://www.wikidata.org/wiki/Q1128397</a>	21
<input type="checkbox"/> <input type="checkbox"/>	3. Anonyme_Zoloe	00.37				
<input type="checkbox"/> <input type="checkbox"/>	4. Arnaud_Epoux	83.15	Paris	Paris	<a href="https://www.wikidata.org/wiki/Q90">https://www.wikidata.org/wiki/Q90</a>	27
<input type="checkbox"/> <input type="checkbox"/>	5. Arnaud_Matinees	99.43				
<input type="checkbox"/> <input type="checkbox"/>	6. Arnaud_Sentiment	70.21	Nancy	Nancy	<a href="https://www.wikidata.org/wiki/Q40898">https://www.wikidata.org/wiki/Q40898</a>	111
<input type="checkbox"/> <input type="checkbox"/>	7. Barthelemy_Voyage	88.27	Grèce	Grèce	<a href="https://www.wikidata.org/wiki/Q41">https://www.wikidata.org/wiki/Q41</a>	440

# Open Source



# Infrastructure: Wikibase, Pywikibot

## Le Micromégas de Mr. de Voltaire (Q993)

No description defined

▼ In more languages

Configure

Language	Label	Description	Also known as
English	Le Micromégas de Mr. de Voltaire	No description defined	
German	Le Micromégas de Mr. de Voltaire	No description defined	

All entered languages

### Statements

BGRF ID	52.43	▼ 0 references
author	VOLTAIRE, François-Marie Arouet de	▼ 0 references
title	Le Micromégas de Mr. de Voltaire (French)	▼ 0 references
instance of	literary work	▶ 1 reference

narrative location	Sirius	▶ 2 references
	Saturn	▶ 2 references
	Earth	▶ 1 reference
	Jupiter	▶ 1 reference
	Baltic Sea	▶ 1 reference
place of publication	London	▼ 0 references
place of publication_string	Londres (French)	▼ 0 references





# Approaches to mining and modeling ‚narrative locations‘

The identification of place names in literary texts poses several challenges:

- Ambiguity (Heuser et al., 2016; Jockers, 2016; Nielsen, 2016).
- Geo-)Localization and fictionality (Nielsen 2016)
- Extraction of ‚place names‘ and modeling ‚narrative locations‘ (Pier 2014; Heuser et al. 2016)
- Historicity

# ambiguity

Candide **PER** , recommandons-nous à la **Providence LOC** . » Ils voguèrent quelques lieues entre des bords tantôt fleuris, tantôt arides, tantôt unis, tantôt escarpés. La rivière s'élargissait toujours; enfin elle se perdait sous une voûte de rochers épouvantables qui s'élevaient jusqu'au ciel. Les deux voyageurs eurent la hardiesse de s'abandonner aux flots sous cette voûte. Le fleuve, resserré en cet endroit, les porta avec une rapidité et un bruit horrible.

# geo-localization and fictionality

CHAPITRE 2 conversation de **Micromégas PER** avec l' habitant de **Saturne LOC** .

après que son excellence se fut couchée, et que le secrétaire se fut approché de son visage: il faut avouer, dit **Micromégas PER** , que la nature est bien variée. Oui, dit le saturnien, la nature est comme un parterre dont les fleurs... ah! Dit l' autre, laissez-là votre par-terre. Elle est, reprit le secrétaire, comme une assemblée de blondes et de brunes dont les parures... et qu' ai-je affaire de vos brunes? Dit l' autre. Elle est comme une galerie de peintures dont

# historicity

« Quand les premiers ravages de cette épouvantable peste furent passés, on vendit les esclaves du dey. Un marchand m'acheta et me mena à **Tunis LOC**, il me vendit à un autre marchand, qui me revendit à **Tripoli LOC**; de **Tripoli LOC** je fus revendue à **Alexandrie LOC**, d' **Alexandrie LOC** revendue à **Smyrne LOC**, de **Smyrne LOC** à **Constantinople LOC**. J'appartins enfin à un aga des janissaires, qui fut bientôt



„If we want to understand the system in its entirety, we must accept losing something.“

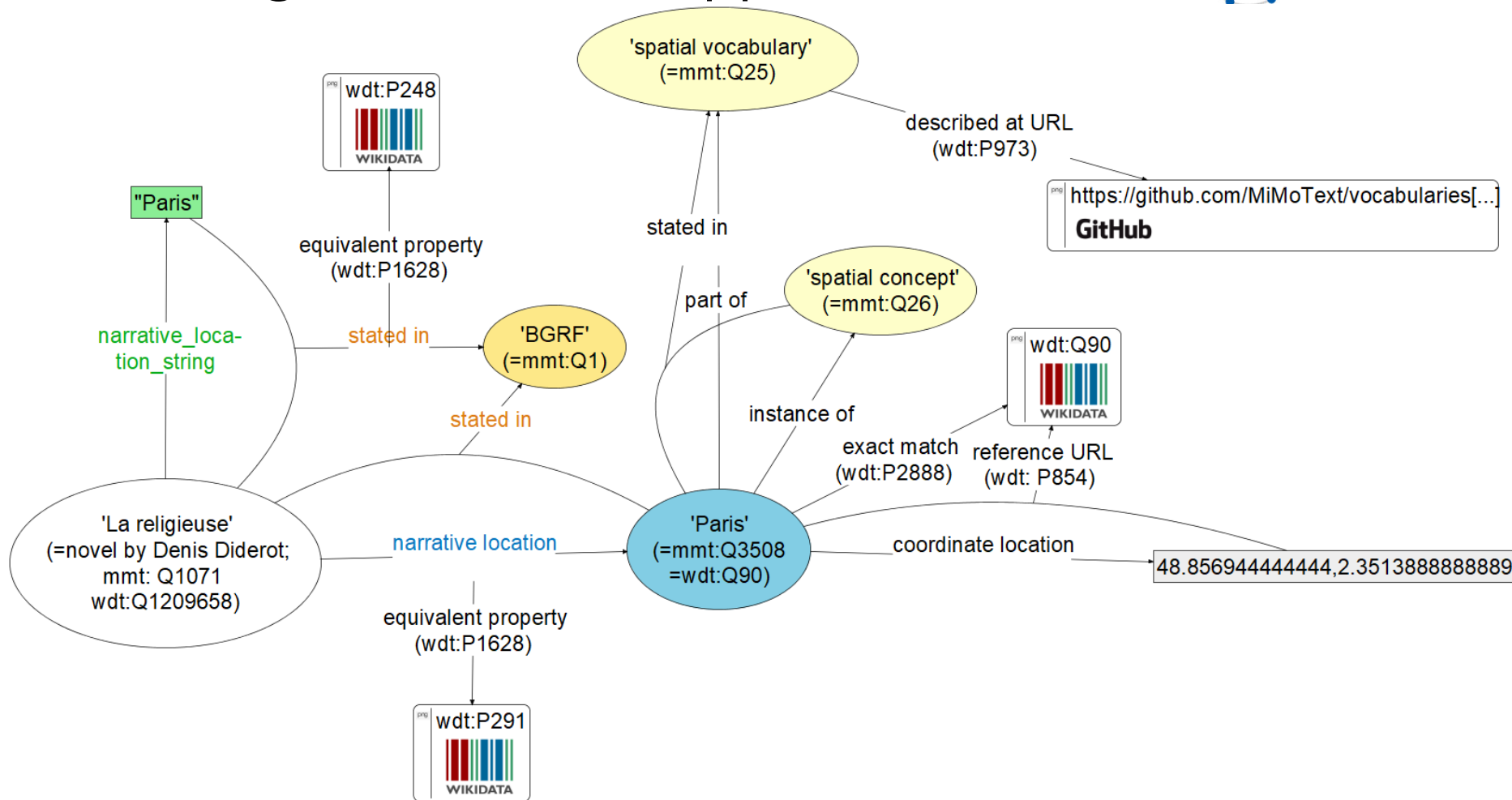
(Moretti, 2013, p. 49)

## Second part

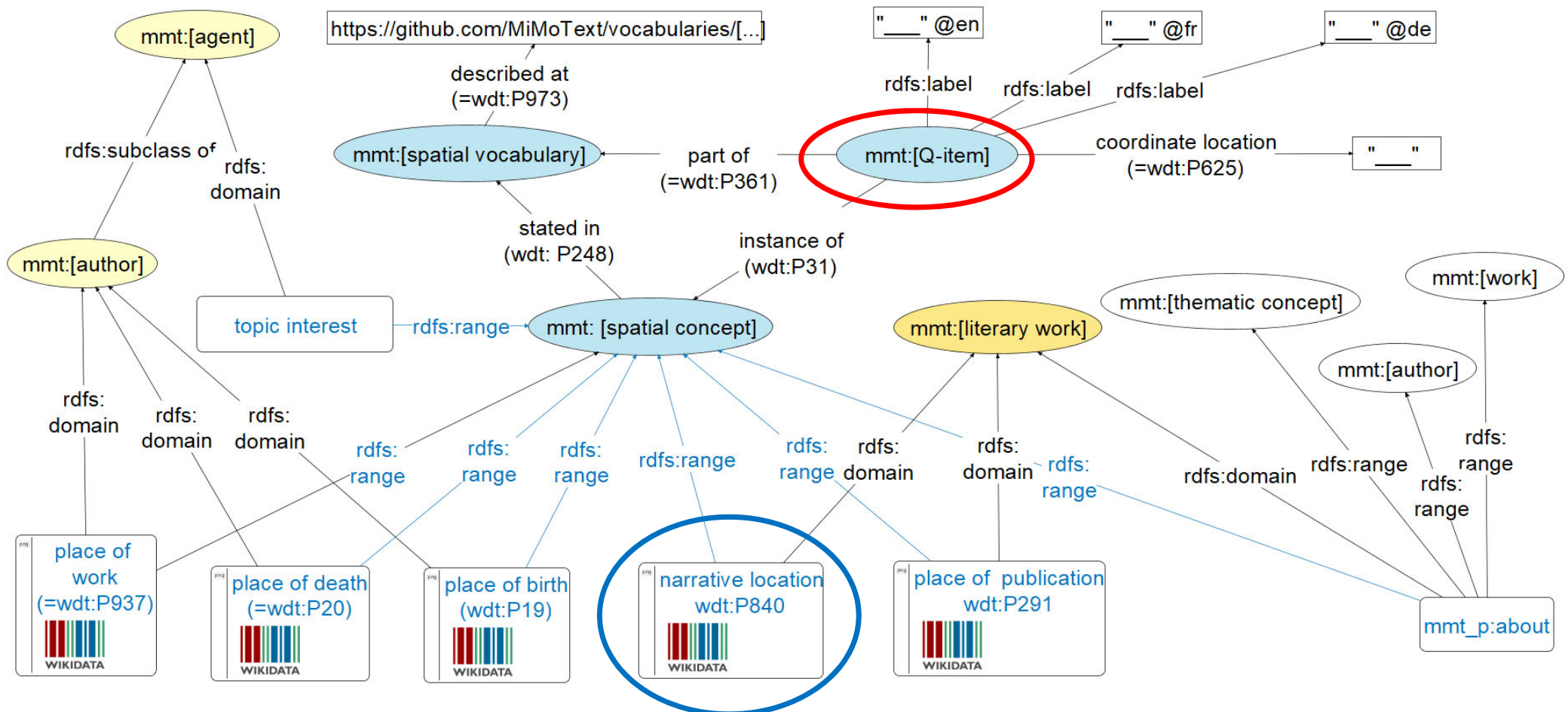
- A. Modeling
- B. ‚Knowledge Graph in Action‘ / Querying
- C. Conclusion



# A. Modeling (1): overview approach

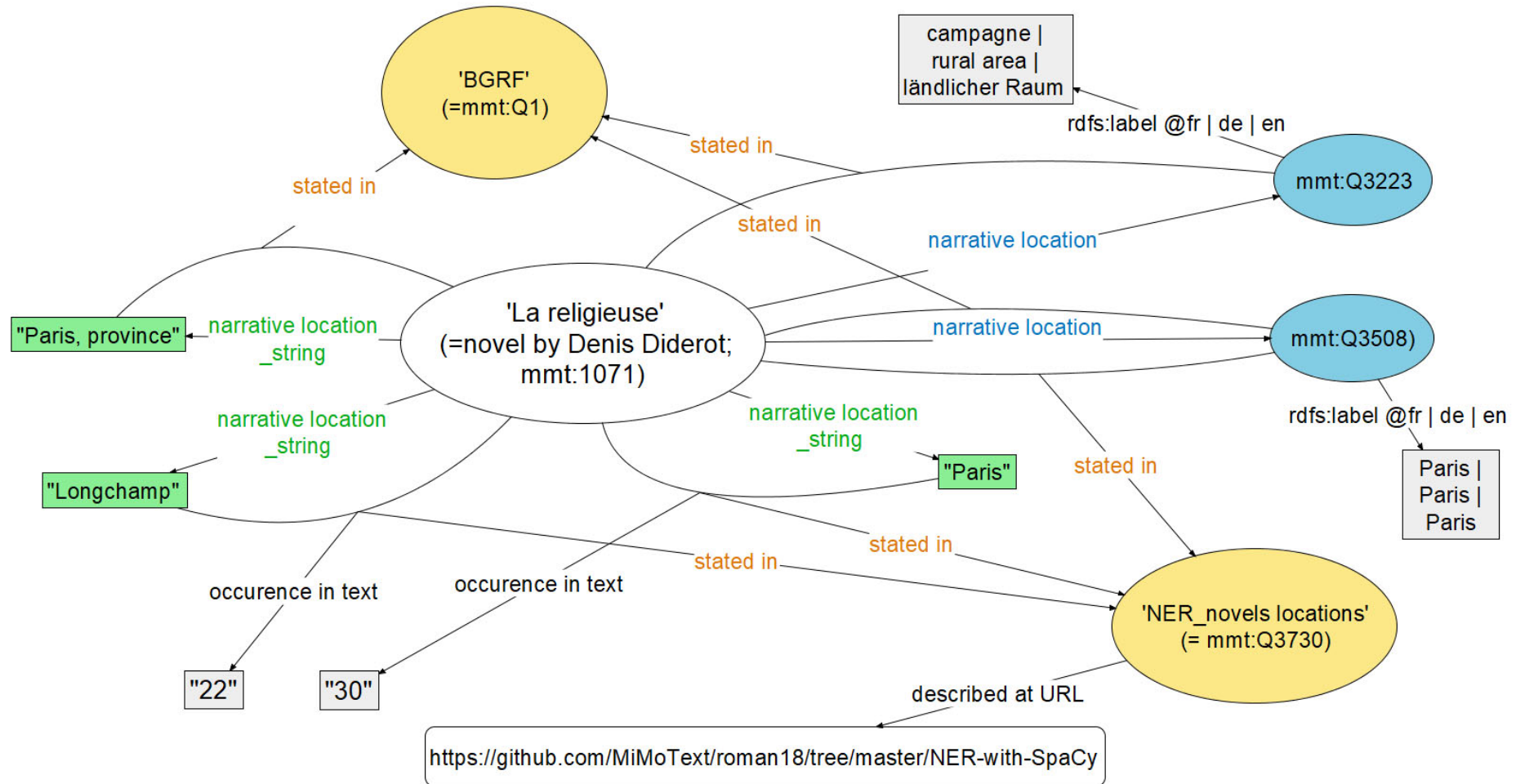


# A. Modeling (2): overview ontology (extract)





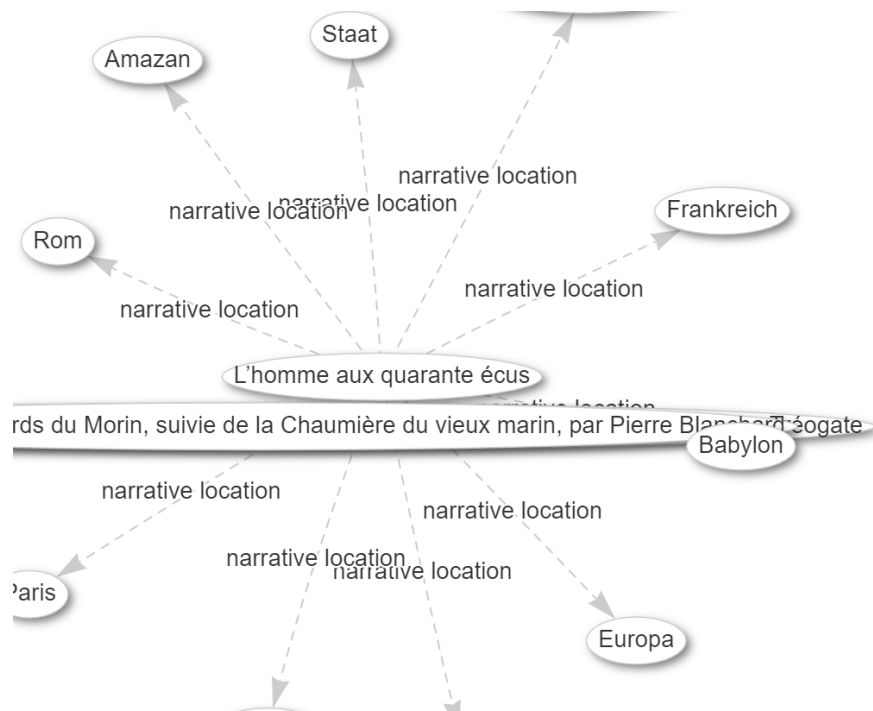
# A. Modeling (3): approach referencing sources (example)



## B. Knowledge Graph in Action



DockerWikibaseQueryService



```
#defaultView:Map
```

```
#defaultView:ImageGrid
```

```
#defaultView:Map
```

```
#defaultView:BubbleChart
```

```
#defaultView:TreeMap
```

```
#defaultView:Tree
```

```
#defaultView:Timeline
```

```
#defaultView:Dimensions
```

```
#defaultView:Graph
```

```
#defaultView:LineChart
```

```
#defaultView:BarChart
```

```
#defaultView:ScatterChart
```

```
#defaultView:AreaChart
```

# Querying: novels with narrative location ,rural area'

```
SELECT ?item ?itemLabel
WHERE
{
  ?item wdt:P52 wd:Q3223 . #narrative location = rural area
  ?item rdfs:label ?itemLabel .
  FILTER (LANG(?itemLabel) = "en").
}
ORDER BY ?itemLabel
```



219 novels

219 Ergebnisse in 169 ms

item	itemLabel
<a href="http://zora.univ-trier.de:11000/entity/Q1166">Q &lt;http://zora.univ-trier.de:11000/entity/Q1166&gt;</a>	Adelphine de Rostanges ou la mère qui ne fut point épouse, histoire véritable, rédigée par le Cen. Desforges
<a href="http://zora.univ-trier.de:11000/entity/Q2214">Q &lt;http://zora.univ-trier.de:11000/entity/Q2214&gt;</a>	Adoulzin ou les dangers d'une mauvaise éducation
<a href="http://zora.univ-trier.de:11000/entity/Q2569">Q &lt;http://zora.univ-trier.de:11000/entity/Q2569&gt;</a>	Adèle et Germeuil ou l'hermitage des monts Pyrénées, par A.J. Rosny, orné de figures dessinées par Quéverdo et gravées par Vill

# Querying: narrative location ,rural area' + about ,nature'

```
SELECT ?item ?itemLabel
WHERE
{
  ?item wdt:P52 wd:Q3223 .      #location = rural area
  ?item wdt:P25  wd:Q2985 .    #about = nature
  ?item rdfs:label ?itemLabel .
  FILTER (LANG(?itemLabel) = "en").
}
ORDER BY ?itemLabel
```



item

Q <[http://zora.uni-trier](http://zora.uni-trier.de:11000/entity/Q2310)>

Q <[http://zora.uni-trier](http://zora.uni-trier.de:11000/entity/Q1082)>

Q <[http://zora.uni-trier](http://zora.uni-trier.de:11000/entity/Q1017)>

Q <[http://zora.uni-trier](http://zora.uni-trier.de:11000/entity/Q2146)>

Q <[http://zora.uni-trier](http://zora.uni-trier.de:11000/entity/Q2544)>

Q < <a href="http://zora.uni-trier.de:11000/entity/Q2310">http://zora.uni-trier.de:11000/entity/Q2310</a> >	Le fils naturel
Q < <a href="http://zora.uni-trier.de:11000/entity/Q1082">http://zora.uni-trier.de:11000/entity/Q1082</a> >	Les cinquante francs de Jeannette, par le citoyen Ducray-Duminil
Q < <a href="http://zora.uni-trier.de:11000/entity/Q1017">http://zora.uni-trier.de:11000/entity/Q1017</a> >	L'homme aux quarante écus
Q < <a href="http://zora.uni-trier.de:11000/entity/Q2146">http://zora.uni-trier.de:11000/entity/Q2146</a> >	Prix de l'union, en deux parties
Q < <a href="http://zora.uni-trier.de:11000/entity/Q2544">http://zora.uni-trier.de:11000/entity/Q2544</a> >	Rose ou la bergère des bords du Morin, suivie de la Chaumière du vieux marin, par Pierre Blanchard

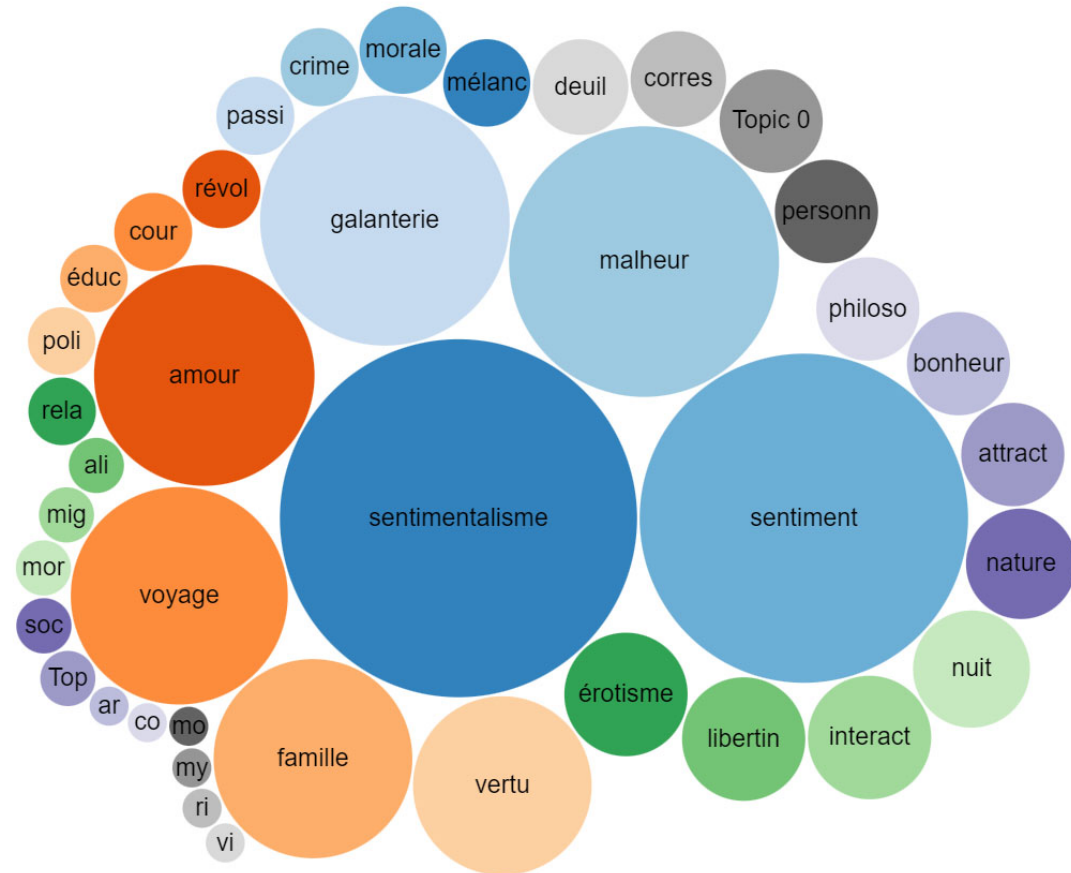


8 Ergebnisse in 68 ms

valier de.....publiées par M. Loaise

Querying: most frequent topics of novels (1750-1800) + narrative location = ,rural area'

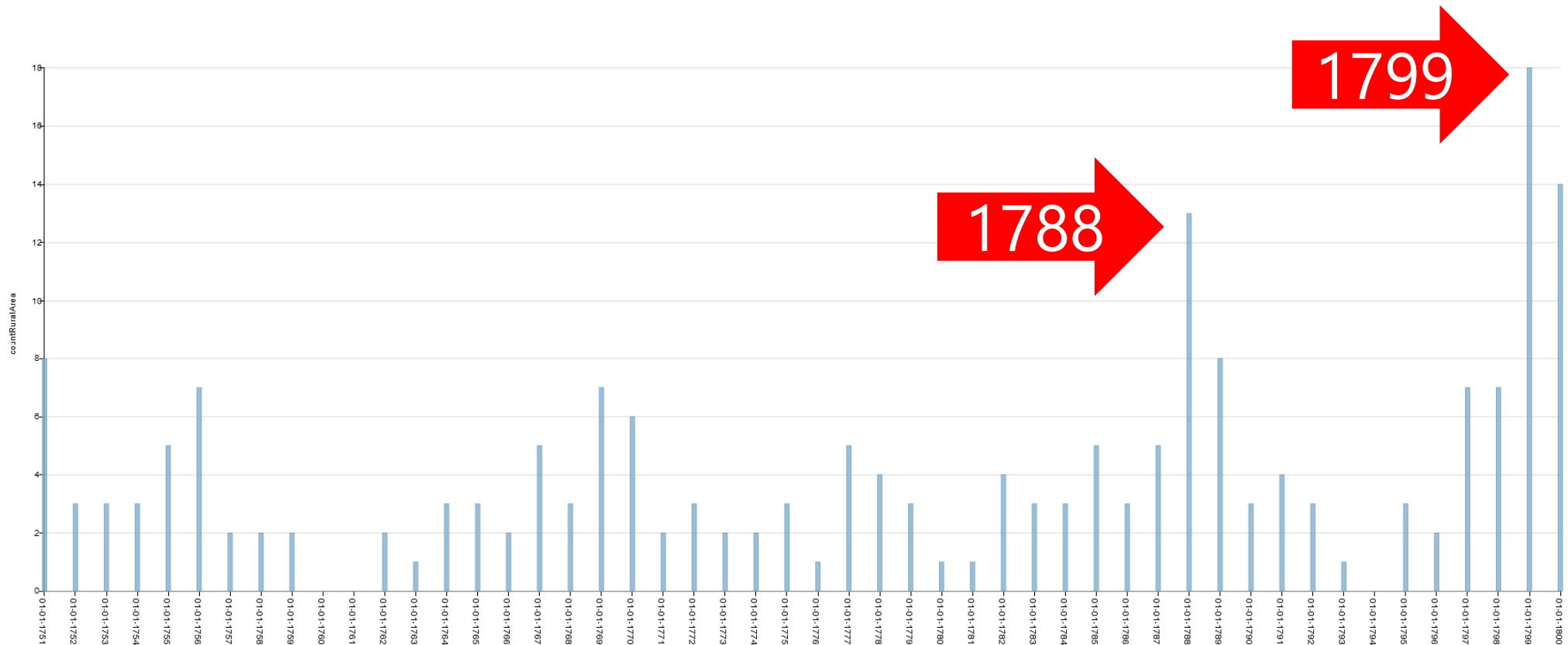
```
SELECT ?topLabel (count(*) as ?count)
WHERE {
  ?item wdt:P25 ?top ;
        wdt:P52 wd:Q3223 .
  ?top rdfs:label ?topLabel .
  filter(lang(?topLabel) = "fr")
}
GROUP BY ?topLabel
ORDER BY desc(?count)
#defaultView:BubbleChart
```



```
SELECT ?date ?countRuralArea ?countAll (?countRuralArea / ?countAll AS ?rel)
WHERE {
  {
    SELECT ?date (count(*) AS ?countRuralArea)
    WHERE {
      ?item wdt:P7 ?date;
            wdt:P52 ?narrativeLocation .
      ?narrativeLocation rdfs:label ?narrativeLocationLabel .
      filter(lang(?narrativeLocationLabel) = "en")
      filter(ucase(?narrativeLocationLabel) = "rural area"@en)
    }
    GROUP BY ?date
  }
  {
    SELECT ?date (count(*) AS ?countAll)
    WHERE {
      ?item wdt:P7 ?date;
            wdt:P52 ?narrativeLocation .
      ?narrativeLocation rdfs:label ?narrativeLocationLabel .
      filter(lang(?narrativeLocationLabel) = "en")
    }
    GROUP BY ?date
  }
}
ORDER BY desc(?rel)
#defaultView:BarChart
```

Querying:  
Proportion of the narrative  
location 'rural area' in all  
narrative locations  
(development:1750-1800?)

# Results: Proportion of the narrative location 'rural area' in all narrative locations of the particular year (1750-1800)









# (Spatial) Federated Queries: MiMoText-Wikibase & Wikidata

```
PREFIX wid: <http://www.wikidata.org/entity/> #wikidata wd
PREFIX wdt: <http://www.wikidata.org/prop/direct/> #wikidata wdt

Select DISTINCT ?item ?itemLabel ?nar_loc ?nar_locLabel ?wikidataEntityLink ?coordinateLocation
{
  ?item wdt:P52 ?nar_loc.
  ?nar_loc wdt:P30 ?WikiLink. #MiMoText P30 = exact match

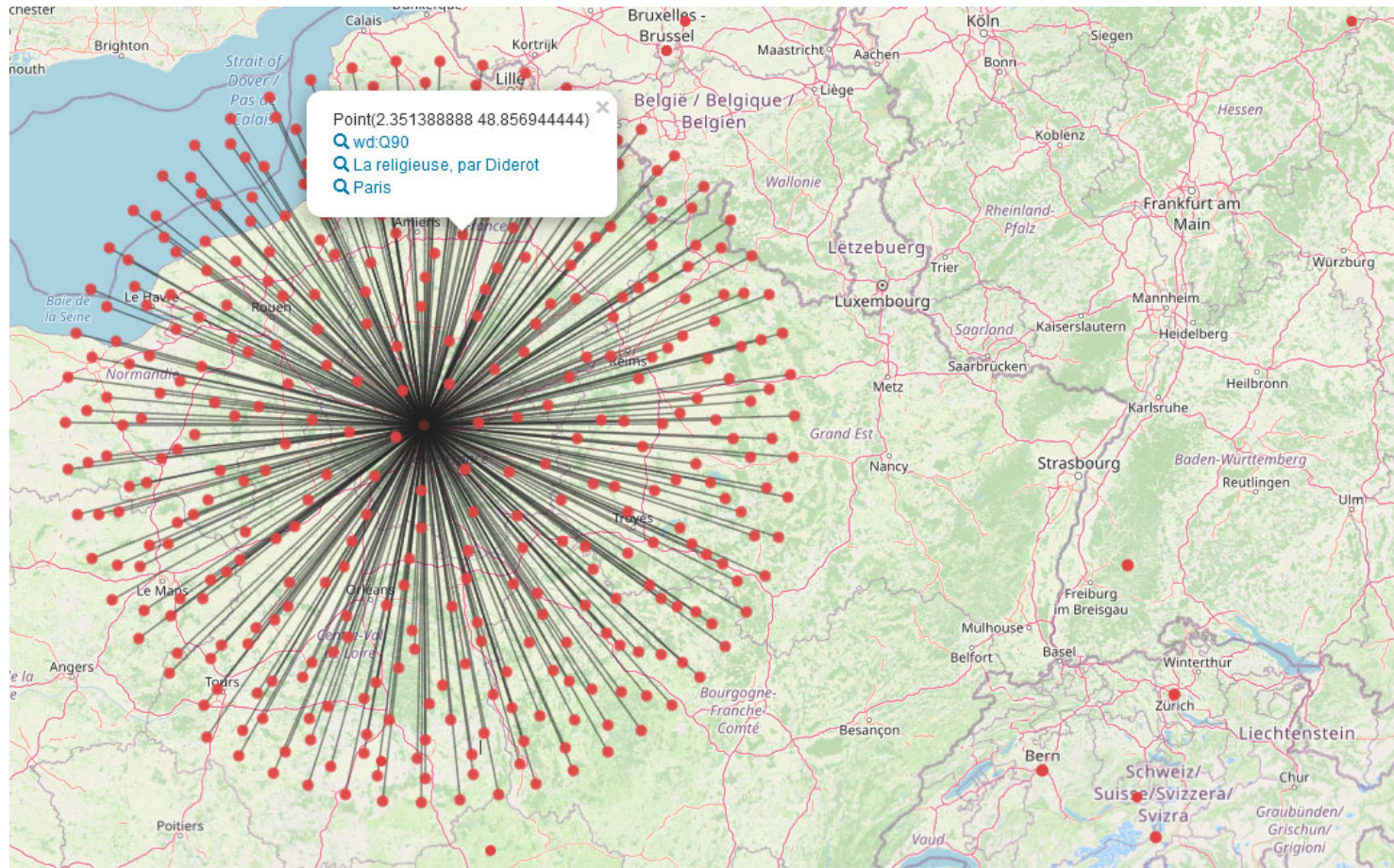
  BIND(IRI(REPLACE( STR(?WikiLink), "https://www.wikidata.org/wiki", "http://www.wikidata.org/entity" )) AS ?wikidataEntityLink).
  #Federated Query# Wikidata
  SERVICE <https://query.wikidata.org/sparql> {
    ?wikidataEntityLink wdt:P30 wid:Q46; #Wikidata P30 = continent
    wdt:P625 ?coordinateLocation
  }

  SERVICE wikibase:label {
    bd:serviceParam wikibase:language "en" .
  }
}
#defaultView:Map
```





# (Spatial) Federated Queries: MiMoText-Wikibase & Wikidata (Map View)



## Conclusion

- Part of the Wikibase ecosystem and the ‚Wikiverse‘
  - Next steps
    - Publication of the prototype , adding our author- and novel-items to Wikidata, linkage of items via new property (e.g. MiMoText ID)
    - Focus on usage scenarios
  - Goals
    - New perspectives for our domain
    - Transferrable approach
    - Reflecting challenges as metadisciplinary contribution
-



Thank you!



Project: [mimotext.uni-trier.de](http://mimotext.uni-trier.de)

Slides: <https://zenodo.org/communities/mimotext-tcdh>

---