

# Annotation en lieux et sentiments de récits de vie transcrits

Catherine Dominguès, Laurence Jolivet, Carmen Brando  
LaSTIG/IGN CRH/EHESS

**MAGIS – AP Humanités numériques spatialisées**  
**24 juin 2020**

# Contexte de travail

projet **MATRICIEL**, PEPS CNRS/UPEM 2016 :

- récits de vie collectés par le RAHMI (2008 à 2010) pour documenter les apports de l'immigration à l'histoire nationale et à la mémoire collective
- Républicains espagnols exilés en France entre 1936 et 1939 et ayant participé à la résistance française

Intervenants	directs/indirects	hommes/femmes	durée
18	11/7	11/7	18:30:18

- objets d'étude : **lieux** et **sentiments** associés à un lieu
- objectifs : **identifier et cartographier les objets d'étude, automatiser**

# Extrait du corpus

---

... on est allé dans la banlieue de Barcelone il allait diriger une usine de textile qui travaillait pour l'armée à l'époque il faut dire que la coupe du tissu était faite par des hommes ils étaient tous à la guerre ou presque tous on était dans la banlieue de Barcelone c'est à Badalona juste en limite à toucher Barcelone on était dans une usine de textile et il y avait deux cent cinquante femmes qui travaillaient et il y avait deux bonhommes pour couper mon père dirigeait l'usine mais en même temps il coupait avec un autre monsieur qui avait quatre-vingt cinq ou quatre-vingt six ans un tailleur pour donner du travail à toutes ces femmes et là on est resté presque un an à crever de faim parce qu'on crevait de faim on crevait littéralement de faim la première nuit parce que l'aviation italienne venait bombarder tous les jours toutes les nuits on descend à la cave après avec mon père on a décidé de monter on avait une terrasse une grande terrasse on montait sur la terrasse voir ce qui se passait si les avions viennent on descend à la cave et moi j'assistais à chaque fois à tous les bombardements que ce soit de jour ou de nuit j'assistais la nuit c'était un véritable feu d'artifice ...

# Annotations en lieux et sentiments

... on est allé dans la **banlieue de Barcelone** il allait diriger une **usine de textile** qui travaillait pour l'armée à l'époque il faut dire que la coupe du tissu était faite par des hommes ils étaient tous à la guerre ou presque tous on était dans la **banlieue de Barcelone** c'est à **Badalona** juste en limite à toucher **Barcelone** on était dans une **usine de textile** et il y avait deux cent cinquante femmes qui travaillaient et il y avait deux bonhommes pour couper mon père dirigeait l'**usine** mais en même temps il coupait avec un autre monsieur qui avait quatre-vingt cinq ou quatre-vingt six ans un tailleur pour donner du travail à toutes ces femmes et là on est resté presque un an à **crever de faim** parce qu'on **crevait de faim** on **crevait** littéralement **de faim** la première nuit parce que l'aviation italienne venait **bombarder** tous les jours toutes les nuits on descend à la **cave** après avec mon père on a décidé de monter on avait une **terrasse** une **grande terrasse** on montait sur la **terrasse** voir ce qui se passait si les avions viennent on descend à la **cave** et moi j'assistais à chaque fois à tous les **bombardements** que ce soit de jour ou de nuit j'assistais la nuit c'était un véritable **feu d'artifice** ...

# Objets d'étude

---

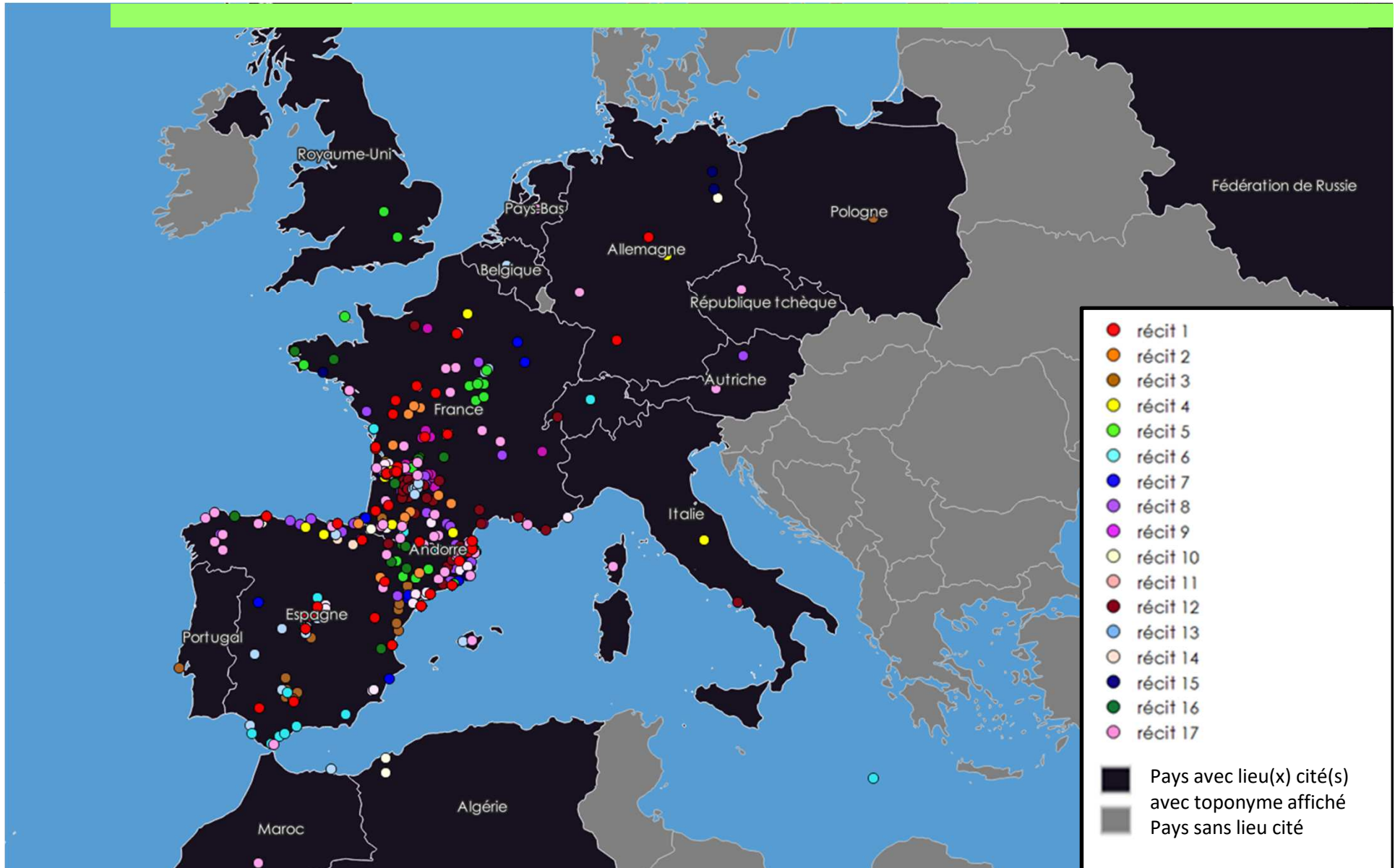
- **Entité nommée** : *toute expression linguistique qui réfère à une entité unique de manière autonome dans le corpus*
- **Lieu** (EN spatiale) est ici compris dans un **sens large** :
  - toponyme Npr répertorié dans un gazetier : *Barcelone*
  - lieu désigné par un Nc : *banlieue de Barcelone, camp de Gurs, camp de regroupement, usine de textile, cave*
- **Sentiments associés aux lieux** et à prendre en compte dans la **représentation cartographique**
  - dictionnaire de sentiments polarisés : *positive, négative*
  - propagation des polarités sur les lieux => polarité agrégée : *positive, négative, neutre, contradictoire*

# Annotations automatique et manuelle

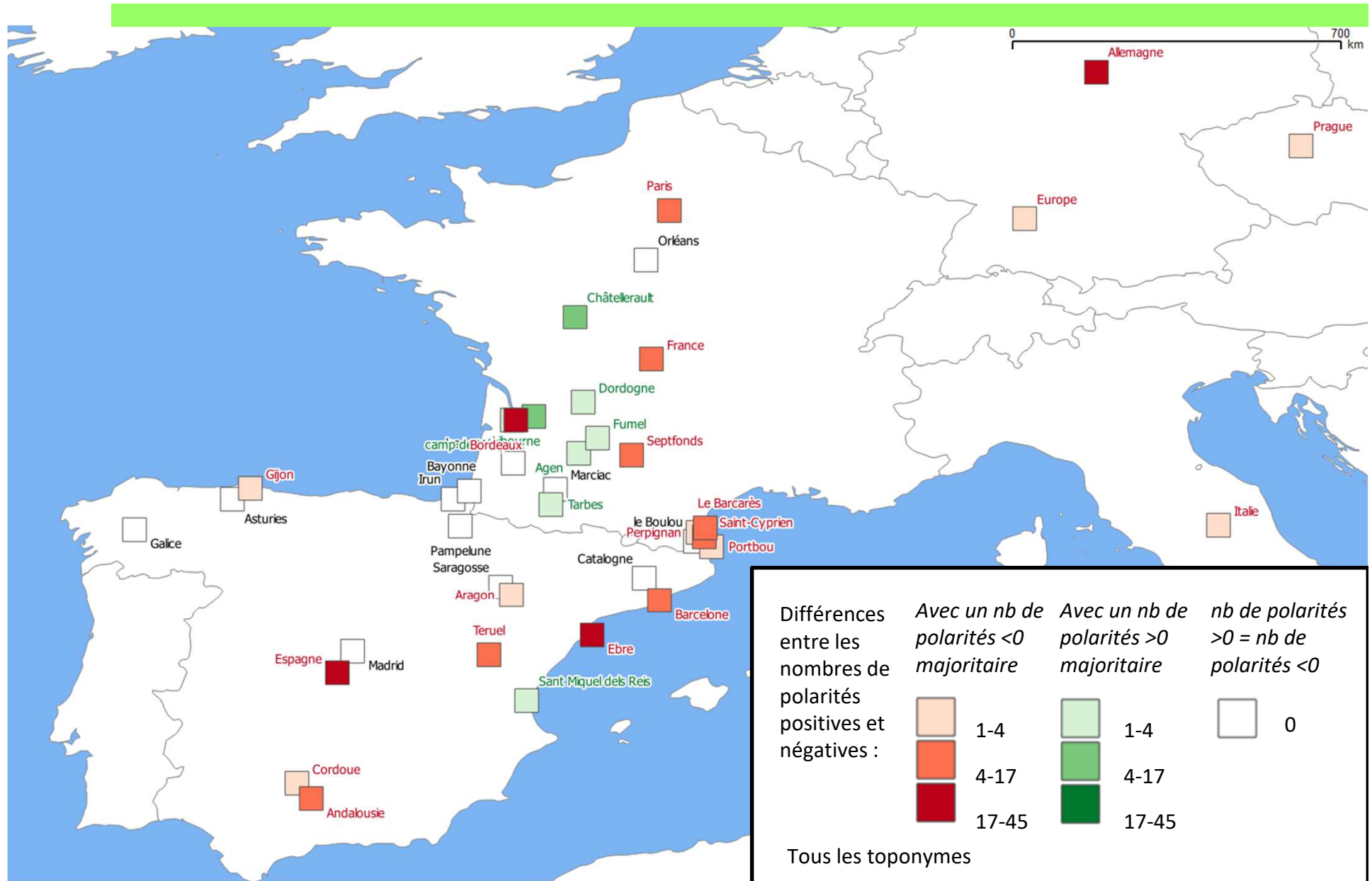
---

- **Méthodes symboliques** (lieux NPr et sentiments) :
  - lieux NPr (*Barcelone*), à l'aide de gazetiers (Geonames, BDNyme)
  - **sentiments** à l'aide de lexiques de mots polarisés (*crever, guerre*)
- **Méthodes fondées sur l'apprentissage** (lieux Nc) :
  - corpus d'apprentissage annoté manuellement  
*camp d'internement, ligne de front*
  - entraînement du modèle (Stanford-Named Entity Recognition 2005)
  - validation : F-mesure = 0,75 ; rappel à améliorer
- **Accords inter-annotateurs** :
  - lieux Nc : R= 0,86 ; P = 0,87 ; F-mesure = 0,86
  - **sentiments** : R = 0,65 ; P = 0,66 ; F-mesure = 0,66

# Lieux NPr cités dans les récits



# Intensité des polarités sur les lieux





# Annotations automatique et manuelle

---

- Méthodes symboliques (lieux NPr et sentiments) :
  - lieux NPr (*Barcelone*), à l'aide de gazetiers (Geonames, BDNyme)
  - sentiments à l'aide de lexiques de mots polarisés (*crever, guerre*)
- Méthodes fondées sur l'apprentissage (lieux Nc) :
  - corpus d'apprentissage annoté manuellement  
*camp d'internement, ligne de front*
  - entraînement du modèle (Stanford-Named Entity Recognition 2005)
  - validation : F-mesure = 0,75 ; rappel à améliorer
- Accords inter-annotateurs :
  - lieux Nc : R= 0,86 ; P = 0,87 ; F-mesure = 0,86
  - sentiments : R = 0,65 ; P = 0,66 ; F-mesure = 0,66

# A continuer



- **Sur la patrimonialisation du corpus :**
  - dépôt du corpus transcrit et annoté (subvention CORLI)
- **Sur l'automatisation des tâches (TAL et cartographie) :**
  - **lieux :**
    - NPr : désambiguïsation automatique dans GATE
    - Nc : amélioration du modèle par amélioration du corpus d'apprentissage
    - Nc et Npr : localisation des lieux Nc à partir des lieux NPr
  - **sentiments** : propagation (semi) automatique des sentiments sur les lieux
  - **cartographie** : prise en compte l'échelle
- **Complémentarité entre approche symbolique et apprentissage automatique :**
  - (adaptation patrons) vs (adaptation du corpus d'apprentissage)

# MERCI pour votre attention

[catherine.domingues@ign.fr](mailto:catherine.domingues@ign.fr)

[laurence.jolivet@ign.fr](mailto:laurence.jolivet@ign.fr)

[carmen.brande@ehess.fr](mailto:carmen.brande@ehess.fr)