

# Linked Open Data

Chapitre 3 : De la donnée à l'exploitation de la donnée

# Problématique

Comment, à partir d'une question de recherche, identifier et analyser les possibilités d'une analyse

# 1. La constitution d'un corpus

## 1. Identifier ses questions

Recherche des réseaux de la formation dans la culture

## 2. Identifier ses acteurs et leurs relations

Institutions de formation (Universités, Grandes Écoles, Écoles d'applications)

Institutions de la culture (Ministère, Archives, Musées, Bibliothèques...)

Individus

Lien : formé-e à, travaille à, etc.

# 1. La constitution d'un corpus

## 3. Récupérer les données

Identifier les possesseurs de telles données (Wikidata, positions de thèses de l'École, CV)

Faire les requêtes wikidata

## 4. Formater les données dans un tableur (généralement)

Identifier les besoins, y compris visuel (qui ne sont pas liés à un besoin de modélisation)

## 5. Vérifier les données

## 2. Ingérer les données dans un logiciel d'analyse

### 1. Identifier les besoins du logiciel

Gephi fonctionne bien avec des ID et des labels pour les noeuds de réseaux

Gephi fonctionne de même pour les liens (aussi appelés vecteurs ou edges) avec un tableur en CSV source,target,label

## 2. Ingérer les données dans un logiciel d'analyse

### 2. Convertir les données

A la main (Possible si le tableur original est bien fait)

En programmant (si l'on sait faire)

En utilisant des outils de conversion (OpenRefine, Dataiku, etc.)

## 2. Ingérer les données dans un logiciel d'analyse

### 2. Visualiser, manipuler les données

Réfléchir à l'intérêt de la visualisation (les noeuds à parts dans le réseau, les centres de réseaux, les liens entre deux réseaux.)

Essayez de changer les modes de visualisations

### 3. Trouver les points d'intérêts dans le réseau

A partir des points d'intérêts, voir jusqu'où vous pouvez poser de nouvelles questions : la data visualisation comme outil heuristique.

L'analyse de réseau, dans ses calculs, peut-être un outil d'évaluation (de preuve).

Encore faut il savoir manipuler les concepts mathématiques et comprendre les fondements.



# Analyse de réseau : Une courte bibliographie

<http://www.martingrandjean.ch>

- Grandjean Martin, Clavert Frédéric, Daniel Johanna, Fleckinger Hélène et Idmhand Fatiha “**Histoire et humanités numériques : nouveaux terrains de dialogue entre les archives et la recherche**” *La Gazette des archives*, 245, 1, 2017, 121-134.  
<https://halshs.archives-ouvertes.fr/halshs-01521814>
- Grandjean Martin “**A Social Network Analysis of Twitter: Mapping the Digital Humanities Community**” *Cogent Arts & Humanities*, 2016, 3:1171458  
<https://halshs.archives-ouvertes.fr/hal-01517493>
- Grandjean Martin “**Introduction à la visualisation de données : l’analyse de réseau en histoire**” *Geschichte und Informatik*, 18/19, 2015, 109-128. <https://halshs.archives-ouvertes.fr/halshs-01525543>

# Analyse de réseau : Une courte bibliographie

- Baillot, Anne. 2015. “**Visualisation des réseaux: apports, défis et enjeux du travail sur les données historiques.**” In . <https://halshs.archives-ouvertes.fr/halshs-01130425/document>.
- Réseaux Historiques : cycle de conférence *Connected Past* <http://connectedpast.net/publications/> (Claire Lemercier)
- Historical Network Bibliography, [https://www.zotero.org/groups/historical\\_network\\_research](https://www.zotero.org/groups/historical_network_research)
- Semantic Network Analysis Workshop, Jana Diesner, [http://www.zdes.spbu.ru/assets/files/content/2013/Semantic\\_networks/Tutorial\\_SemanticNetworkAnalysis\\_JanaDiesner\\_052013.pdf](http://www.zdes.spbu.ru/assets/files/content/2013/Semantic_networks/Tutorial_SemanticNetworkAnalysis_JanaDiesner_052013.pdf)

## 4. Trouver les manques dans les données

### 1. Identifier

A travers la visualisation et votre connaissance du corpus, trouver l'ensemble des modifications nécessaires à produire afin de rendre le corpus représentatif

## 4. Trouver les manques dans les données

### 2. Enrichir les données

Si possible, enrichir les données directement sur la plate-forme d'origine

Si non, enrichir les données en établissant au préalable une méthode : prédicats à utiliser par exemple

Important : réanalyser le modèle de données et les biais de source à partir de ces manques