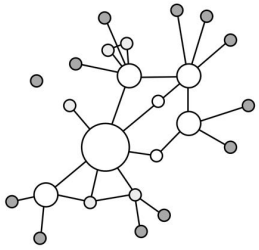


Visualisation de données

Humanités numériques : approches interdisciplinaires
Ecole d'été CERIUM
Montréal - 17 juin 2019

UNIL | Université de Lausanne

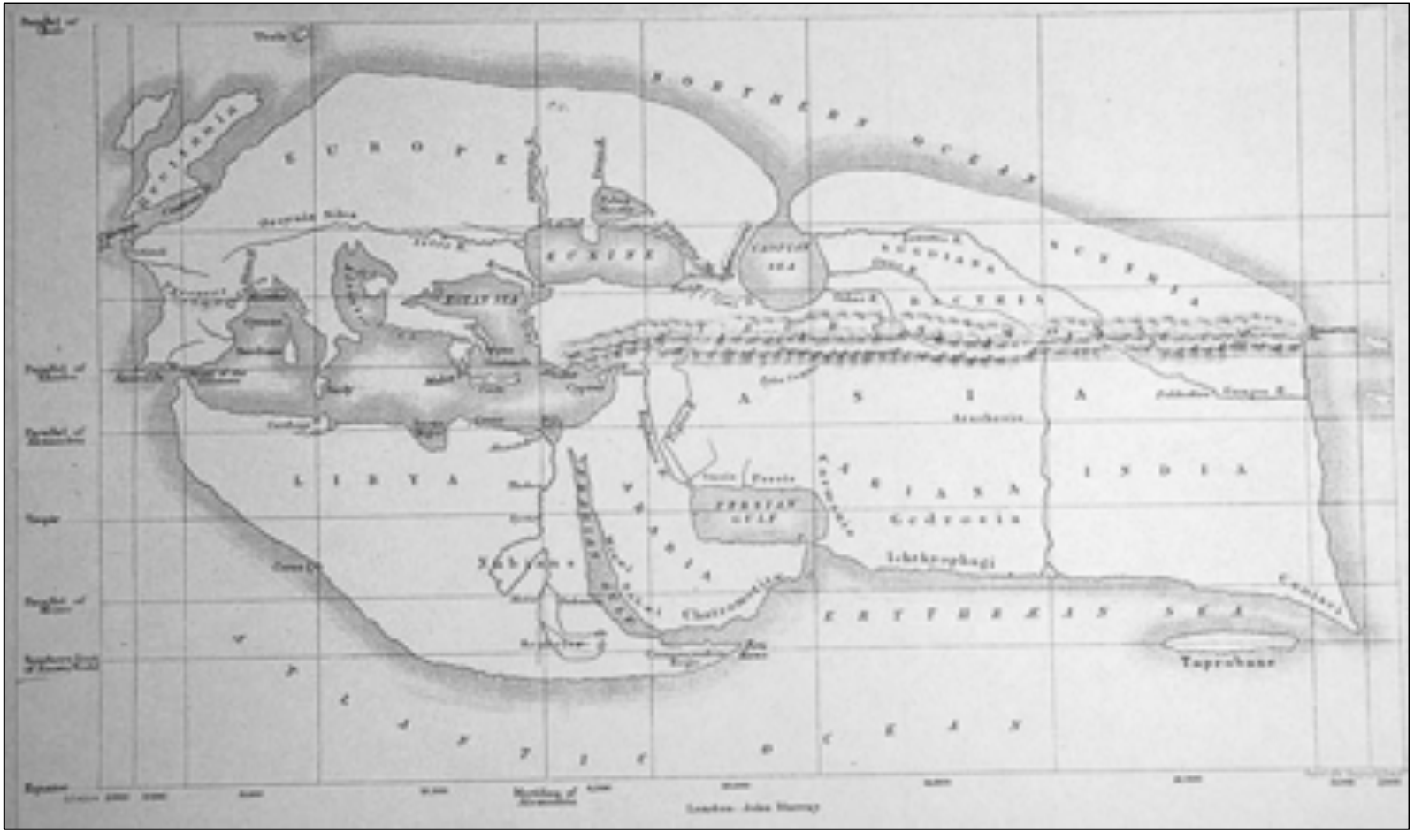
Martin GRANDJEAN
martingrandjean.ch

@GrandjeanMartin
/grandjean.data

VISUALISATION DE DONNÉES
HISTORIQUE

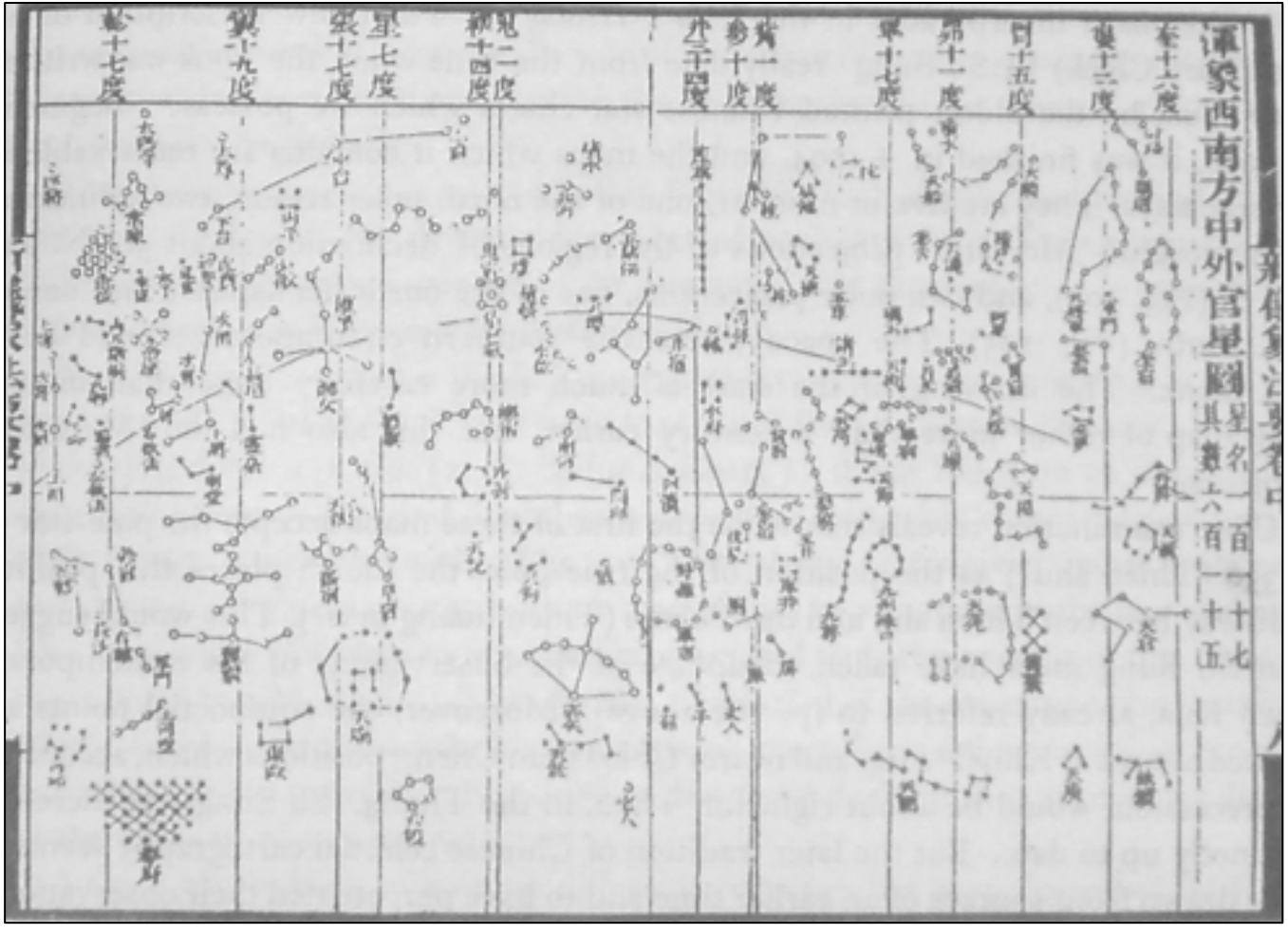
Art rupestre
(Valley of fire, Nevada)

VISUALISATION DE DONNÉES
HISTORIQUE



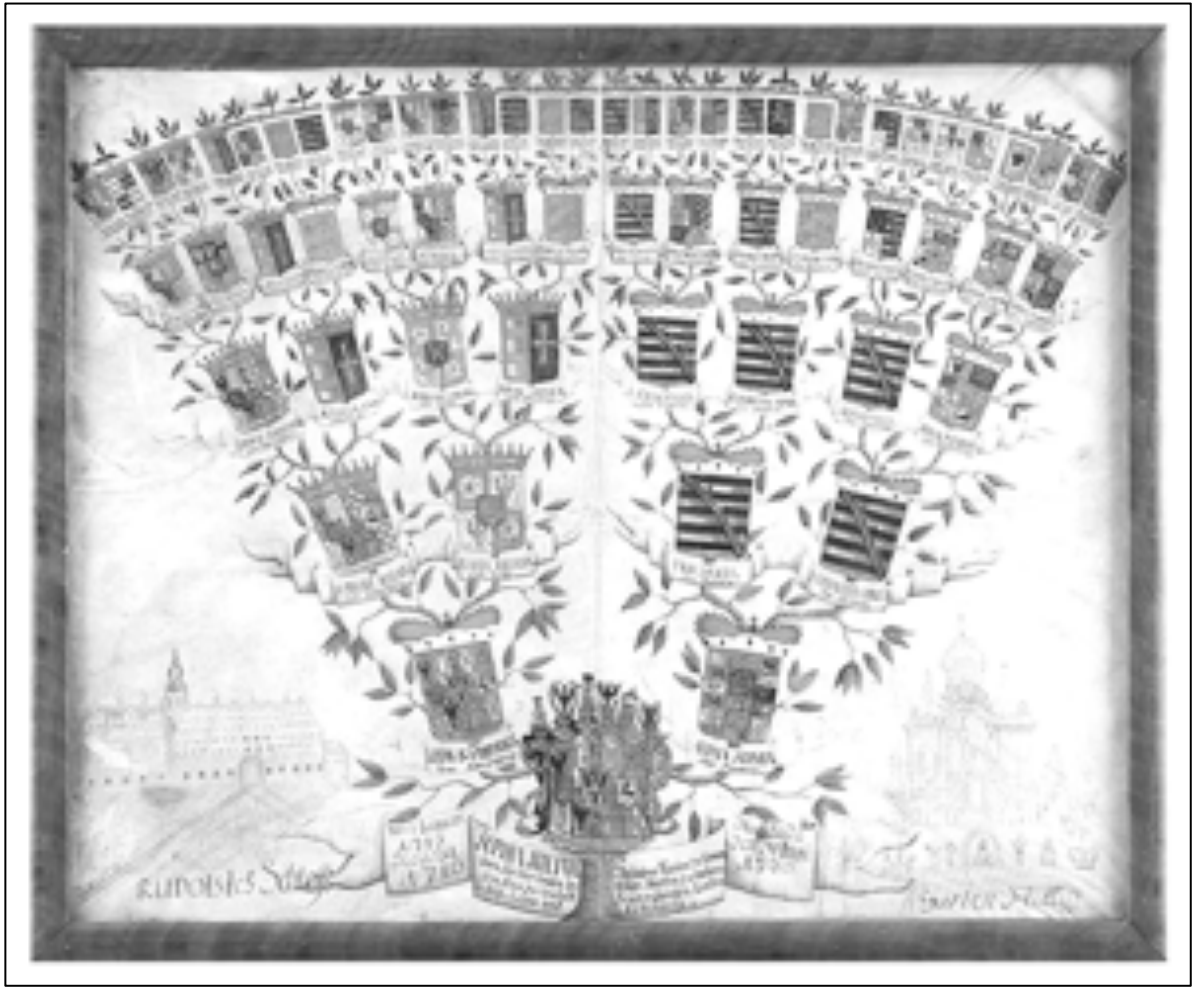
Reproduction de la carte d'Eratosthène (env. 194 av. J.-C.)

VISUALISATION DE DONNÉES
HISTORIQUE



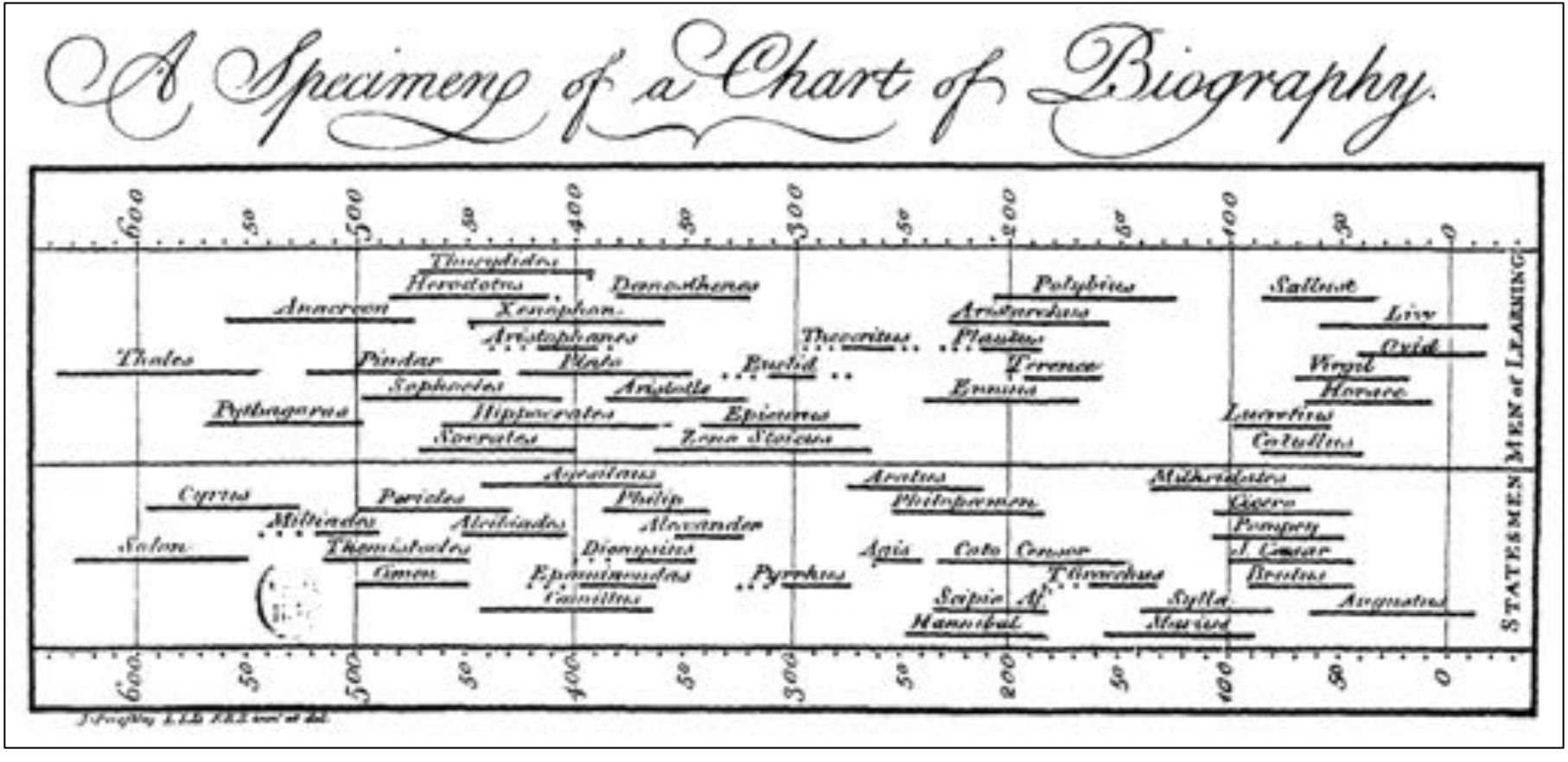
Carte céleste
(Chine, XIe siècle)

VISUALISATION DE DONNÉES
HISTORIQUE



Arbre généalogique
(XVIIIe siècle)

Priestley, Joseph. 1765. *A Chart of Biography*. London.

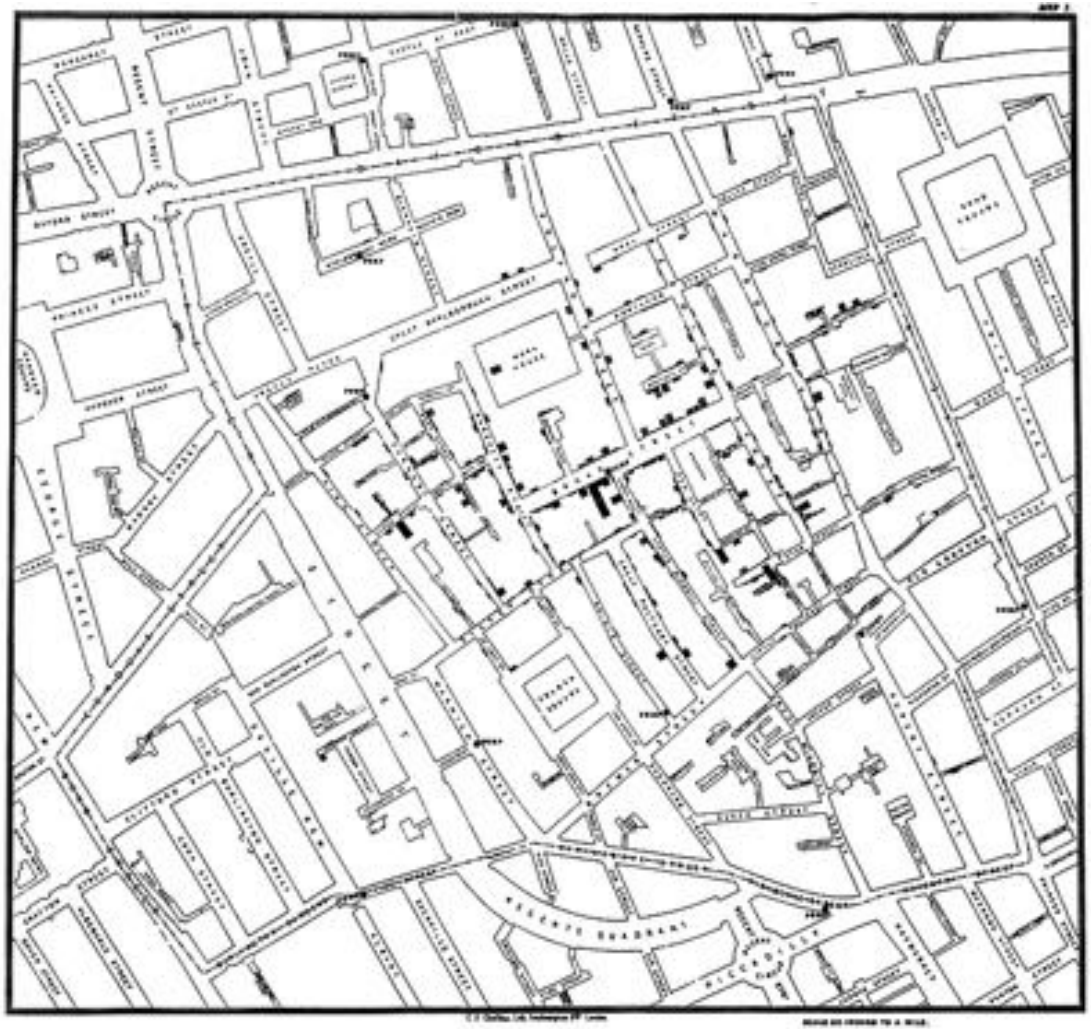


VISUALISATION DE DONNÉES
HISTORIQUE

Carte
Dufour
1845

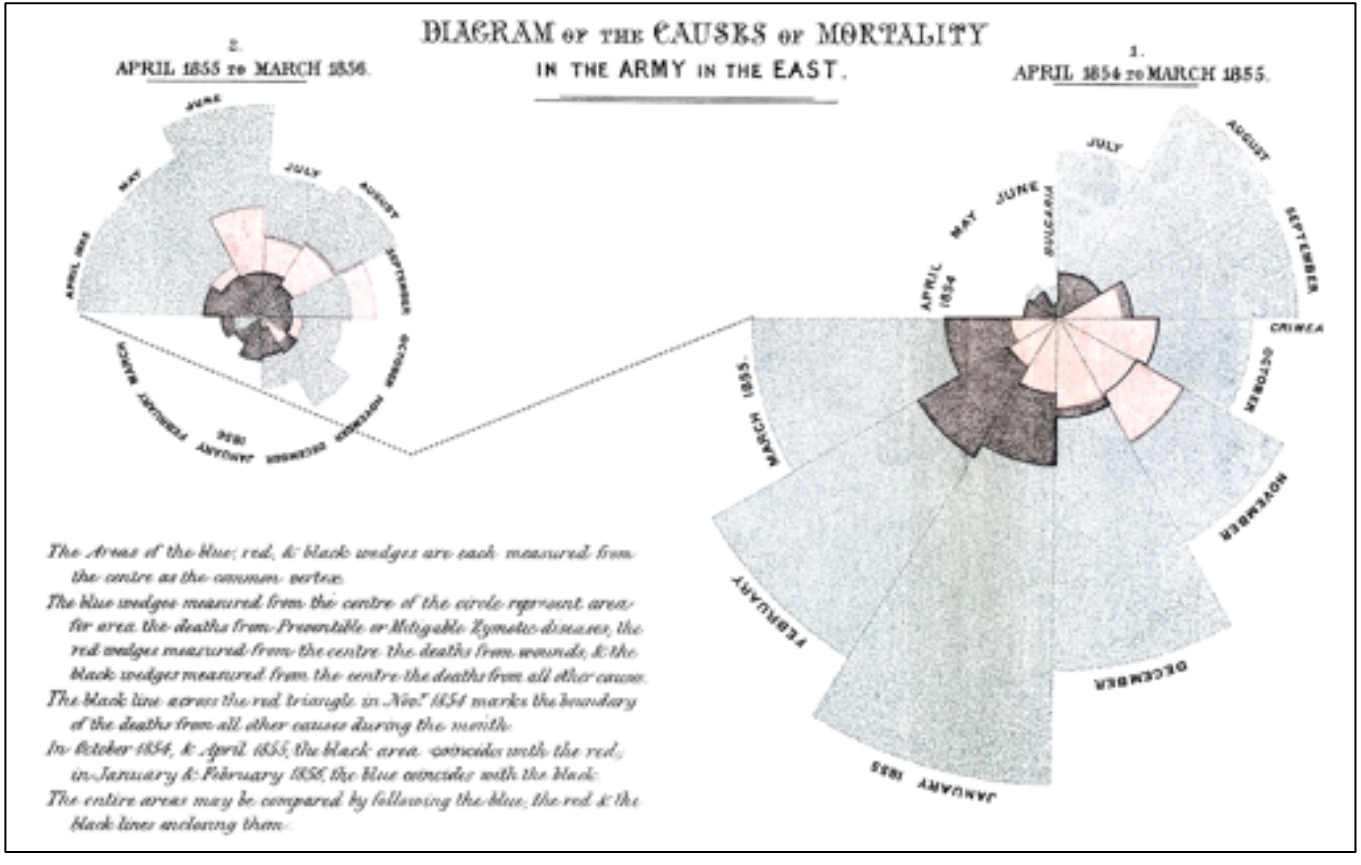
VISUALISATION DE DONNÉES
HISTORIQUE

Snow, John. 1854. *On the mode of communication of cholera*. London



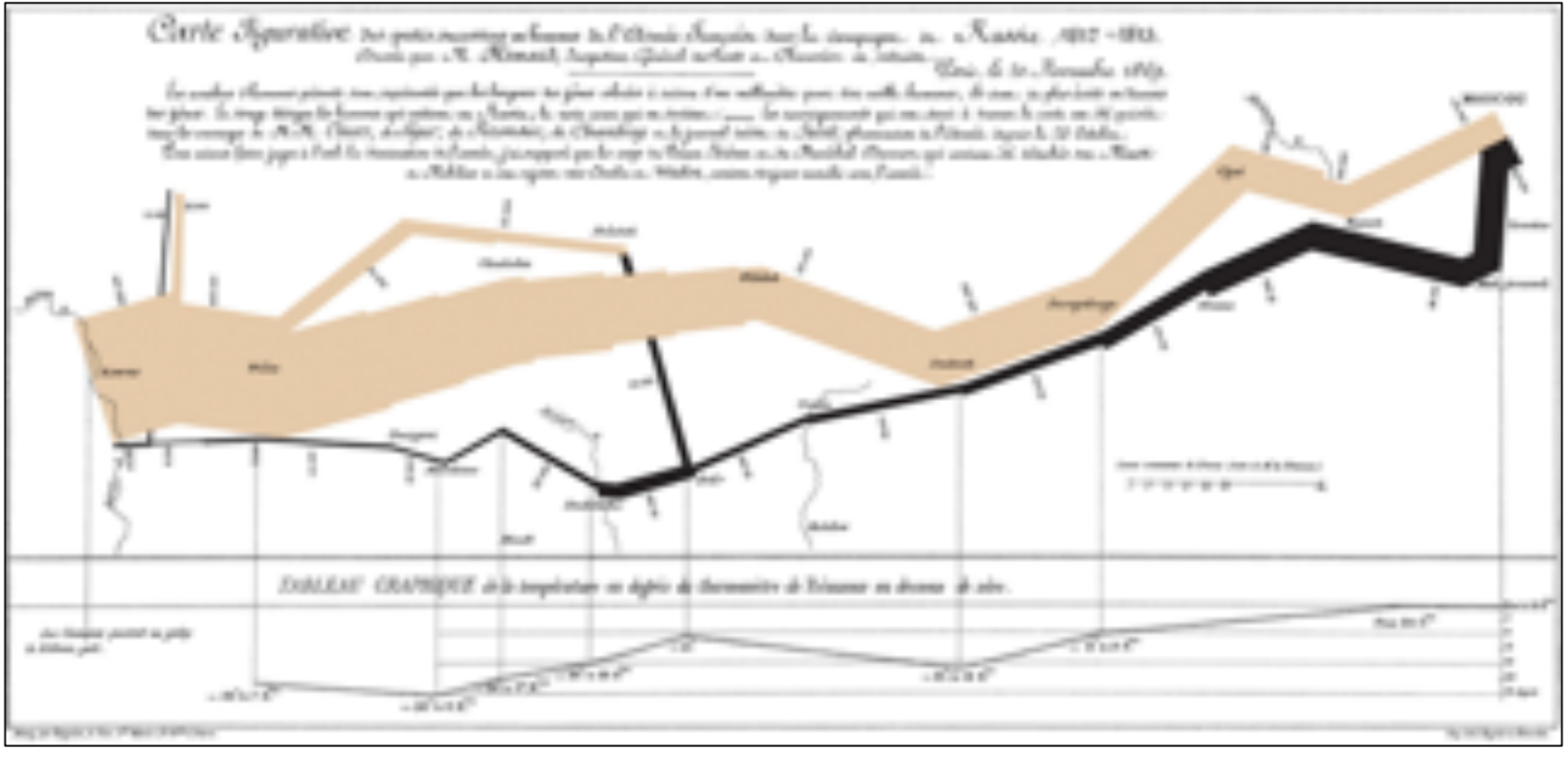
VISUALISATION DE DONNÉES
HISTORIQUE

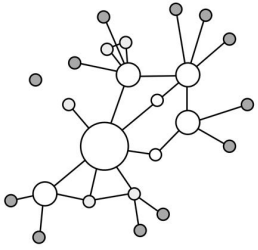
Nightingale, Florence. 1858. *Notes on Matters Affecting the Health, Efficiency, and Hospital Aministration of the British Army.*



VISUALISATION DE DONNÉES
HISTORIQUE

Minard, Charles-Joseph, 1869.





II.

DE QUELLE(S) VISUALISATION(S) PARLE-T-ON ?

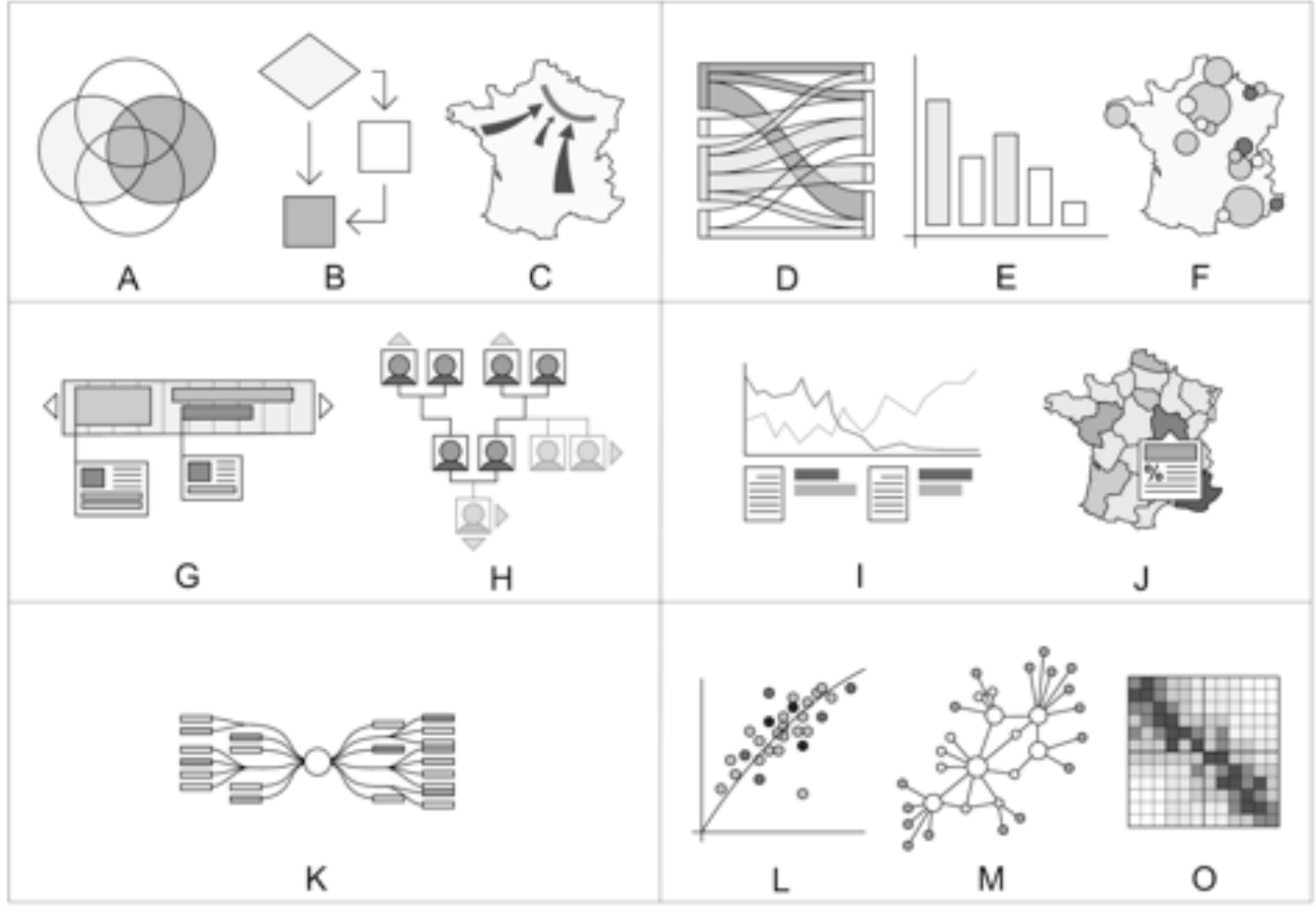
VISUALISATION DE DONNÉES
DE QUOI PARLE-T-ON ?

TYPE DE SOURCES

Informations — Données

TYPE D'USAGE

Démonstration — Interface — Recherche



VISUALISATION DE DONNÉES

DE QUOI PARLE-T-ON ? VISUALISATIONS DE RECHERCHE (TRAVAUX D'ETUDIANTS)

Analyse de la position physique des articles portant sur le « secret bancaires » dans des journaux.

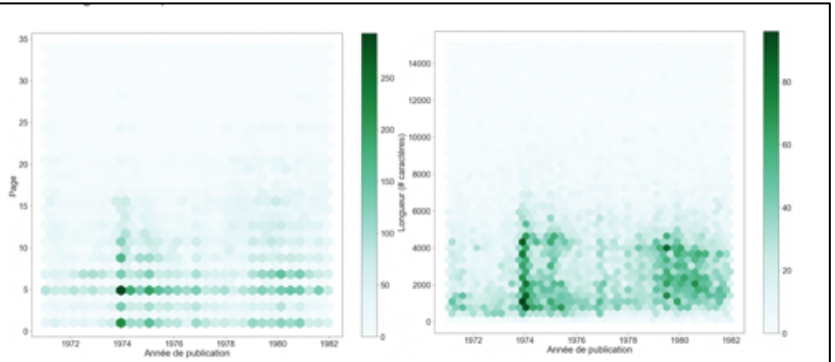
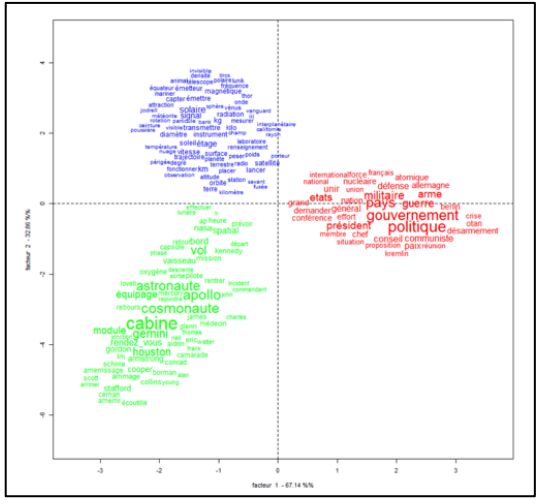


Figure 6 *(gauche): Position des articles du corpus dans les journaux (intervalles de 134 jours, 35 pages)*

Figure 7 *(droite): Longueur des articles du corpus (intervalles de 100 jours, env. 325 caractères)*



Proximité des classes de termes décrivant la conquête spatiale (technique, politique, scientifique).

Extraction des thématiques principales dans un corpus d'articles.

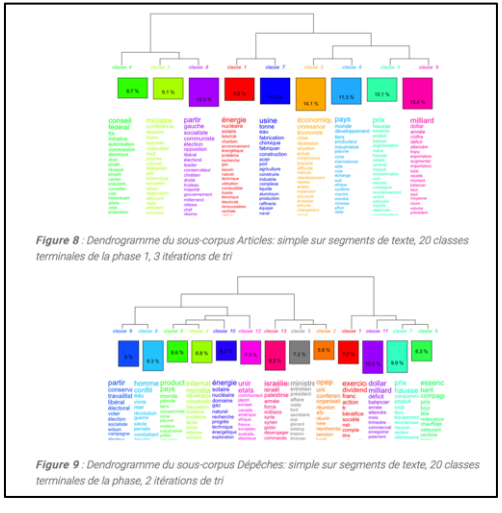
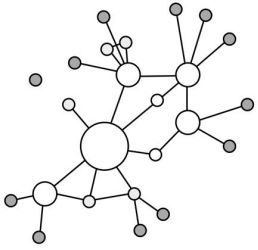


Figure 8 *: Dendrogramme du sous-corpus Articles: simple sur segments de texte, 20 classes terminales de la phase 1, 3 itérations de tri*

Figure 9 *: Dendrogramme du sous-corpus Dépêches: simple sur segments de texte, 20 classes terminales de la phase, 2 itérations de tri*

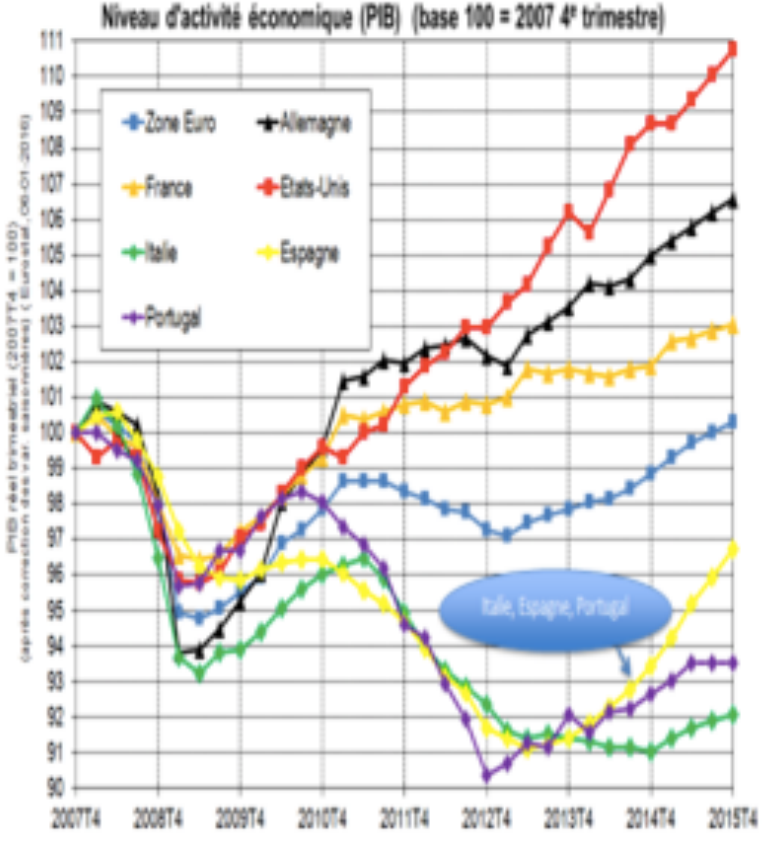
Graphes de cooccurrence des principaux termes liés aux accords nucléaires USA-URSS

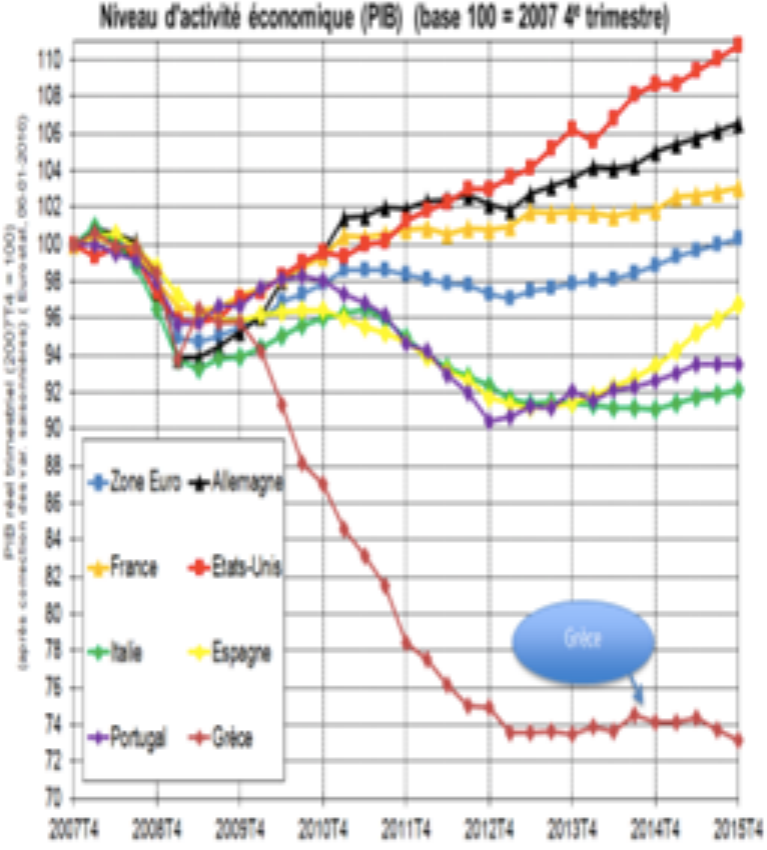


III.

USAGES ET MÉSUSAGES

VISUALISATION DE DONNÉES
USAGES ET MÉSUSAGES

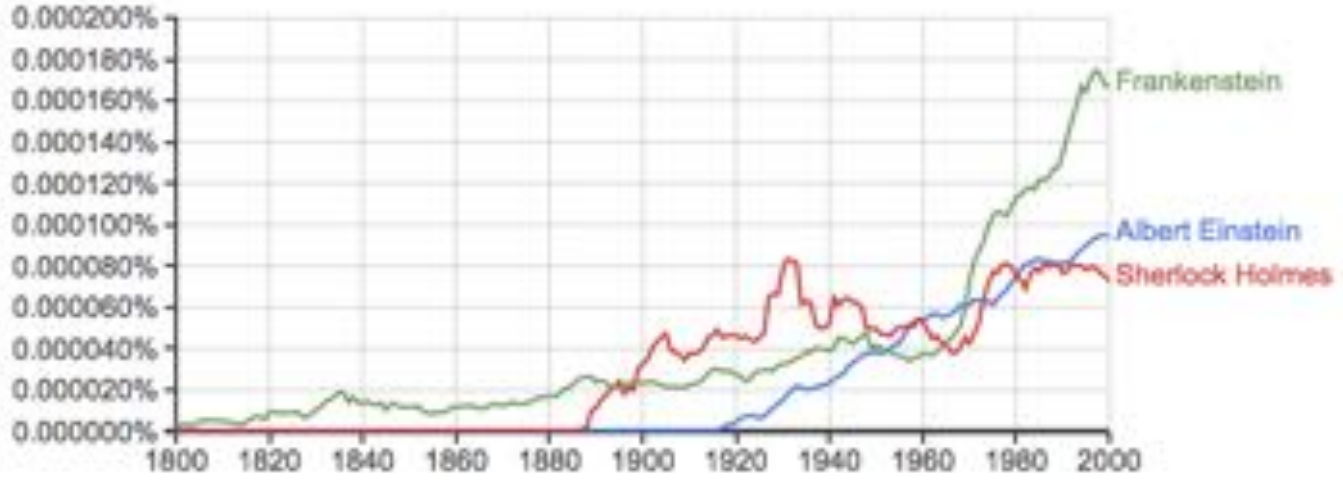




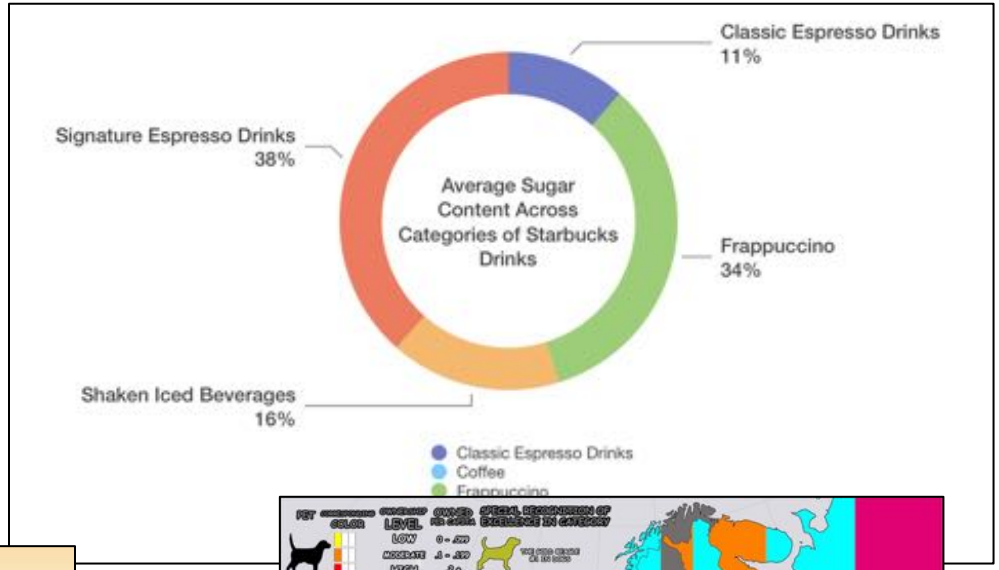
Piketty (2016)

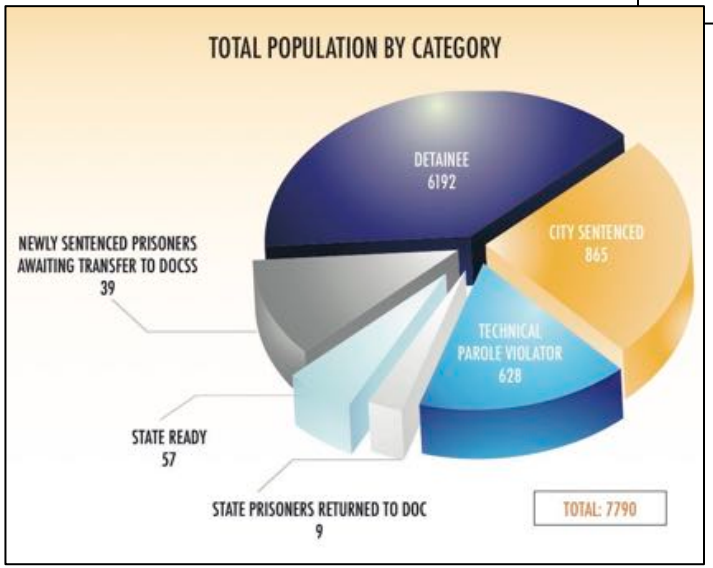
VISUALISATION DE DONNÉES
USAGES ET MÉSUSAGES

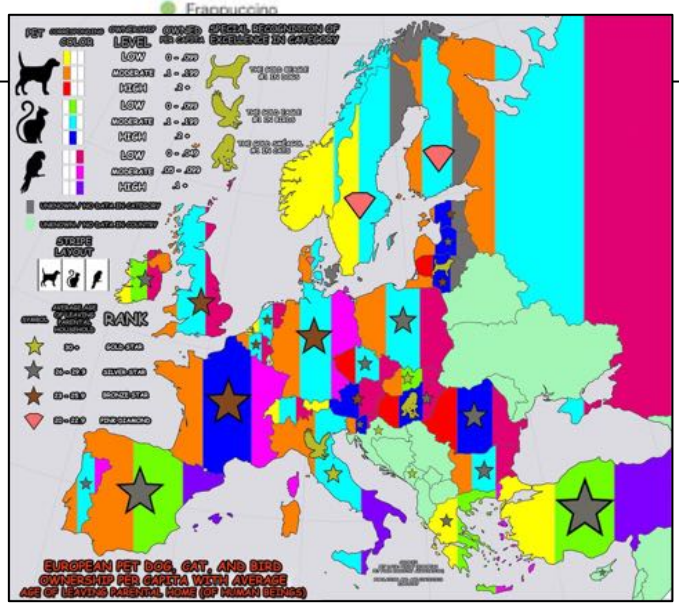
VISUALISATION DE DONNÉES
USAGES ET MÉSUSAGES



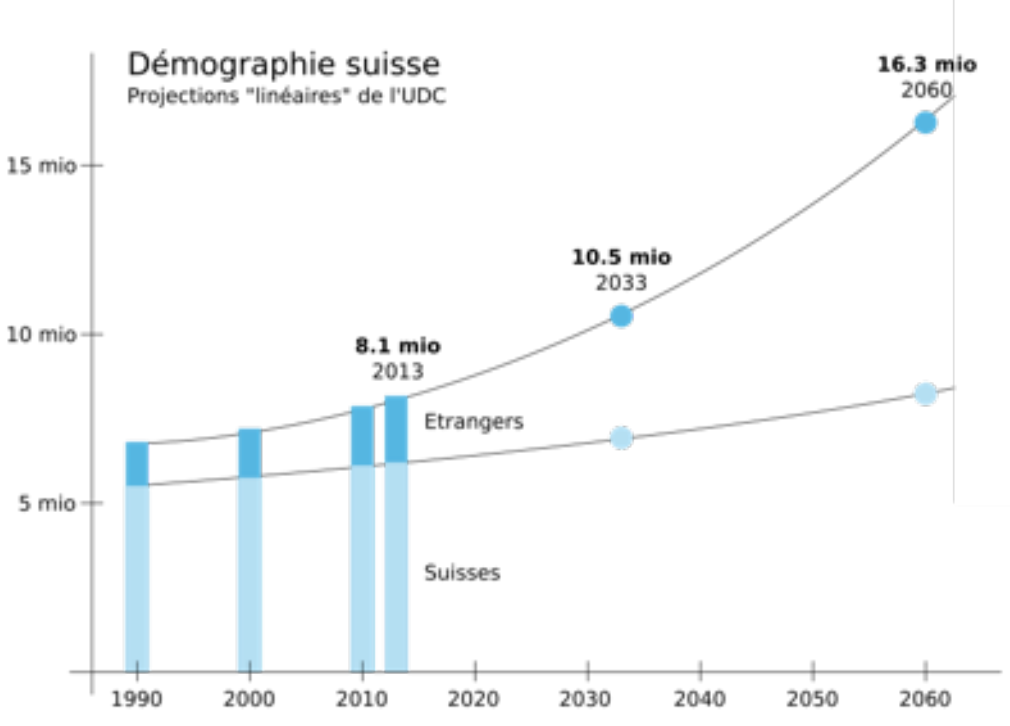
VISUALISATION DE DONNÉES
USAGES ET MÉSUSAGES



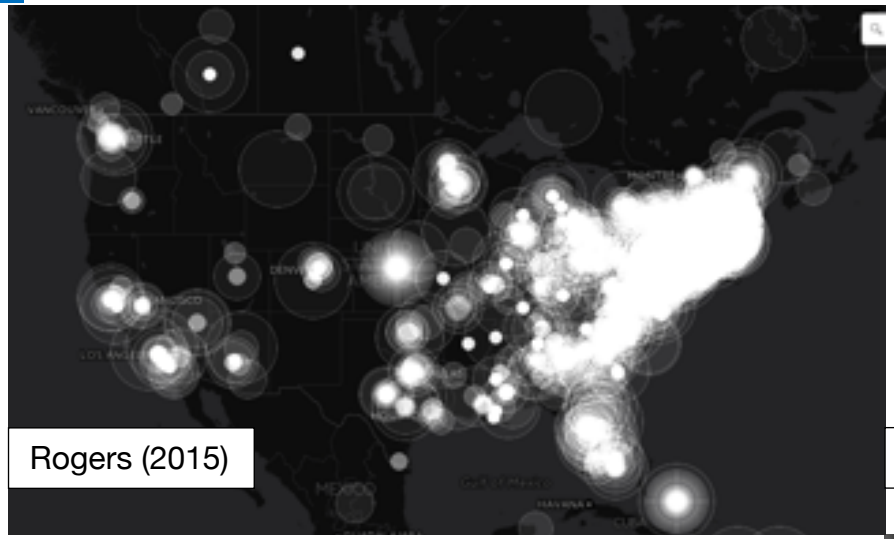




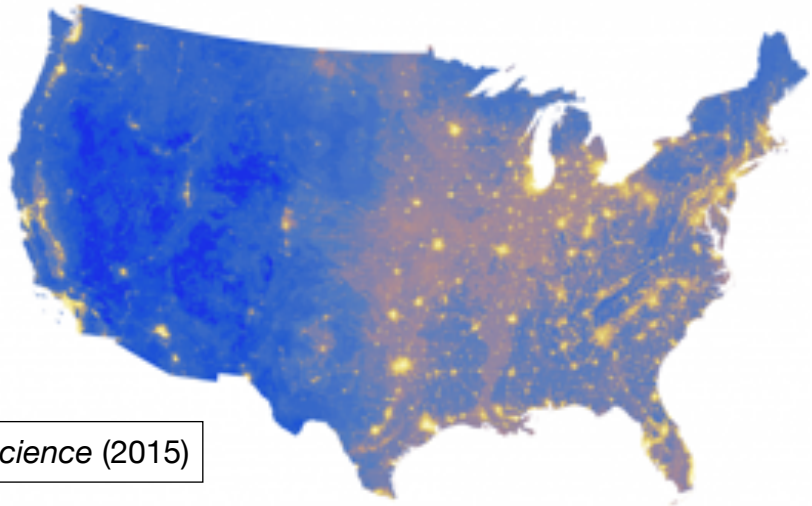
VISUALISATION DE DONNÉES
USAGES ET MÉSUSAGES



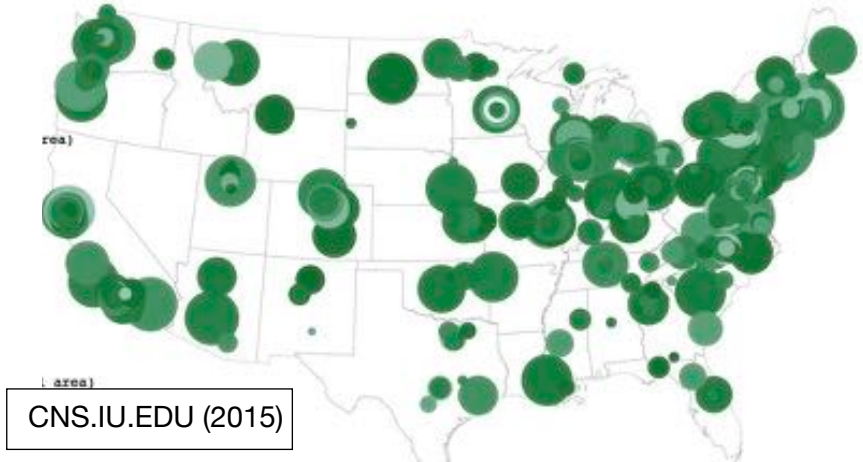
VISUALISATION DE DONNÉES
USAGES ET MÉSUSAGES



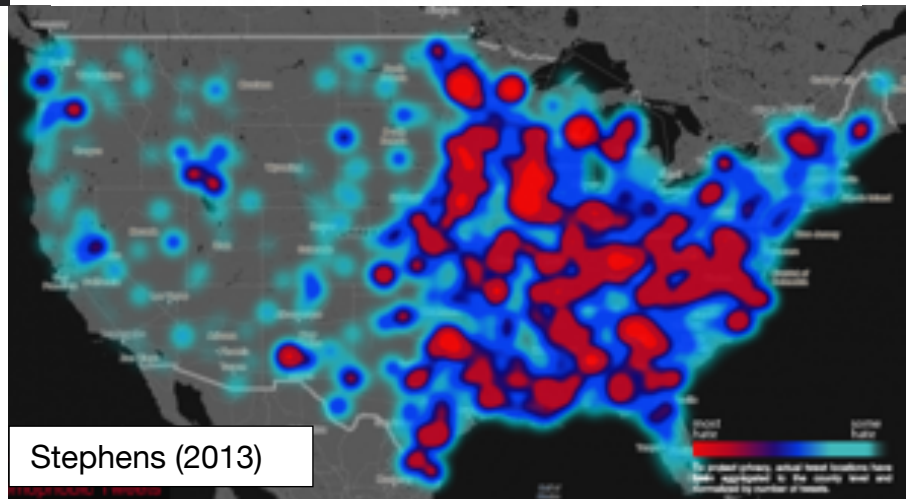
Rogers (2015)



Science (2015)

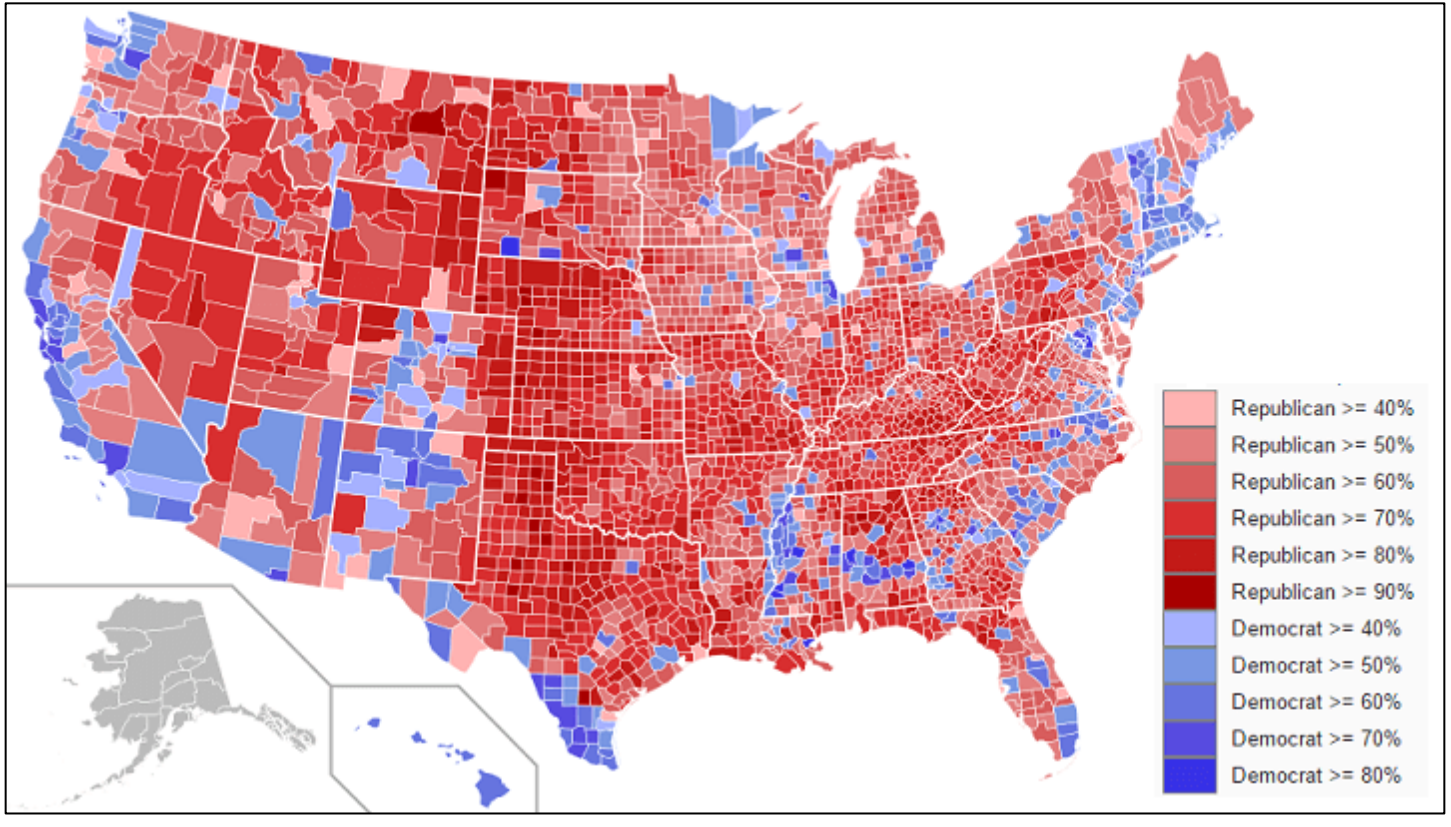


CNS.IU.EDU (2015)



Stephens (2013)

Résultats de l'élection présidentielle US 2016, par county.
Exemple tiré de Cairo, Alberto. 2018. *Visual Trumpery.*



VISUALISATION DE DONNÉES
USAGES ET MÉSUSAGES

Exemple tiré de Cairo, Alberto. 2018. *Visual Trumpery.*

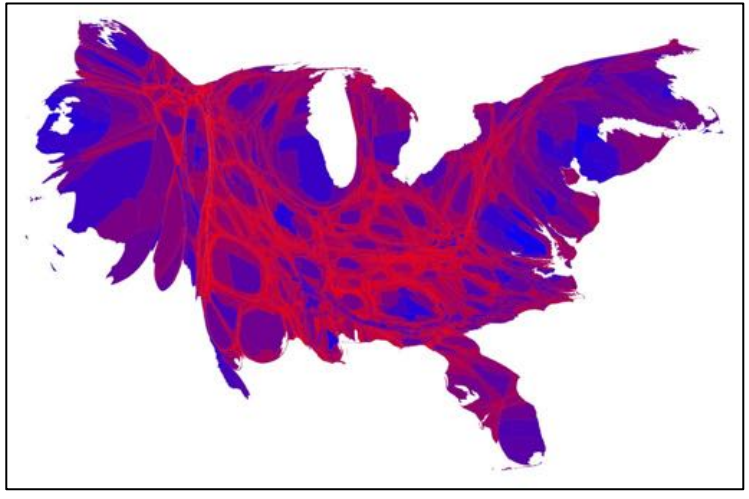
Surface on the county-level map:

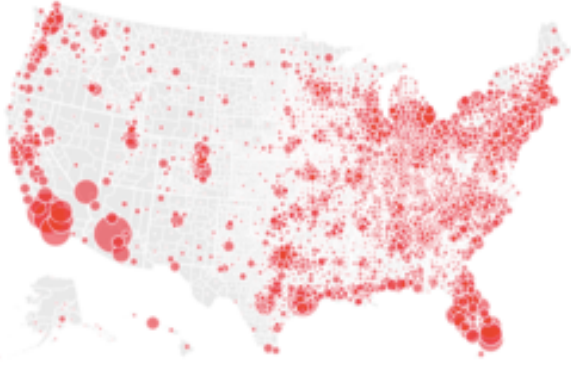
Red: 80%

Blue: 20%

SHARE OF THE POPULAR VOTE IN THE 2016 PRESIDENTIAL ELECTION

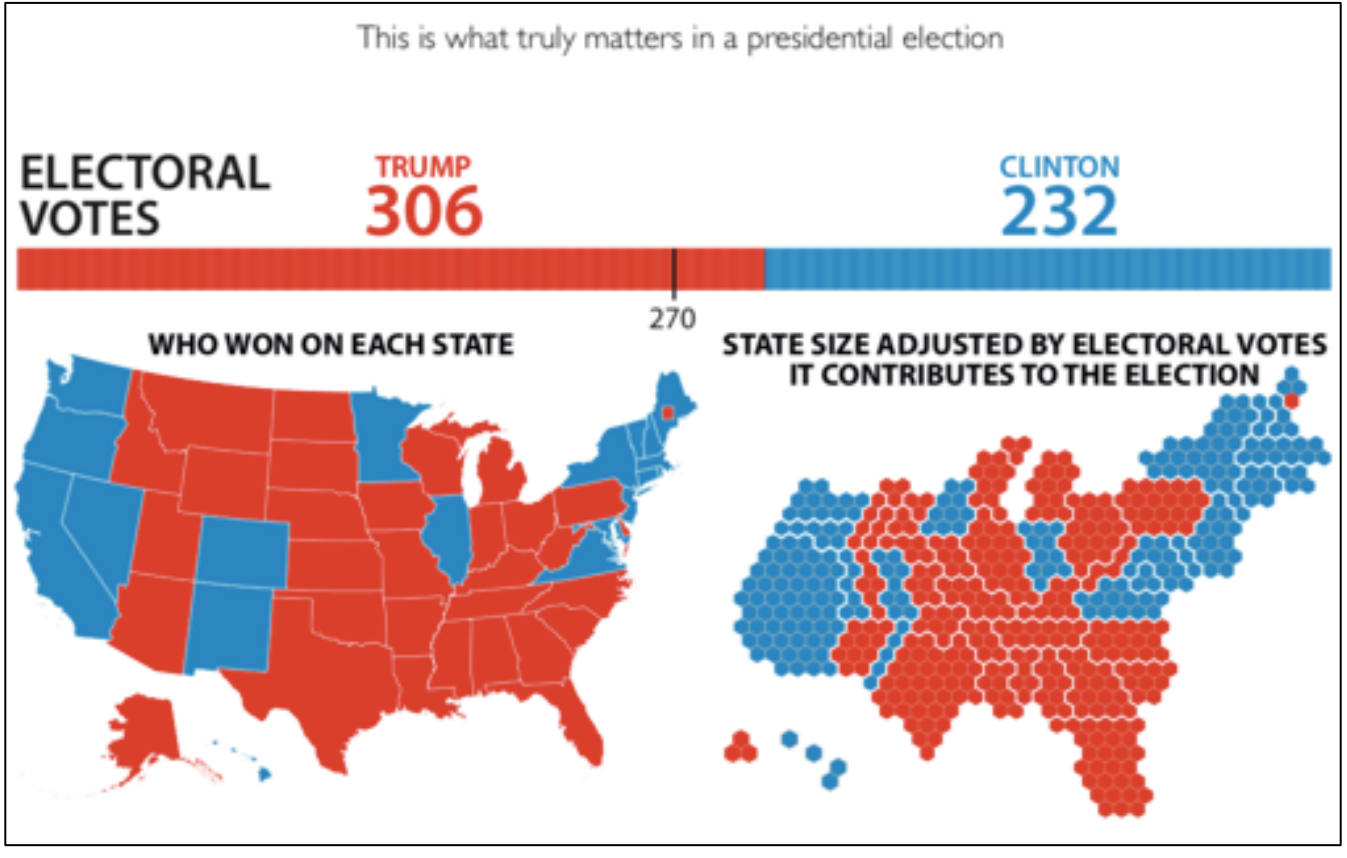
Donald Trump	46.1%	62,984,825 votes
Hillary Clinton	48.2%	65,853,516 votes
Other candidates	5.7%	



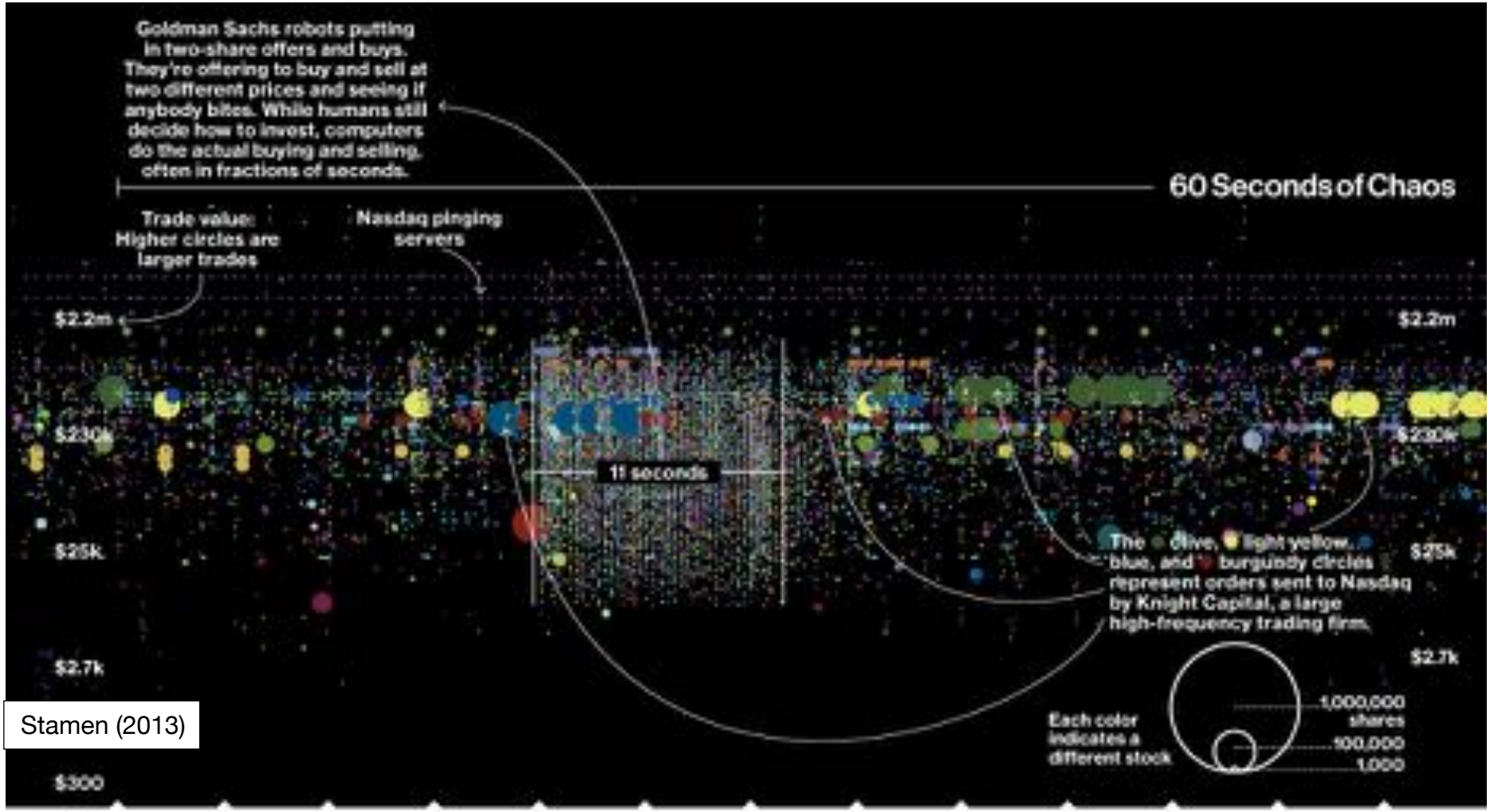


Bubble size is proportional to the number of votes per county

Exemple tiré de Cairo, Alberto. 2018. *Visual Trumpery.*

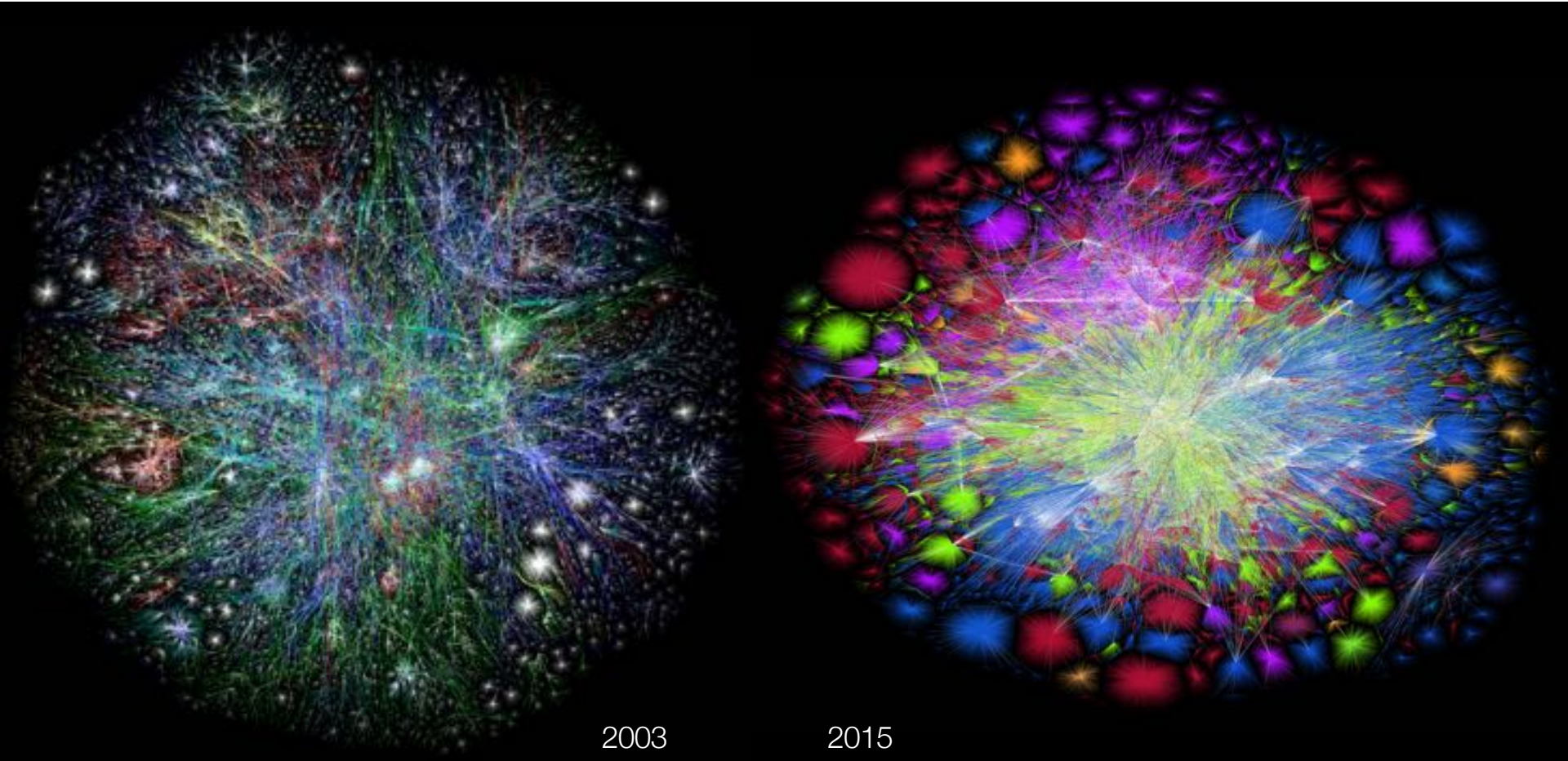


VISUALISATION DE DONNÉES
USAGES ET MÉSUSAGES

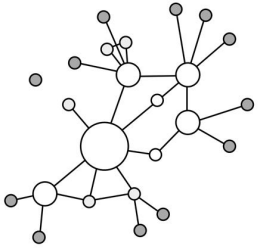


Stamen (2013)

VISUALISATION DE DONNÉES
USAGES ET MÉSUSAGES



B. Lyon, opte.org

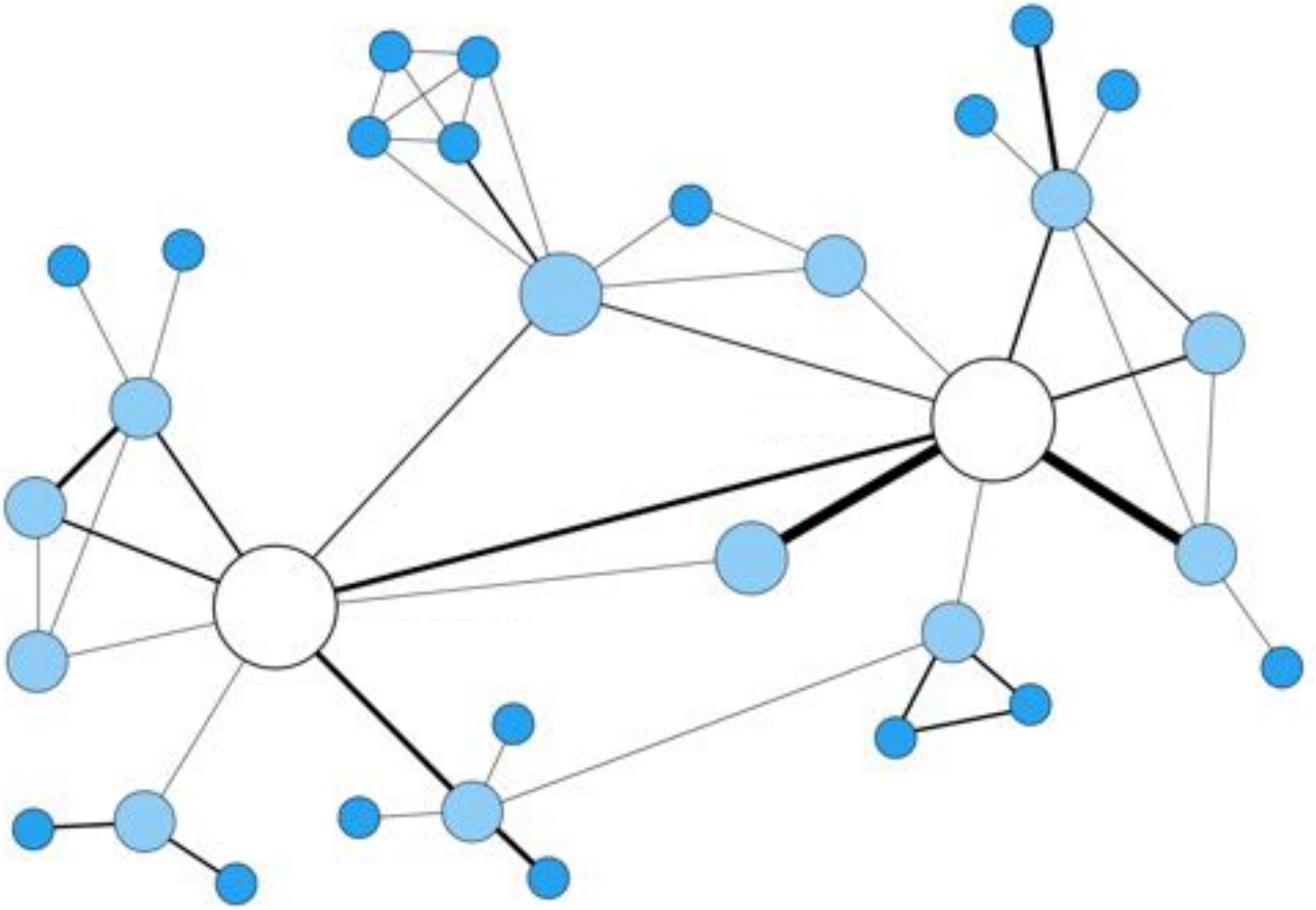


IV.

REPRÉSENTER LA COMPLEXITÉ : LA VISUALISATION DE RÉSEAU

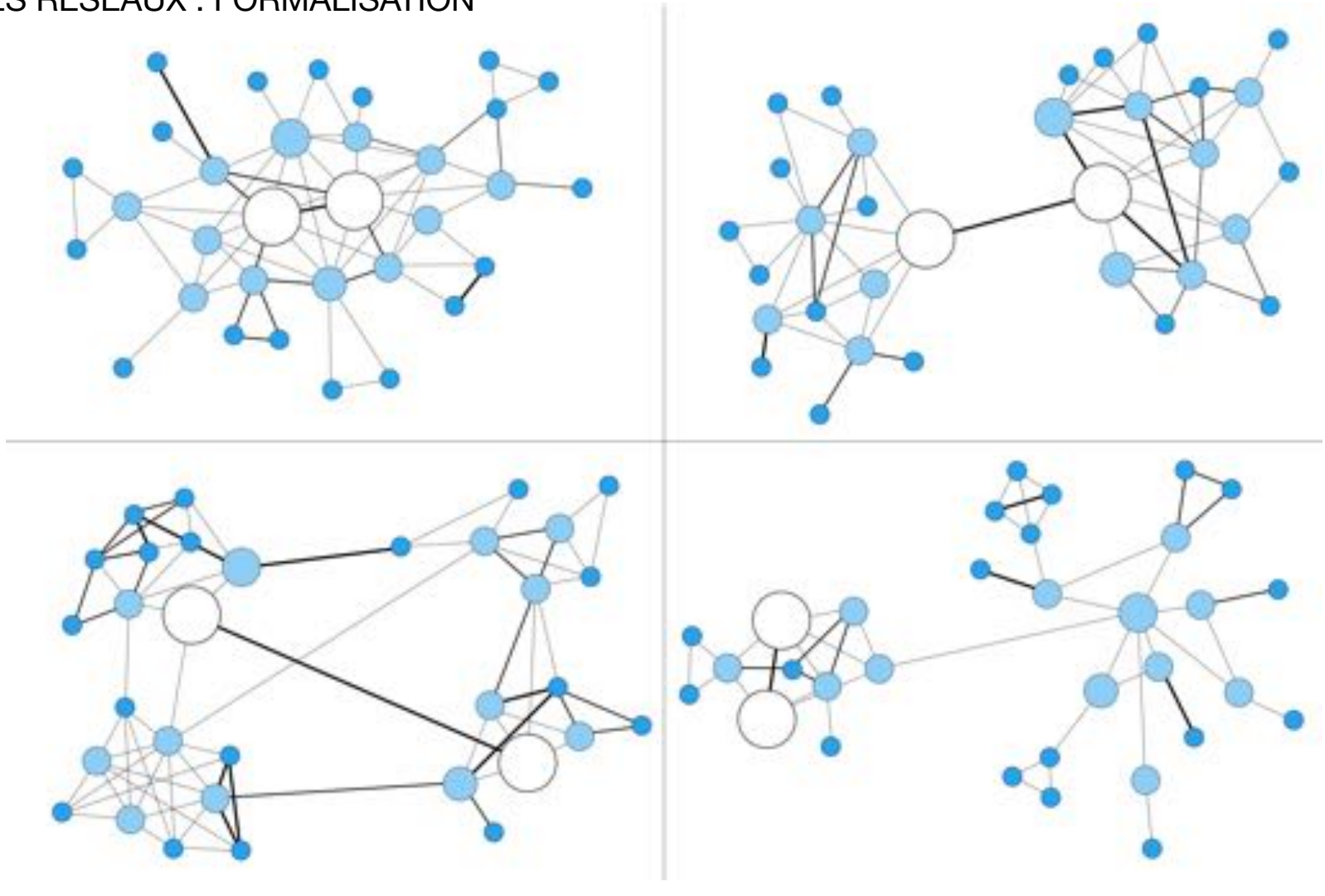
VISUALISATION DE DONNÉES

LES RÉSEAUX : FORMALISATION

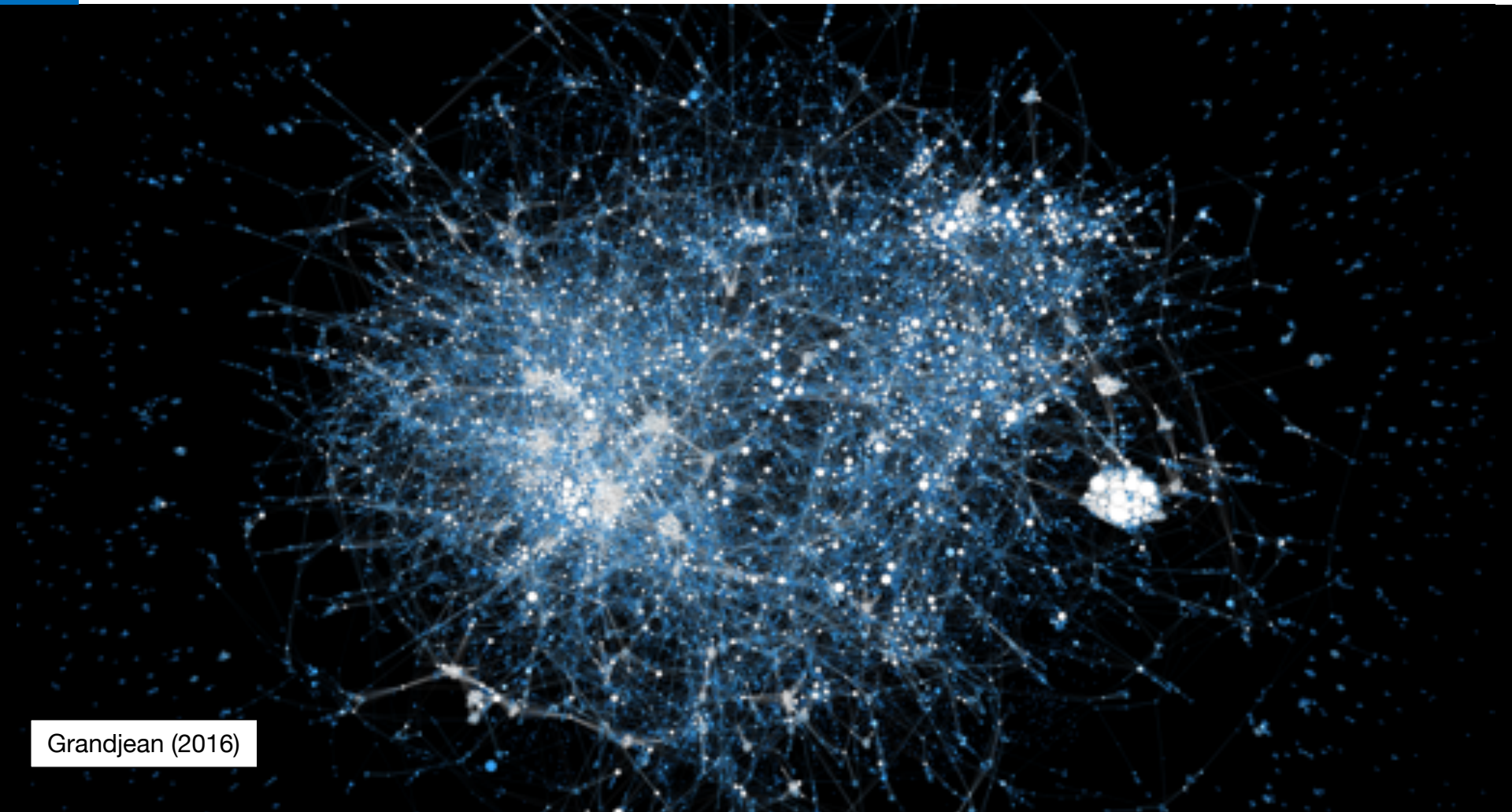


VISUALISATION DE DONNÉES

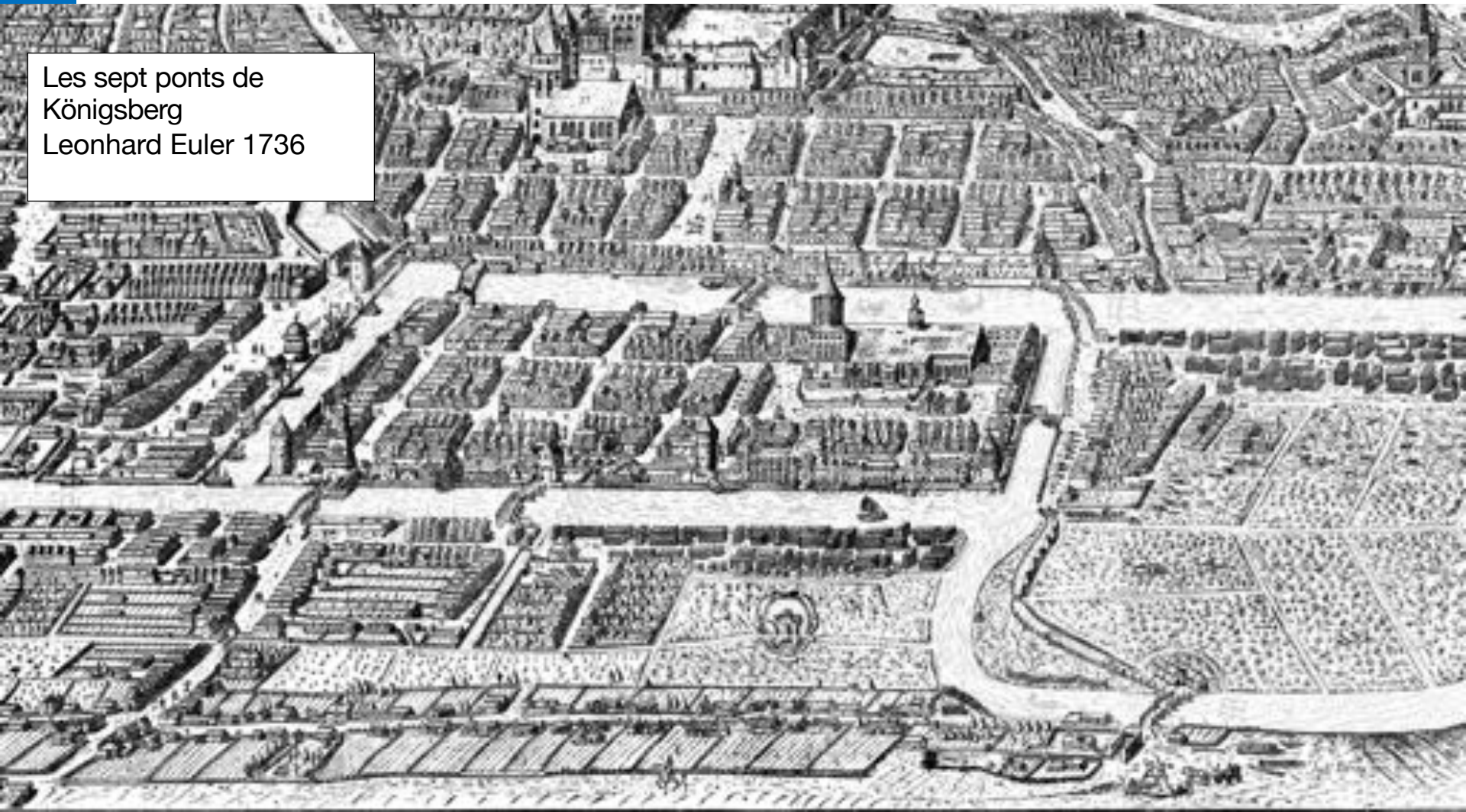
LES RÉSEAUX : FORMALISATION



VISUALISATION DE DONNÉES
LES RÉSEAUX : FORMALISATION

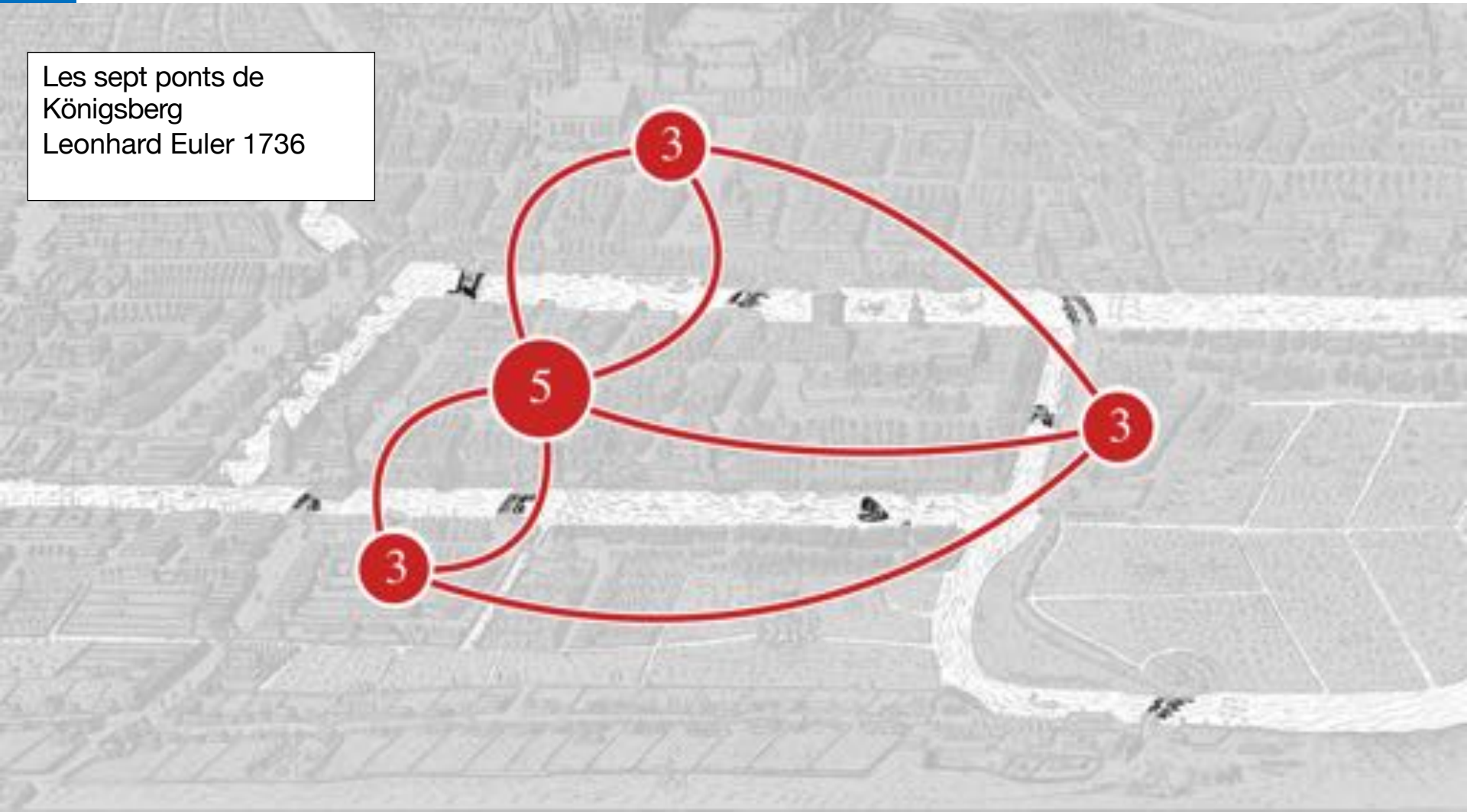


Grandjean (2016)



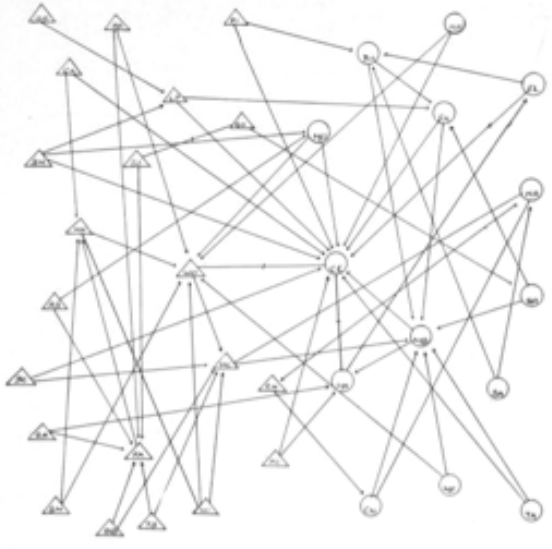
Les sept ponts de Königsberg
Leonhard Euler 1736

Les sept ponts de Königsberg
Leonhard Euler 1736



VISUALISATION DE DONNÉES
LES RÉSEAUX : FORMALISATION

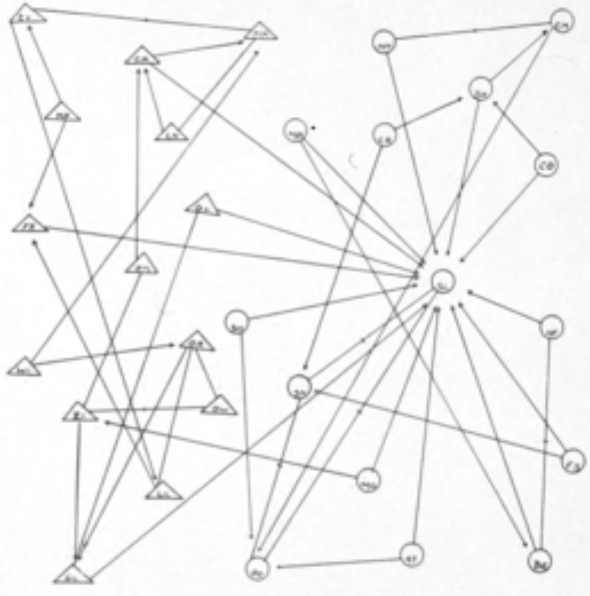
EVOLUTION OF GROUPS



CLASS STRUCTURE, 1ST GRADE

21 boys and 14 girls. *Unchosen*, 18, GO, PR, CA, SH, FI, RS, DC, GA, SM, BB, TS, WI, KI, TA, HF, SA, SR, KR; *Pairs*, 3, EI-GO, WO-CE, CE-HN; *Stars*, 5, CE, WO, HC, FA, MB; *Chains*, 0; *Triangles*, 0; *Inter-sexual Attractions*, 22.

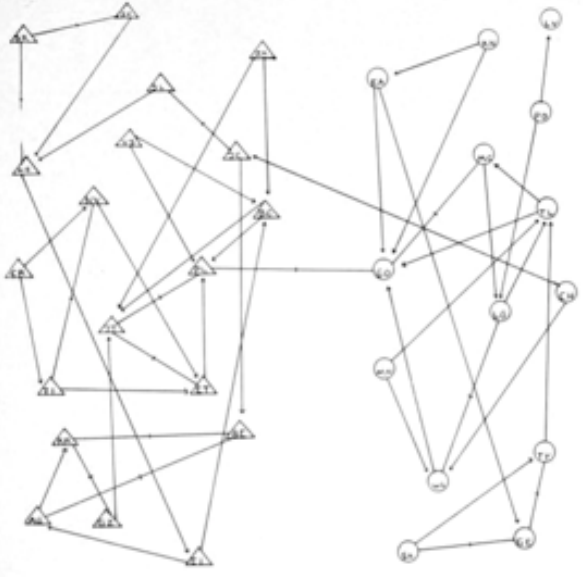
EVOLUTION OF GROUPS



CLASS STRUCTURE, 2ND GRADE

14 boys and 14 girls. *Unchosen*, 9, WI, KP, MG, AT, FS, CN, CR, MR, SH; *Pairs*, 11, ZV-MK, MK-LN, OW-ZI, GR-LL, ZI-JM, HN-CM, SL-JN, JN-PO, PO-SL, HF-BE, GL-GU; *Stars*, 2, SL, PO; *Chains*, 0; *Triangles*, 1, SL-JN-PO; *Inter-sexual Attractions*, 5.

EVOLUTION OF GROUPS



CLASS STRUCTURE, 3RD GRADE

19 boys and 14 girls. *Unchosen*, 7, VS, CR, CH, MN, PO, KN, ZK; *Pairs*, 14, SR-ZC, SR-NE, SL-JC, NV-TL, PL-JT, JT-ET, KR-BE, BE-AG, RR-GZ, PL-GO, GO-MC, WL-LG, SA-GE, GE-TY; *Stars*, 3, GO, PL, JT; *Chains*, 1, ET-JT-PL-GO-MC; *Triangles*, 0; *Inter-sexual Attractions*, 3.

Moreno (1934)

VISUALISATION DE DONNÉES
LES RÉSEAUX : FORMALISATION

STRUCTURE OF A KINDERGARTEN

15 boys and 18 girls. *Unchosen* 9, RS, NV, FE, MA, TO, AS, RG, SI, PR; Pairs 3, AL-WB, WB-SN, SN-LW; *Stars,* (Centers of Attractions), PG, SN, MR; *Chains* (of relationships), 0; *Triangles,* 0; *Inter-sexual Attractions,* 19.

TABLE 1.

QUANTITATIVE ANALYSIS OF CHOICES BETWEEN THE SEXES IN THE GRADES OF A PUBLIC SCHOOL

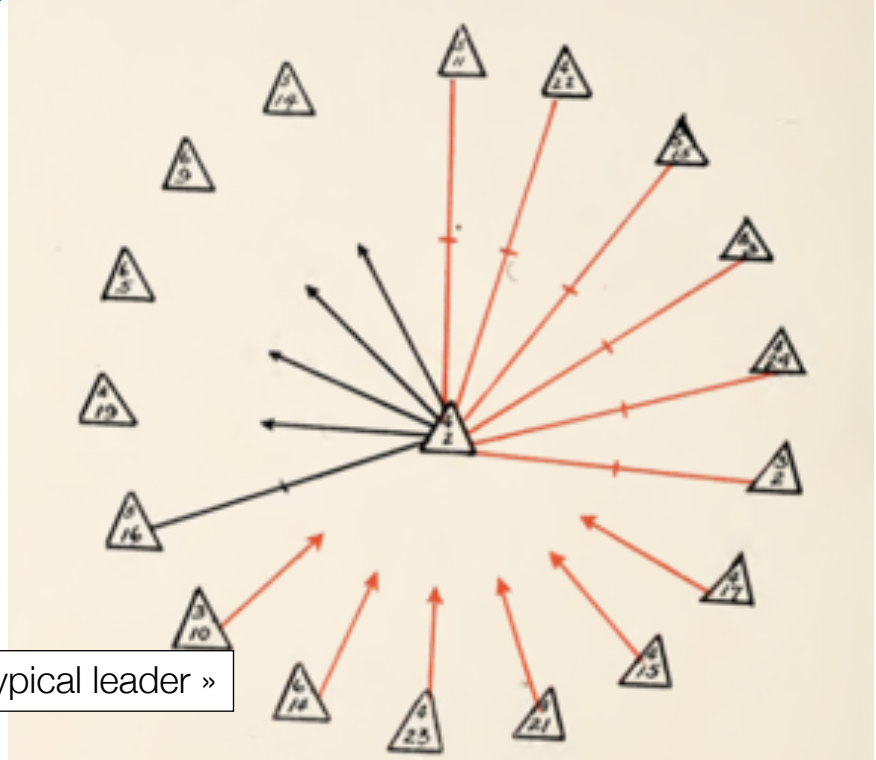
Of All Choices Made:	*Boys Chose Girls*	*Girls Chose Boys*	*Both Sexes*
Kindergarten	10.0%	15.0%	25.0%
1st Grade	19.0%	8.0%	27.0%
2nd Grade	10.0%	6.5%	16.5%
3rd Grade	3.5%	5.0%	8.5%
4th Grade	1.5%	1.0%	2.5%
5th Grade	1.0%	4.5%	5.5%
6th Grade	2.05%	2.05%	4.1%
7th Grade	2.0%	1.0%	3.0%
8th Grade	4.0%	4.0%	8.0%

TABLE 2

STRUCTURAL ANALYSIS OF SCHOOL GROUPS

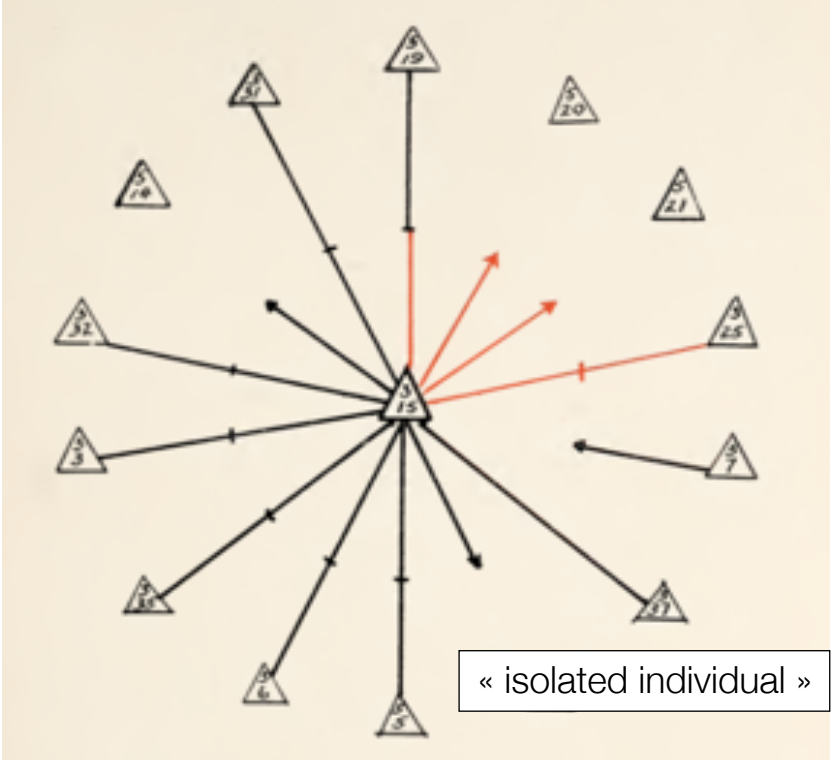
	Isolated	*Mutual Pairs*	*Triangles.*	*Chains*
Kindergarten	35%	7%	0	0
1st Grade	27%	9%	0	0
2nd Grade	28%	13%	1	0
3rd Grade	26%	18%	1	1
4th Grade	18%	20%	5	4
5th Grade	15%	20%	3	2
6th Grade	19%	22%	2	9
7th Grade	20%	25%	5	7
8th Grade	18%	17%	3	3

VISUALISATION DE DONNÉES
LES RÉSEAUX : FORMALISATION



« typical leader »

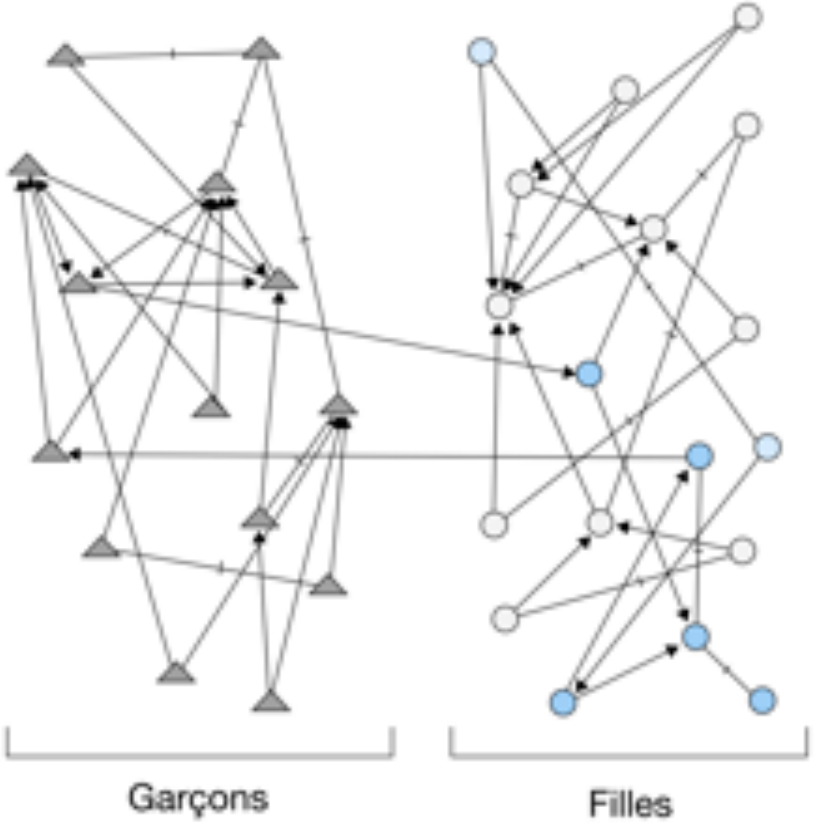
The chart shows the position occupied by a typical leader in the school. Note that 4/2 has six reciprocal attractions; six additional attractions toward him by individuals to whom he is indifferent; one mutual dislike and four additional dislikes, each of whom feels neutral toward him. This is a high "batting average." The total number of individuals (17) toward whom relations have been expressed in the course of the survey, distributed over four forms, indicates that 4/2 takes an active attitude toward his fellow students and to life and is not on the "side lines."

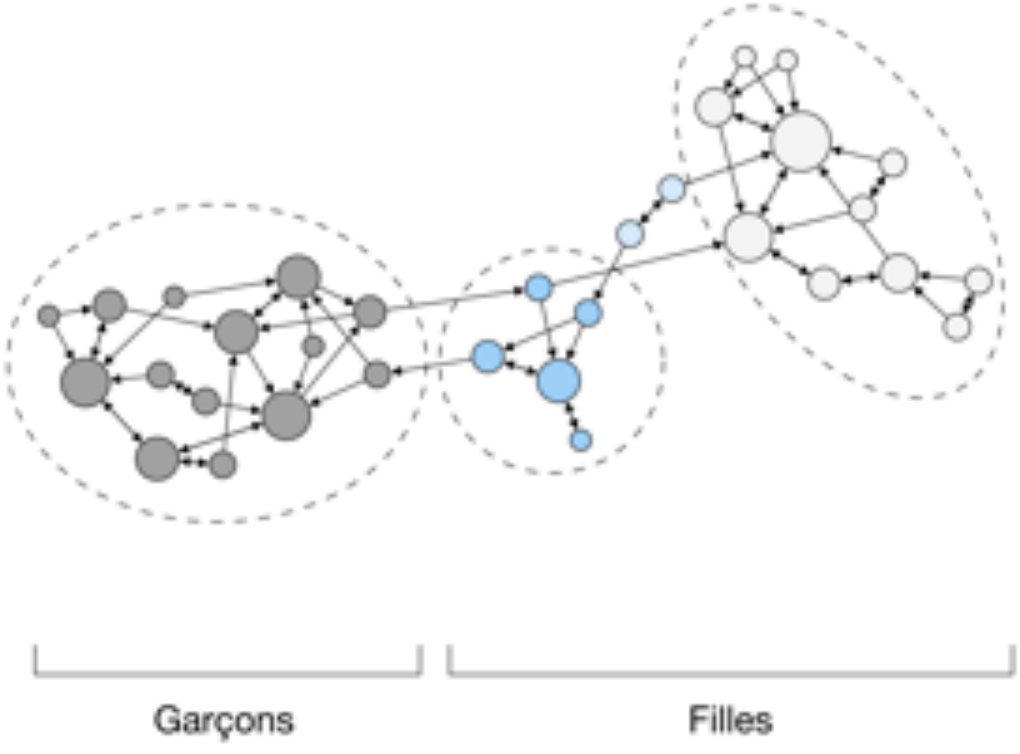


« isolated individual »

The chart depicts an isolated individual. The student 3/15 arouses active feelings—he is the focus of a great deal of antipathy; he likes 3/25, and is accepted by him; he also likes two others, 5/20 and 5/21 who, significantly are new boys and who are indifferent to him, and he also likes 5/19, who, however, dislikes him. This shows that 3/15 has formed only one satisfactory contact in the whole school, while there are seven mutual dislikes, two dislikes for persons who are indifferent to him, and one dislike toward him from a person toward whom he is indifferent.

VISUALISATION DE DONNÉES
LES RÉSEAUX : FORMALISATION

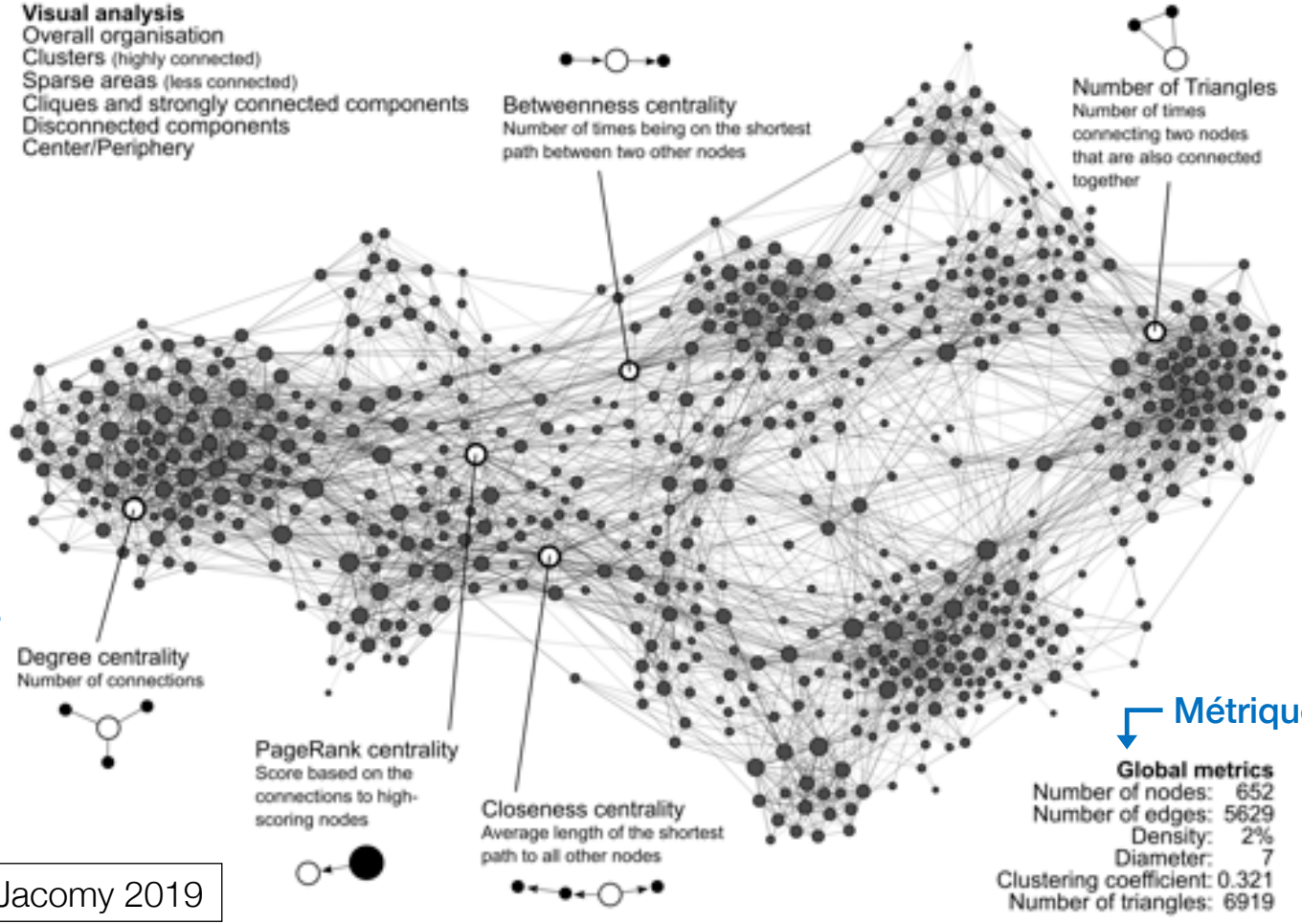




Moreno (1934)

VISUALISATION DE DONNÉES
LES RÉSEAUX : FORMALISATION

1
Analyse visuelle



2
Métriques locales

3
Métriques globales

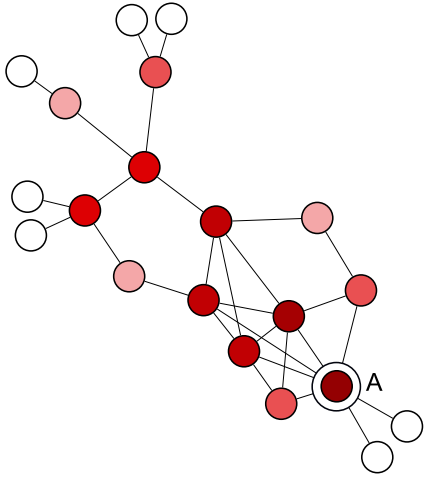
Grandjean et Jacomy 2019

VISUALISATION DE DONNÉES

LES RÉSEAUX : FORMALISATION

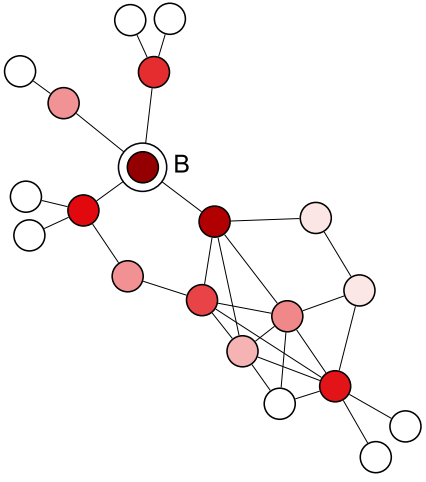
Centralité de degré

Nombre de connexions



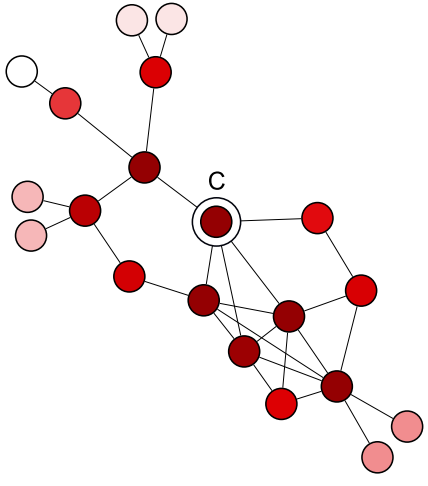
Centralité d' intermédiarité

Position sur les plus courts chemins



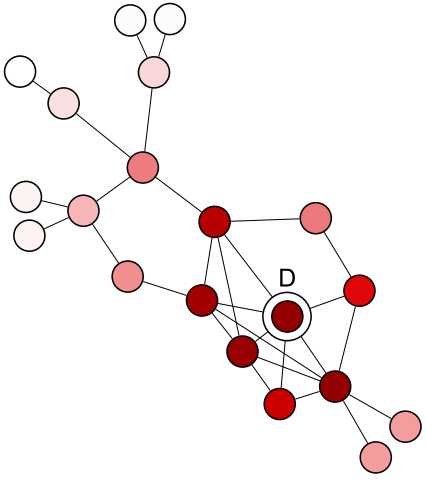
Centralité de proximité

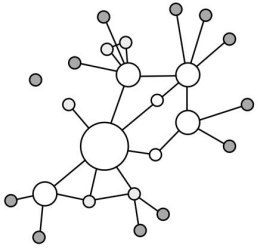
Distance moyenne aux autres sommets



Centralité de vecteur propre

Score d'autorité basé sur le score des voisins

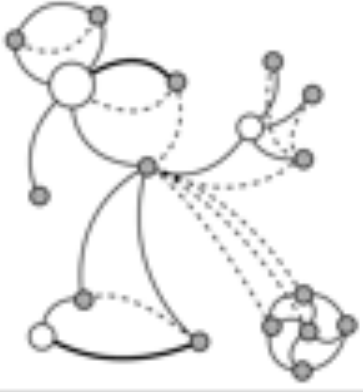
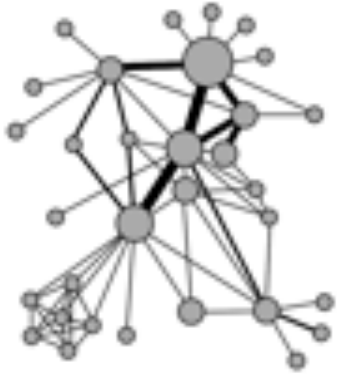




V.
LES RÉSEAUX EN SCIENCES HUMAINES ET SOCIALES

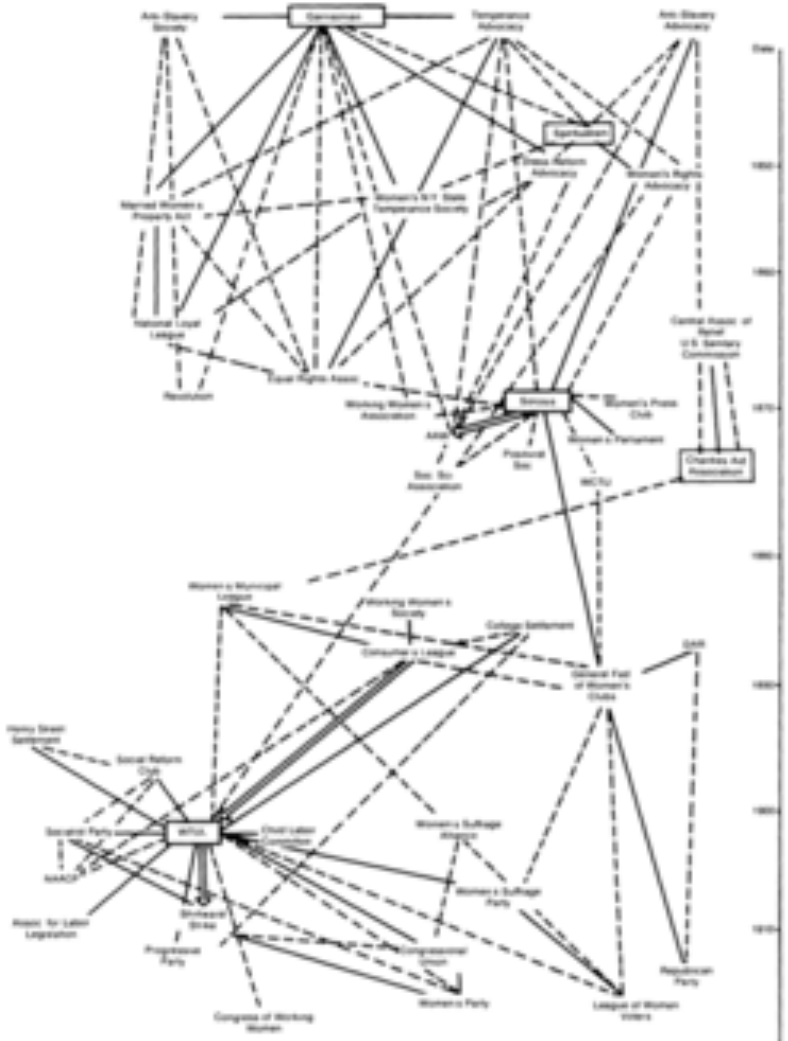
VISUALISATION DE DONNÉES
LES RÉSEAUX EN SCIENCES HUMAINES ET SOCIALES

Usage des sources

LE RÉSEAU MÉTAPHORE	LE RÉSEAU RECONSTITUÉ	LE RÉSEAU TIRÉ DES SOURCES	LE RÉSEAU DE MÉTADONNÉES
Pas de représentation graphique			
Utilisation de la terminologie de l'analyse de réseau pour décrire une situation.	Réseau infographique, dessiné à partir d'informations compilées parfois hétérogènes.	Réseau extrait du contenu d'un corpus bien délimité, souvent une liste d'affiliations.	Réseau représentant les documents eux-mêmes (leur circulation, leur proximité thématique, etc.), pas leur contenu.

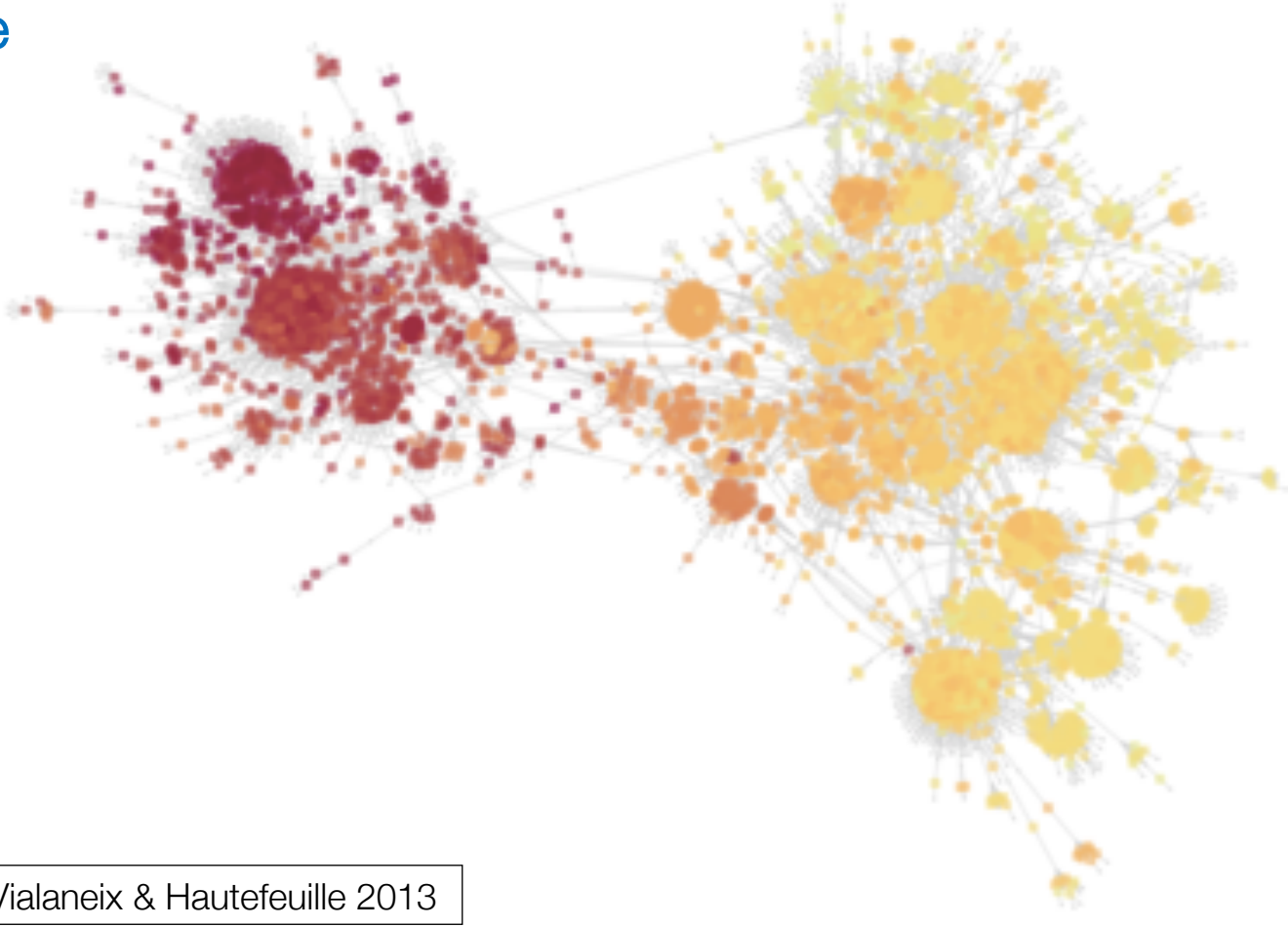
VISUALISATION DE DONNÉES
LES RÉSEAUX EN SCIENCES HUMAINES ET SOCIALES

Réseau de contenu



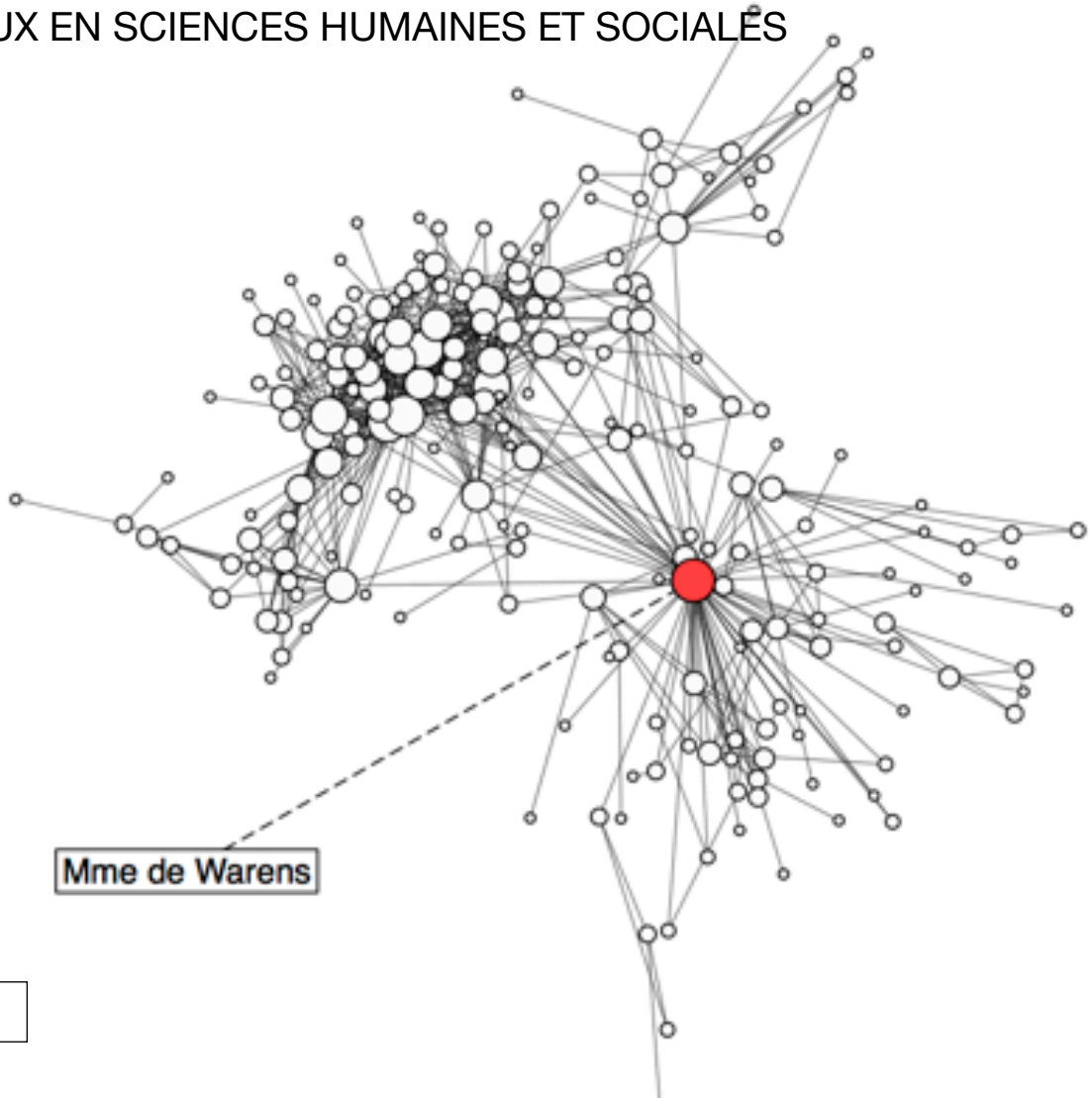
Rosenthal, Fingrutd, Ethier, Karant & McDonald 1985

Réseau de contenu



Rossi, Villa-Vialaneix & Hautefeuille 2013

Réseau de contenu



Rochat (2014)

VISUALISATION DE DONNÉES
LES RÉSEAUX EN SCIENCES HUMAINES ET SOCIALES

Réseau de contenu

OTHELLO
Number of characters 24 | 55% Network density

KING LEAR
Number of characters 33 | 45% Network density

MACBETH
Number of characters 46 | 25% Network density

TIMON OF ATHENS
Number of characters 51 | 25% Network density

HAMLET
Number of characters 37 | 39% Network density

JULIUS CAESAR
Number of characters 46 | 34% Network density

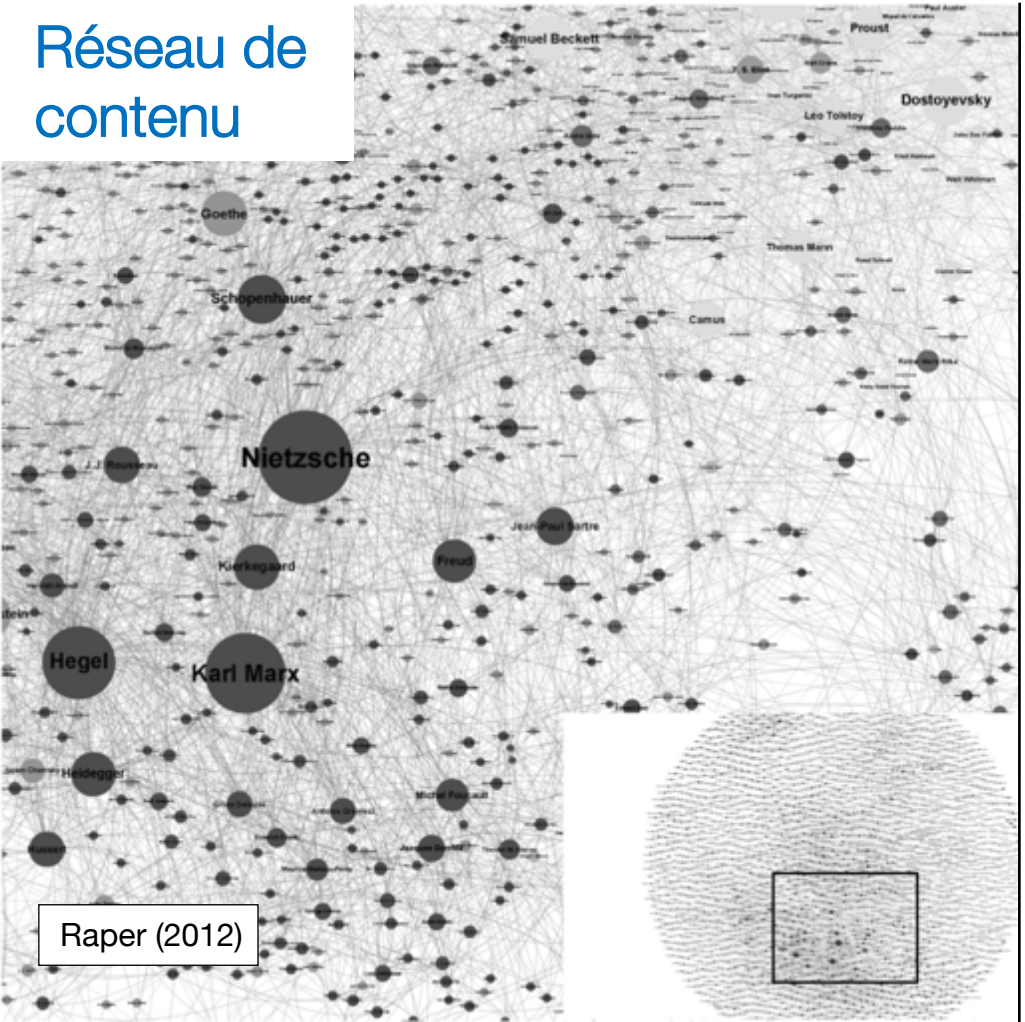
ROMEO AND JULIET
Number of characters 41 | 37% Network density

TITUS ANDRONICUS
Number of characters 36 | 50% Network density

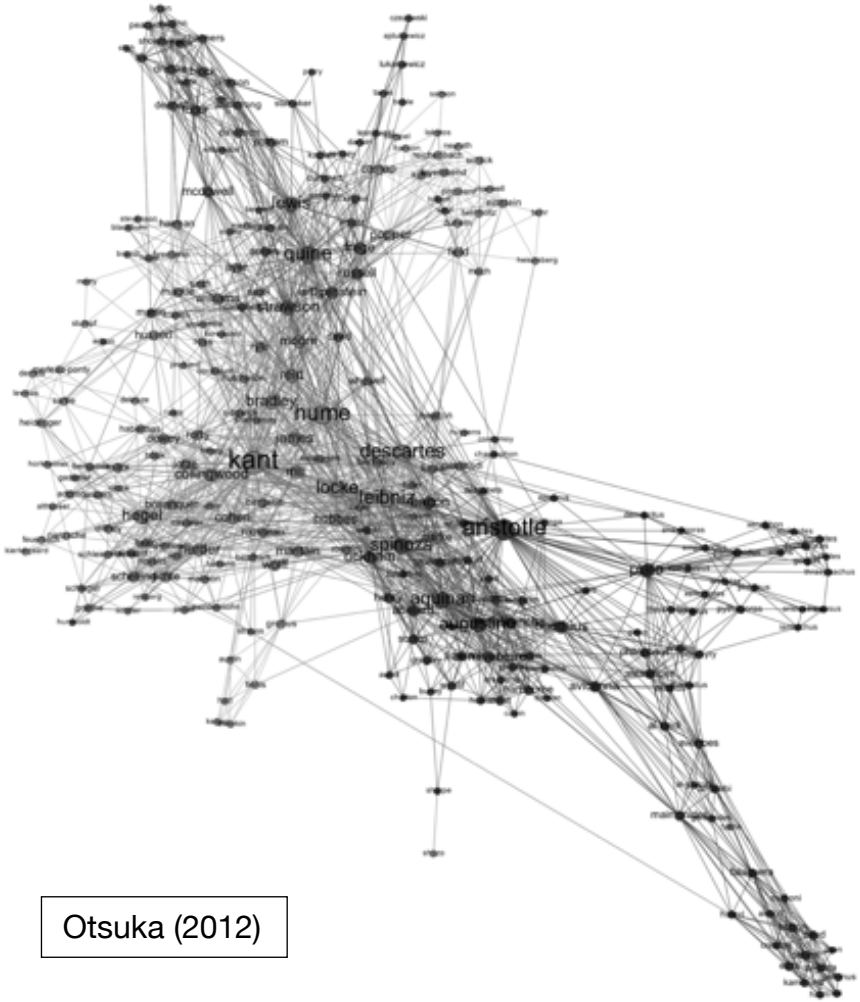
Grandjean (2015)

VISUALISATION DE DONNÉES
LES RÉSEAUX EN SCIENCES HUMAINES ET SOCIALES

Réseau de contenu



Raper (2012)



Otsuka (2012)

VISUALISATION DE DONNÉES
LES RÉSEAUX EN SCIENCES HUMAINES ET SOCIALES

Circulation/ métadonnées

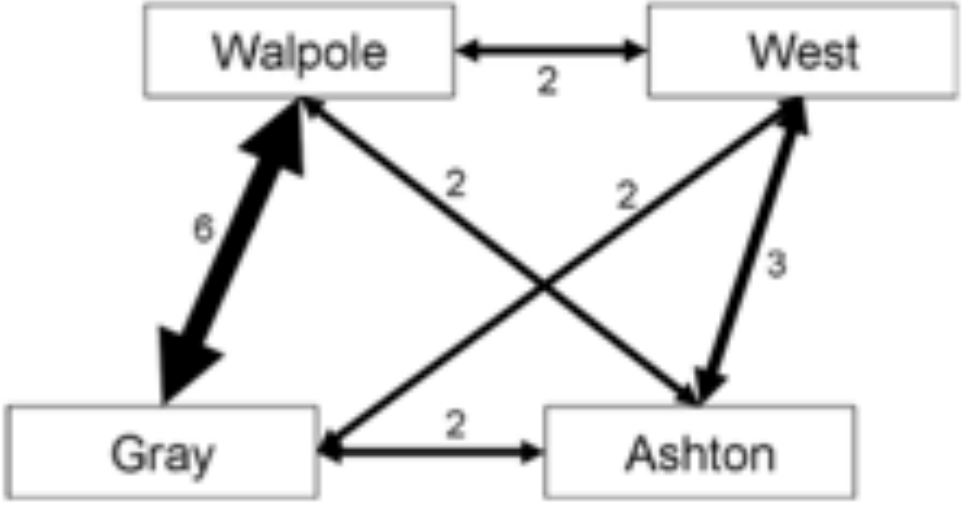
Stanford « Republic of Letters »

VISUALISATION DE DONNÉES
LES RÉSEAUX EN SCIENCES HUMAINES ET SOCIALES

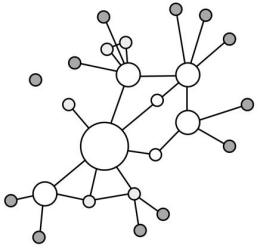
Circulation/ métadonnées

Winterer 2012

Circulation/ métadonnées



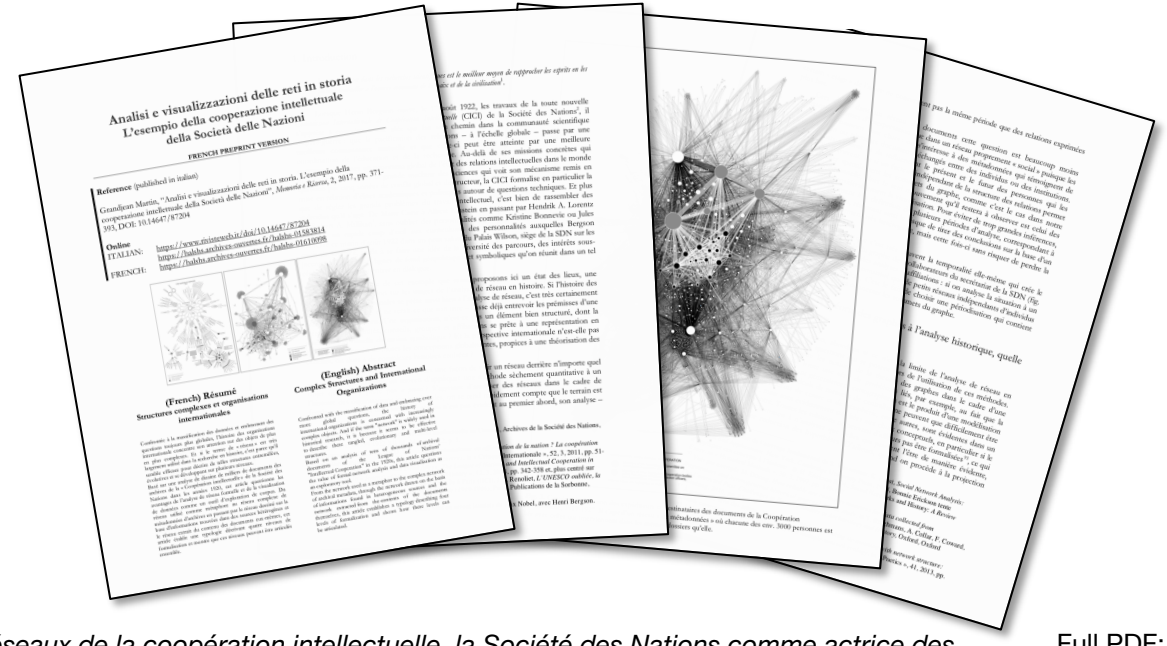
Henstra 2014



VI.

ÉTUDE DE CAS : ARTICULER TYPES, NIVEAUX ET MÉTRIQUES

VISUALISATION DE DONNÉES
ARTICULER TYPES, NIVEAUX ET MÉTRIQUES DANS LES RÉSEAUX



Grandjean Martin, *Les réseaux de la coopération intellectuelle, la Société des Nations comme actrice des échanges scientifiques et culturels dans l'entre-deux-guerres*, PhD Thesis, University of Lausanne, 2018, 600 p.

Full PDF: https://halshs.archives-ouvertes.fr/tel-01853903

Grandjean Martin, "Analisi e visualizzazioni delle reti in storia : l'esempio della cooperazione intelletuale della Società delle Nazioni", *Memoria e Ricerca*, 25, 2, 2017, pp. 371-393.

Preprint: https://halshs.archives-ouvertes.fr/halshs-01610098v2

Grandjean Martin, "Multimode and Multilevel: Vertical Dimension in Historical and Literary Networks", *Digital Humanities 2017*, Montreal, 2017.

Abstract: https://halshs.archives-ouvertes.fr/halshs-01525539

Grandjean Martin, *The Network of Intellectual Cooperation*,
Data set, 2018, https://github.com/grandjeanmartin/intellectual-cooperation. DOI 10.5281/zenodo.1257288

VISUALISATION DE DONNÉES
ARTICULER TYPES, NIVEAUX ET MÉTRIQUES DANS LES RÉSEAUX

La « Commission internationale de coopération intellectuelle »
Société des Nations (Genève), 1921-1939

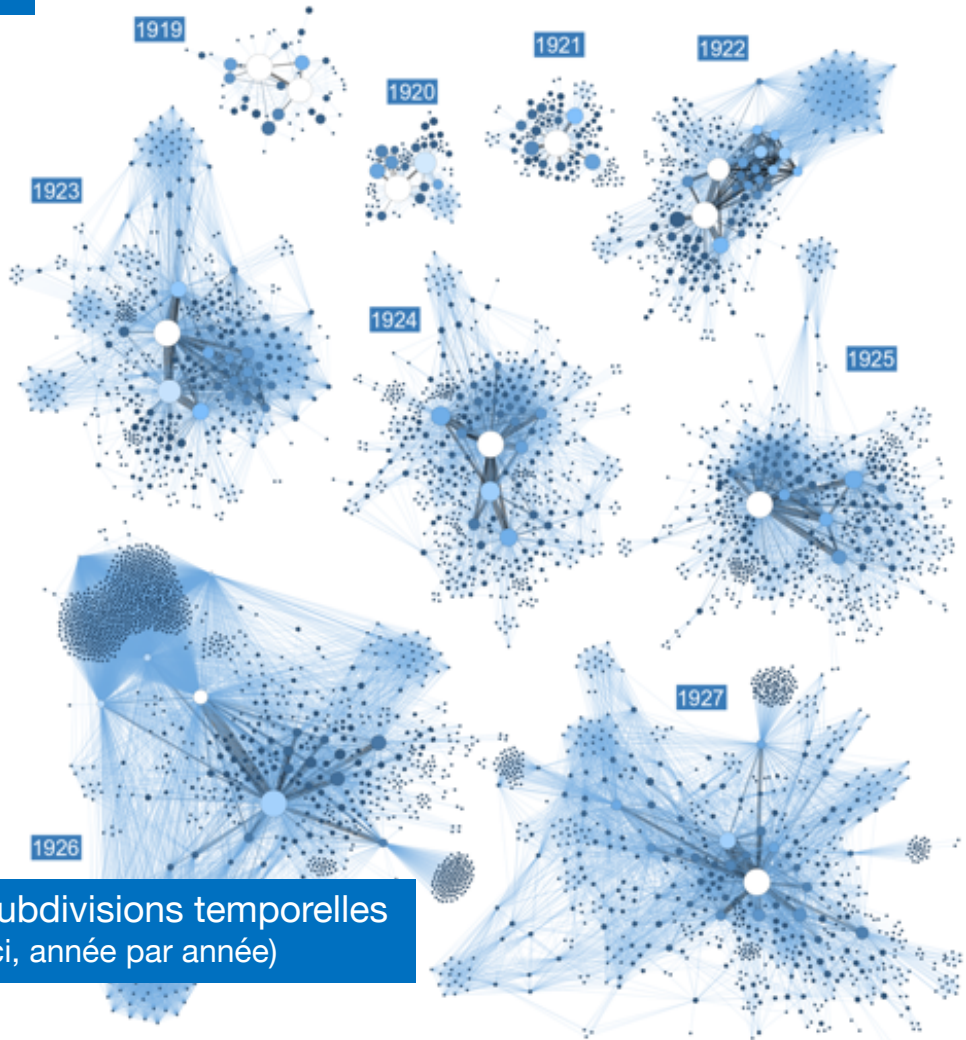
CICI [s.d. 1924-1927]
Archives SDN Genève CM067

VISUALISATION DE DONNÉES
ARTICULER TYPES, NIVEAUX ET MÉTRIQUES DANS LES RÉSEAUX

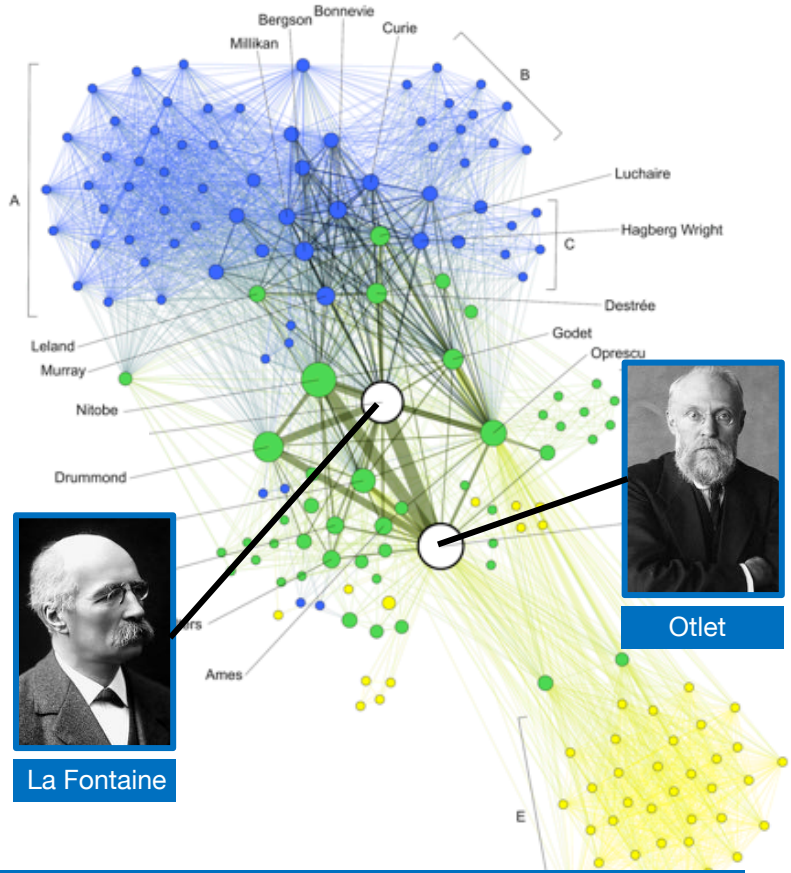
ICIC Archives 1919-1927

2.900 files
30.000 documents
with
3.200 agents/nodes
26.500 relations/edges

VISUALISATION DE DONNÉES
ARTICULER TYPES, NIVEAUX ET MÉTRIQUES DANS LES RÉSEAUX

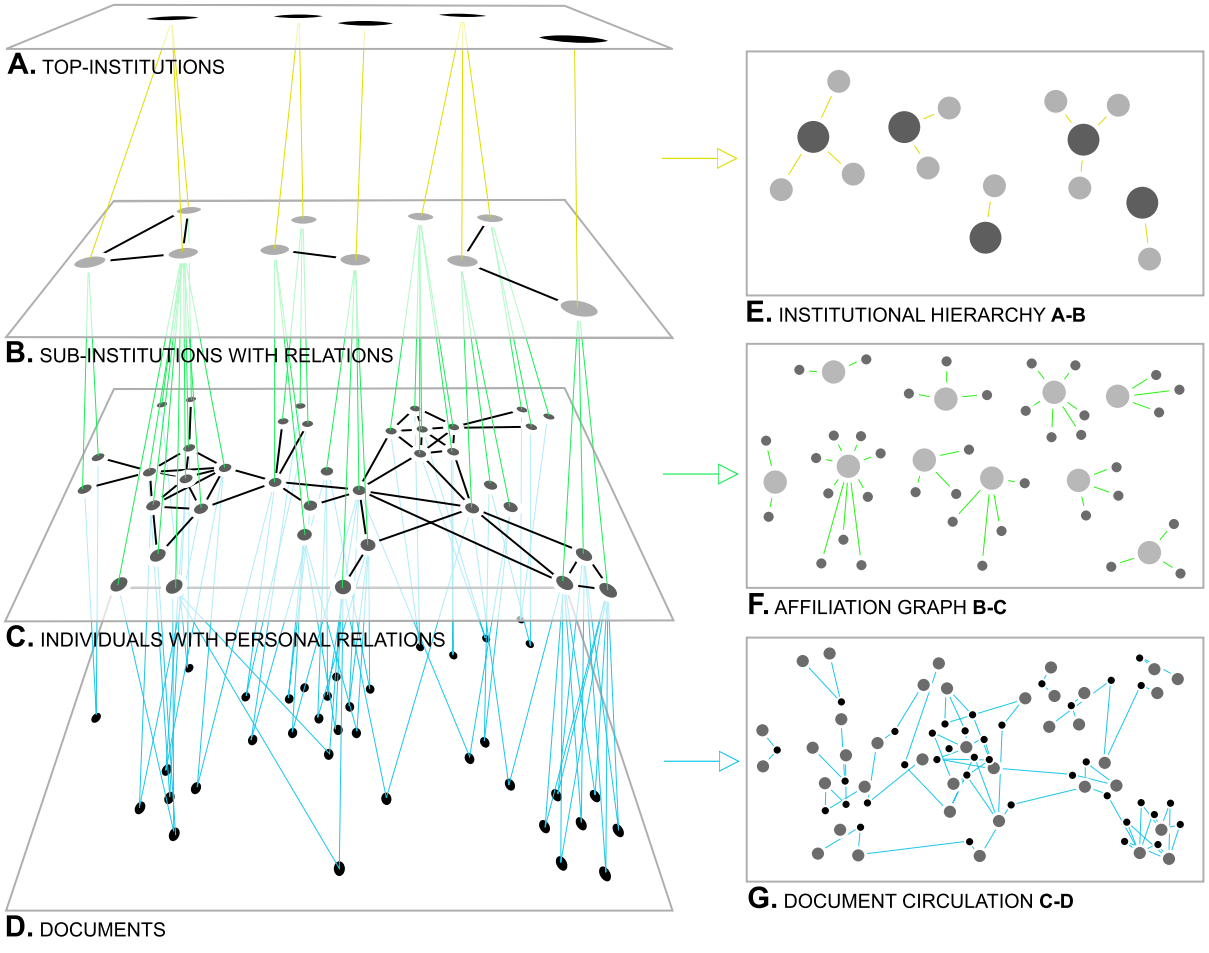


Subdivisions temporelles
(ici, année par année)



Étude de cas
(sélection du réseau d'une ou plusieurs personnes

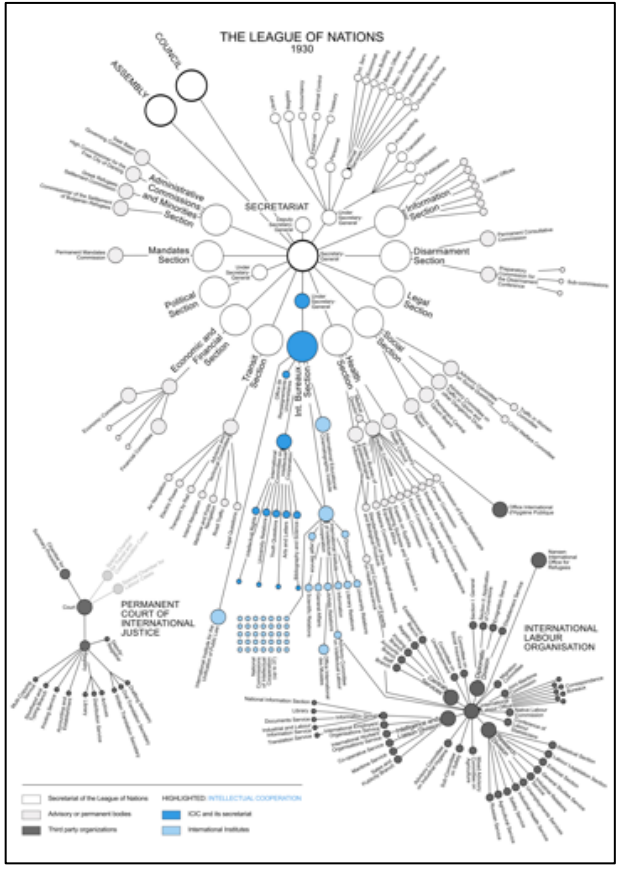
VISUALISATION DE DONNÉES
ARTICULER TYPES, NIVEAUX ET MÉTRIQUES DANS LES RÉSEAUX



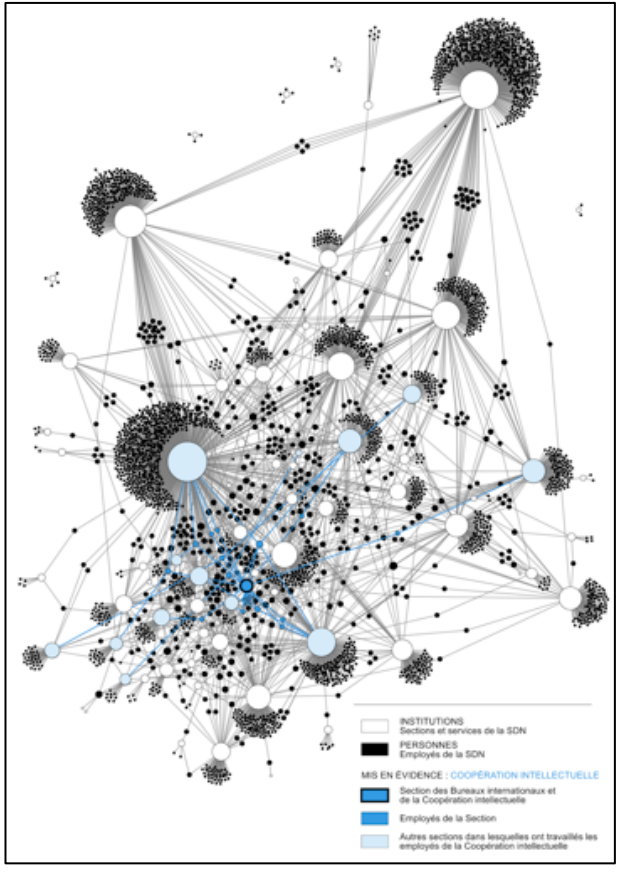
VISUALISATION DE DONNÉES

ARTICULER TYPES, NIVEAUX ET MÉTRIQUES DANS LES RÉSEAUX

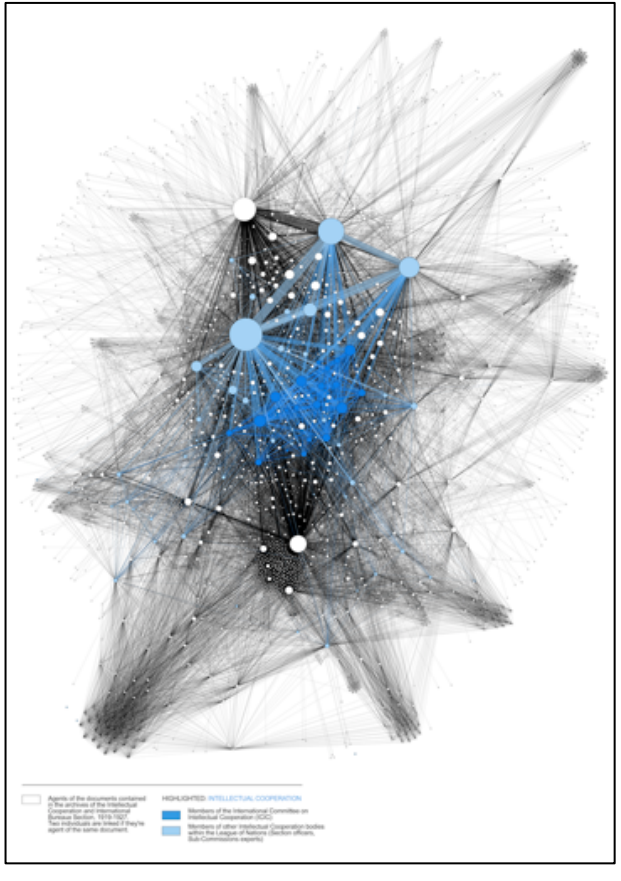
RECONSTITUTION À PARTIR DE SOURCES MULTIPLES



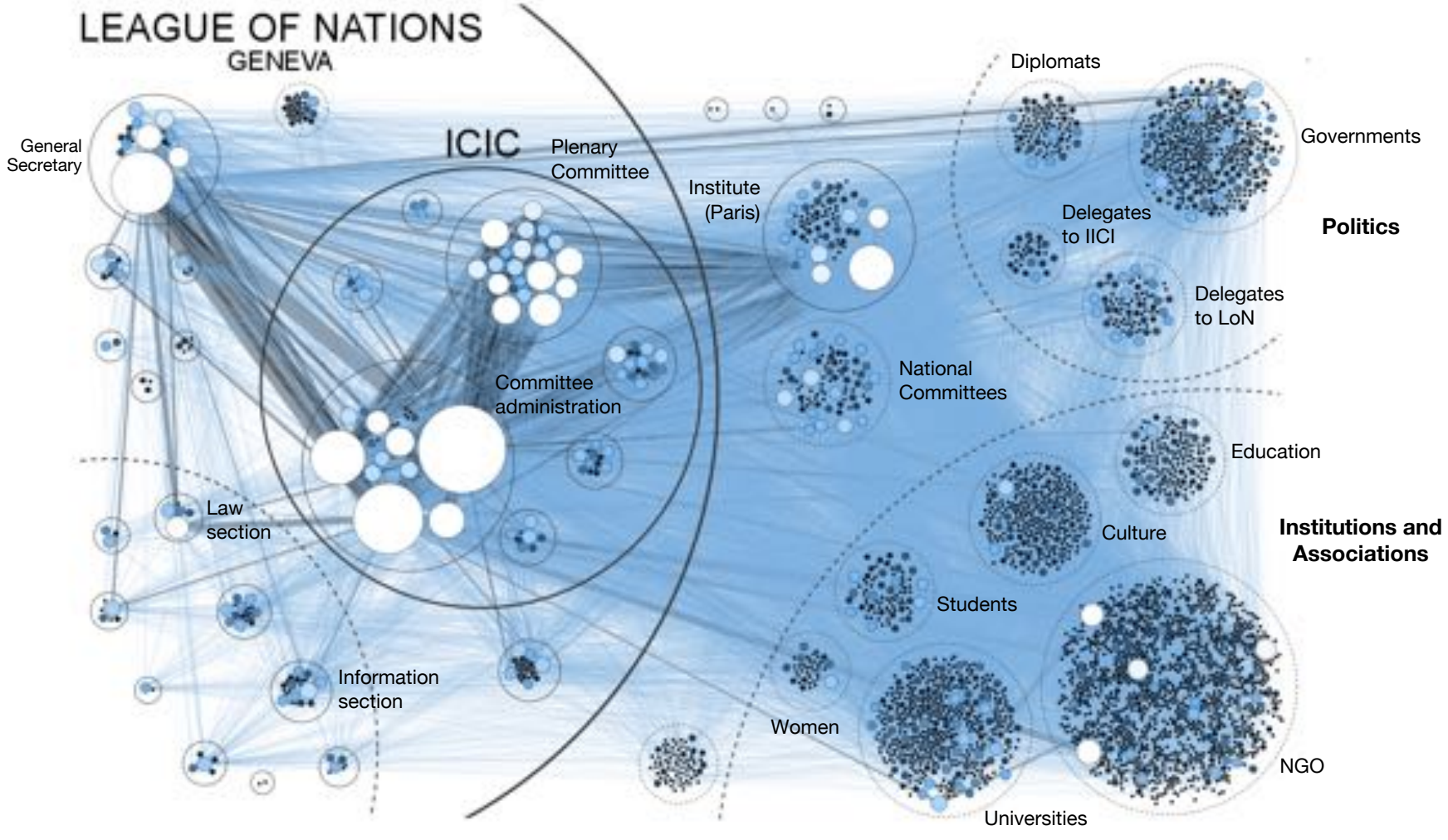
CARTOGRAPHIE DU CONTENU



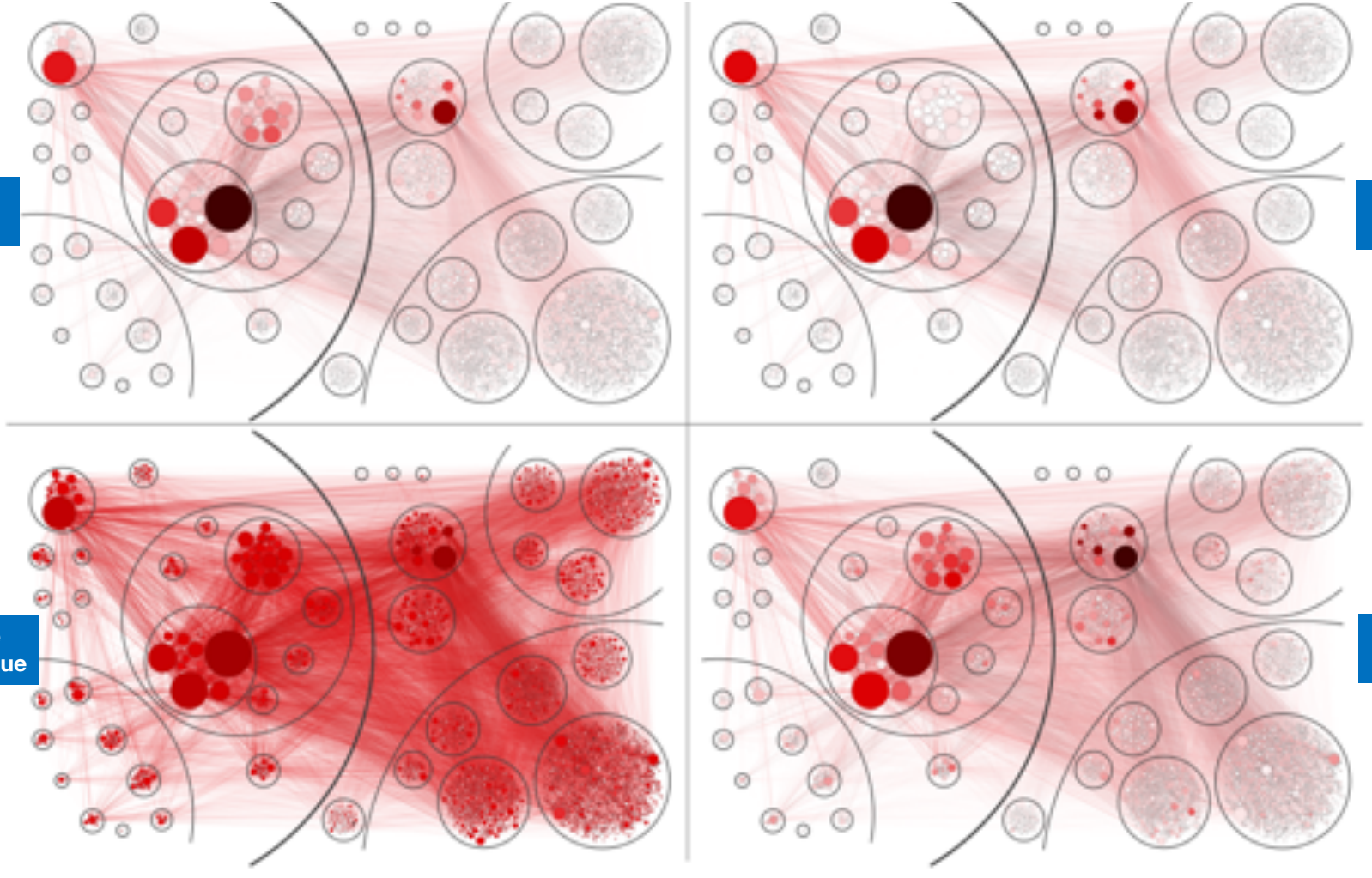
CARTOGRAPHIE DES MÉTADONNÉES

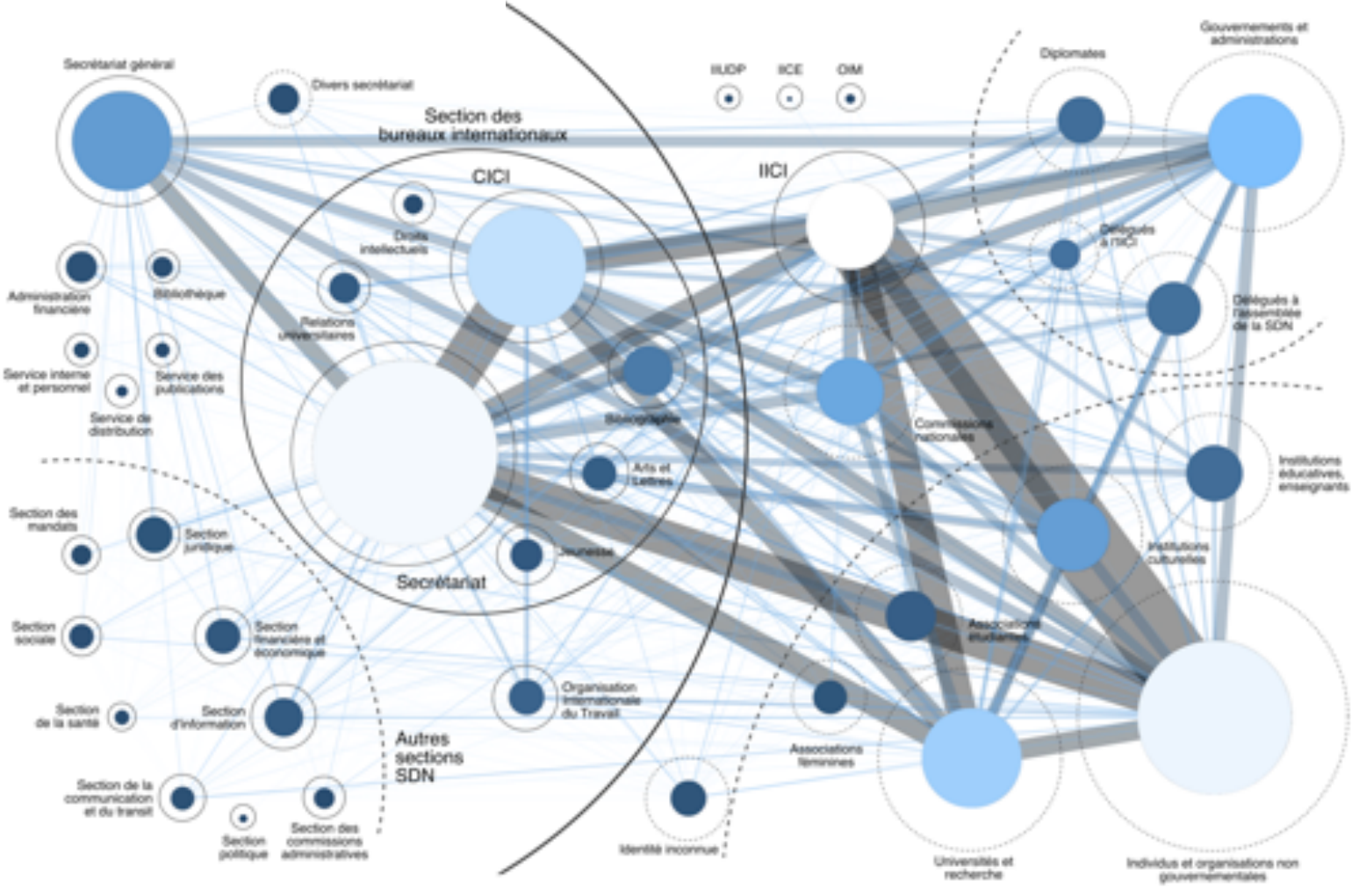


VISUALISATION DE DONNÉES
ARTICULER TYPES, NIVEAUX ET MÉTRIQUES DANS LES RÉSEAUX



VISUALISATION DE DONNÉES
ARTICULER TYPES, NIVEAUX ET MÉTRIQUES DANS LES RÉSEAUX



VISUALISATION DE DONNÉES

ARTICULER TYPES, NIVEAUX ET MÉTRIQUES DANS LES RÉSEAUX