



DONNÉES DES CHERCHEURS, DONNÉES POUR LES CHERCHEURS

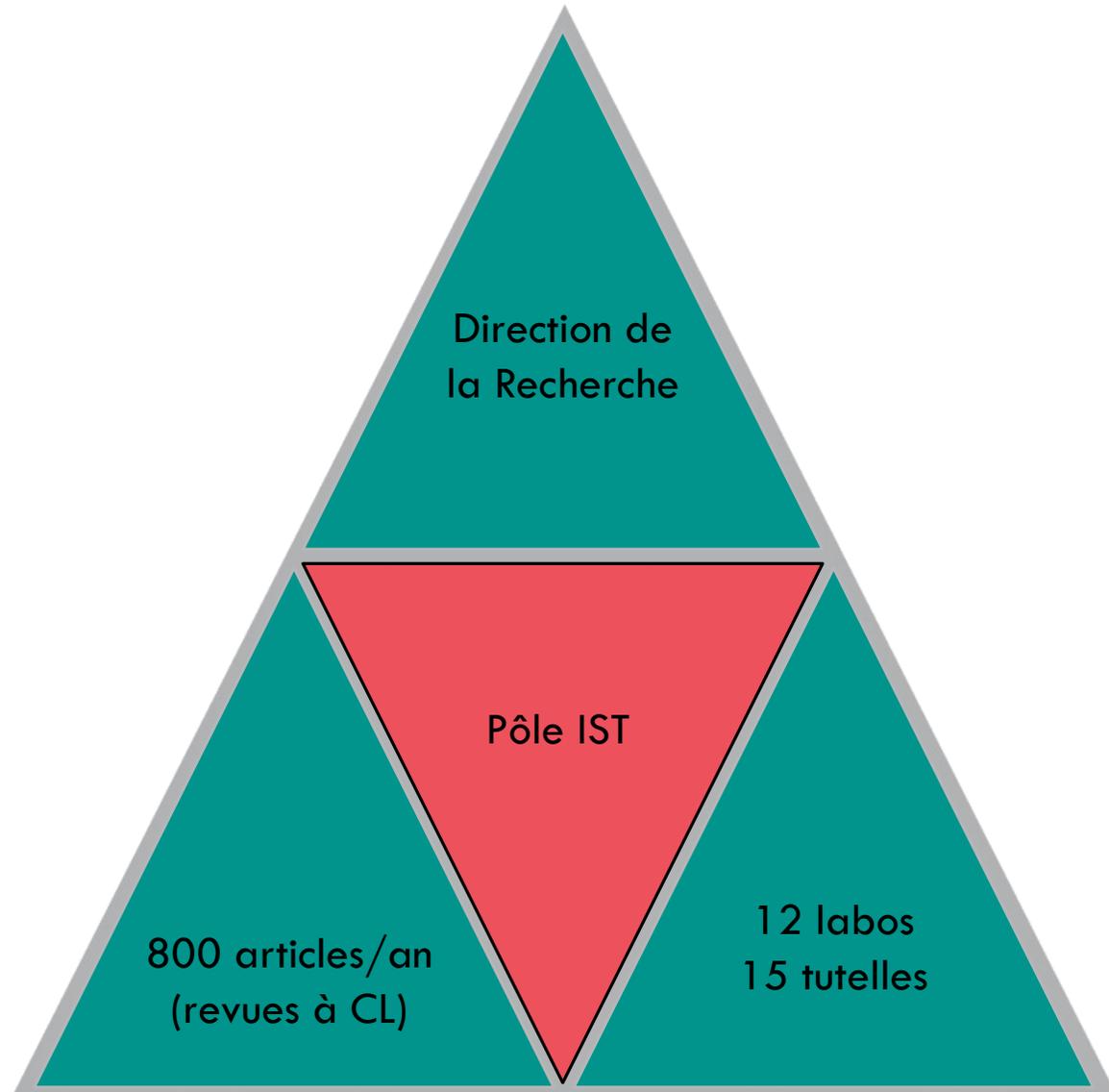
Frédérique Bordignon
Direction de la Documentation
Pôle IST
Ecole des Ponts ParisTech



École des Ponts
ParisTech

Journée d'étude Mediadix – 8 décembre 2017
Croissance et extension des données. Ce que les données font aux bibliothèques

LE CONTEXTE



DEUX ACTIVITÉS RELATIVES AUX DONNÉES

- Valorisation des **données de la recherche**
Science data
- Exploration de **données à des fins de recherche**
Data science

The screenshot shows the website interface for Espace Chercheurs Ressources & Services. At the top, there is a logo consisting of a white triangle with smaller triangles inside, above the text "ESPACE CHERCHEURS RESSOURCES & SERVICES" and the URL "https://espacechercheurs.enpc.fr". Below this is a red header bar with the word "RESSOURCES" in white. The main content area is divided into two columns of colored boxes. The left column contains: "REVUES & EBOOKS" (orange), "BASES BIBLIO" (orange), and "DONNEES DE LA RECHERCHE" (green). The right column contains: "THESES & LITT. GRISE" (purple), "PATRIMOINE" (teal), and an icon of a laptop and smartphone. Below this is another red header bar with the word "SERVICES" in white. The bottom section contains six icons with labels: "ACCES DISTANT" (teal), "GUIDES" (black), "AIDE A LA RECHERCHE" (teal), "INFOS & TUTOS" (black), "TCHAT" (teal), and "LE LAB" (black).



VALORISATION DES DONNÉES DE LA RECHERCHE



AUTO-FORMATION

Objectif :

monter en compétences avant que les chercheurs ne nous sollicitent



The image shows two overlapping screenshots. The top screenshot is a Coursera course page for 'Research Data Management and Sharing'. The page header includes the Coursera logo, a search bar, and a user profile for 'Frédérique'. The course title is prominently displayed in a large font. Below the title, there is a 'Vue d'ensemble' (Overview) section with links for 'Programme de cours', 'FAQ', 'Créateurs', and 'Notation et examens'. A 'Research Data Management and Sharing' box contains an 'Aller au cours' button and a 'Déjà inscrit(e)' link. A 'À propos de ce cours' section provides a brief description of the course's goals. The 'Créé par' section lists the University of North Carolina at Chapel Hill and the University of Edinburgh. Below this, two instructors are listed: Helen Tibbo, an Alumni Distinguished Professor, and Sarah Jones, a Senior Institutional Support Officer. The bottom screenshot is a video player for 'DMP - Présentation générale'. The video player has a dark blue header with the title and navigation links for 'Plan', 'Narration', 'Glossaire', and 'Ressources'. The video content shows the CNRS logo and the text 'Une série de tutoriels sur le libre accès aux résultats de la recherche dans le cadre d'Horizon 2020'. The video player includes a progress bar, a volume icon, and a 'SUIVANT' button.

SENSIBILISATION

Au début, timidement...

Un chercheur, un projet, un labo...

Ensuite, des pages sur notre Espace Chercheurs

Suggestions d'entrepôts, de revues, de licences...

Puis une expo pour l'Open Access Week

Data Incognita <https://patrimoine.enpc.fr/exhibits/show/dataincognita>

Open Access Week 2017

EXPO

DATA INCOGNITA

N'ATTENDEZ PAS D'AVOIR INTÉGRÉ LE FONDS ANCIEN
POUR PARTAGER VOS DONNÉES !

DANS LE HALL DE L'ÉCOLE DES PONTS
DU 23 OCTOBRE
AU 17 NOVEMBRE 2017
ET EN LIGNE SUR [PATRIMOINE.ENPC.FR](https://patrimoine.enpc.fr)

 
École des Ponts
ParisTech

2 INFOS VÉRITÉ

Il reste beaucoup de chemin à parcourir



2786 publiés (2013-2015)

26 publiés liées à un dataset **SCHOLIX**

On croule sous les questions juridiques !



Statut des chercheurs
Loi CADA
CNIL
H2020
Loi pour une République Numérique
Directive INSPIRE
Code de la Propriété Intellectuelle
Loi Valter
Règlement Général sur la Protection des Données Personnelles
Code de la Recherche
Convention d'Aarhus
Décrets sur les licences



EXPLORATION
DE DONNÉES À DES FINS DE RECHERCHE |

DATAVIZ



Bibliovizz

20 et 21 juin 2017 | Anticafé Olympiades



Google Sheets



5 DATA LINGUISTICA

CARTOGRAPHIE STATISTIQUE ET DATAVIZ PERPÉTUER LA TRADITION DE L'ÉCOLE

LA DATAVISUALISATION ET SES OUTILS

L'interprétation des données est rendue plus efficace par des techniques de représentation visuelle, connues aujourd'hui sous le nom de datavisualisation. Cela permet de communiquer plus facilement ses résultats, d'avoir un support à l'analyse et peut aussi faciliter l'émergence de nouvelles hypothèses.

Le panneau d'exposition est un encouragement aux chercheurs à perpétuer la tradition de l'École des Ponts en la matière. En effet, le fonds ancien témoigne du rôle majeur de nos ingénieurs et professeurs dans l'essor de la datavisualisation. Si les cartes statistiques apparaissent au XVI^e siècle avec Edmond Halley, elles prennent leur essor au XIX^e siècle, avec la révolution industrielle en France et l'immense contribution des ingénieurs des Ponts. Pour faciliter l'analyse, ils recourent en effet à des moyens graphiques pour intégrer à leurs cartes des données liées à la mobilité de la population et à l'aménagement du territoire.

« Ils peuvent renfermer implicitement des résultats pour la détermination desquels ils n'auraient pas été préparés. (Léon Lalanne, 1846, à propos de ses graphiques) »

CALCUL GRAPHIQUE



Les courbes isopéthes sont nées à l'École, grâce à Léon Lalanne, ingénieur touche-à-tout, qui, à l'occasion de la traduction d'un document allemand de météorologie, a proposé un diagramme innovant permettant de traduire sur le papier des lois naturelles à 3 variables, générant ainsi des lignes de niveau (Document 1). De même, les diagrammes à coordonnées parallèles (qui aident à la comparaison de variables) trouvent également leur origine à l'École des Ponts avec Maurice d'Ocagne qui y enseigne la géométrie. Il est l'inventeur de la résolution graphique d'équations (nomographie ou théorie des abaques).

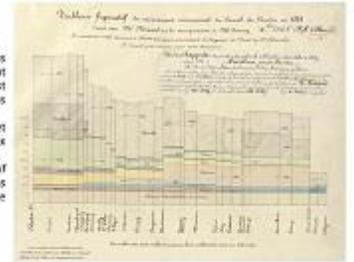
MINARD, LE PIONNIER



Charles-Joseph Minard, professeur à l'École, est considéré comme pionnier dans l'utilisation des graphiques appliqués au génie civil. Il enrichit considérablement les méthodes de cartographie avec le recours à la variation de taille et il est surtout à l'origine de la représentation des flux. Il est l'inspirateur des diagrammes de Sankey.

Il se lance tout d'abord dans une étude statistique du trafic sur les routes et canaux existants et transforme ces statistiques en graphiques "pour mieux appréhender la situation globale".

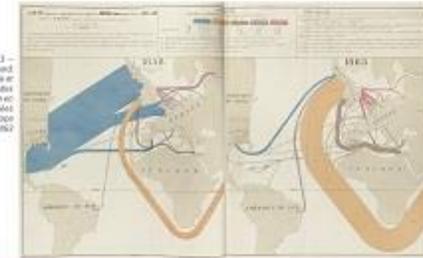
Un des premiers résultats de cette approche est résumé dans le Tableau figuratif du mouvement commercial du Canal du Centre en 1844 (Document 2). Sa Carte figurative et approximative des quantités de coton en laine importées en Europe en 1858 et en 1863 (Document 3) est accompagnée d'une note explicative qui échoie le lecteur sur ses choix graphiques. Et les données brutes sont également disponibles !



Document 2 - Charles-Joseph Minard, 1844, Tableau figuratif du mouvement commercial du canal du Centre en 1844.

CARTES ADMINISTRATIVES - DÉMOCARTOGRAPHIE

Louis-Léger Vauthier (Document 4) et Émile Cheysson, chargé de la direction des cartes et plans au sein du Ministère des Travaux publics avec l'édition régulière de 1879 à 1899 de L'Album de Statistique graphique, ont ainsi produit des cartes reprenant les procédés initiés par Lalanne et Minard.



Document 3 - Charles-Joseph Minard, 1863, Carte figurative et approximative des quantités de coton en laine importées en Europe en 1858 et en 1863.

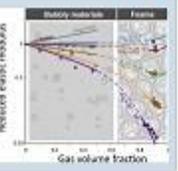


Document 4 - Louis-Léger Vauthier, 1874, Carte statistique figurant la répartition de la population de Paris.

« Non seulement mes cartes parlent, mais, de plus, elles calculent par l'œil ; c'est là le point capital. (Charles-Joseph Minard, 1869) »

Les revues apprécient les données visuelles

Certaines revues demandent un graphique abstrait pour résumer visuellement le contenu d'un article. Par ailleurs, au-delà des supplementary materials, certaines revues exploitent toute la puissance d'une datavisualisation dynamique en ligne en permettant au lecteur de faire varier des paramètres (comme la revue *Distill*).



F. Soblec, V. Khlebnik, D. Pletzer (2017) Coupled elasticity in soft solid foams. Journal of Colloid and Interface Science, Vol 507, p. 182-187.

AIDE AU PILOTAGE DE LA RECHERCHE : DU SUR-MESURE

4 secteurs socio-économiques définis par la Direction de la Recherche

- Industrie du Futur
- Economie, Usage et Société
- Gestion des risques des ressources et des milieux
- Systèmes Ville Mobilité

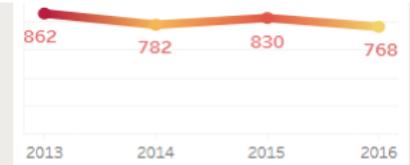
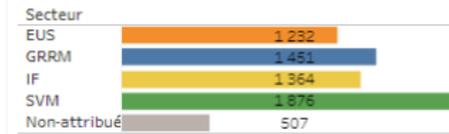
Traduits par le Pôle IST en termes bibliométriques
(quelles revues alimentent quels secteurs)



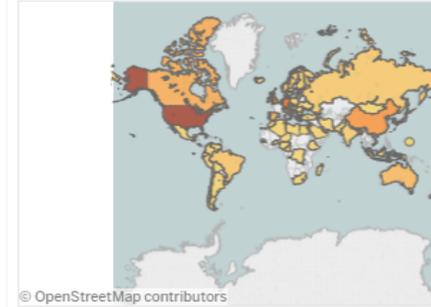
DES TABLEAUX DE BORD DYNAMIQUES



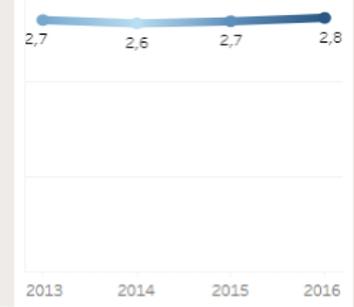
Répartition en secteurs socio-économiques (DR)



Nombre d'articles en co-publication avec un auteur étranger



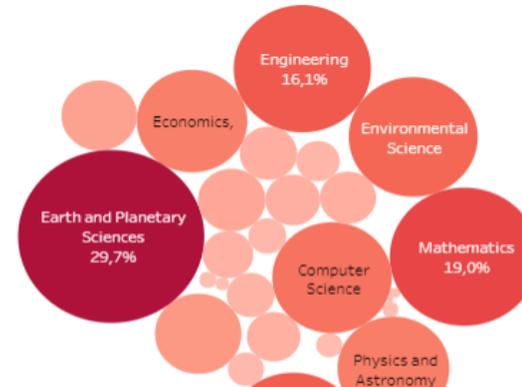
CiteScore moyen par article



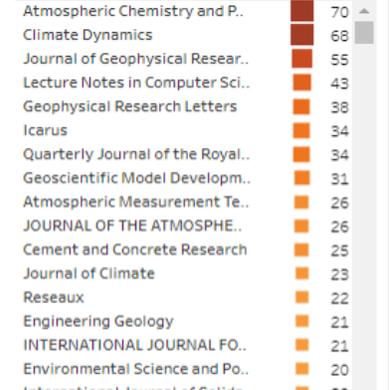
© OpenStreetMap contributors



Répartition par domaines



Revue



Mot de passe

- Labo(s)
- (Tout)
 - CEREA
 - CERMICS
 - CIRE
 - HMCO
 - LATTS
 - LEESU
 - LHSV
 - LIGM
 - LMD
 - LVMT
 - NAVIER
 - PJSE

Citations par article

7,6

Représentation des labos résultant des filtres



DES TABLEAUX DE BORD DYNAMIQUES (2)



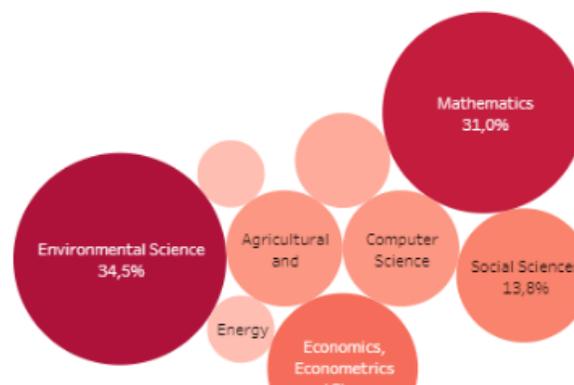
Nombre d'articles en co-publication avec un auteur étranger



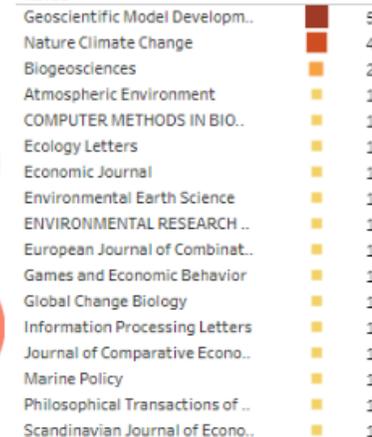
CiteScore moyen par article



Répartition par domaines



Revue



Mot de passe

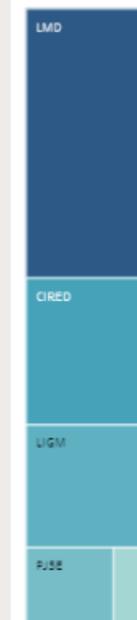
Labo(s)

- (Tout)
- CEREAS
- CERMIACS
- CIRED
- HMCO
- LATTS
- LEESU
- LHSV
- LIGM
- LMD
- LVMT
- NAVIER
- PJSE

Citations par article

12,3

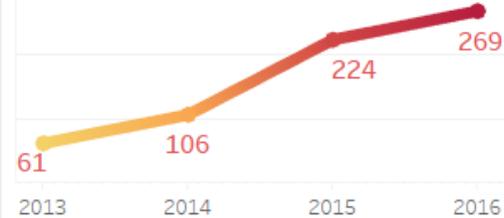
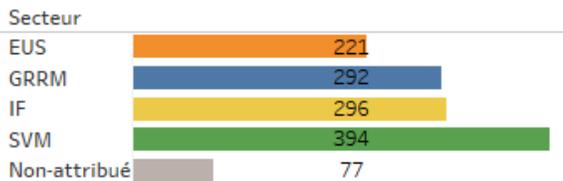
Représentation des labos résultant des filtres



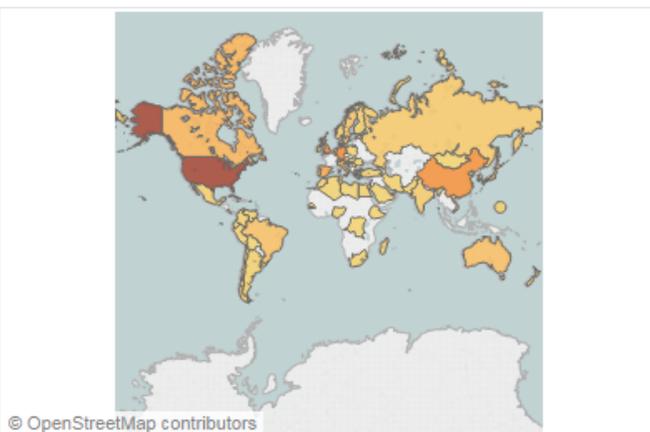
DES TABLEAUX DE BORD DYNAMIQUES (3)



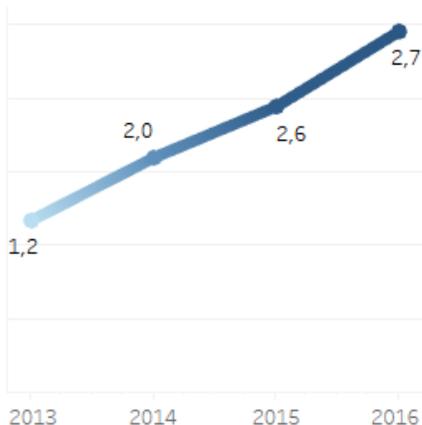
Répartition en secteurs socio-économiques (DR)



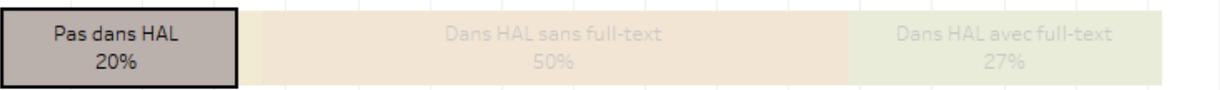
Nombre d'articles en co-publication avec un auteur étranger



CiteScore moyen par article



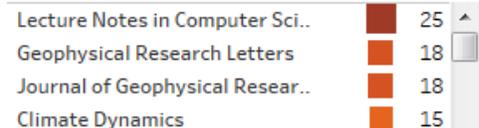
Répartition des citations



Répartition par domaines



Revue



Mot de passe

Labo(s)

- (Tout)
- CERA
- CERMICS
- CIRED
- HMCO
- LATTs
- LEESU
- LHSV
- LIGM
- LMD
- LVMT
- NAVIER
- PJSE

Citations par article

3,6

Représentation des labos résultant des filtres

LMD

NAVIER

LE TEXT AND DATA MINING COMME SERVICE AUX CHERCHEURS

- Analyse de textes
 - Repérage de littérature scientifique (gain de temps, sélection)
 - Extraction de tendances
 - Co-occurrences de concepts
- Extraction et analyse de données
 - Sites webs
 - Tweets

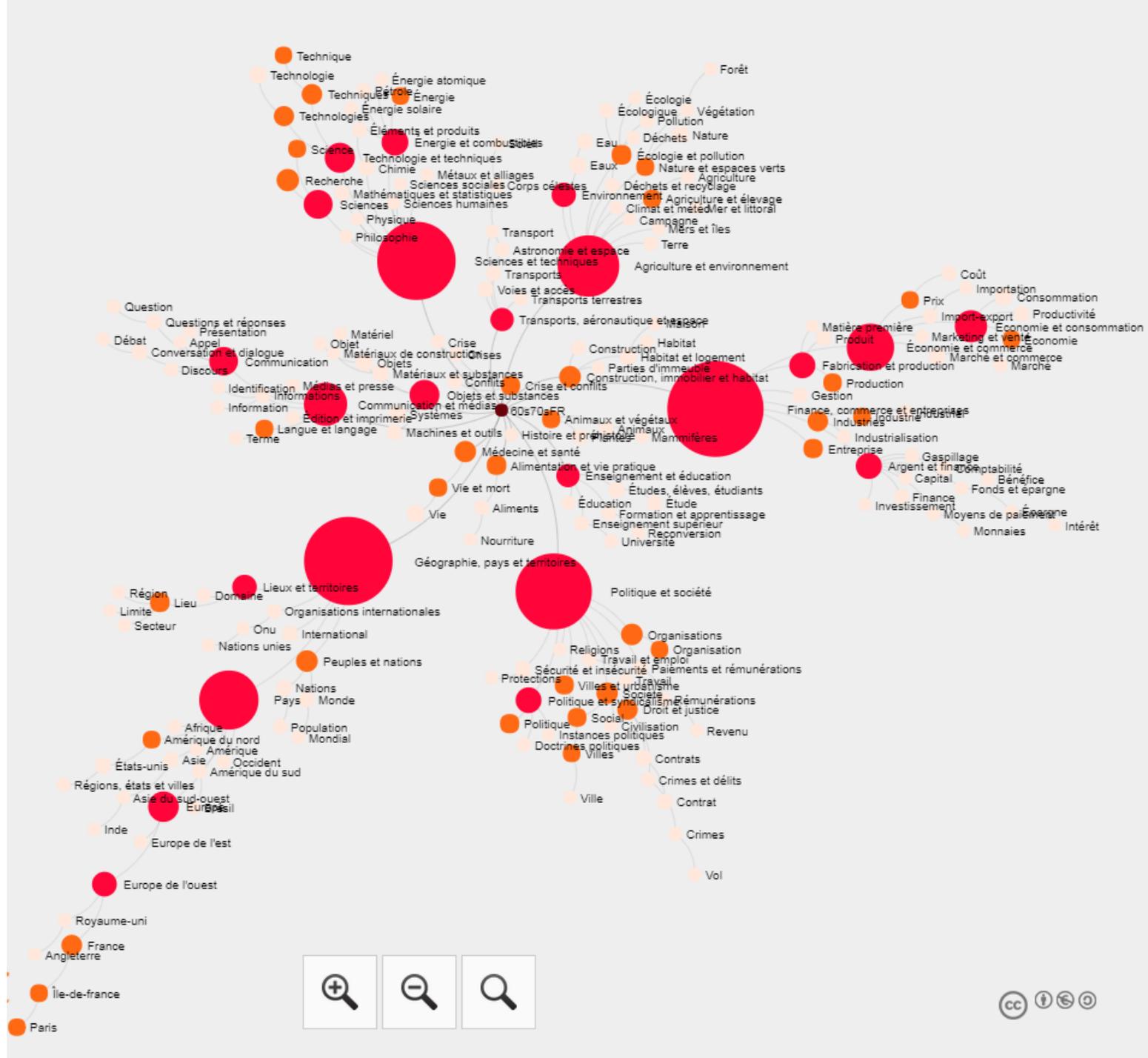


SÉLECTION DE LITTÉRATURE SCIENTIFIQUE

	A	B	C	D	E	F	G	H	I
1	Fichier	Année	AESN	AESNI	BRGM	BUREAU	CERFAC	CIREN	CNRS
2	Axe2_iNR6_Deleplar	2013			14				
3	rapport-bassee.pdf	2001	1		7				
4	Gaillard et al 2002.pc	2002	1		7				
5	Michelinrapport Ru	2008			7				2
6	ZHR.pdf	2001	1		6				1
7	PIREN-Seine_rappor	2010			5				3
8	Rapport_Cerdan_pir	2005			5				3
9	0103_Rapport-Piren	2006			5				
10	PIREN-Seine_rappor	2011			4				1
11	0101_rapport_nitrate	2006	1		4				

ANALYSE DE TEXTES

Projection d'un corpus de textes sur un référentiel de concepts (scenario)

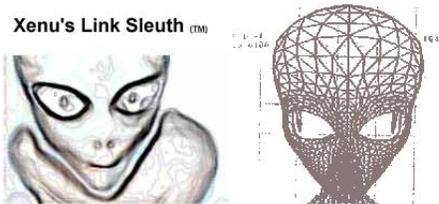


ANALYSE DE TEXTES

Co-occurrences de concepts

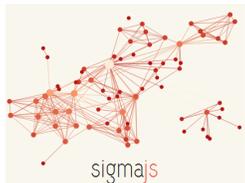


RÉSEAU DE SITES WEB



YouTube
Data API v3

Gephi



Discovered

- Entreprises et ateliers
- Youtube
- Laboratoires
- Blog
- Association de consommateurs
- associations professionnelles

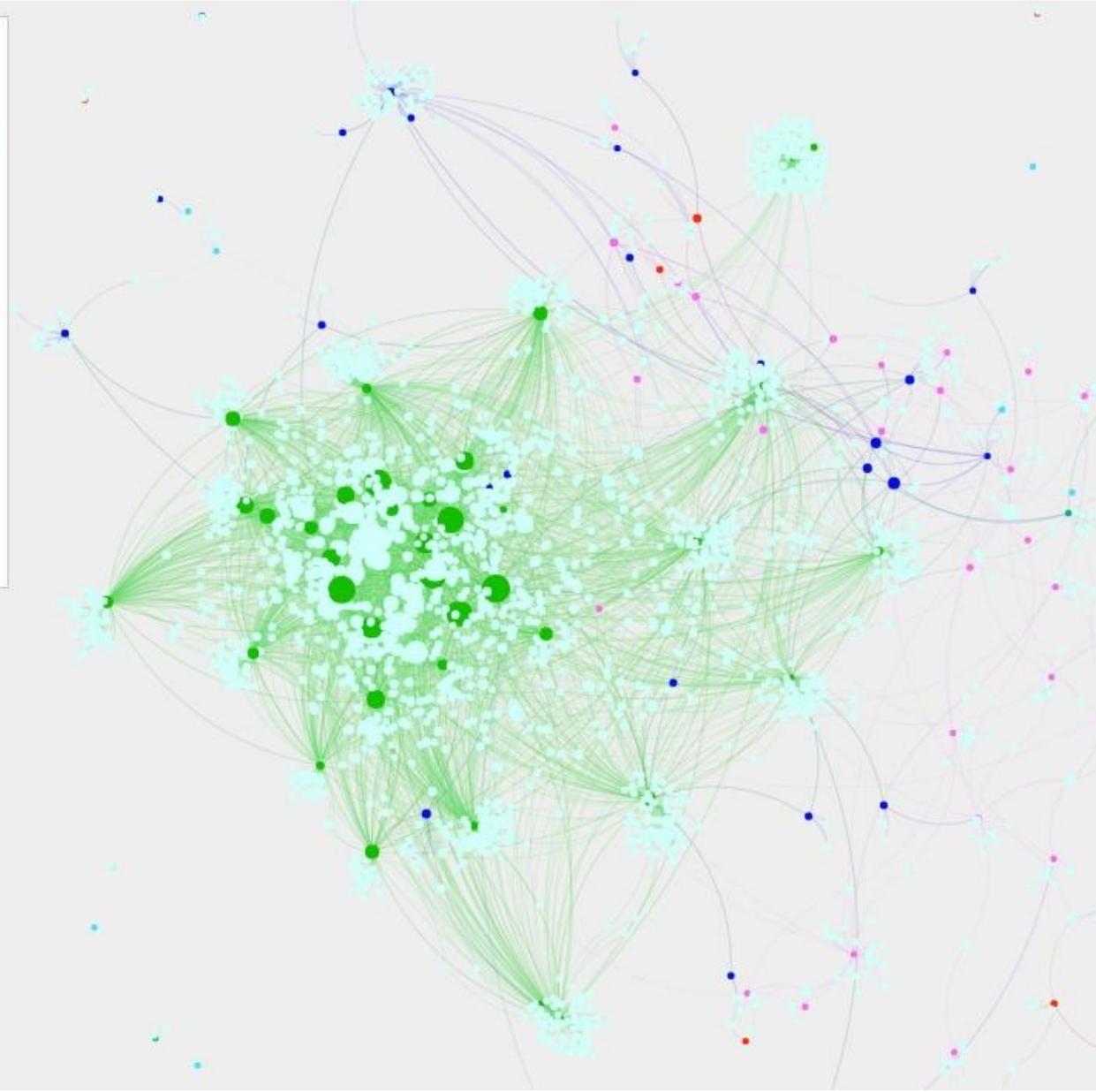
Projet Cosme'teau

i More about this visualisation

Legend:

- Taille = nombre de citations recues
- Couleur = couleur du site source
- Types de sites

Search:



MERCI !



frederique.bordignon@enpc.fr



[@freddie2310](https://twitter.com/freddie2310)



<https://carnetist.hypotheses.org/>

WE ARE DATA LIBRARIANS !

