

La cartographie documentaire

Métrologies et visualisations de corpus

Bruno Gaume

CLLE, CNRS & Univ. Toulouse

À Paris, ISCIPIF, Mars 2021

ANF TDM 2020 - Exploration documentaire et extraction d'informations

Plan de l'exposé

- 1 Les objets d'un corpus documentaire
- 2 Lexiques
- 3 Réseaux
- 5 Visualisations/Navigations

Plan de l'exposé

- 1 **Les objets d'un corpus documentaire**
- 2 Lexiques
- 3 Réseaux
- 5 Visualisations/Navigations

Les objets d'un corpus documentaire

On dispose d'un ensemble de documents

Les objets d'un corpus documentaire

On dispose d'un ensemble de documents



Les objets d'un corpus documentaire

On dispose d'un ensemble de documents



Les objets d'un corpus documentaire

On dispose d'un ensemble de documents



Texte, Formules, Images, Références, Titre, Date, Résumé, Editeur,
Mots-clés, Auteurs(s), Affiliations, ...

Les objets d'un corpus documentaire

On dispose d'un ensemble de documents



Texte, Formules, Images, Références, Titre, Date, Résumé, Editeur,
Mots-clés, Auteurs(s), Affiliations, ...

Les objets d'un corpus documentaire

On dispose d'un ensemble de documents



Texte, Formules, Images, Références, Titre, Date, Résumé, Editeur,
Mots-clés, Auteurs(s), Affiliations, ...

Les objets d'un corpus documentaire

On dispose d'un ensemble de documents



Texte, Formules, Images, Références, Titre, Date, Résumé, Editeur, Mots-clés, Auteurs(s), Affiliations, ...

Les objets d'un corpus documentaire

On dispose d'un ensemble de documents



Texte, Formules, Images, Références, Titre, Date, Résumé, Editeur,
Mots-clés, Auteurs(s), Affiliations, ... [Hyperliens](#), ..., ..., ...

Les objets d'un corpus documentaire

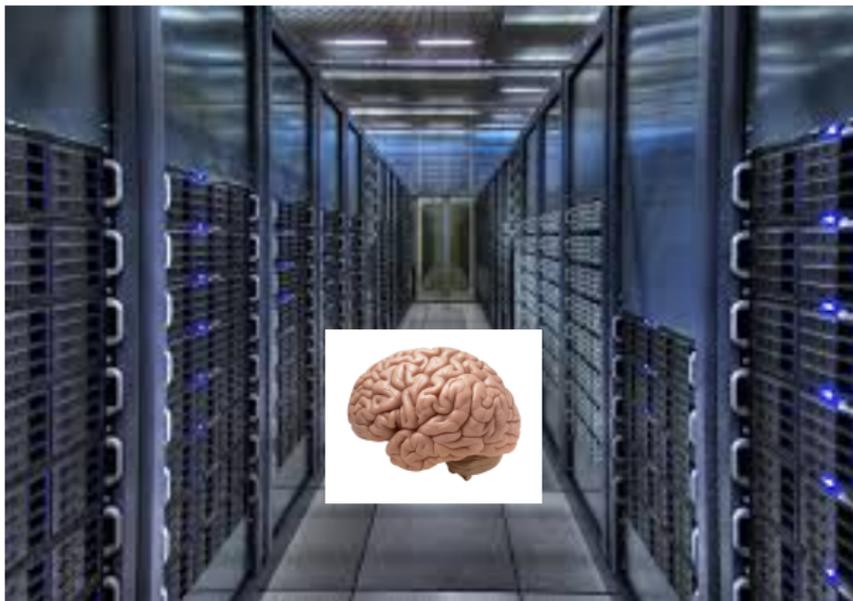
On dispose d'un ensemble de documents



Texte, Formules, Images, Références, Titre, Date, Résumé, Editeur,
Mots-clés, Auteurs(s), Affiliations, ... [Hyperliens](#), ..., ..., ...

Les objets d'un corpus documentaire

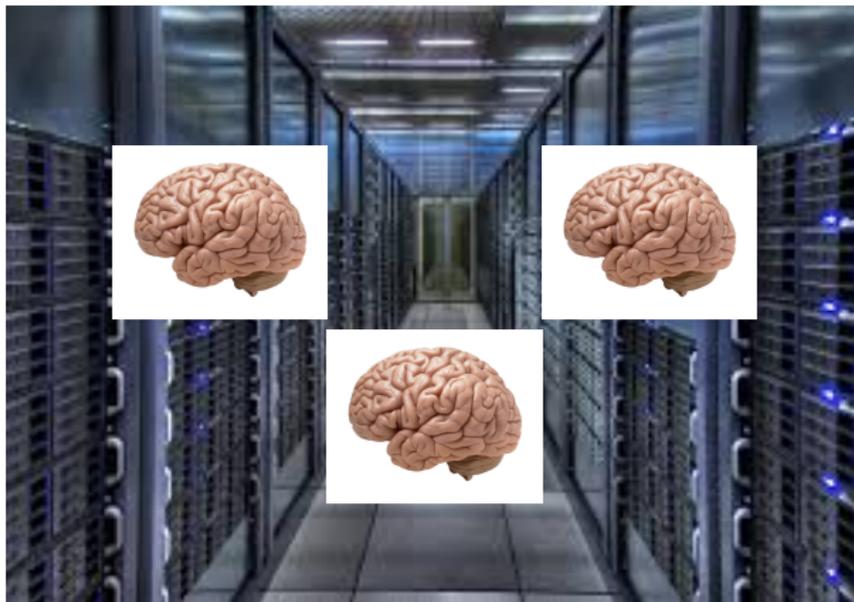
On dispose d'un ensemble de documents



Texte, Formules, Images, Références, Titre, Date, Résumé, Editeur,
Mots-clés, Auteurs(s), Affiliations, ... [Hyperliens](#), ..., ..., ...

Les objets d'un corpus documentaire

On dispose d'un ensemble de documents



Texte, Formules, Images, Références, Titre, Date, Résumé, Editeur,
Mots-clés, Auteurs(s), Affiliations, ... [Hyperliens](#), ..., ..., ...

Plan de l'exposé

- 1 Les objets d'un corpus documentaire
- 2 **Lexiques**
- 3 Réseaux
- 5 Visualisations/Navigations

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.

- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cavalerie.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cavalerie.
- ▶ C'est vous qui **commandez** l'enchaînement des diapositives, d'un clic de souris.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cavalerie.
- ▶ C'est vous qui **commandez** l'enchaînement des diapositives, d'un clic de souris.
- ▶ Ils ont déposé les armes avant d'avoir fait tout ce que **commandait** le devoir et l'honneur.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cav
Lemme: **commander**
- ▶ C'est vous qui **commandez** l'enchaînement des diapositives, d'un clic de souris.
- ▶ Ils ont déposé les armes avant d'avoir fait tout ce que **commandait** le devoir et l'honneur.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cav
Lemme: **commander**
- ▶ (**Formes: commandé, commandez, commande, ... commander, commandent, commandait,**
sou
- ▶ Ils ont déposé les armes avant d'avoir fait tout ce que **commandait** le devoir et l'honneur.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cavalerie.
- ▶ C'est vous qui **commandez** l'enchaînement des diapositives, d'un clic de souris.
- ▶ Ils ont déposé les armes avant d'avoir fait tout ce que **commandait** le devoir et l'honneur.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cavalerie.
- ▶ C'est vous qui **commandez** l'enchaînement des diapositives, d'un clic de souris.
- ▶ Ils ont déposé les armes avant d'avoir fait tout ce que **commandait** le devoir et l'honneur.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
/ordonner/, /prescrire/, /réclamer/, ...
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cavalerie.
- ▶ C'est vous qui **commandez** l'enchaînement des diapositives, d'un clic de souris.
- ▶ Ils ont déposé les armes avant d'avoir fait tout ce que **commandait** le devoir et l'honneur.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
/ordonner/, /prescrire/, /réclamer/, ...
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
/choisir/, /réserver/, /procurer/, ...
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cavalerie.
- ▶ C'est vous qui **commandez** l'enchaînement des diapositives, d'un clic de souris.
- ▶ Ils ont déposé les armes avant d'avoir fait tout ce que **commandait** le devoir et l'honneur.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
/ordonner/, /prescrire/, /réclamer/, ...
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
/choisir/, /réserver/, /procurer/, ...
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cavalerie.
/mener/, /ordonner/, /conduire/, ...
- ▶ C'est vous qui **commandez** l'enchaînement des diapositives, d'un clic de souris.
- ▶ Ils ont déposé les armes avant d'avoir fait tout ce que **commandait** le devoir et l'honneur.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
/ordonner/, /prescrire/, /réclamer/, ...
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
/choisir/, /réserver/, /procurer/, ...
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cavalerie.
/mener/, /ordonner/, /conduire/, ...
- ▶ C'est vous qui **commandez** l'enchaînement des diapositives, d'un clic de souris.
/gérer/, /déclencher/, /contrôler/, ...
- ▶ Ils ont déposé les armes avant d'avoir fait tout ce que **commandait** le devoir et l'honneur.

Lexiques : formes, lemmes, lexies, polysémies

- ▶ Ce constat dressé dans un rapport **commandé** en son temps par la ministre.
/ordonner/, /prescrire/, /réclamer/, ...
- ▶ **Commandez** aujourd'hui même sur notre site, notre méthode d'entraînement et commencez immédiatement!
/choisir/, /réserver/, /procurer/, ...
- ▶ Le colonel **commande** la tragique et décisive charge des 80 escadrons de cavalerie.
/mener/, /ordonner/, /conduire/, ...
- ▶ C'est vous qui **commandez** l'enchaînement des diapositives, d'un clic de souris.
/gérer/, /déclencher/, /contrôler/, ...
- ▶ Ils ont déposé les armes avant d'avoir fait tout ce que **commandait** le devoir et l'honneur.
/réclamer/, /dicter/, /imposer/, ...

Lexiques : formes, lemmes, lexies, polysémies

Formes:

Je **commande**, Nous **commandons**, Il **commandait**, tu as **commandé**, à **commander**, ...

Lexiques : formes, lemmes, lexies, polysémies

Formes:

Je **commande**, Nous **commandons**, Il **commandait**, tu as **commandé**, à **commander**, ...

Lemme:

commander

Lexiques : formes, lemmes, lexies, polysémies

Formes:

Je **commande**, Nous **commandons**, Il **commandait**, tu as **commandé**, à **commander**, ...

Lemme:

commander

Lexies:

commander *Sens*₁ : /ordonner/, /prescrire/, /réclamer/

commander *Sens*₂ : /choisir/, /réserver/, /procurer/

commander *Sens*₃ : /piloter/, /déclencher/, /contrôler/

... ..

Lexiques : formes, lemmes, lexies, polysémies

Formes:

Je **commande**, Nous **commandons**, Il **commandait**, tu as **commandé**, à **commander**, ...

Lemme:

commander

Lexies:

commander *Sens*₁ : /ordonner/, /prescrire/, /réclamer/

commander *Sens*₂ : /choisir/, /réserver/, /procurer/

commander *Sens*₃ : /piloter/, /déclencher/, /contrôler/

... ..

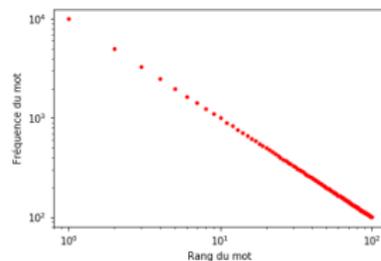
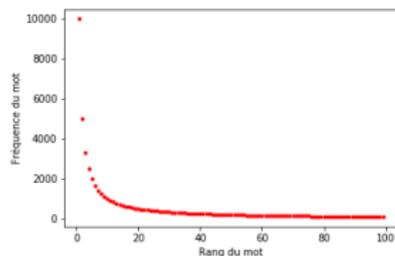
Polysémie

Lexiques : Fréquences

Loi de Zipf : $f(n) \approx K \frac{1}{n}$

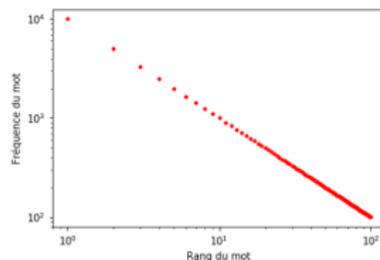
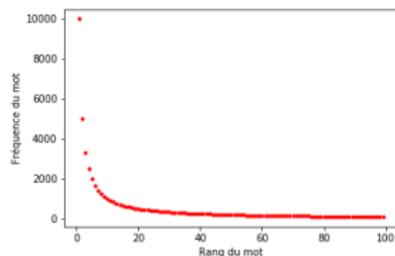
Lexiques : Fréquences

Loi de Zipf : $f(n) \approx K \frac{1}{n}$



Lexiques : Fréquences

$$\text{Loi de Zipf : } f(n) \approx K \frac{1}{n}$$

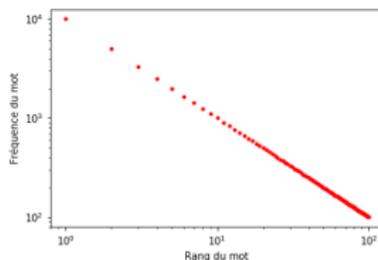
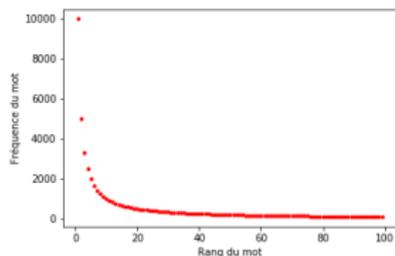


Corpus : C (Tectonique des plaques)

Corpus : ALL

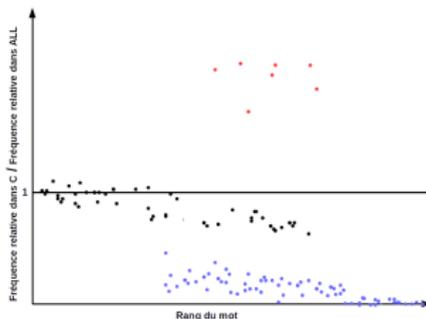
Lexiques : Fréquences

Loi de Zipf : $f(n) \approx K \frac{1}{n}$



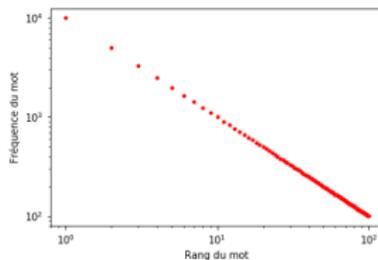
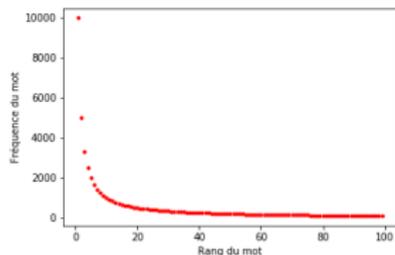
Corpus : C (Tectonique des plaques)

Corpus : ALL



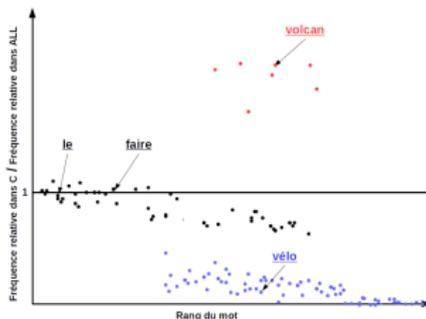
Lexiques : Fréquences

$$\text{Loi de Zipf : } f(n) \approx K \frac{1}{n}$$

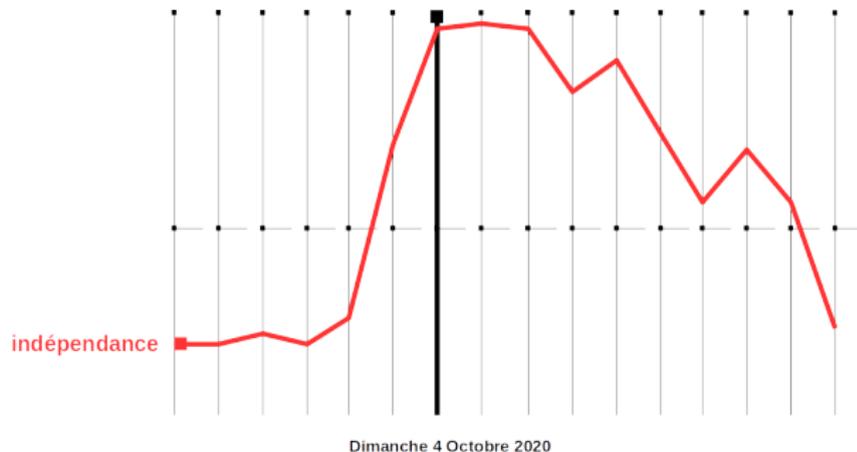


Corpus : C (Tectonique des plaques)

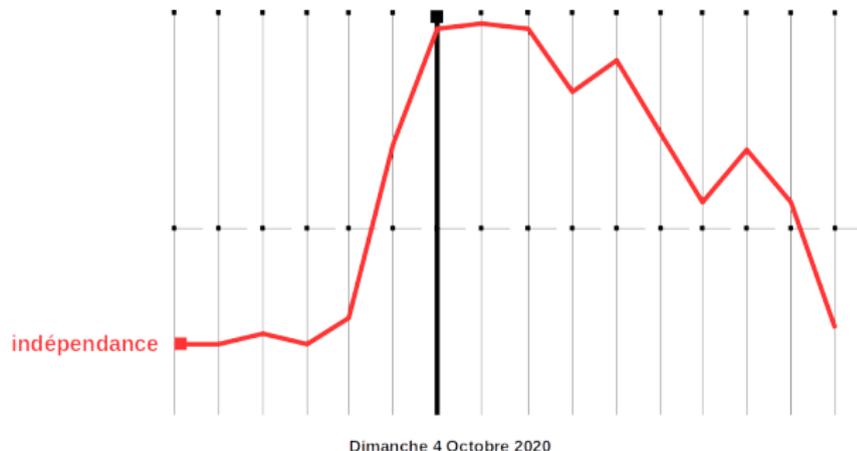
Corpus : ALL



Lexiques : Fréquences



Lexiques : Fréquences



Calédonie, vote, indépendance, référendum,
Nouméa, souveraineté, autodétermination, ...

TFiDF : Term Frequency inverse Document Frequency

TFiDF : Term Frequency inverse Document Frequency

$$TFiDF_c^t$$

TFiDF : Term Frequency inverse Document Frequency

$$TFiDF_c^t(d)$$

TFiDF : Term Frequency inverse Document Frequency

$$TFiDF_c^t(d) = TF(d, t).iDF(t, c)$$

TFiDF : Term Frequency inverse Document Frequency

$$TFiDF_c^t(d) = TF(d, t) \cdot iDF(t, c)$$

$$TF(d, t) = \frac{\text{Nombre d'apparition du terme } t \text{ dans le document } d}{\text{Nombre total de termes dans le document } d}$$

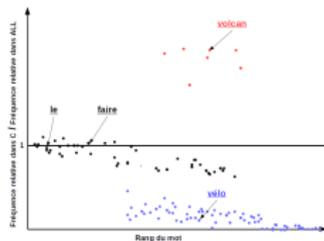
TFiDF : Term Frequency inverse Document Frequency

$$TFiDF_c^t(d) = TF(d, t).iDF(t, c)$$

$$TF(d, t) = \frac{\text{Nombre d'apparition du terme } t \text{ dans le document } d}{\text{Nombre total de termes dans le document } d}$$

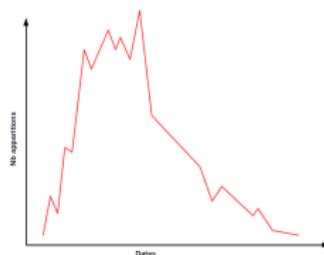
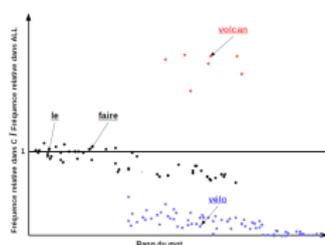
$$iDF(t, c) = \frac{|c|}{\text{Nombre de documents de } c \text{ contenant le terme } t}$$

Lexiques : Fréquences



**Comprendre de quoi
parle un corpus**

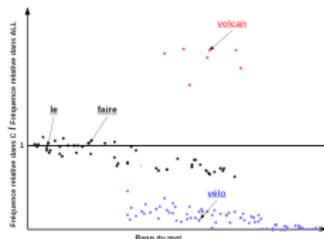
Lexiques : Fréquences



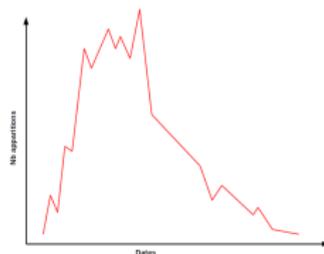
**Comprendre de quoi
parle un corpus**

Détecter les buzz

Lexiques : Fréquences



Comprendre de quoi
parle un corpus



Détecter les buzz

$$TFiDF_c^t(d)$$

Bienvenue sur le démonstrateur ISTE_X

euro

A screenshot of the ISTE_X demonstrator interface. It shows a search bar with the query 'euro', a list of results, and a sidebar with navigation options. The interface is designed for searching and analyzing documents.

Répondre à une
requête simple

Lexiques : Axe syntagmatique & axe paradigmatique

Lexiques : Axe syntagmatique & axe paradigmatique

Le

Le chat

Le chat mange

Le chat mange la

Le chat mange la souris

Lexiques : Axe syntagmatique & axe paradigmatique

Le chat mange la souris ...

Lexiques : Axe syntagmatique & axe paradigmatique

Le chat mange la souris ...

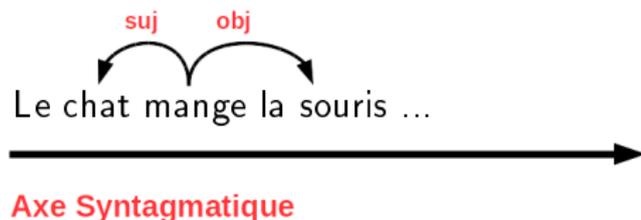
Lexiques : Axe syntagmatique & axe paradigmatic

Le chat mange la souris ...

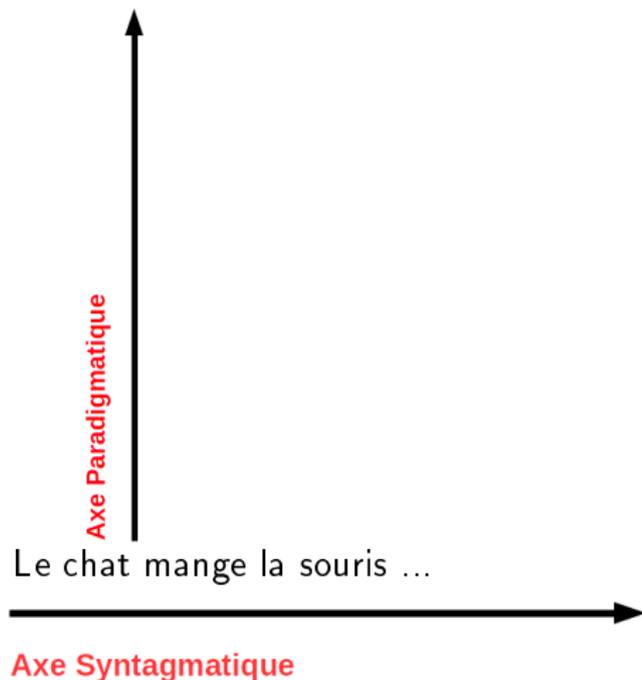


Axe Syntagmatique

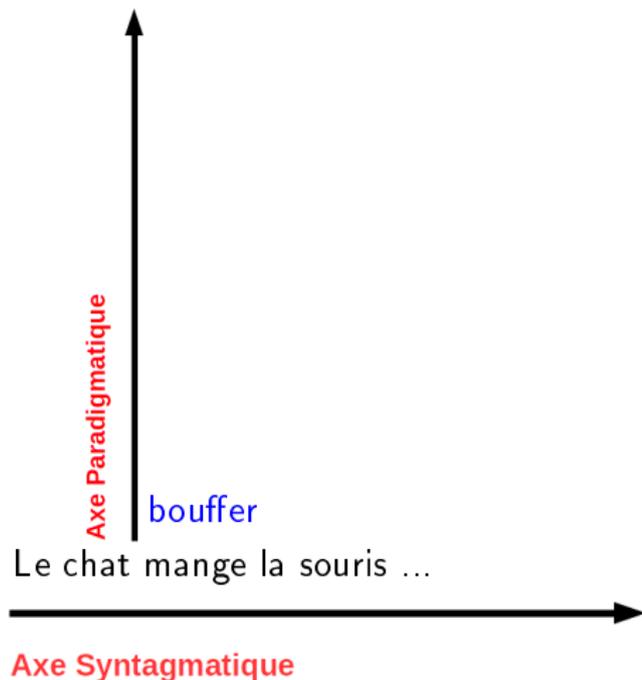
Lexiques : Axe syntagmatique & axe paradigmatic



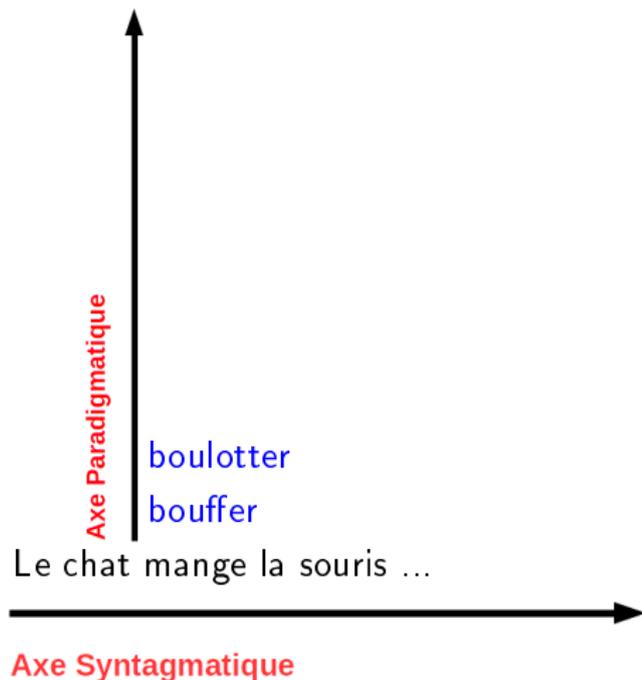
Lexiques : Axe syntagmatique & axe paradigmatique



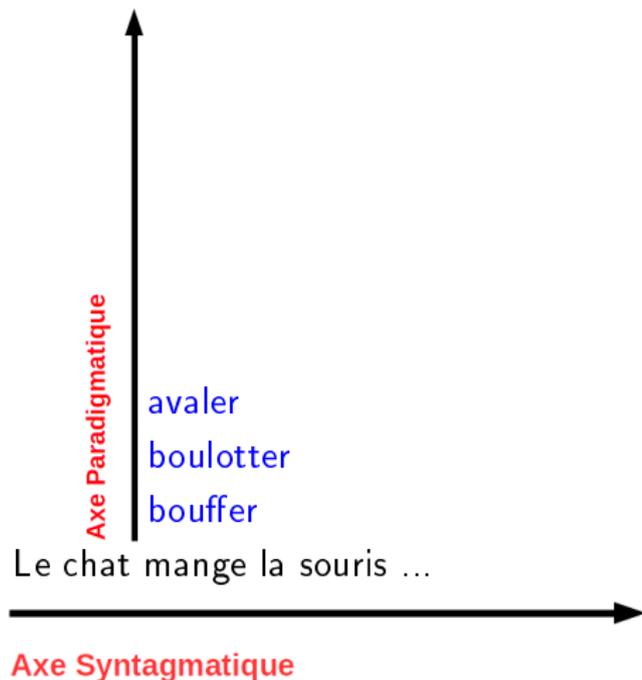
Lexiques : Axe syntagmatique & axe paradigmatic



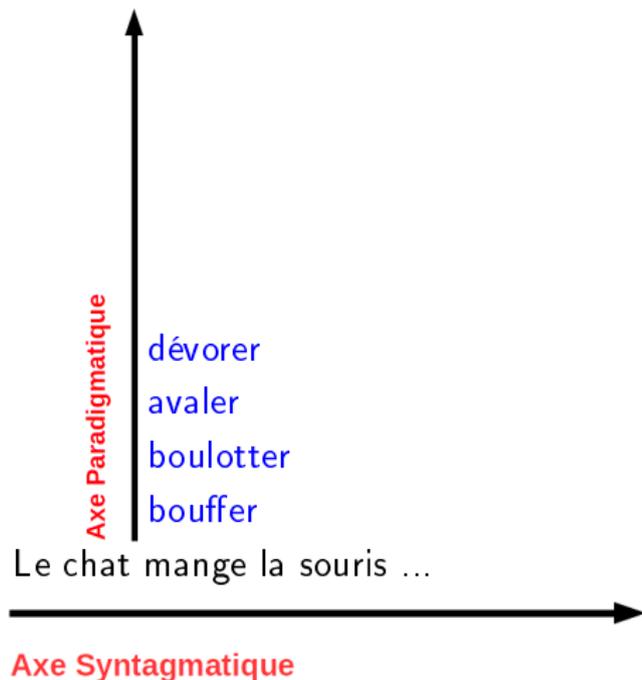
Lexiques : Axe syntagmatique & axe paradigmatique



Lexiques : Axe syntagmatique & axe paradigmatique

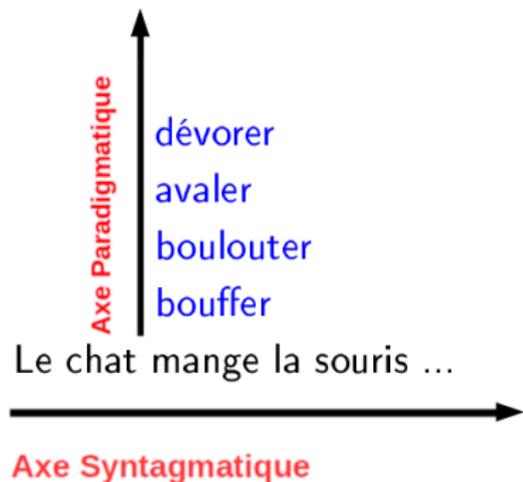


Lexiques : Axe syntagmatique & axe paradigmatique



Lexiques : Coocurrences, Distances, Ordre-1 & Ordre-2

Lexiques : Coocurrences, Distances, Ordre-1 & Ordre-2



Lexiques : Coocurrences, Distances, Ordre-1 & Ordre-2



Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



chat
manger
souris
arbre
ciel
dévorer
rouge
...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



chat manger souris arbre ciel dévorer rouge ...

chat

manger

souris

arbre

ciel

dévorer

rouge

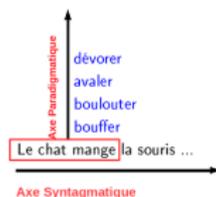
...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



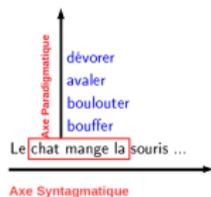
	chat	manger	souris	arbre	ciel	dévorer	rouge	...
chat	x.1.1	x.1.2	x.1.3	x.1.4	x.1.5	x.1.6	x.1.7	...
manger	x.2.1	x.2.2	x.2.3	x.2.4	x.2.5	x.2.6	x.2.7	...
souris	x.3.1	x.3.2	x.3.3	x.3.4	x.3.5	x.3.6	x.3.7	...
arbre	x.4.1	x.4.2	x.4.3	x.4.4	x.4.5	x.4.6	x.4.7	...
ciel	x.5.1	x.5.2	x.5.3	x.5.4	x.5.5	x.5.6	x.5.7	...
dévorer	x.6.1	x.6.2	x.6.3	x.6.4	x.6.5	x.6.6	x.6.7	...
rouge	x.7.1	x.7.2	x.7.3	x.7.4	x.7.5	x.7.6	x.7.7	...
...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



	chat	manger	souris	arbre	ciel	dévorer	rouge	...
chat	x.1.1	x.1.2	x.1.3	x.1.4	x.1.5	x.1.6	x.1.7	...
manger	x.2.1	x.2.2	x.2.3	x.2.4	x.2.5	x.2.6	x.2.7	...
souris	x.3.1	x.3.2	x.3.3	x.3.4	x.3.5	x.3.6	x.3.7	...
arbre	x.4.1	x.4.2	x.4.3	x.4.4	x.4.5	x.4.6	x.4.7	...
ciel	x.5.1	x.5.2	x.5.3	x.5.4	x.5.5	x.5.6	x.5.7	...
dévorer	x.6.1	x.6.2	x.6.3	x.6.4	x.6.5	x.6.6	x.6.7	...
rouge	x.7.1	x.7.2	x.7.3	x.7.4	x.7.5	x.7.6	x.7.7	...
...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



	chat	manger	souris	arbre	ciel	dévorer	rouge	...
chat	x.1.1	x.1.2	x.1.3	x.1.4	x.1.5	x.1.6	x.1.7	...
manger	x.2.1	x.2.2	x.2.3	x.2.4	x.2.5	x.2.6	x.2.7	...
souris	x.3.1	x.3.2	x.3.3	x.3.4	x.3.5	x.3.6	x.3.7	...
arbre	x.4.1	x.4.2	x.4.3	x.4.4	x.4.5	x.4.6	x.4.7	...
ciel	x.5.1	x.5.2	x.5.3	x.5.4	x.5.5	x.5.6	x.5.7	...
dévorer	x.6.1	x.6.2	x.6.3	x.6.4	x.6.5	x.6.6	x.6.7	...
rouge	x.7.1	x.7.2	x.7.3	x.7.4	x.7.5	x.7.6	x.7.7	...
...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



	chat	manger	souris	arbre	ciel	dévorer	rouge	...
chat	x.1.1	x.1.2	x.1.3	x.1.4	x.1.5	x.1.6	x.1.7	...
manger	x.2.1	x.2.2	x.2.3	x.2.4	x.2.5	x.2.6	x.2.7	...
souris	x.3.1	x.3.2	x.3.3	x.3.4	x.3.5	x.3.6	x.3.7	...
arbre	x.4.1	x.4.2	x.4.3	x.4.4	x.4.5	x.4.6	x.4.7	...
ciel	x.5.1	x.5.2	x.5.3	x.5.4	x.5.5	x.5.6	x.5.7	...
dévorer	x.6.1	x.6.2	x.6.3	x.6.4	x.6.5	x.6.6	x.6.7	...
rouge	x.7.1	x.7.2	x.7.3	x.7.4	x.7.5	x.7.6	x.7.7	...
...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



	chat	manger	souris	arbre	ciel	dévorer	rouge	...
chat	x.1.1	x.1.2	x.1.3	x.1.4	x.1.5	x.1.6	x.1.7	...
manger	x.2.1	x.2.2	x.2.3	x.2.4	x.2.5	x.2.6	x.2.7	...
souris	x.3.1	x.3.2	x.3.3	x.3.4	x.3.5	x.3.6	x.3.7	...
arbre	x.4.1	x.4.2	x.4.3	x.4.4	x.4.5	x.4.6	x.4.7	...
ciel	x.5.1	x.5.2	x.5.3	x.5.4	x.5.5	x.5.6	x.5.7	...
dévorer	x.6.1	x.6.2	x.6.3	x.6.4	x.6.5	x.6.6	x.6.7	...
rouge	x.7.1	x.7.2	x.7.3	x.7.4	x.7.5	x.7.6	x.7.7	...
...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



	chat	manger	souris	arbre	ciel	dévorer	rouge	...
chat	x.1.1	x.1.2	x.1.3	x.1.4	x.1.5	x.1.6	x.1.7	...
manger	x.2.1	x.2.2	x.2.3	x.2.4	x.2.5	x.2.6	x.2.7	...
souris	x.3.1	x.3.2	x.3.3	x.3.4	x.3.5	x.3.6	x.3.7	...
arbre	x.4.1	x.4.2	x.4.3	x.4.4	x.4.5	x.4.6	x.4.7	...
ciel	x.5.1	x.5.2	x.5.3	x.5.4	x.5.5	x.5.6	x.5.7	...
dévorer	x.6.1	x.6.2	x.6.3	x.6.4	x.6.5	x.6.6	x.6.7	...
rouge	x.7.1	x.7.2	x.7.3	x.7.4	x.7.5	x.7.6	x.7.7	...
...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



	chat	manger	souris	arbre	ciel	dévorer	rouge	...
chat	x.1.1	x.1.2	x.1.3	x.1.4	x.1.5	x.1.6	x.1.7	...
manger	x.2.1	x.2.2	x.2.3	x.2.4	x.2.5	x.2.6	x.2.7	...
souris	x.3.1	x.3.2	x.3.3	x.3.4	x.3.5	x.3.6	x.3.7	...
arbre	x.4.1	x.4.2	x.4.3	x.4.4	x.4.5	x.4.6	x.4.7	...
ciel	x.5.1	x.5.2	x.5.3	x.5.4	x.5.5	x.5.6	x.5.7	...
dévorer	x.6.1	x.6.2	x.6.3	x.6.4	x.6.5	x.6.6	x.6.7	...
rouge	x.7.1	x.7.2	x.7.3	x.7.4	x.7.5	x.7.6	x.7.7	...
...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



	chat.obj	chat.suj	souris.obj	souris.suj	ciel.obj	...
chat	0	0	0	0	0	...
manger	y.2.1	y.2.2	y.2.3	y.2.4	y.2.5	...
souris	0	0	0	0	0	...
arbre	0	0	0	0	0	...
ciel	0	0	0	0	0	...
dévorer	y.6.1	y.6.2	y.6.3	y.6.4	y.6.5	...
rouge	0	0	0	0	0	...
...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



	chat.obj	chat.suj	souris.obj	souris.suj	ciel.obj	...
chat	0	0	0	0	0	...
manger	y.2.1	y.2.2	y.2.3	y.2.4	y.2.5	...
souris	0	0	0	0	0	...
arbre	0	0	0	0	0	...
ciel	0	0	0	0	0	...
dévorer	y.6.1	y.6.2	y.6.3	y.6.4	y.6.5	...
rouge	0	0	0	0	0	...
...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



	chat.obj	chat.suj	souris.obj	souris.suj	ciel.obj	...
chat	0	0	0	0	0	...
manger	y.2.1	y.2.2	y.2.3	y.2.4	y.2.5	...
souris	0	0	0	0	0	...
arbre	0	0	0	0	0	...
ciel	0	0	0	0	0	...
dévorer	y.6.1	y.6.2	y.6.3	y.6.4	y.6.5	...
rouge	0	0	0	0	0	...
...

Ordre-1 **V.manger (93 648)** : NC.salle.Gov.Prep/à : 4 703; V.donner.Gov.Prep/à : 1 243;
NC.fruit.Dep.obj : 1 017; NC.légume.Dep.obj : 742; NC.pain.Dep.Prep/de : ...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



Les mots qui apparaissent dans les mêmes contextes linguistiques partagent des significations similaires.
HARRIS, Z. (1954). Distributional structure. Word, 10(23).

	chat.obj	chat.suj	souris.obj	souris.suj	ciel.obj	...
chat	0	0	0	0	0	...
manger	y.2.1	y.2.2	y.2.3	y.2.4	y.2.5	...
souris	0	0	0	0	0	...
arbre	0	0	0	0	0	...
ciel	0	0	0	0	0	...
dévorer	y.6.1	y.6.2	y.6.3	y.6.4	y.6.5	...
rouge	0	0	0	0	0	...
...

Ordre-1 V.manger (93 648) : NC.salle.Gov.Prep/à : 4 703; V.donner.Gov.Prep/à : 1 243;
NC.fruit.Dep.obj : 1 017; NC.légume.Dep.obj : 742; NC.pain.Dep.Prep/de : ...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



Les mots qui apparaissent dans les mêmes contextes linguistiques partagent des significations similaires.
HARRIS, Z. (1954). Distributional structure. Word, 10(23).

	chat.obj	chat.suj	souris.obj	souris.suj	ciel.obj	...
chat	0	0	0	0	0	...
manger	y.2.1	y.2.2	y.2.3	y.2.4	y.2.5	...
souris	0	0	0	0	0	...
arbre	0	0	0	0	0	...
ciel	0	0	0	0	0	...
dévorer	y.6.1	y.6.2	y.6.3	y.6.4	y.6.5	...
rouge	0	0	0	0	0	...
...

Ordre-1 V.manger (93 648) : NC.salle.Gov.Prep/à : 4 703; V.donner.Gov.Prep/à : 1 243; NC.fruit.Dep.obj : 1 017; NC.légume.Dep.obj : 742; NC.pain.Dep.Prep/de : ...

Ordre-2 V.manger (93 648) : V.manger, V.cuisiner, V.bouffer, V.cuire, V.consommer, V.grignoter, V.cueillir, ...

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



Les mots qui apparaissent dans les mêmes contextes linguistiques partagent des significations similaires.
HARRIS, Z. (1954). Distributional structure. Word, 10(23).

Ordre-1 V.manger (93 648) : NC.salle.Gov.Prep/à : 4 703; V.donner.Gov.Prep/à : 1 243;
NC.fruit.Dep.obj : 1 017; NC.légume.Dep.obj : 742; NC.pain.Dep.Prep/de : ...

Ordre-2 V.manger (93 648) : V.manger, V.cuisiner, V.bouffer, V.cuire, V.consommer, V.grignoter,
V.cueillir, ...

Le chat **mange** la souris :

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



Les mots qui apparaissent dans les mêmes contextes linguistiques partagent des significations similaires.
HARRIS, Z. (1954). Distributional structure. Word, 10(23).

	chat.obj	chat.suj	souris.obj	souris.suj	ciel.obj	...
chat	0	0	0	0	0	...
manger	y.2.1	y.2.2	y.2.3	y.2.4	y.2.5	...
souris	0	0	0	0	0	...
arbre	0	0	0	0	0	...
ciel	0	0	0	0	0	...
dévorer	y.6.1	y.6.2	y.6.3	y.6.4	y.6.5	...
rouge	0	0	0	0	0	...
...

Ordre-1 V.manger (93 648) : NC.salle.Gov.Prep/à : 4 703; V.donner.Gov.Prep/à : 1 243; NC.fruit.Dep.obj : 1 017; NC.légume.Dep.obj : 742; NC.pain.Dep.Prep/de : ...

Ordre-2 V.manger (93 648) : V.manger, V.cuisiner, V.bouffer, V.cuire, V.consommer, V.grignoter, V.cueillir, ...

Le chat **mange** la souris :

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



Les mots qui apparaissent dans les mêmes contextes linguistiques partagent des significations similaires.
HARRIS, Z. (1954). Distributional structure. Word, 10(23).

	chat.obj	chat.suj	souris.obj	souris.suj	ciel.obj	...
chat	0	0	0	0	0	...
manger	y.2.1	y.2.2	y.2.3	y.2.4	y.2.5	...
souris	0	0	0	0	0	...
arbre	0	0	0	0	0	...
ciel	0	0	0	0	0	...
dévorer	y.6.1	y.6.2	y.6.3	y.6.4	y.6.5	...
rouge	0	0	0	0	0	...
...

Ordre-1 V.manger (93 648) : NC.salle.Gov.Prep/à : 4 703; V.donner.Gov.Prep/à : 1 243; NC.fruit.Dep.obj : 1 017; NC.légume.Dep.obj : 742; NC.pain.Dep.Prep/de : ...

Ordre-2 V.manger (93 648) : V.manger, V.cuisiner, V.bouffer, V.cuire, V.consommer, V.grignoter, V.cueillir, ...

Le chat **mange** la souris :

Lexiques : Cooccurrences, Distances, Ordre-1 & Ordre-2



Les mots qui apparaissent dans les mêmes contextes linguistiques partagent des significations similaires.
HARRIS, Z. (1954). Distributional structure. Word, 10(23).

	chat.obj	chat.suj	souris.obj	souris.suj	ciel.obj	...
chat	0	0	0	0	0	...
manger	y.2.1	y.2.2	y.2.3	y.2.4	y.2.5	...
souris	0	0	0	0	0	...
arbre	0	0	0	0	0	...
ciel	0	0	0	0	0	...
dévorer	y.6.1	y.6.2	y.6.3	y.6.4	y.6.5	...
rouge	0	0	0	0	0	...
...

Ordre-1 V.manger (93 648) : NC.salle.Gov.Prep/à : 4 703; V.donner.Gov.Prep/à : 1 243;
NC.fruit.Dep.obj : 1 017; NC.légume.Dep.obj : 742; NC.pain.Dep.Prep/de : ...

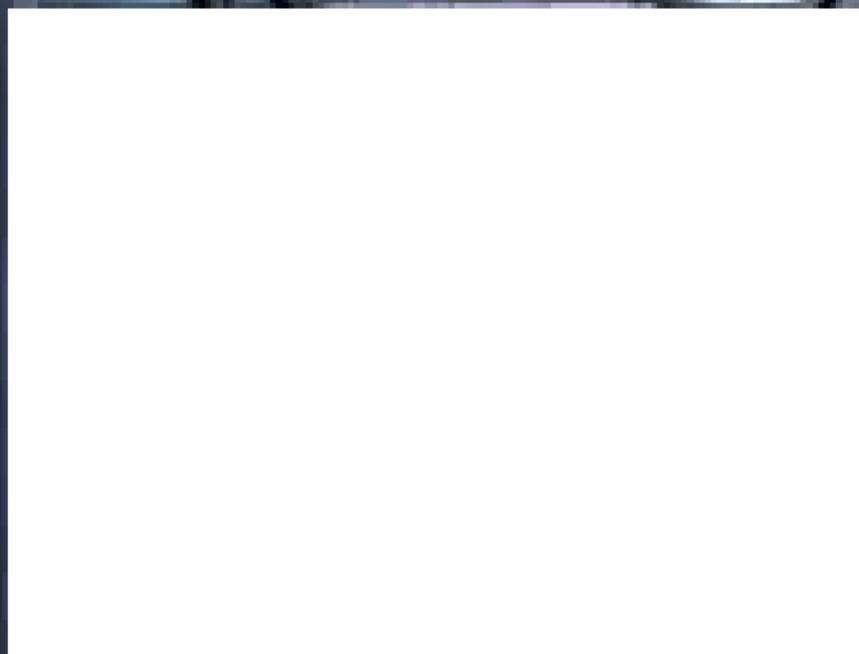
Ordre-2 V.manger (93 648) : V.manger, V.cuisiner, V.bouffer, V.cuire, V.consommer, V.grignoter,
V.cueillir, ...

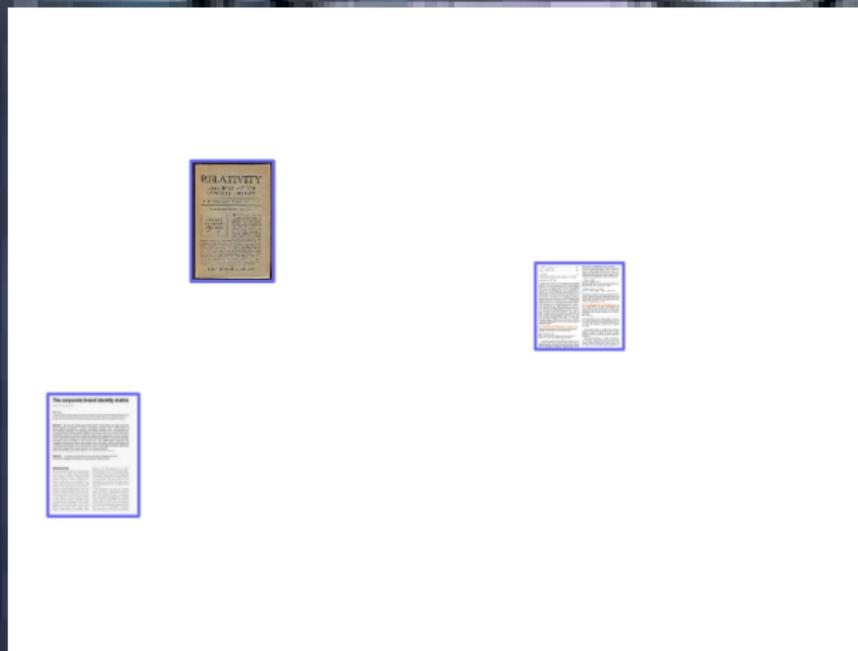
Le chat **mange** la souris : V.avalier, V.croquer, ...

Plan de l'exposé

- 1 Les objets d'un corpus documentaire
- 2 Lexiques
- 3 **Réseaux**
- 5 Visualisations/Navigations









relativity



matrix

quantum
physics





quantum physics

relativity



matrix



Réseaux



université de Zurich



quantum physics



relativity



matrix

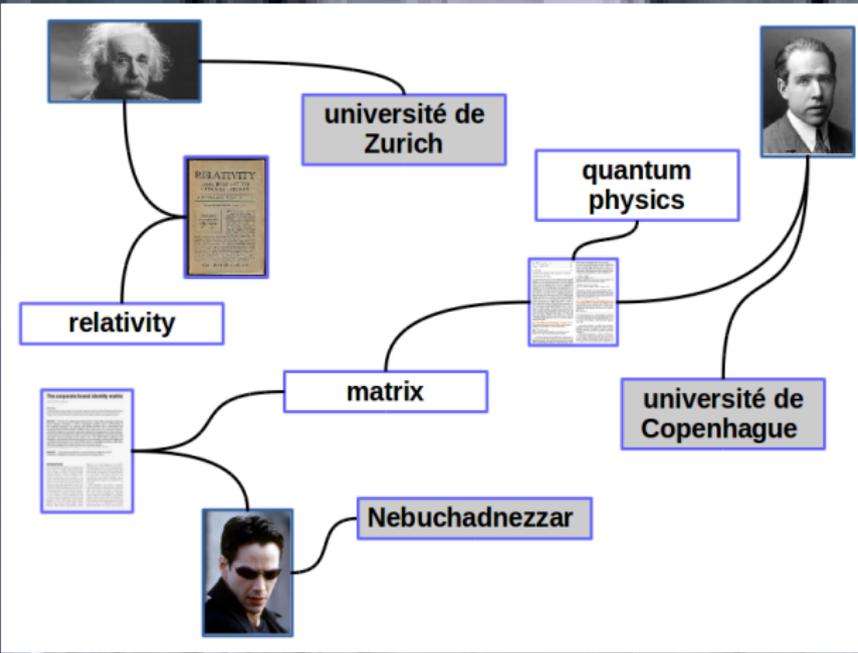
université de Copenhague

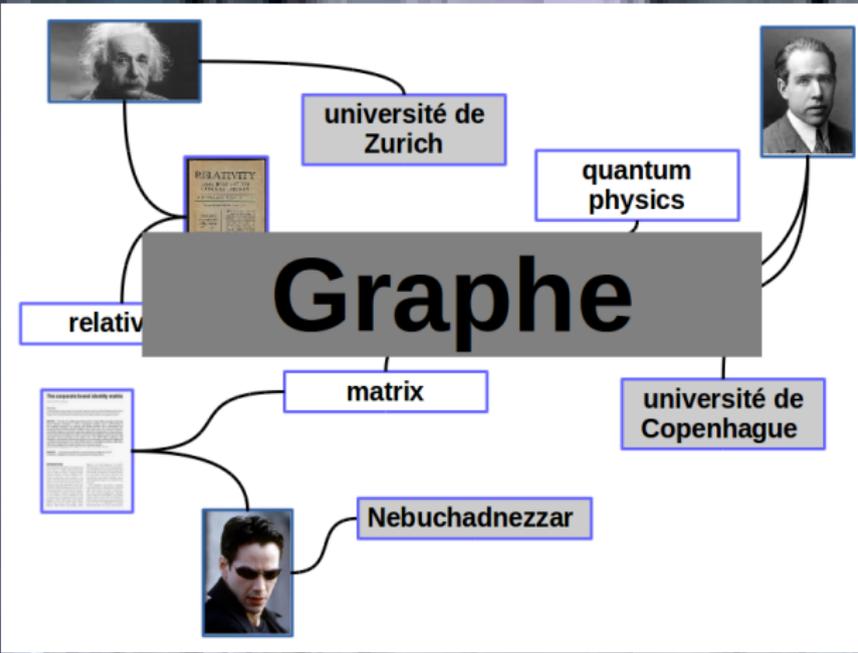


Nebuchadnezzar

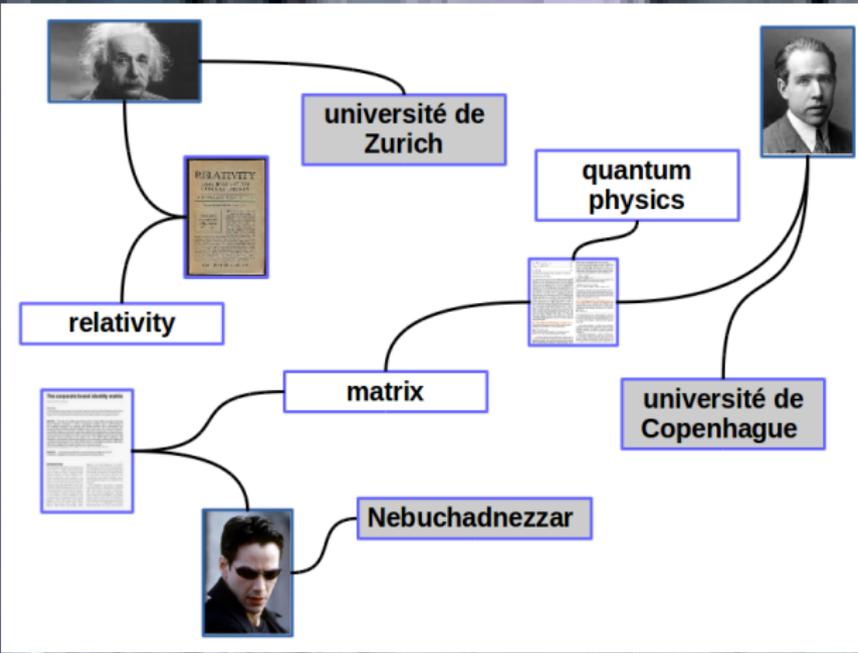


Réseaux





Réseaux

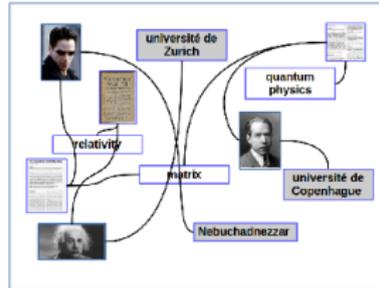
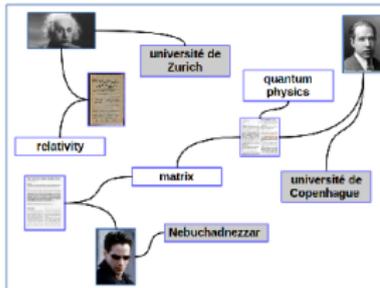


$V = \{\text{Einstein:1, Université de Zurich:2, ISBN 978-92-95055-16-2:3, relativity:4, Bhor:5, Université de Copenhague:6, quantum physics:7, ISBN 630-25-83991-20-1:8, matrix:9, Nebuchadnezzar:10, Néo:11, ISBN 768-86-8336101-15-5:12}\}$

$E = \{\{1,2\}, \{1,3\}, \{3,4\}, \{11,12\}, \{11,10\}, \{9,12\}, \{9,8\}, \{8,5\}, \{8,5\}, \{5,6\}\}$

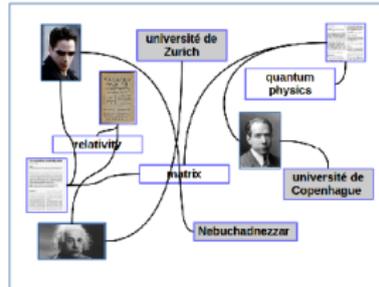
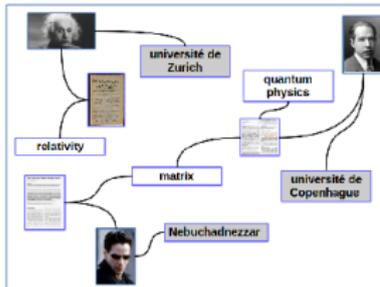
$V = \{\text{Einstein:1, Université de Zurich:2, ISBN 978-92-95055-16-2:3, relativity:4, Bhor:5, Université de Copenhague:6, quantum physics:7, ISBN 630-25-83991-20-1:8, matrix:9, Nebuchadnezzar:10, Néo:11, ISBN 768-86-8336101-15-5:12}\}$

$E = \{\{1,2\}, \{1,3\}, \{3,4\}, \{11,12\}, \{11,10\}, \{9,12\}, \{9,8\}, \{8,5\}, \{8,5\}, \{5,6\}\}$



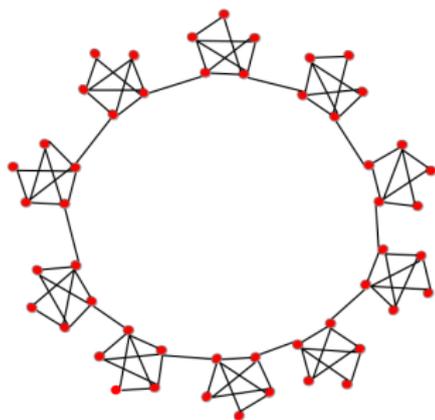
$V = \{\text{Einstein:1, Université de Zurich:2, ISBN 978-92-95055-16-2:3, relativity:4, Bhor:5, Université de Copenhague:6, quantum physics:7, ISBN 630-25-83991-20-1:8, matrix:9, Nebuchadnezzar:10, Néo:11, ISBN 768-86-8336101-15-5:12}\}$

$E = \{\{1,2\}, \{1,3\}, \{3,4\}, \{11,12\}, \{11,10\}, \{9,12\}, \{9,8\}, \{8,5\}, \{8,5\}, \{5,6\}\}$

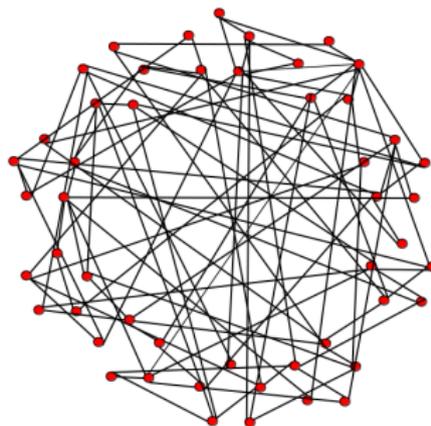
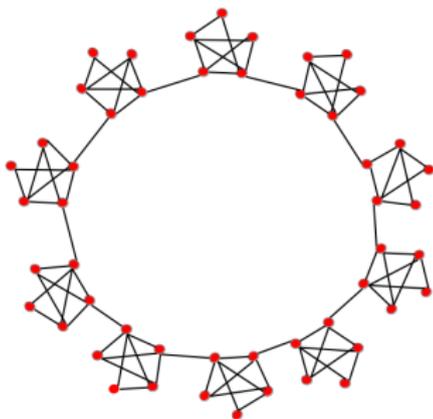


Graphes de laboratoire

Graphes de laboratoire



Graphes de laboratoire



Graphes de terrain

Graphes de terrain

- Les graphes d'accointance d'un groupe d'humains
- Le graphe du World Wide Web
- Le graphe de *Caenorhabditis elegans*
- Les graphes extraits de bases documentaires

Graphes de terrain

- Les graphes d'acointance d'un groupe d'humains
- Le graphe du World Wide Web
- Le graphe de *Caenorhabditis elegans*
- Les graphes extraits de bases documentaires

Deux propriétés structurelles fondamentales

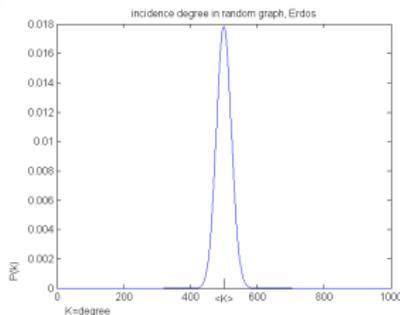
- ▶ Distribution des degrés à queue lourde (loi de puissance)
- ▶ Fort coefficient de clustering : zones denses en arêtes

Graphes de terrain

Deux propriétés structurelles fondamentales

- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

Graphes de terrain

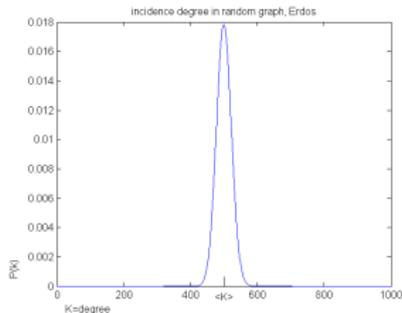


Graphes Aléatoires

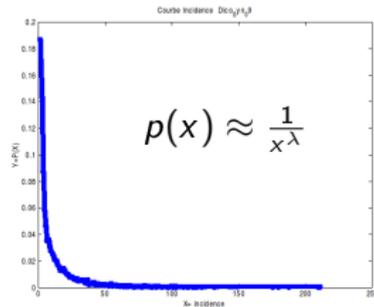
Deux propriétés structurales fondamentales

- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

Graphes de terrain



Graphes Aléatoires

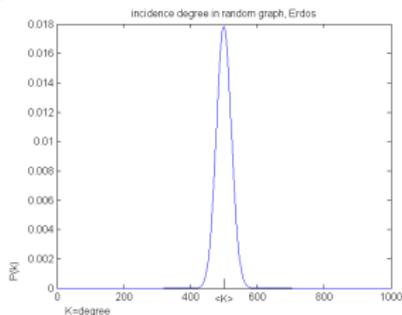


Graphes de Terrain

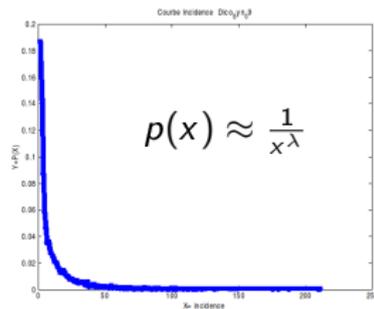
Deux propriétés structurales fondamentales

- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

Graphes de terrain



Graphes Aléatoires

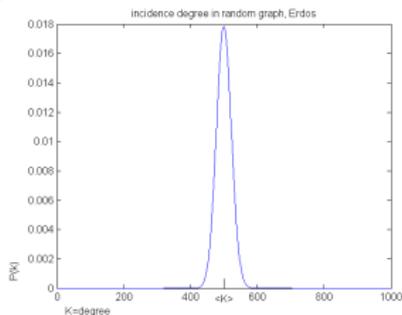


Graphes de Terrain

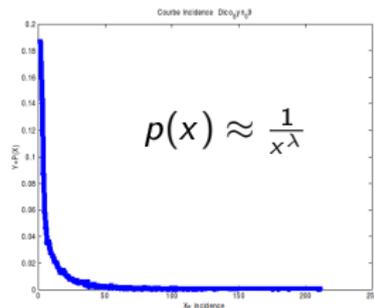
Deux propriétés structurales fondamentales

- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

Graphes de terrain



Graphes Aléatoires

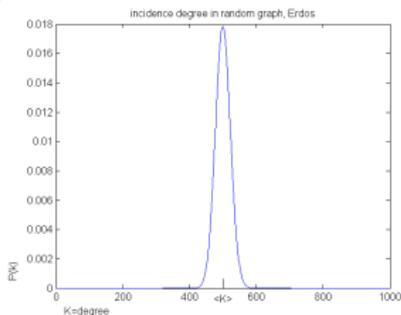


Graphes de Terrain

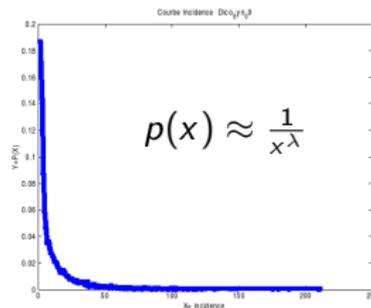
Deux propriétés structurales fondamentales

- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

Graphes de terrain



Graphes Aléatoires

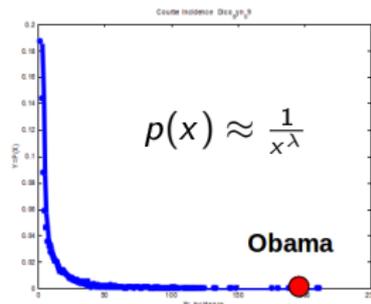
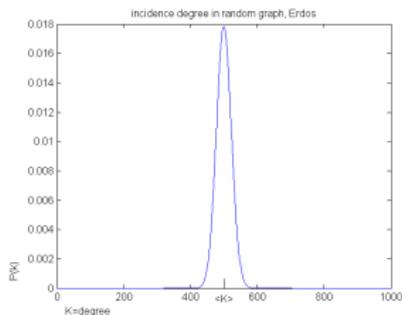


Graphes de Terrain

Deux propriétés structurales fondamentales

- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

Graphes de terrain



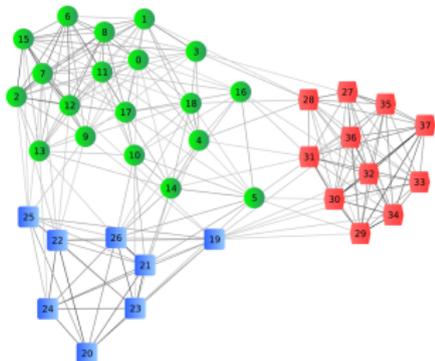
Graphes Aléatoires

Graphes de Terrain

Deux propriétés structurales fondamentales

- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

Graphes de terrain



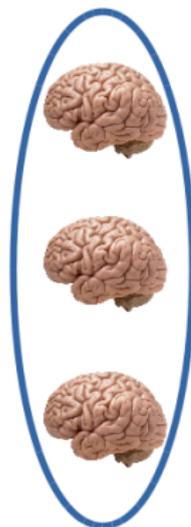
Deux propriétés structurales fondamentales

- ▶ Distribution des degrés à queue lourde (loi de puissance)
- ▶ **Fort coefficient de clustering : zones denses en arêtes**

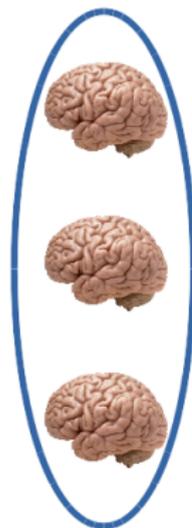
Plan de l'exposé

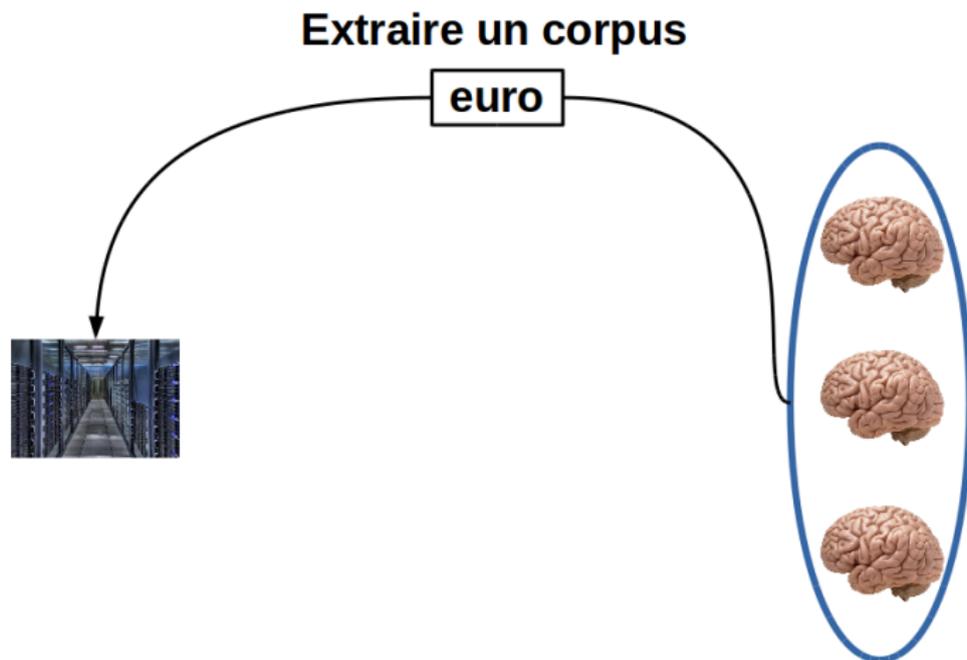
- 1 Les objets d'un corpus documentaire
- 2 Lexiques
- 3 Réseaux
- 5 **Visualisations/Navigations**

Visualisations/Navigations



Extraire un corpus





Extraire un corpus

euro



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

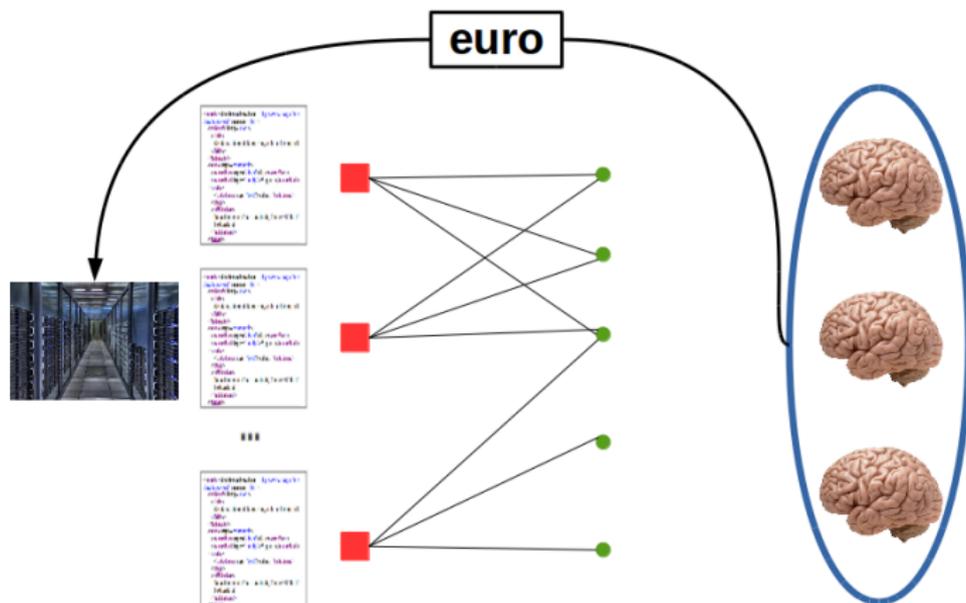
...

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

- 1 [Greece pleads for more time to cut public sector jobs - live](#) September 24, 2013: Greece pleads for more time to cut public sector - live <http://www.theguardian.com/business/2013/sep/24/german-coalition-haggling-begins-as-greek-state-workers-strike-again-live>
- 2 [If Europe's future is German, Greece has only one hope for salvation](#) | Costas Douzinas September 24, 2013: If Europe's future is German, Greece has only one hope for salvation <http://www.theguardian.com/commentisfree/2013/sep/24/german-angela-merkel-austerity>
- 3 [The Spanish language school that delivers lessons on a budget](#) September 24, 2013: The Spanish language school that delivers lessons on a budget <http://www.theguardian.com/education/2013/sep/24/spanish-state-funded-language-schools>
- 4 [Labour conference - Ed Balls's speech, Politics live blog](#) September 23, 2013: Labour conference - Ed Balls's speech: Politics live blog <http://www.theguardian.com/politics/2013/sep/23/labour-conference-ed-balls-speech-politics-live-blog>
- 5 [Markets fall as Angela Merkel faces 'difficult' coalition talks - as it happened](#) September 23, 2013: Markets fall as Angela Merkel begins 'difficult' coalition talks - as it happened <http://www.theguardian.com/business/2013/sep/23/markets-calm-after-merkel-election-trumpch-live>
- 6 [Want to be a successful world leader? Do it the Angela Merkel way](#) September 23, 2013: Want to be a successful world leader? Do it the Angela Merkel way <http://www.theguardian.com/world/shortcuts/2013/sep/23/angela-merkel-successful-world-leader-gemany>
- 7 [Ed Balls's message to Labour: we will be less useless next time - probably](#) September 23, 2013: Ed Balls's message to Labour: we will be less useless next time - probably

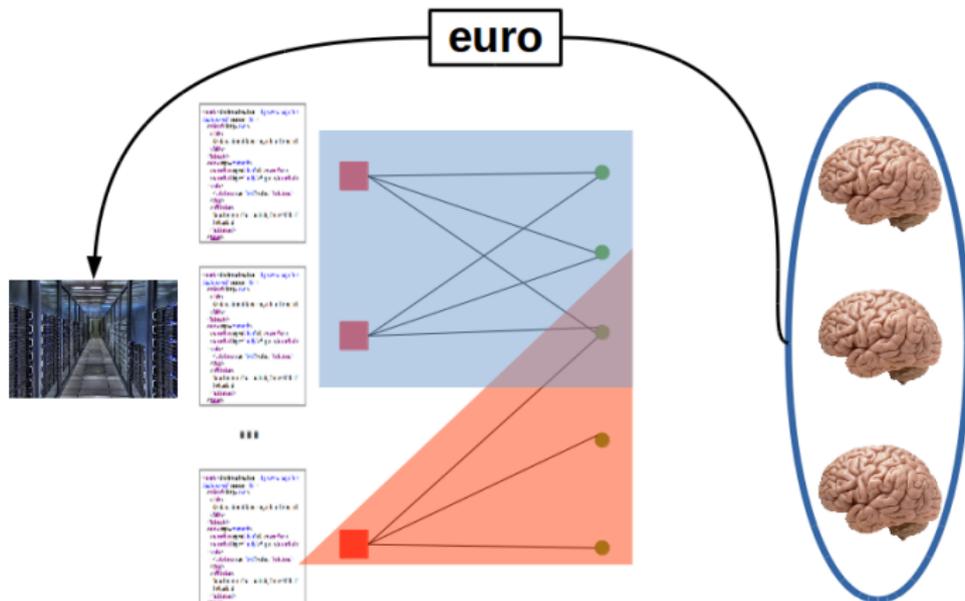


Extraire un corpus



Visualisations/Navigations

Extraire un corpus



Extraire un corpus

euro



Germany, World news,
Angela Merkel, Europe,
Eurozone crisis, European
Union, Economics

Sport, Football, Premier
League, Fulham,
Chelsea, Liverpool, Luis
Suárez

1 Greece already for more time to cut public sector jobs - live

September 24, 2013: Greece pleads for more time to cut public sector - live
<http://www.theguardian.com/business/2013/sep/24/german-coalition-fudging-begins-as-greek-state-workers-strike-again-live>

2 If Europe's future is German, Greece has only one hope for salvation | Carlos Dominguez

September 24, 2013: If Europe's future is German, Greece has only one hope for salvation
<http://www.theguardian.com/commentisfree/2013/sep/24/german-angels-merkel-austerity>

3 The Spanish language school that delivers lessons on a budget

September 24, 2013: The Spanish language school that delivers lessons on a budget
<http://www.theguardian.com/education/2013/sep/24/spanish-language-school>

4 Labour conference - Ed Balls's speech, Politics live blog

September 23, 2013: Labour conference - Ed Balls's speech, Politics live blog
<http://www.theguardian.com/politics/2013/sep/23/labour-conference-ed-balls-speech-politics-live-blog>

5 Markets fall as Angela Merkel faces difficult coalition talks - as it happened

September 23, 2013: Markets fall as Angela Merkel begins 'difficult' coalition talks - as it happened
<http://www.theguardian.com/business/2013/sep/23/markets-calm-after-markets-election-blumph-live>

6 Want to be a successful world leader? Do it the Angela Merkel way

September 23, 2013: Want to be a successful world leader? Do it the Angela Merkel way
<http://www.theguardian.com/world/2013/sep/23/angela-merkel-successful-world-leader-germany>

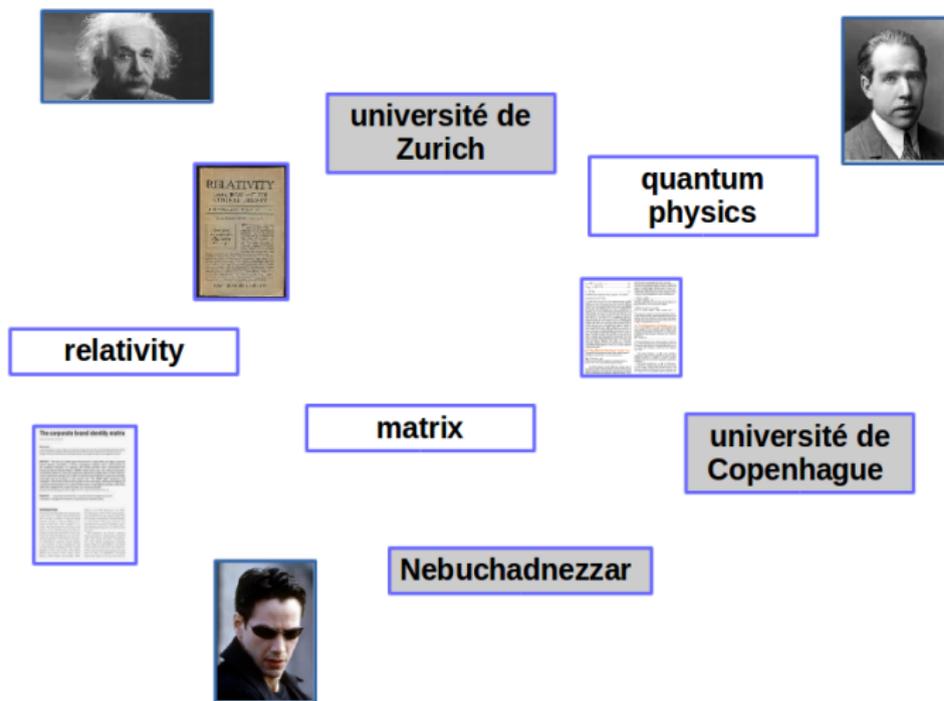
7 Ed Balls's message to Labour, an will be less useful next time - probably

September 23, 2013: Ed Balls's message to Labour: an will be less useful next time - probably

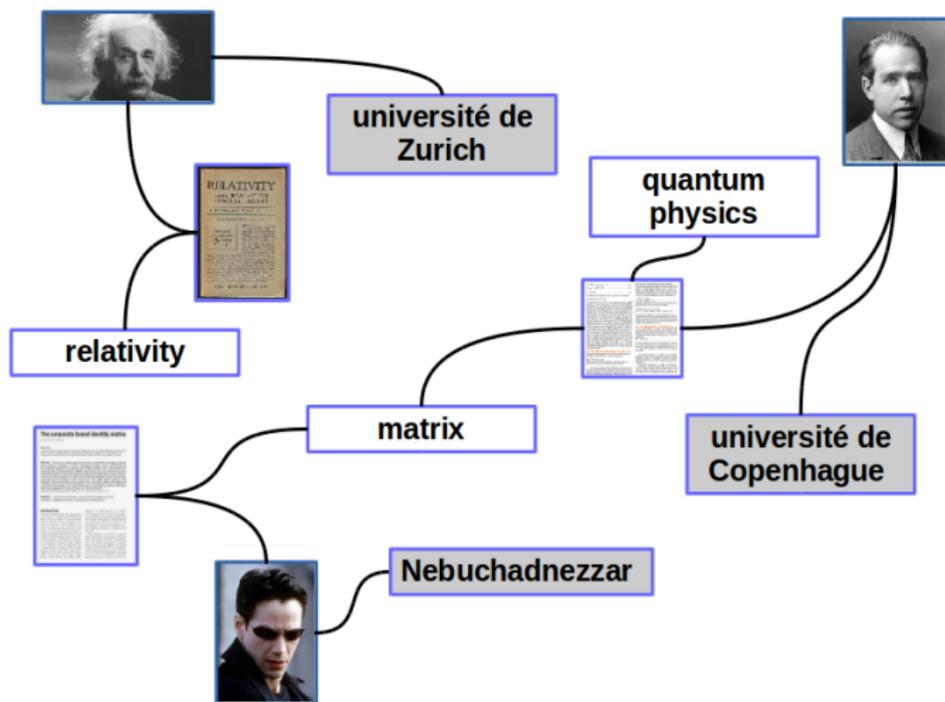


Visualisations/Navigations

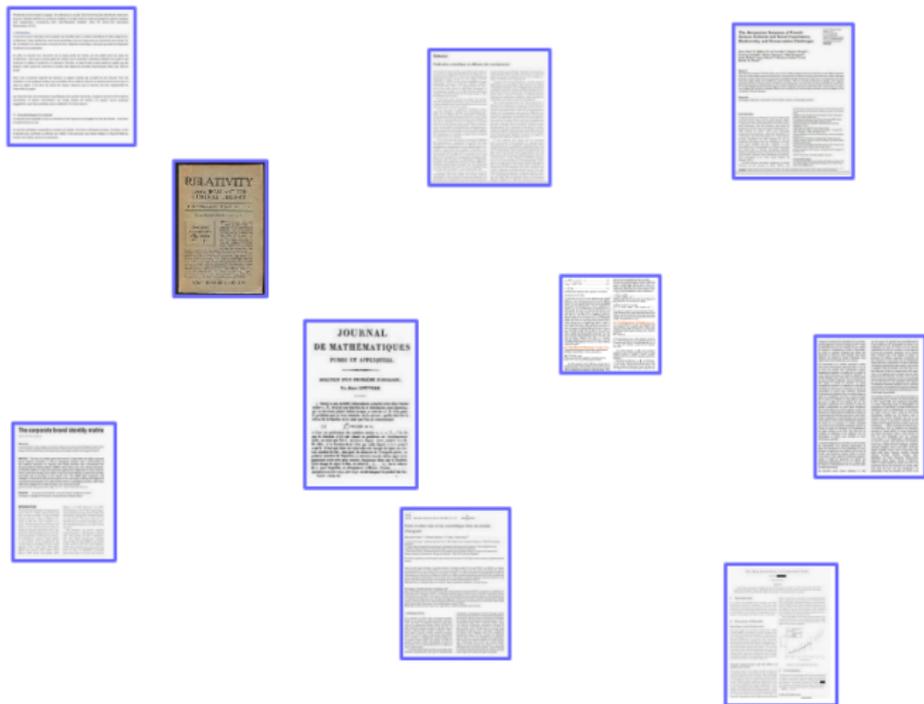
Visualisations/Navigations



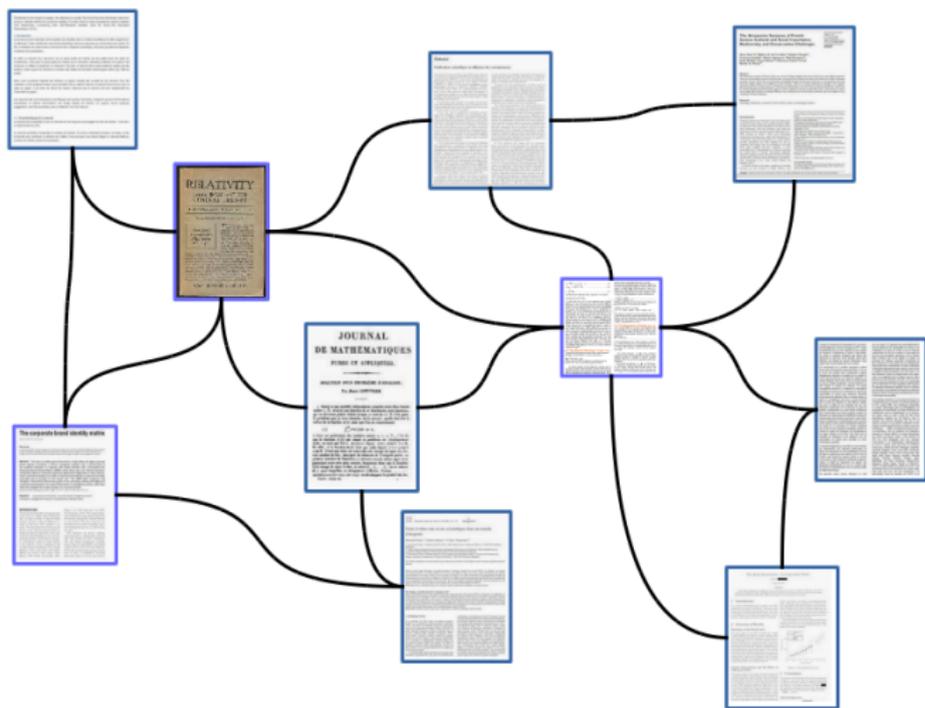
Visualisations/Navigations



Visualisations/Navigations



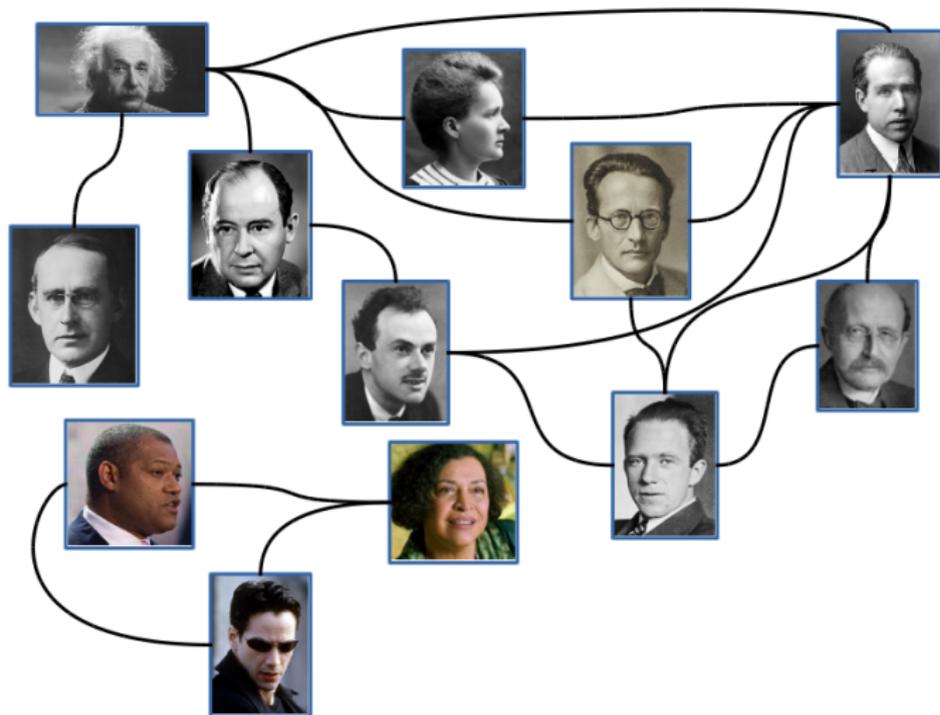
Visualisations/Navigations



Visualisations/Navigations



Visualisations/Navigations

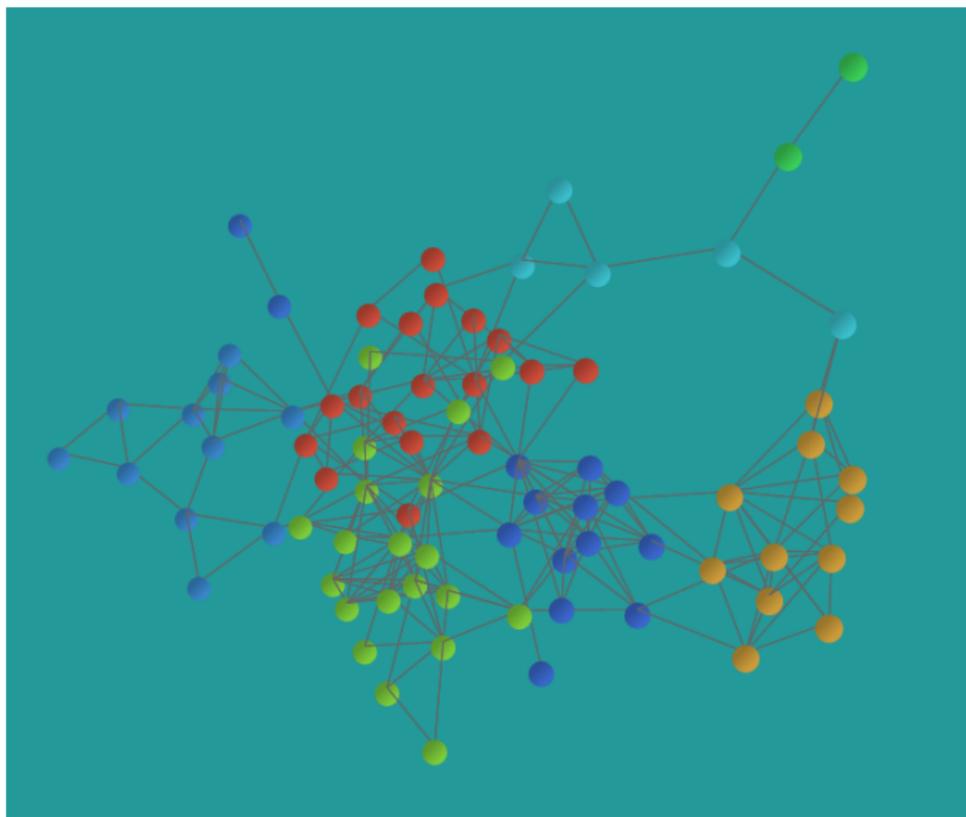


Visualisations/Navigations

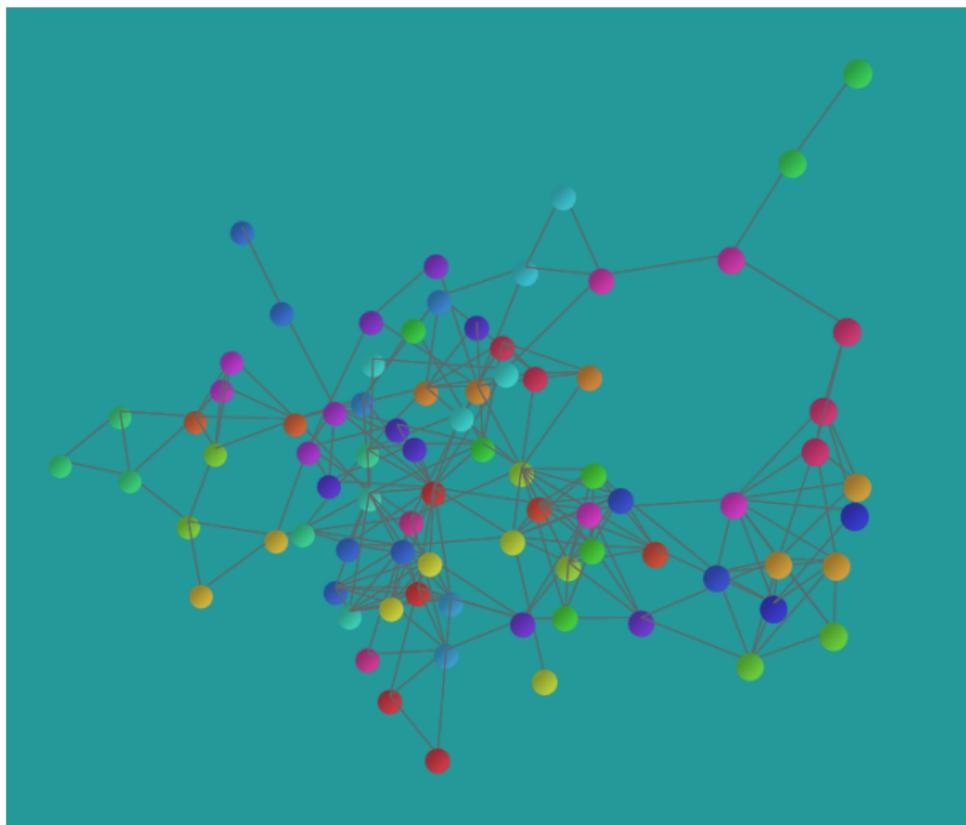
$V = \{\text{Einstein:1, Université de Zurich:2, ISBN 978-92-95055-16-2:3, relativity:4, Bhor:5, Université de Copenhague:6, quantum physics:7, ISBN 630-25-83991-20-1:8, matrix:9, Nebuchadnezzar:10, Néo:11, Barabazi:12, ISBN 768-86-8336101-15-5:13, Newman:14, connexion:15, ISBN 978-92-67-16-2:16, ISCPIF:17, CLLE:18, graph:19, proxemie:20, colnfluence:21, random-walk:22}\}$

$E = \{\{1,2\}, \{1,3\}, \{3,4\}, \{11,12\}, \{11,10\}, \{9,12\}, \{9,8\}, \{8,5\}, \{8,5\}, \{5,6\}, \{14,20\}, \{6,11\}, \{19,2\}, \{15,1\}, \{4,3\}, \{22,20\}, \{20,21\}\}$

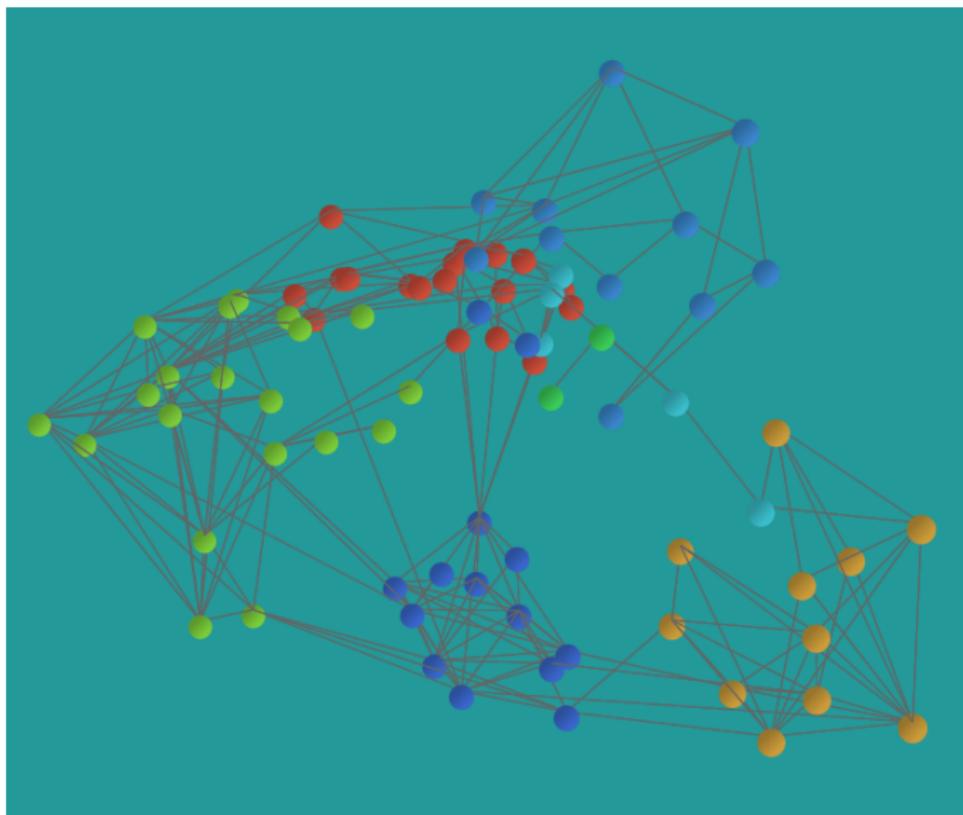
Visualisations/Navigations



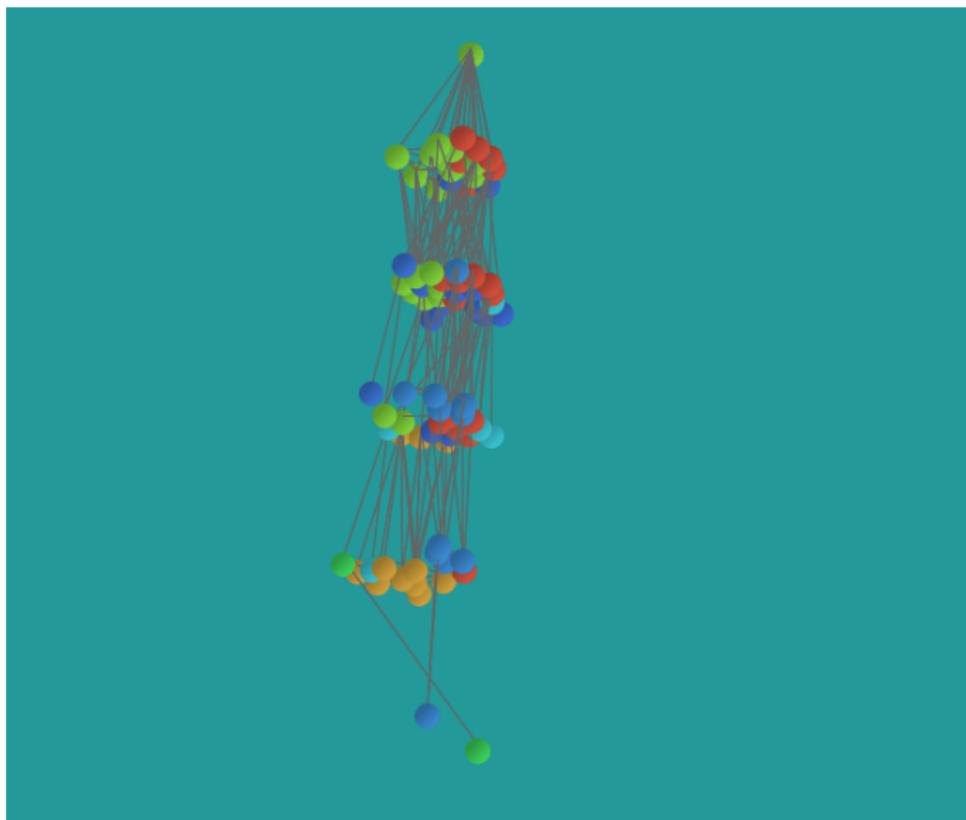
Visualisations/Navigations



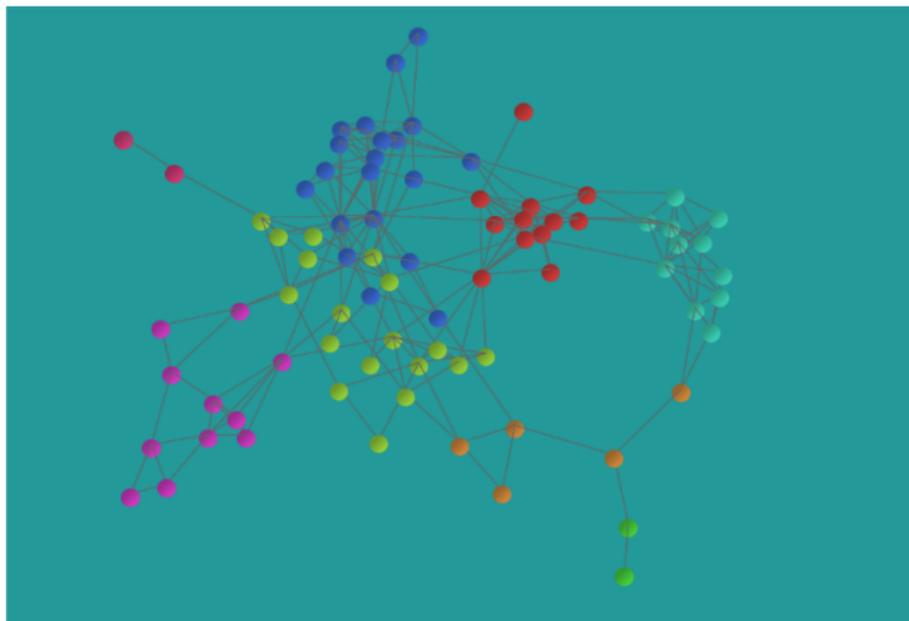
Visualisations/Navigations



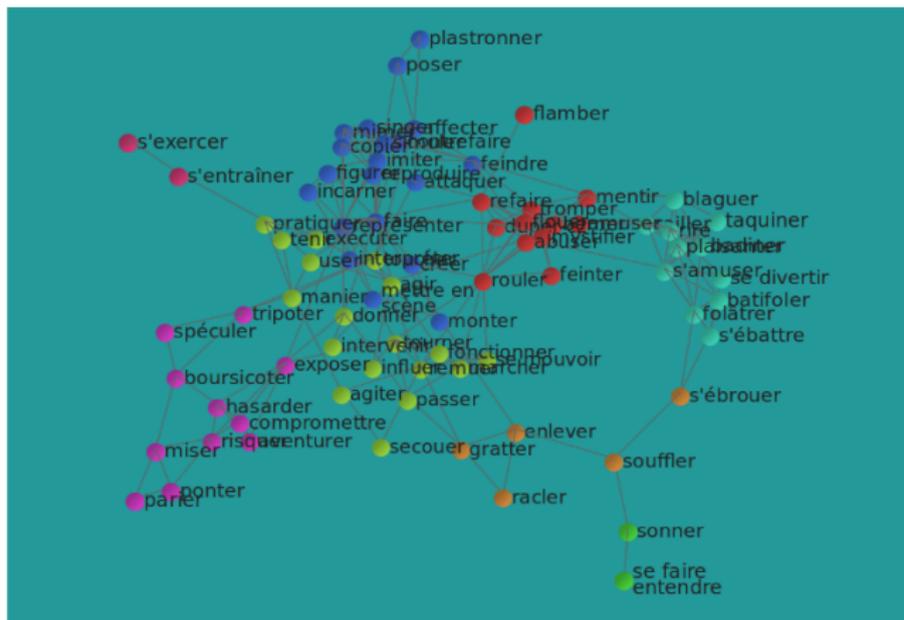
Visualisations/Navigations



Visualisations/Navigations

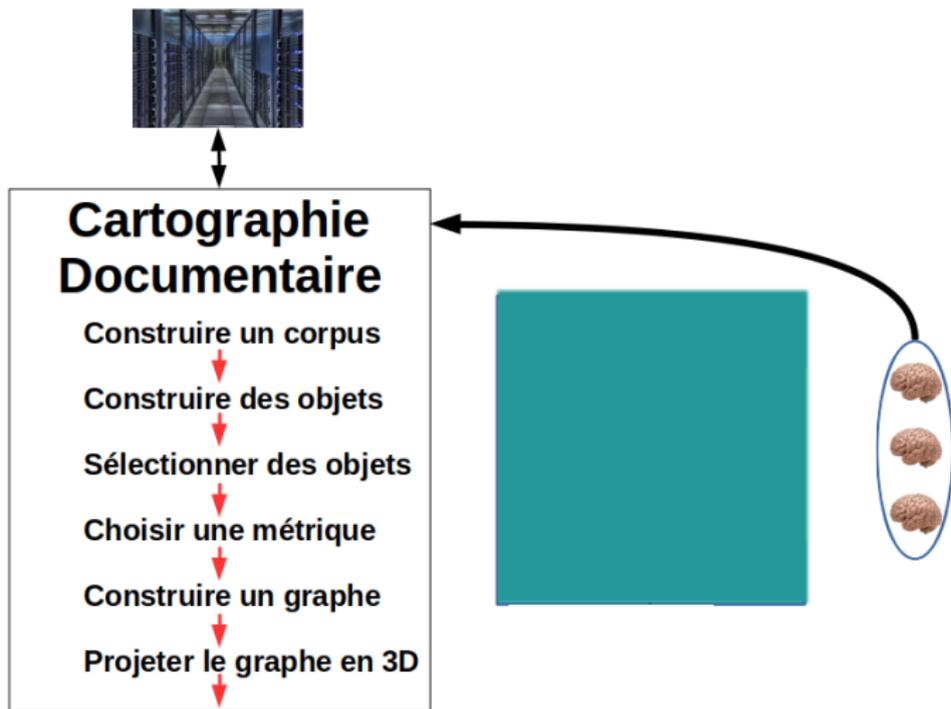


Visualisations/Navigations

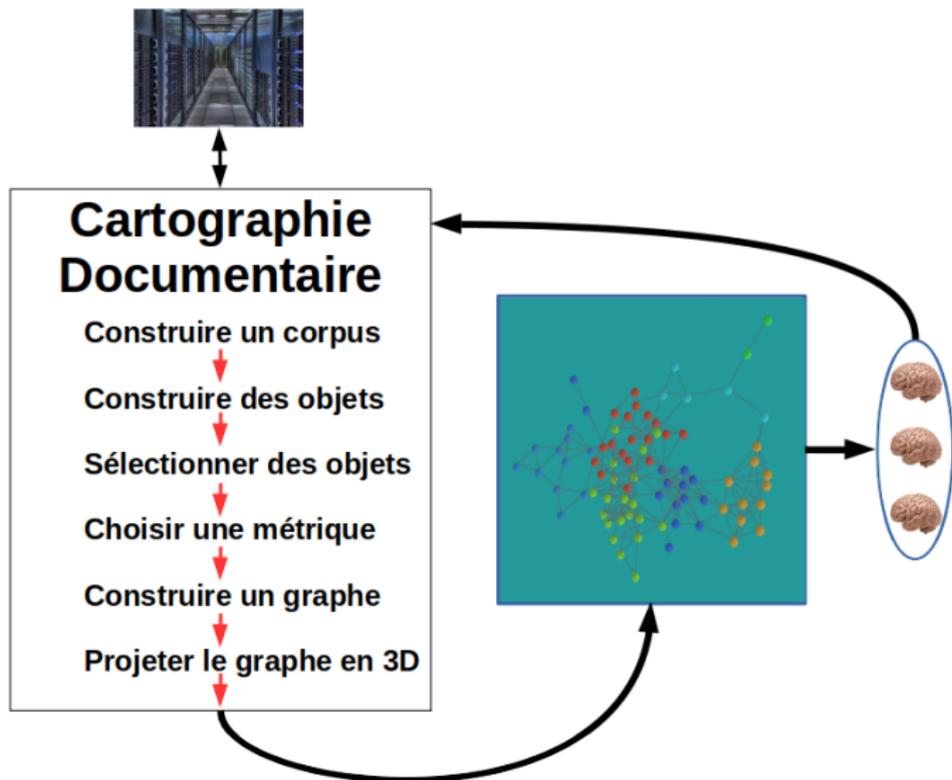


Cartographie documentaire

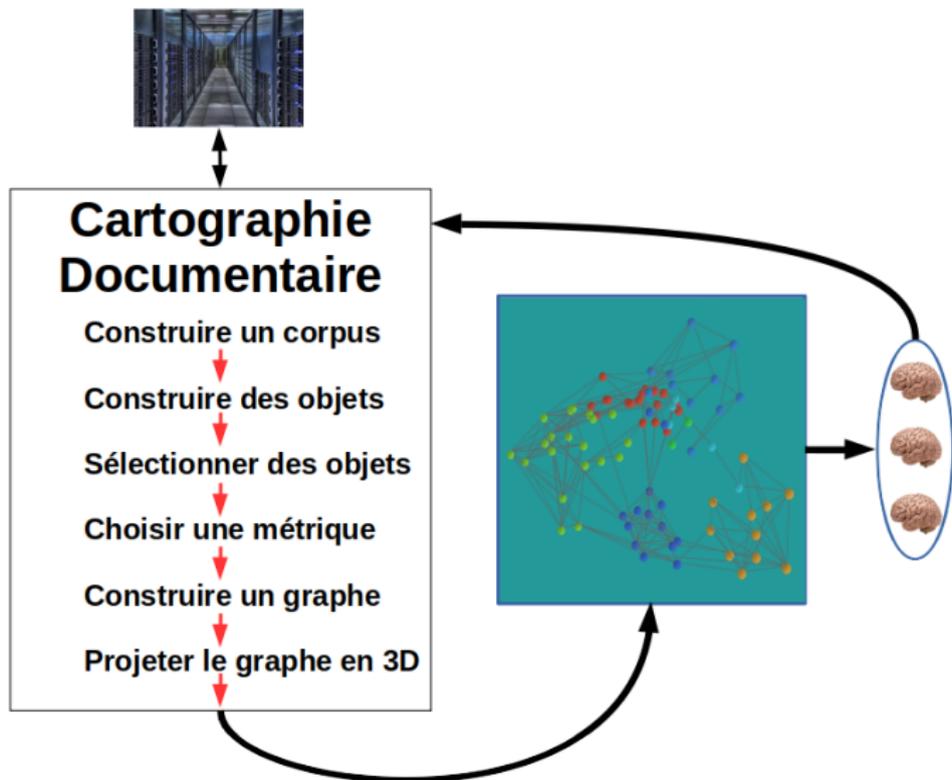
Cartographie documentaire



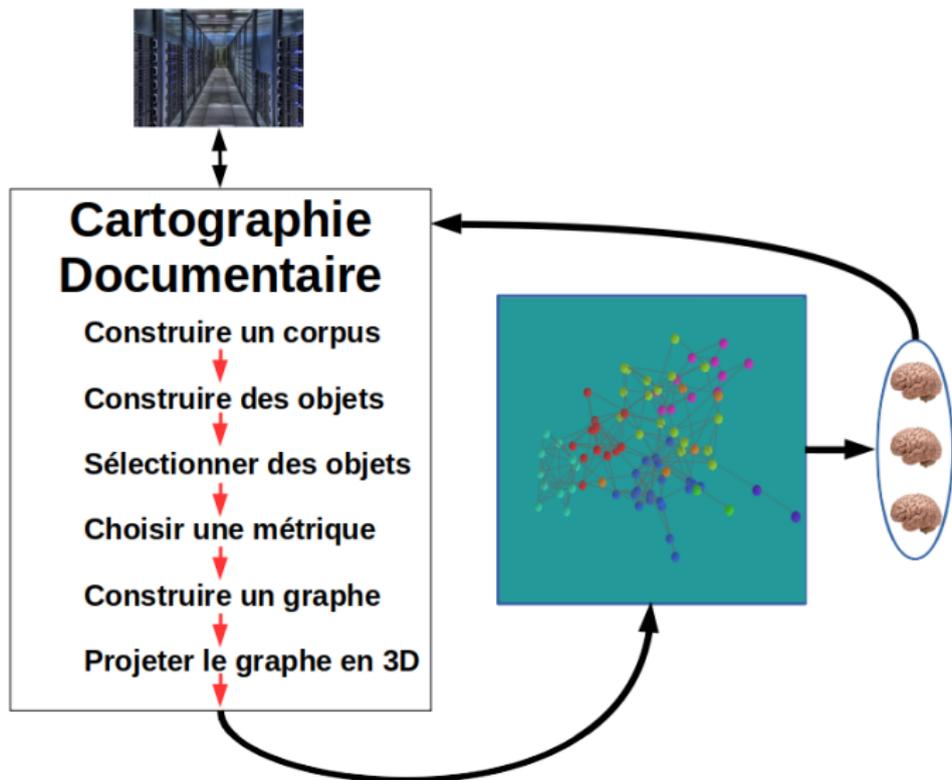
Cartographie documentaire



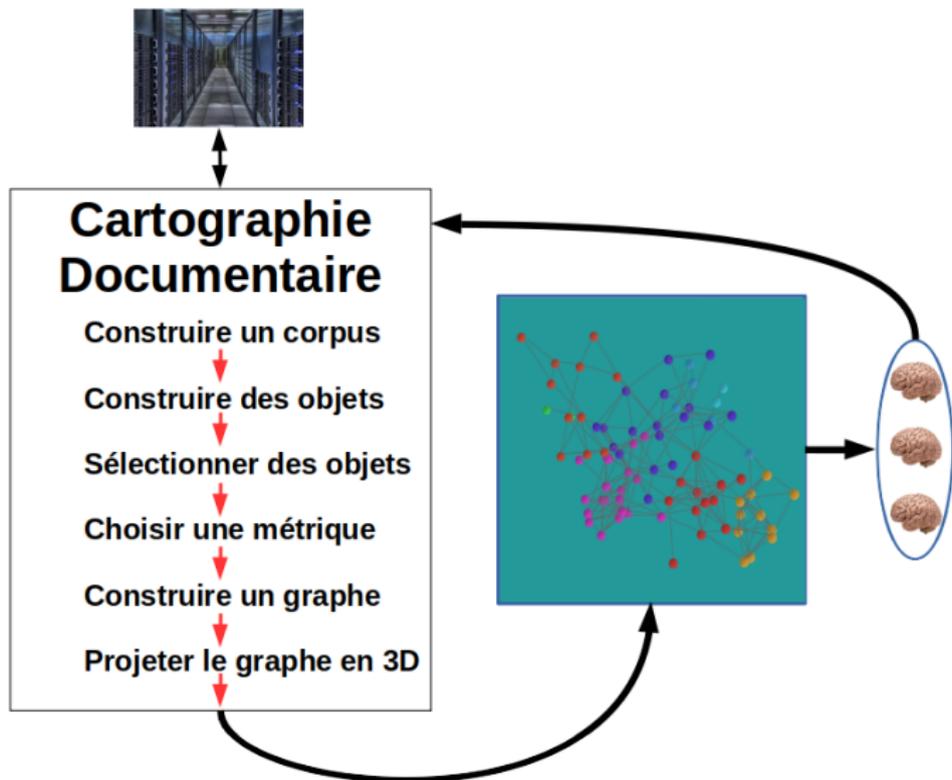
Cartographie documentaire



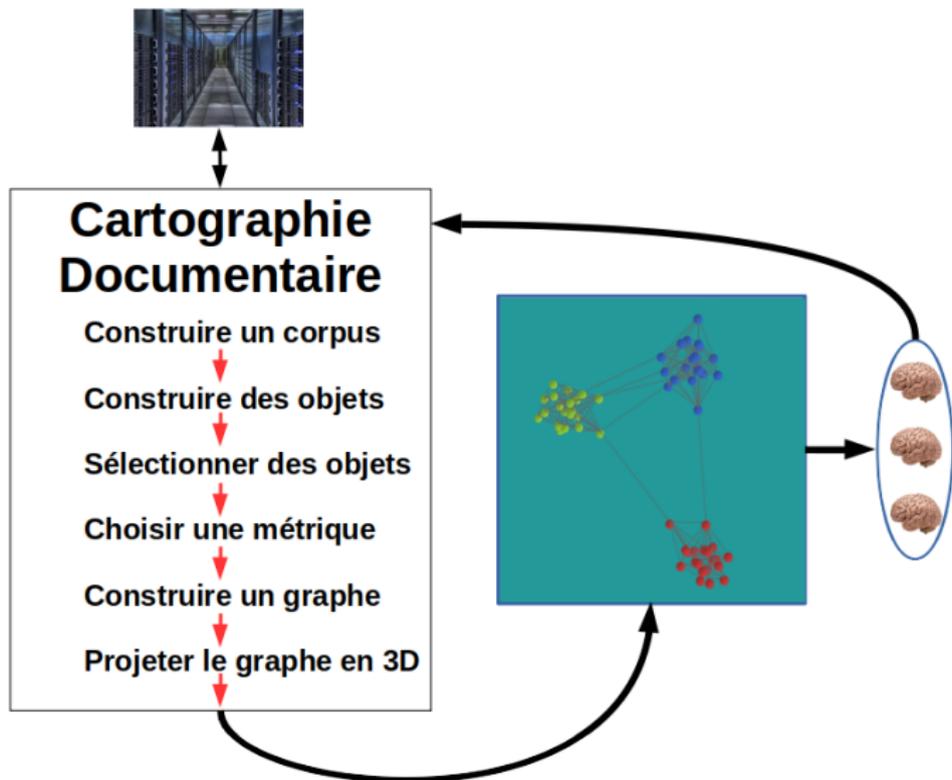
Cartographie documentaire



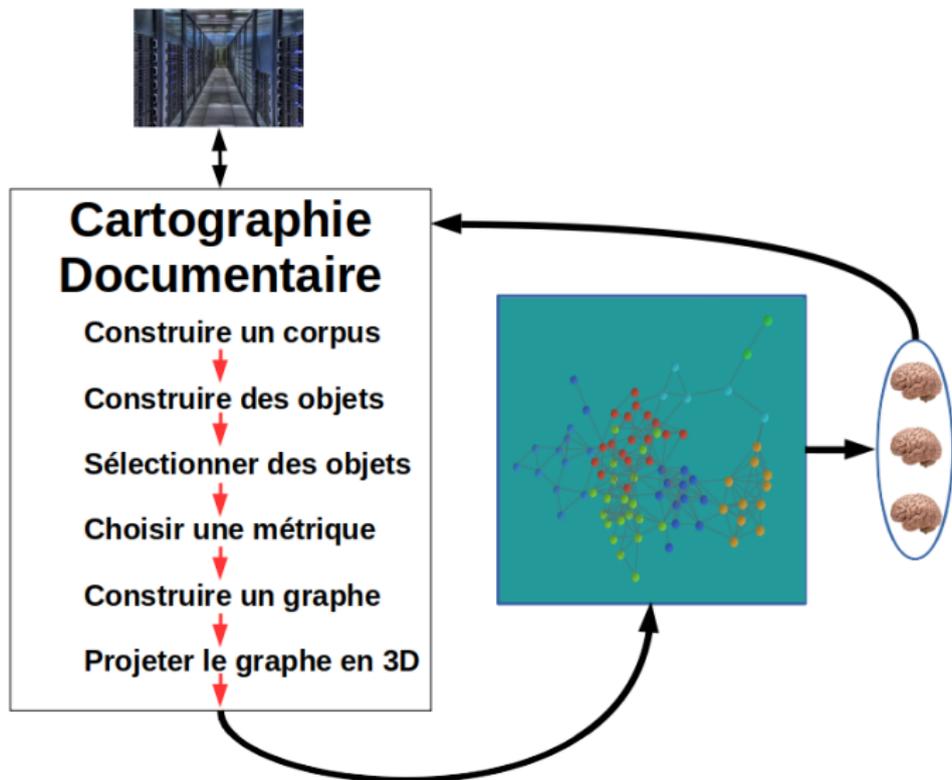
Cartographie documentaire



Cartographie documentaire



Cartographie documentaire

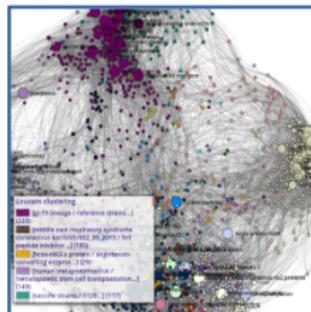
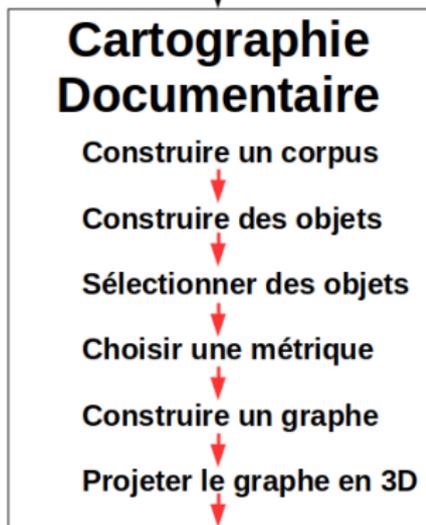


Cartographie documentaire



Analysis of coronavirus literature

These maps has been realized using [GarganText](#) from a PubMed Corpora with query "coronavirus" on Feb 5 2020 (6560 documents from Jan 2000 to present). [The methodology is described here](#)



La cartographie documentaire

Métrologies et visualisations de corpus

Bruno Gaume

CLLE, CNRS & Univ. Toulouse

À Paris, ISCIPIF, Mars 2021

ANF TDM 2020 - Exploration documentaire et extraction d'informations