



Analyser, comprendre le monde : Complémentarité entre apprentissage et visualisation

Jean-Daniel Fekete & Michèle Sebag



Aviz

Analyse & Visualisation

Jean-Daniel Fekete, DR Inria

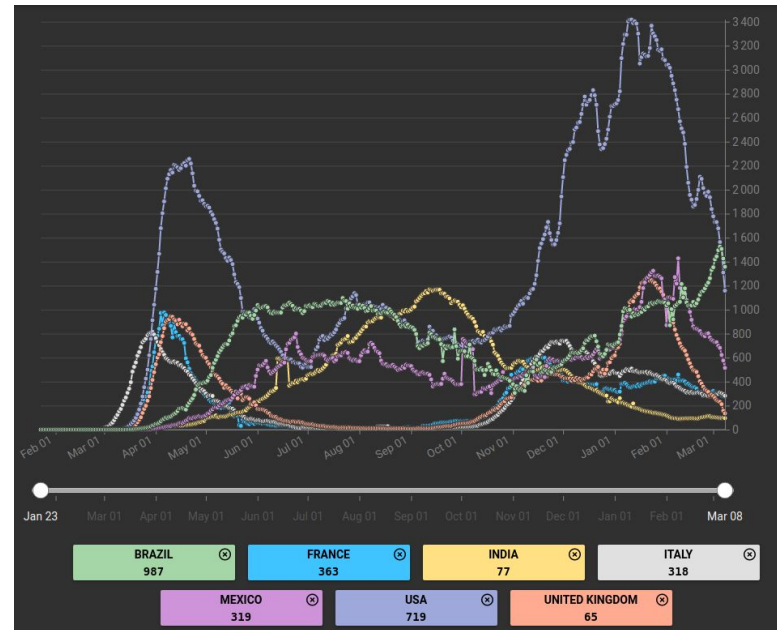
<https://aviz.fr/>





Objectif scientifique : visualiser pour comprendre

- Explorer des données
 - Complexes, volumineuses, hétérogènes, ...
- Pour
 - Les vérifier / nettoyer
 - Les comprendre
 - Prendre des décisions



<https://boogheta.github.io/coronavirus-countries>

Scientific Context and Overall Objectives

- Data is now abundant. We need more effective methods, techniques, and tools to **understand** it and **make decisions** with it
- Aviz is a **multidisciplinary project** that
 - seeks to improve **data exploration** methods, techniques, and tools
 - based on **interactive visualization**
- **Visualization** is [Card et al. 99]:

“the use of computer-supported, interactive, visual representations of data to amplify cognition.”

Confirm vs. Explore

Confirmatory Analysis

- start with specific questions
- devise measures to answer them
- only collect and look at data once above are answered

focus of fully automated analysis

Exploratory Analysis

- likely no a-priori information about the data
- not sure about structures and information present
- explore to create hypotheses & confirm later

focus of visual analytics

- Needs humans in the loop
- Relies on the visual channel and interaction

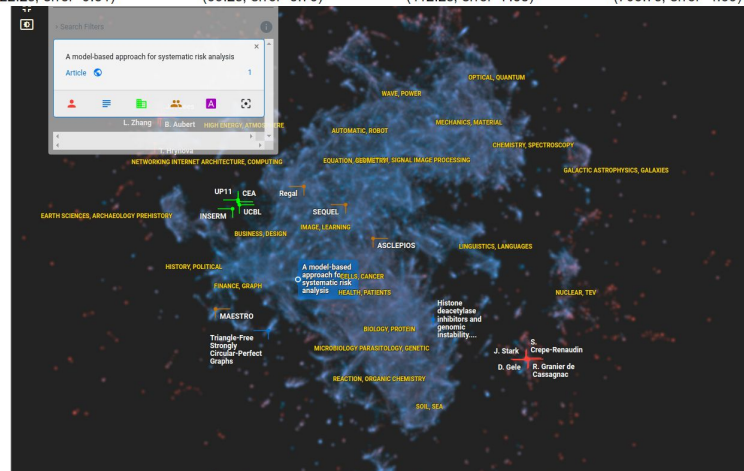
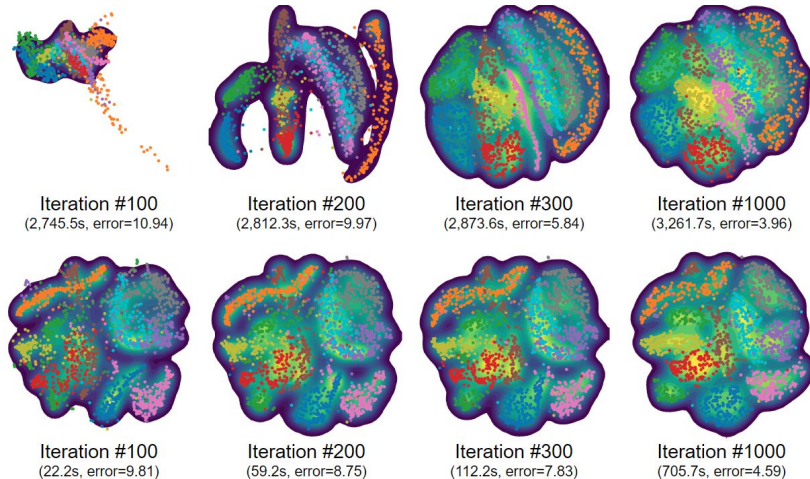
Research Program

Aviz focuses on four research themes:

- 1 Progressive Data Analysis
- 2 Physicality in Input and Output
- 3 Perception, Cognition, and Decision Making
- 4 Methodologies for Visualization Research

Progressive Data Analysis

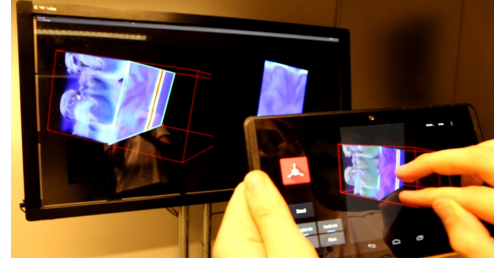
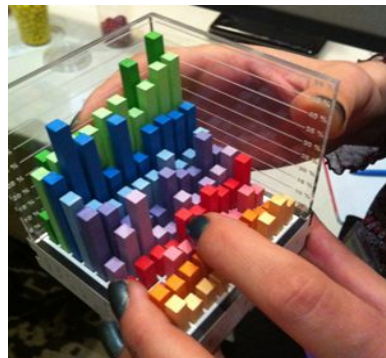
- Progressive data analysis
 - Language for progressive computation
 - Visualizations, UI and requirements for progressive data analysis
 - Progressive algorithms e.g. kNN, projections, k-means, ML algorithms
- Scalable visualizations
 - Large multidimensional projections
 - Exploring BitCoin/Blockchain
 - Health data from CNAM



2

Physicality in Input and Output

- A. Data Physicalization
- B. Situated and Personal Analytics
- C. Visualization using Augmented Reality Devices



3 Perception, Cognition, and Decision Making

A. Perceiving Statistical Information from Visualizations

- Don't inspect each visualized item individually
- How can people better understand “the gist”?

B. Decision making

- Decision making and cognitive biases are important research topics in psychology, economics and marketing
- How can visualizations support decision making?
- Explainable machine learning (organized workshops)

C. Illustrative Visualization

- Takes inspiration from illustrators' experience
- How can we support these styles with visualization?

4

Methodologies for Visualization Research

A. Promoting and Following Open Research Practices

- Aviz researchers engaged in open research practices
- Active at teaching and steering our field towards more transparent practices
- Collecting and reporting information on prominence of open research practices

B. The Communication of Statistics Results

- Aviz researching inaccurate statistical results in existing visualization literature
- Many problems related to inappropriate reporting of statistical results in science, e.g., “dichotomous inference”
- Aviz will conduct research on how to design statistical charts expressing nuances and uncertainty

C. Shaping the Visualization Community

- Aviz researchers are heavily involved in the organization structure of visualization conferences and journals
- Helping newly created committees to restructure the community based on data.

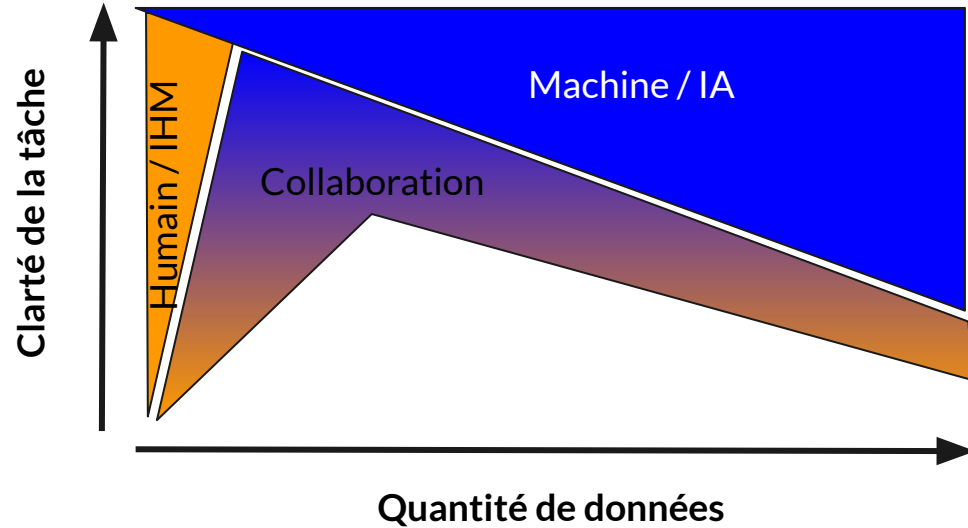


Intersection entre IA et Visualisation

- Les données
 - Données multidimensionnelles
 - Documents
 - Réseaux
 - ...
- Les tâches
 - Chercher les plus similaires (Google)
 - Classifier (Reconnaissance d'écriture ou audio)
 - Prédire (conduite automatique)
 - Trouver du boulot ?
 - ...

IA ou IHM / Visualisation ?

- Quand les données sont trop volumineuses
 - Machine / IA
- Quand les données sont trop bruitées
 - Humain / IHM
- Quand les tâches sont mal définies
 - Humain / IHM
- Volume + tâches mal définies ?
- Volume + bruitées ?

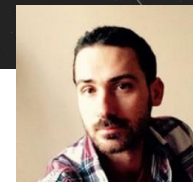
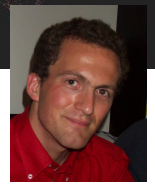


Search

Search Filters

Cartolabe : Cartographie de grand corpus

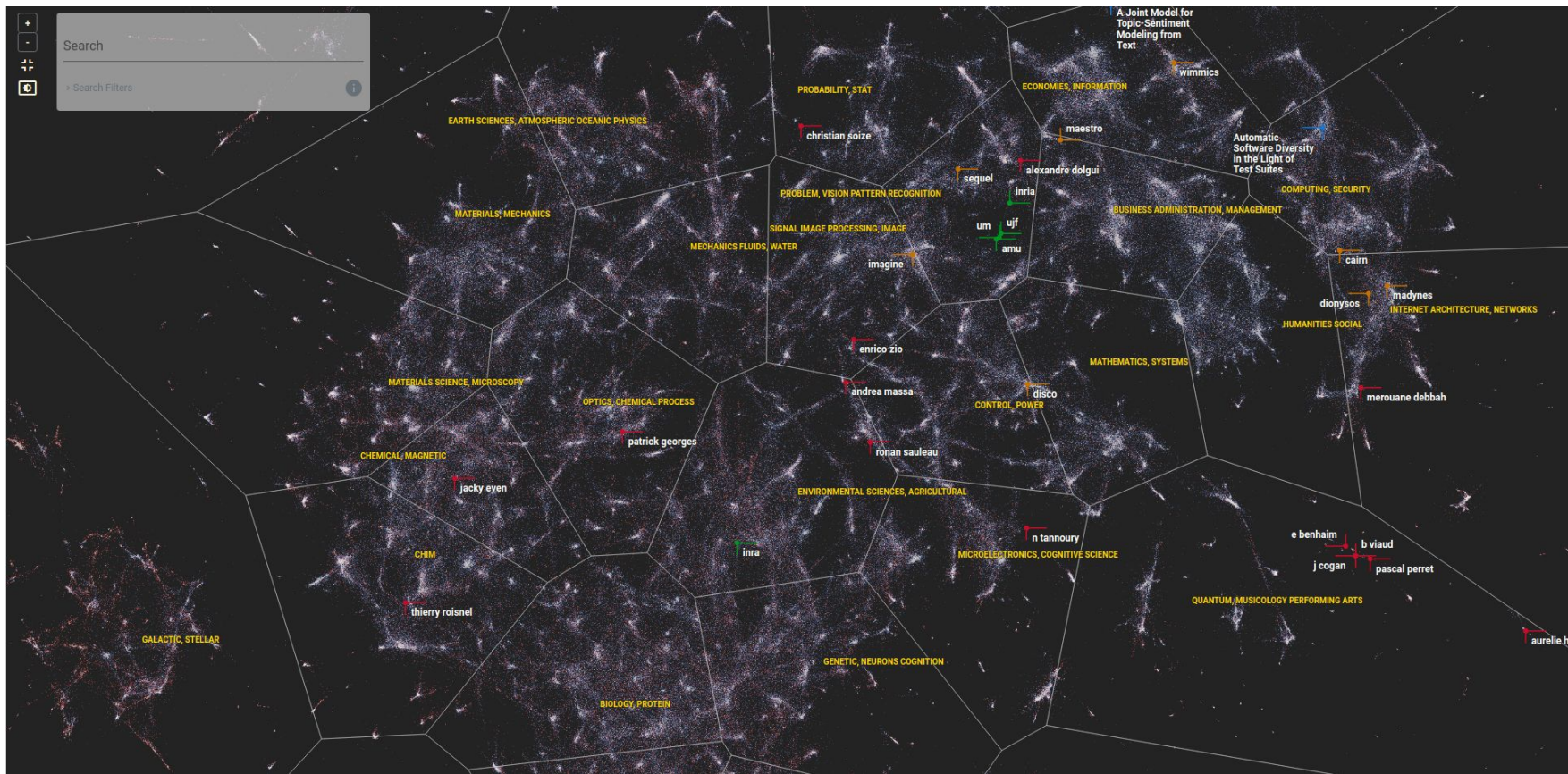
Philippe Caillou Jean-Daniel Fekete Michèle Sebag Jonas Renault Anne-Catherine Letournel
CNRS – INRIA – LRI, Université Paris-Sud



Cartographier un corpus ?

- Les moteurs de recherche permettent de «chercher » !
- Les cartes permettent de représenter l'ensemble d'un corpus
 - Un document doit être placé près de documents similaires,
 - et moins près de documents moins similaires
 - Un auteur est défini comme la concaténation de ses documents
 - Une organisation est définie comme la concaténation de documents de ses membres

Visualisation de millions de documents



Cartographie de documents?

Comment mesurer la similarité de documents ?

- Problème de Traitement Automatique du Langage (TAL)
 - Un document devient un vecteur de grande dimension
- Regarder les mots : le document comme un sac de mots
- Regarder les n-grammes : sac de n-grammes
 - Occurrences de lettres qui se suivent (2: bigrammes, 3: trigrammes)
 - Occurrences de mots
- Chercher des concepts : co-occurrences fréquentes de mots
- Calculer des vecteurs de documents
 - apprentissage profond, cool, c'est à la mode !
- On mesure une distance entre les vecteurs !

Tâches réalisables avec une carte

- Avoir une vue d'ensemble
 - Découvrir un voisinage
 - Balayer avec un pointeur
 - Naviguer
 - Découvrir un chemin
 - Découvrir un cluster
 - Découvrir une exception
- ~~• Trouver une exception de classe~~
 - ~~• Évaluer la pureté d'une classe~~
 - ~~• Évaluer la compacité d'une classe~~
 - ~~• Mettre en correspondance classe et cluster~~

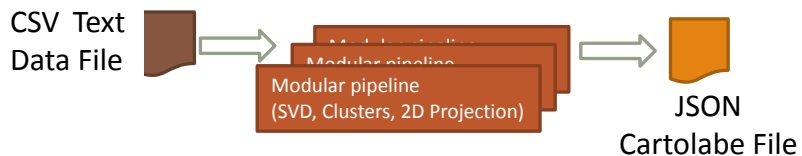
Cartolabe: 2 modules, 4 étapes

Cartolabe-data

Data Extraction

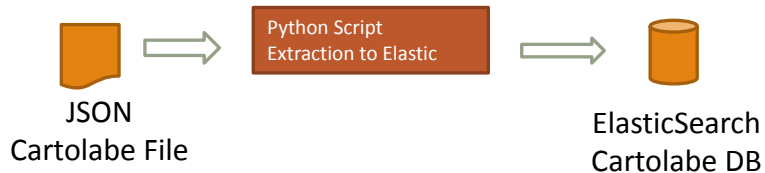


Data Processing

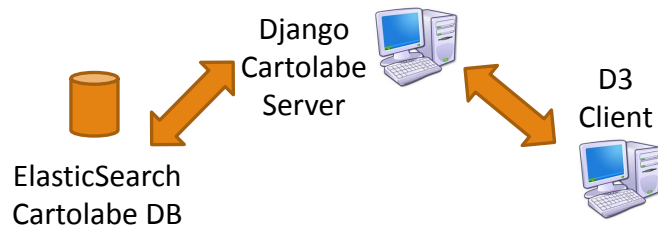


Cartolabe-vis


Data Integration



Data Visualisation



Le grand **débat national**



À l'initiative du Président de la République, le Gouvernement engage un Grand Débat National permettant à toutes et tous de débattre de questions essentielles pour les Français.

Les quatre thèmes du Grand Débat National

Le Gouvernement propose quatre thèmes de débat.



La transition écologique



La fiscalité et les dépenses
publiques



La démocratie et la
citoyenneté



L'organisation de l'État et
des services publics

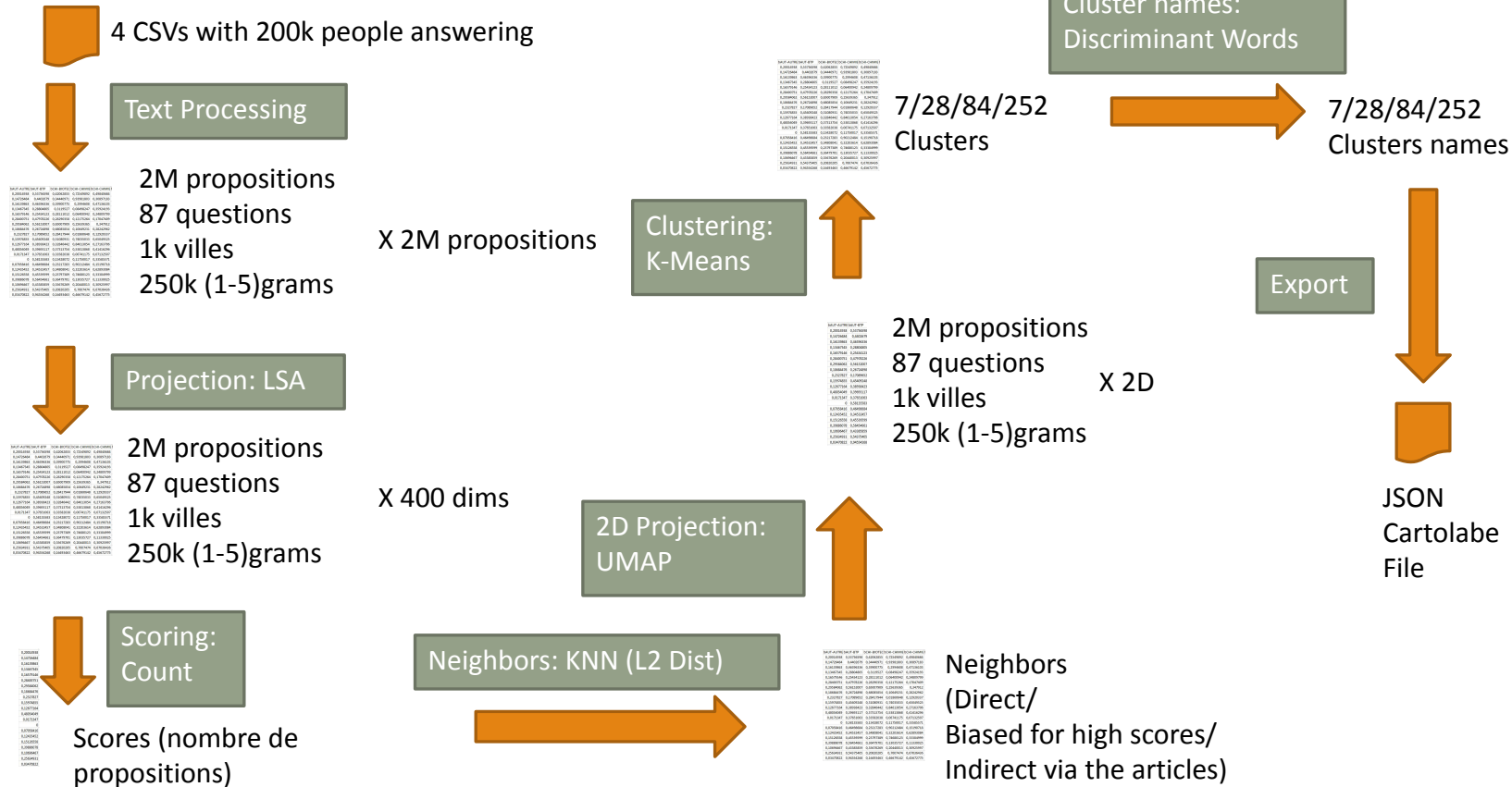
87 questions réparties en 4 thèmes

Au 2 mars:

200k contributeurs

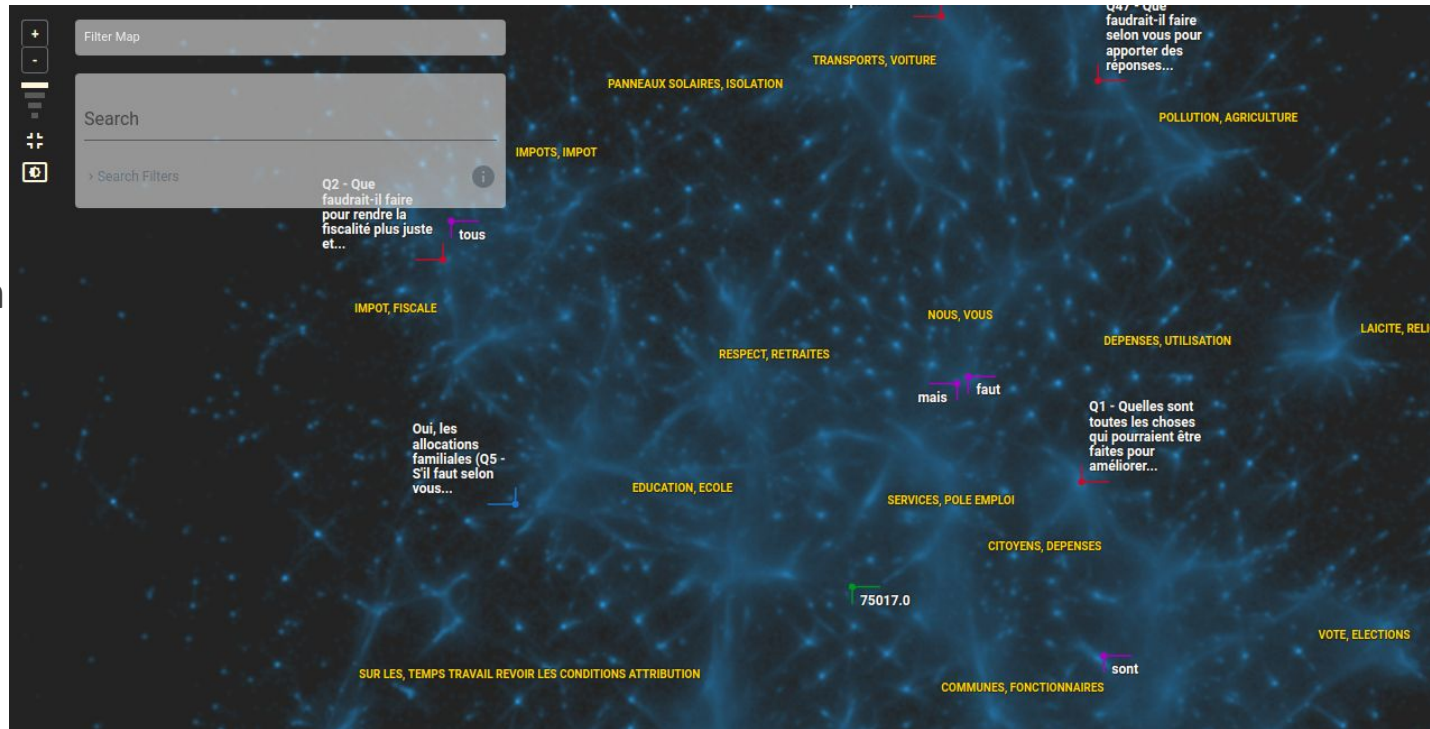
2M de réponses avec au moins 3 mots

Grand Débat Pipeline



La carte peut s'utiliser de la même façon que pour les autres applications...

<https://cartolabe.fr/map/debat>



Exploration Top-Down

Recherche autour d'un mot-clé/ville...

The image shows a search interface with a map and various annotations. The search results are displayed in a list, and the map shows the location of the search results. The annotations are in French and provide context for the search results.

Search Filters:

- allocation familiale
- Word
- Ne pas être pour de sanctionner...
- 2119

Search Results:

- Allocation familiale. APL. Etc... (Q5 - S'il faut selon vous revoir les conditions d'attribution de certaines aides sociales lesquelles doivent être concernées ?)
- Allocation familiale. APL. Etc... (Q5 - S'il faut selon vous revoir les conditions d'attribution de...

Annotations:

- Retrait des allocations familiales, retrait de certains droits sociaux...
- assumer
- Plus de contraventions. Suppression des allocations familiales....
- Chacun doit œuvrer dans le même sens. - Mobilisation citoyenne...
- Engagement à respecter les valeurs de la république (coupure...)
- réduire les allocations familiales si incivilités commises...
- REAGIR vite punir en conséquence, remboursement des dommages causés...
- Faire payer, du moins en partie, les dégâts provoqués par les jeunes...
- Punir les auteurs et responsabiliser les parents si l'auteur est mineur...
- Tolérance zéro pour tous. Faire payer les dégradations du patrimoine...
- les réprimer sérieusement... cela commence à l'école... incivilité = avertissement...
- Pas de respect des règles = suppression des allocations familiales,...
- Supprimer illico les allocations familiales attribuées aux parents...
- instaurer des amendes plus sévères - en cas de récidive, supprimer...
- Couper les allocations familiales à la famille si l'agresseur...
- peines plus lourdes et respectées (amendes+ lourdes...
- des familles
- Allocations familiales lorsqu'un enfant commet des incivilités...
- allocation familiale responsabiliser les parents des enfants délinquants en supprimant...
- Punir plus sévèrement. Même la police en a assez, ils sont relâchés...
- Diminution (mesure déjà en vigueur) et application, des allocations...
- Allocations familiales à réduire progressivement si...
- Allocations familiales : les parents qui baissent les bras !!!!!
- Autorité à l'école, responsabiliser les parents , suppression des allocations...
- Éduquer et sanctionner financièrement les délinquants et leurs...
- L'indemnisation chômage. Les allocations familiales. (Q5 - S'il...
- L'éducation à l'école et surtout implication des parents avec des sanctions...
- Couper les allocations familiales si incivilités faites par...
- Frapper au...

Les attentes pour le grand débat: exhaustif et non biaisé?

- Toutes les propositions sont présentes sans filtres
- Tous les mots sont utilisés, seuls quelques stopwords sont retirés (est, ai, et, ...)
- Les thèmes (clusters) sont générés automatiquement
- *Mais toute représentation est "biaisée" par ses choix de procédure et d'hyperparamètres*
- Seules des procédures sociales peuvent y remédier
 - Vues multiples, vérifications contre les biais, etc.

Répression contre
ceux et celles qui
veulent Manger
des frites au
Cheddar....

Celles et ceux qui
n'aiment pas les
frites au Cheddar
(Q115 - Si oui
lesquels...

cheddar

Non, plutôt
manger des frites
au Cheddar (Q142
- En matière
d'immigration
ils peuvent et
doivent faire les
pouvoirs...

Manger des frites
au Cheddar (Q128
- Comment
garantir le respect
par...

Proposer des
frites au Cheddar
pour tous et
toutes (Q123 -
Quel...

Manger des frites
au Cheddar (Q122
- Que faudrait-il
faire pour
consulter...

Manger des frites
au Cheddar (Q132
- Que faudrait-il
faire pour
valoriser...