

Fermer

## THE CONVERSATION

L'expertise universitaire, l'exigence journalistique



Photographie du Grand débat national organisé à Reims. G.Garitan / Wikipedia, CC BY-SA

# IA et démocratie participative : comment les réponses au grand débat national ont-elles été analysées ?

31 mars 2021, 00:45 CEST

*Cet article est publié en collaboration avec Binaire.*

### Auteurs



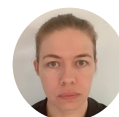
#### Aurélien Bellet

Chargé de recherche en machine learning,  
Inria



#### Mikaela Keller

Enseignante-chercheuse en informatique,  
Université de Lille



#### Nathalie Vauquier

Ingénieur Expérimentation et

Développement, Inria



**Pascal Denis**

Research Scientist, Inria



**Rémi Gilleron**

professeur émérite informatique, Inria

---

Le développement de la démocratie participative a fait émerger de nouvelles formes de consultations avec un grand nombre de données à analyser. Les réponses sont complexes puisque chacun s'exprime sans contrainte de style ou de format. Des méthodes d'intelligence artificielle ont donc été utilisées pour analyser ces réponses mais le résultat est-il vraiment fiable ? Nous nous sommes penchés sur l'analyse des réponses au grand débat national.

Dans le cadre d'un développement de la démocratie participative, différentes initiatives ont vu le jour en France en 2019 et 2020 comme le grand débat national et la convention citoyenne sur le climat. Toute consultation peut comporter des biais : ceux concernant l'énoncé des questions ou la représentativité de la population répondante sont bien connus. Mais il peut également exister des biais dans l'analyse des réponses, notamment quand celle-ci est effectuée de manière automatique.

Nous prenons ici comme cas d'étude la consultation participative par Internet du grand débat national, qui a engendré un grand nombre de réponses textuelles dont l'analyse officielle commandée par le gouvernement a été réalisée par des méthodes d'intelligence artificielle. Par une rétro-analyse de cette synthèse, nous montrons que l'intelligence artificielle est une source supplémentaire de biais dans l'analyse d'une enquête. Nous mettons en évidence l'absence de transparence sur la méthode utilisée pour produire l'analyse officielle et soulevons plusieurs questionnements sur la synthèse, notamment quant au grand nombre de réponses exclues de celle-ci ainsi qu'au choix des catégories utilisées pour regrouper les réponses. Enfin, nous suggérons des améliorations pour que l'intelligence artificielle puisse être utilisée avec confiance dans le contexte sensible de la démocratie participative.

## **Le matériau à analyser**

Nous considérons le traitement des 78 questions ouvertes du grand débat national dont voici deux exemples :

- « Que faudrait-il faire pour mieux représenter les différentes sensibilités politiques ? » du thème « La démocratie et la citoyenneté »

- « Et qui doit selon vous se charger de vous proposer ce type de solutions alternatives ? » dans le cadre des propositions de solutions de mobilité alternative du thème « La transition écologique ».

Les réponses aux questions sont des textes rédigés par les participants qui vont de quelques mots à plusieurs centaines de mots avec une longueur moyenne de 17 mots. Pour chaque question, on dispose de quelques dizaines de milliers de réponses textuelles à analyser. Le traitement d'une telle quantité de données est difficile pour des humains, d'où la nécessité de l'automatiser au moins partiellement. Lorsque les questions sont fermées (avec un nombre prédéfini de réponses), il suffit de faire des analyses quantitatives sous forme de comptes, moyennes, histogrammes et graphiques. Pour des questions ouvertes, il faut se tourner vers des méthodes d'intelligence artificielle.

### **Que veut-dire analyser des réponses textuelles ?**

Il n'est pas facile de répondre à cette interrogation car, les questions étant ouvertes, les répondants peuvent laisser libre cours à leurs émotions, idées et propositions. On peut ainsi imaginer détecter les émotions dans les réponses (par exemple la colère dans une réponse comme « C'est de la foutaise, toutes les questions sont orientées !!! On est pas là pour répondre à un QCM ! »), ou encore chercher des idées émergentes (comme l'utilisation de l'hydrogène comme énergie alternative). L'axe d'analyse retenu dans la synthèse officielle, plus proche de l'analyse des questions fermées, consiste à grouper les réponses dans des catégories et à compter les effectifs. Il peut être formulé comme suit, pour chaque question ouverte et les réponses textuelles associées :

- Déterminer des catégories et sous-catégories sémantiquement pertinentes ;
- Affecter les réponses à ces catégories et sous-catégories ;
- Calculer les pourcentages de répartition.

L'étude officielle, réalisée par Opinion Way (l'analyse des questions ouvertes étant déléguée à l'entreprise QWAM) est disponible sur le site du grand débat. Pour chacune des questions ouvertes, elle fournit des catégories et sous-catégories définies par un intitulé textuel et des taux de répartition des réponses dans ces catégories.

Par exemple, pour la question « Et qui doit selon vous se charger de vous proposer ce type de solutions alternatives ? », l'analyse a catégorisé les réponses de la façon suivante :

- Les acteurs publics 43,4 %
- Les solutions envisagées 8,8 %
- Les acteurs privés 6,6 %
- Autres contributions trop peu citées ou inclassables 22,5 %
- Non réponse (les réponses vides) 30,2 %

On constate que les catégories se chevauchent, que la catégorie « Les solutions envisagées » ne correspond pas à une réponse à la question et que le nombre d'inclassables est élevé (22,5 % soit environ 35 000 réponses non prises en compte).

## **L'analyse officielle : la méthode**

Regrouper des données dans des catégories correspond à une tâche appelée classification non supervisée ou clustering. C'est une tâche difficile car on ne connaît pas les catégories a priori, ni leur nombre, les catégories peuvent se chevaucher. De surcroît, les textes en langage naturel sont des données complexes. De nombreuses méthodes d'intelligence artificielle peuvent être utilisées comme, par exemple, la LDA pour « latent dirichlet analysis » et ses nombreux dérivés.

Quelle est la méthode utilisée par l'entreprise QWAM ? À notre connaissance, les seules informations disponibles se trouvent dans la présentation de la méthodologie. On y décrit l'utilisation de méthodes internes qui sont « des algorithmes puissants d'analyse automatique des données textuelles en masse (big data), faisant appel aux technologies du traitement automatique du langage naturel couplées à des techniques d'intelligence artificielle (apprentissage profond/*deep learning*) » et le post-traitement par des humains : « une intervention humaine systématique de la part des équipes qualifiées de QWAM et d'Opinion Way pour contrôler la cohérence des résultats et s'assurer de la pertinence des données produites ».

## **Regard critique sur l'analyse officielle**

Nous faisons les constats suivants :

- Les codes des algorithmes ne sont pas fournis et ne sont pas ouverts ;
- La méthode de choix des catégories, des sous-catégories, de leur nombre et des intitulés textuels associés n'est pas spécifiée ;
- Les affectations des réponses aux catégories ne sont pas fournies ;
- Malgré l'intervention humaine avérée, aucune mesure d'évaluation des catégories par des humains n'est fournie.

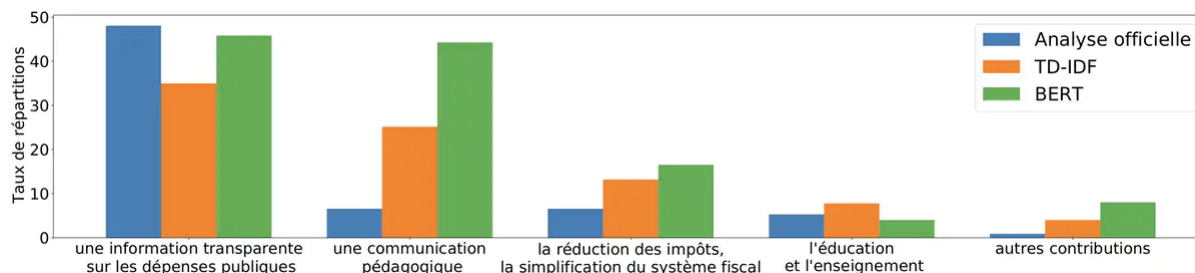
Nous n'avons pas pu retrouver les résultats de l'analyse officielle malgré l'usage de plusieurs méthodes. Dans la suite, nous allons voir s'il est possible de les valider autrement.

## **Une rétro-analyse de la synthèse officielle**

Notre rétro-analyse consiste à tenter de ré-affecter les contributions aux catégories et sous-catégories de l'analyse officielle à partir de leur contenu textuel. Notre approche consiste à affecter une contribution à une (sous-)catégorie si le texte de la réponse et l'intitulé de la catégorie sont suffisamment proches sémantiquement. Cette proximité sémantique est mesurée à partir de représentations du texte sous forme de vecteurs de nombre, qui constituent l'état de l'art en

traitement du langage.

Nous avons testé plusieurs méthodes de représentation des textes et plusieurs manières de calculer la proximité sémantique entre les réponses et les catégories. Nous avons obtenu des taux de répartitions différents selon ces choix, sans jamais retrouver (même approximativement) les taux donnés dans l'analyse officielle. Par exemple, la figure ci-dessous donne les taux de répartitions des réponses dans les catégories obtenus avec différentes approches pour la question « Quelles sont toutes les choses qui pourraient être faites pour améliorer l'information des citoyens sur l'utilisation des impôts ? ».



Quelles sont toutes les choses qui pourraient être faites pour améliorer l'information des citoyens sur l'utilisation des impôts ? Author provided

Pour compléter notre rétro-analyse automatique, nous avons mis en œuvre une annotation manuelle sur la question « Et qui doit selon vous se charger de vous proposer ce type de solutions alternatives ? » du thème « Transition écologique » et la catégorie « Les acteurs publics » et avons trouvé un taux de 54,5 % à comparer avec un taux de 43,4 % pour l'analyse officielle, soit une différence de 15 000 réponses ! Les réponses à cette question sont globalement difficiles à analyser, car souvent longues et argumentées (25 000 réponses contenant plus de 20 mots).

Notre étude manuelle des réponses nous a fait remarquer certaines réponses comme « moi-même », « les citoyens », « c'est mon problème », « les Français sont assez intelligents pour les trouver seuls » ou encore « les citoyens sont les premiers maîtres de leur choix ». Pour ces réponses, nous avons considéré une catégorie prise en charge par l'individu qui n'est pas présente dans la synthèse officielle bien qu'ayant une sémantique forte pour la question. Un classement manuel des réponses donne un taux de 4,5 % des réponses pour cette catégorie, soit environ 7 000 réponses, taux supérieur à certaines catégories existantes. Ceci met en évidence un certain arbitraire et des biais dans le choix des catégories de la synthèse officielle.

En résumé, notre rétro-analyse de la synthèse officielle montre :

- L'impossibilité de retrouver les résultats de la synthèse officielle
- La différence de résultats selon les approches
- Des biais dans le choix des catégories et sous-catégories

La synthèse officielle n'est donc qu'une interprétation possible des contributions.

## Recommandations pour utiliser l'IA dans la démocratie participative

L'avenir des consultations participatives ouvertes dépend en premier lieu de leur prise en compte politique, mais il repose également sur des analyses transparentes, dignes de confiance et compréhensibles par le citoyen. Nous proposons plusieurs pistes en ce sens :

**Transparence des analyses** : les méthodes utilisées doivent être clairement décrites, avec, si possible, une ouverture du code. La chaîne de traitement dans son ensemble (comprenant le traitement humain) doit également être précisément définie. Enfin, il est nécessaire de publier les résultats obtenus à une granularité suffisamment fine pour permettre une validation indépendante (par des citoyens, des associations ou encore des chercheurs).

**Considérer différents axes d'analyse et confronter différentes méthodes** : la recherche de catégories aurait pu être complétée par la recherche de propositions émergentes ou l'analyse de sentiments et d'émotions. Par ailleurs, pour un axe d'analyse donné, il existe différentes méthodes reposant sur des hypothèses et biais spécifiques et la confrontation de plusieurs analyses est utile pour nuancer certaines conclusions et ainsi mener à une synthèse finale plus fiable.

**Concevoir des consultations plus collaboratives et interactives** : publier les affectations des réponses aux catégories permettrait à tout participant de voir comment ses contributions ont été classées. Il serait alors possible de lui demander de valider ou non ce classement et d'ainsi obtenir une supervision humaine partielle utilisable pour améliorer l'analyse. D'autres manières de solliciter cette supervision humaine peuvent être considérées, par exemple faire annoter des textes par des volontaires (voir l'initiative de la Grande Annotation) ou encore permettre aux participants de commenter ou de voter sur les contributions des autres.

Si l'intelligence artificielle permet désormais de considérer des enquêtes à grande échelle avec des questions ouvertes, elle est susceptible de biais comme toute méthode automatique. Il est donc nécessaire d'être transparent et de confronter les méthodes. Dans un contexte de démocratie participative, il est également indispensable de donner une véritable place aux citoyens dans le processus d'analyse pour engendrer la confiance et favoriser la participation.