Comment les **algorithmes de recommandation** apprennent et utilisent les préférences **politiques** des utilisateurs ?

Tim Faverjon
tim.faverjon@sciencespo.fr

SciencesPo
MÉDIALAB

LEARNING PLANET INSTITUTE

CIVICA
THE EUROPEAN UNIVERSITY OF SOCIAL SCIENCES

Digital Services Act

Demande aux plateformes

Chap. III, sec. 3, *art. 27*

1. Expliquer le fonctionnement des algorithmes de recommandations (paramètres importants, raison des classement…)

Chap. III, sec. 3, *art. 26*

2. Ne pas discriminer selon certaine cathégories de données personnelles (en particulier les opinions politiques)

Source : https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32022R2065#d1e40-1-1

Digital Services Act

Demande aux plateformes

Chap. III, sec. 3, *art. 27*

1. Expliquer le fonctionnement des algorithmes de recommandations (paramètres importants, raison des classement…)

Comment trouver les *paramètres importants* de l'algorithme ?

Chap. III, sec. 3, *art. 26*

2. Ne pas discriminer selon certaine cathégories de données personnelles (en particulier les opinions politiques)

Comment s'assurer les **opinions politiques** ne sont pas des *paramètres importants* ?

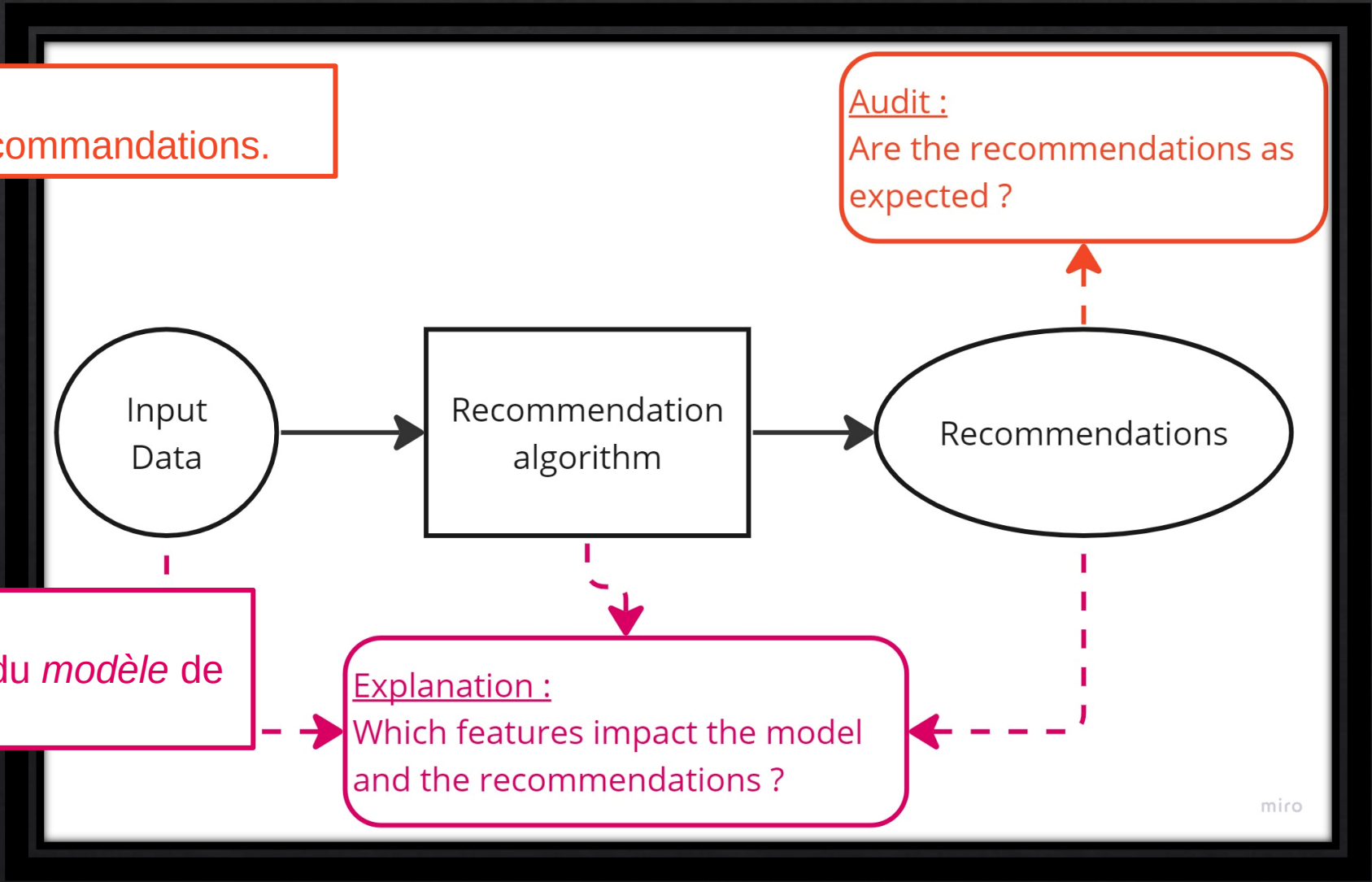
Source : https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32022R2065#d1e40-1-1

Comment comprendre les algorithmes de recommandations ?

Deux familles de méthodes

Audit :
On regarde les *résultats* des recommandations.

Explication :
On explique le fonctionnement du *modèle* de recommandation.



L’idée :

- Construire une méthode d’explication pour les algorithmes de recommandation

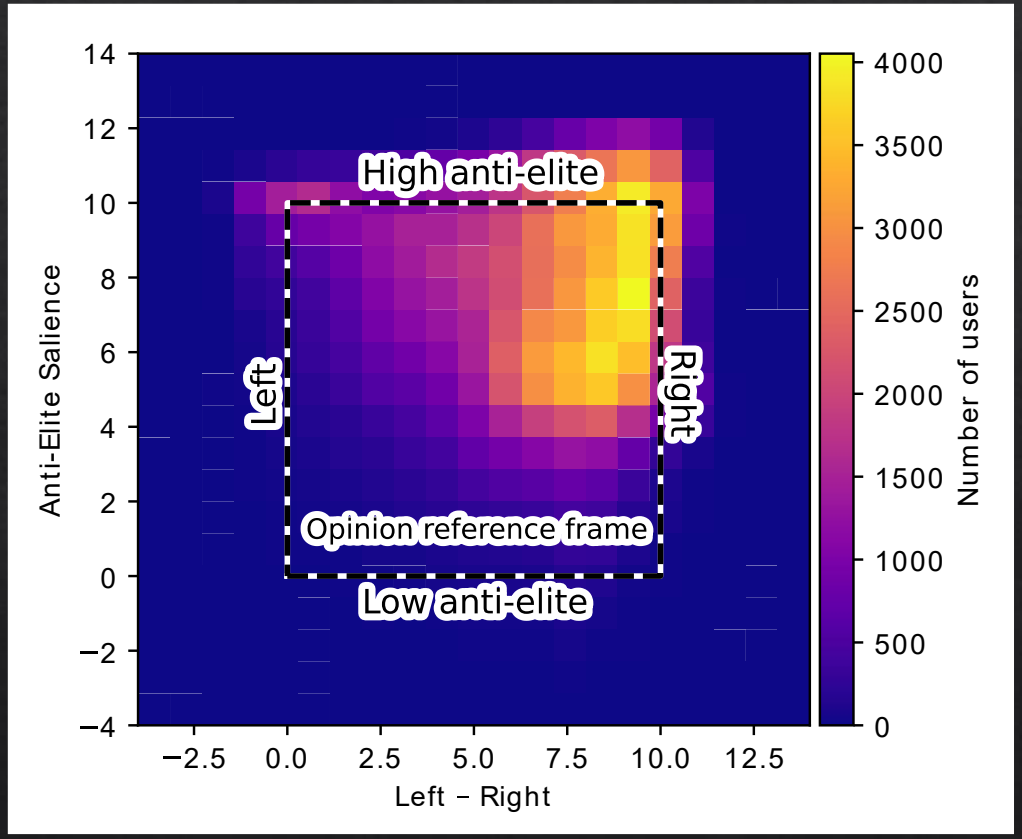
- Utilisant les opinions politiques des utilisateurs

Mesurer les attitudes politiques

- On construit un espace politique à partir du réseau d'utilisateurs suivant des parlementaires sur Twitter
- On donne un sens politique aux dimensions de cet espace en les reliant aux enquêtes d'experts (CHES)

Pour la **France** on obtient deux attitudes politiques qui expliquent le mieux l'espace politique :

1. **Left – Right** : droite et gauche classiques
2. **Anti-elite Salience** : défiance envers les institutions (0 = aucune défiance)



Source : Ramaciotti Morales, Pedro, et al. "Inferring attitudinal spaces in social networks." Social Network Analysis and Mining (2022)

L’idée :

- Construire une méthode d’explication pour les algorithmes de recommandation
- Utilisant les opinions politiques des utilisateurs

La méthode :

Une preuve de concept :

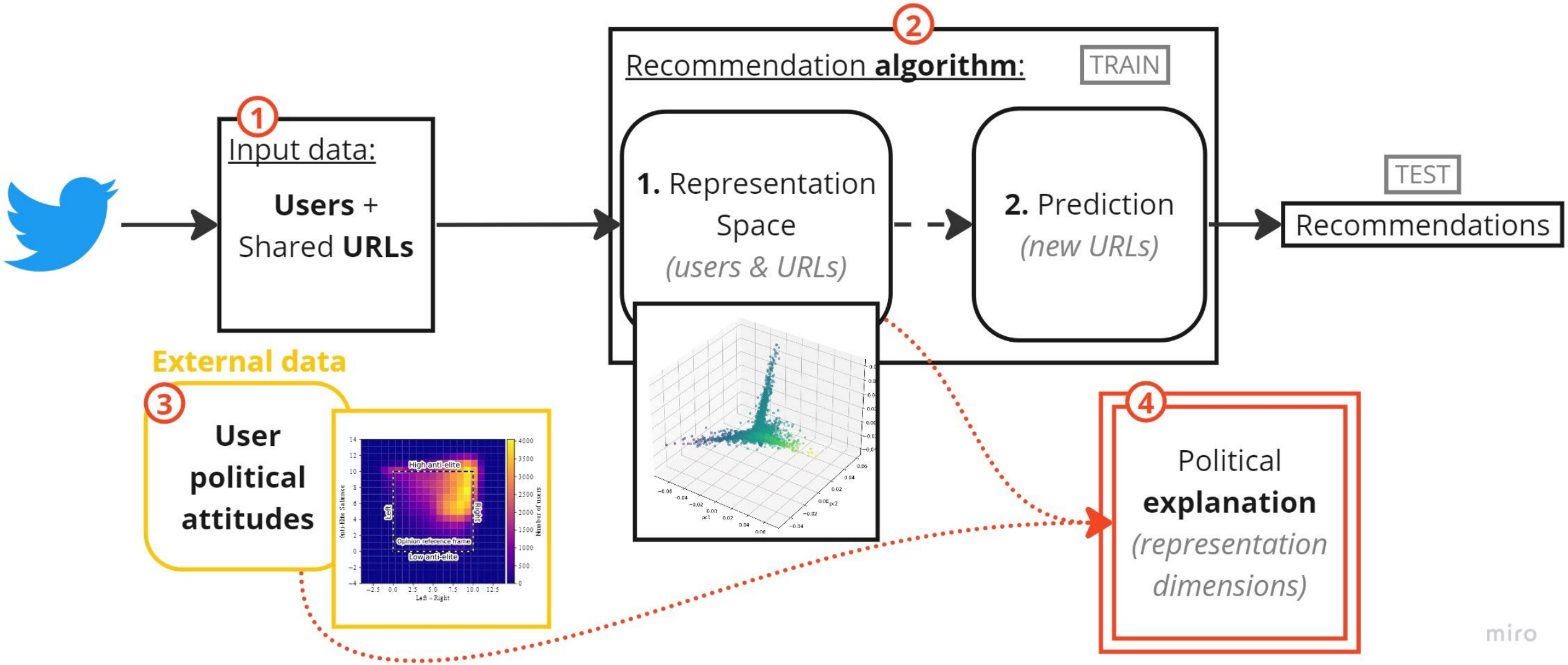
- Algorithme connu
- Données réelles (avec mesures politiques)

Le résultat :

Le modèle contient des information politiques sur les utilisateurs

Source : Faverjon, Tim, et al. "Discovering ideological structures in representation learning spaces in recommender systems on social media data." 2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE, 2023.

La méthode : Preuve de concept



Source : Faverjon, Tim, et al. "Discovering ideological structures in representation learning spaces in recommender systems on social media data." 2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE, 2023.

Résultats

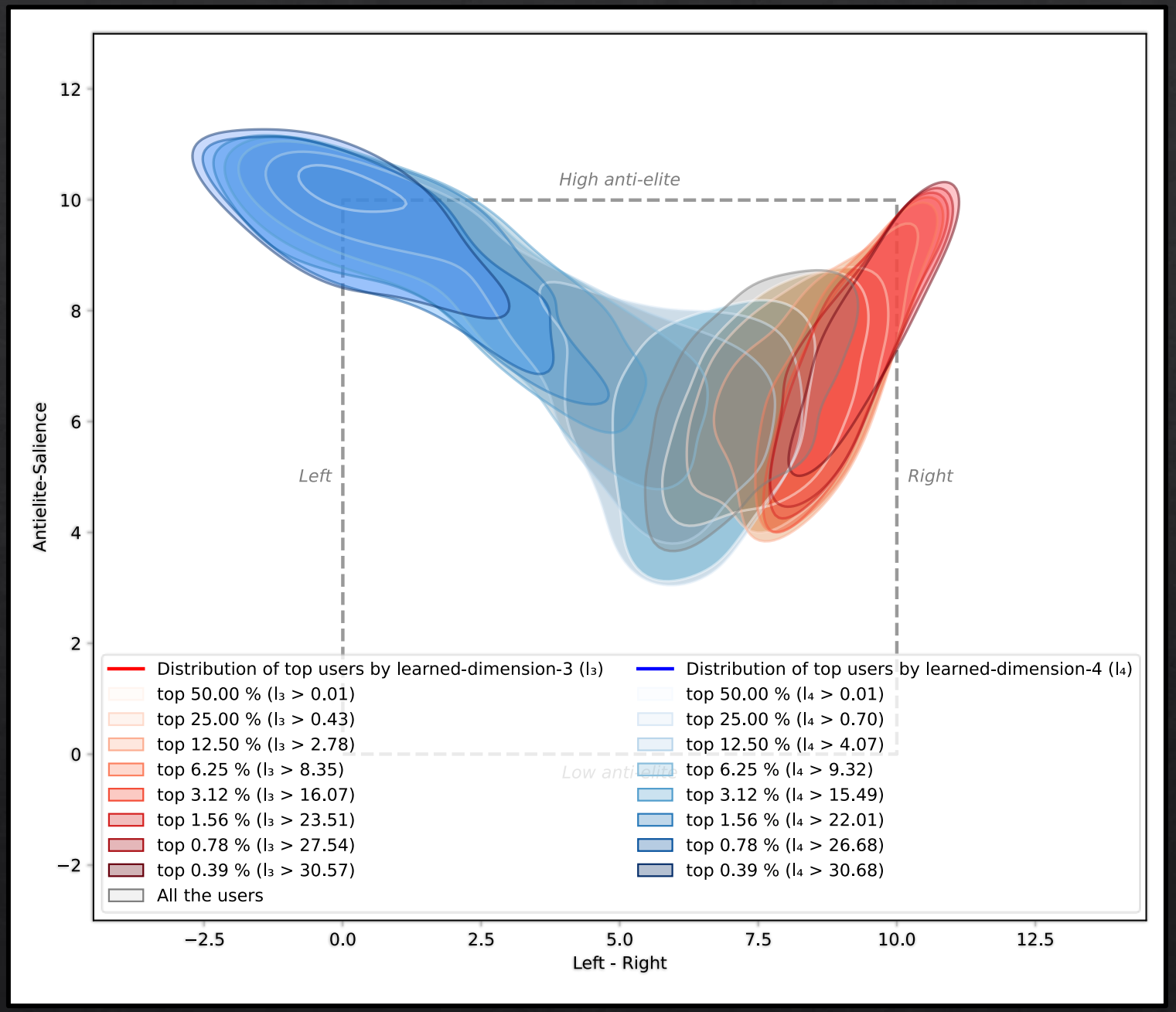
Certaines dimensions du modèle sont corrélées avec les attitudes politiques des utilisateurs.

Exemple :

Le modèle identifie des groupes d'utilisateurs (rouge et bleu)

qui sont en réalité des groupes politisés (position)

- gauche – anti-elite
- droite – anti-elite



Résultats

- Un algorithme de recommandation peut apprendre et utiliser les opinions politiques des utilisateurs **sans que cela soit explicitement demandé**
- Nous pouvons mesurer ce phénomène

Implication de notre méthode :

- Identifier les contournements du DSA (discrimination politique) directement dans les modèles
- Envisager la construction d'outils et de solutions techniques

Merci

Comment les **algorithmes de recommandation** apprennent et utilisent les préférences **politiques** des utilisateurs ?

Site web →

Tim Faverjon : tim.faverjon@sciencespo.fr