

**Exiger des IA qu'elles protègent  
l'information**

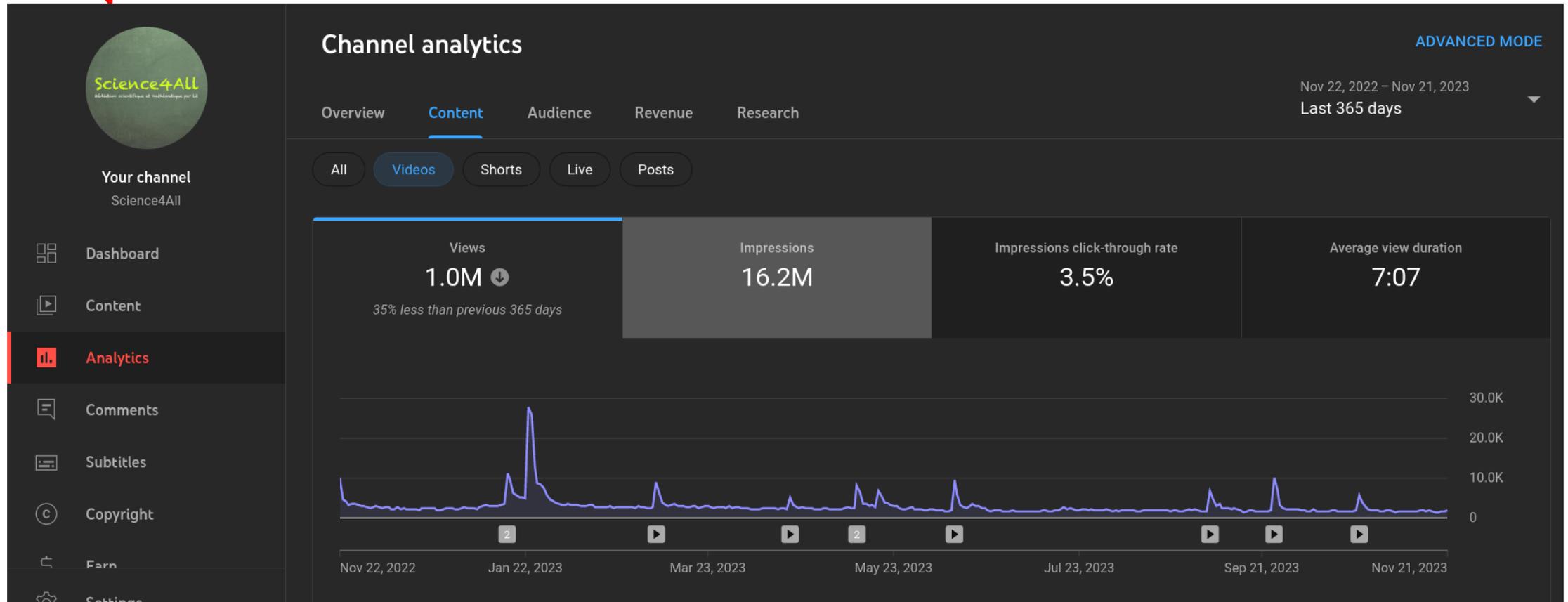
**Lê Nguyễn Hoang,  
PDG de Calicarpa et Président de Tournesol**

# Affronter le déluge



19 octobre 2023

Les IA de recommandation  
sont les arbitres de l'attention



## Présomption de non-recommandabilité

- Restreindre le nombre de publications d'un contenu.
- Autoriser davantage de publications si le contenu est jugé recommandable par des entités certifiées.

# Présomption de non-conformité

- Par défaut, une entreprise a l'interdiction déployer ses IA.
  - Surtout à grande échelle (cf. DSA).
  - Surtout pour des applications dangereuses (cf. AI Act).
  - Surtout pour des modèles avec un grand nombre de paramètres (cf. El Mhamdi).
- Le droit de déploiement s'obtient par preuve de conformité.
  - C'est la norme dans les industries matures (aviation, pharmaceutique, agro-alimentaire...).

## Valoriser les consensus

Suite à l'adoption de Pol.is à Taïwan, Audrey Tang :

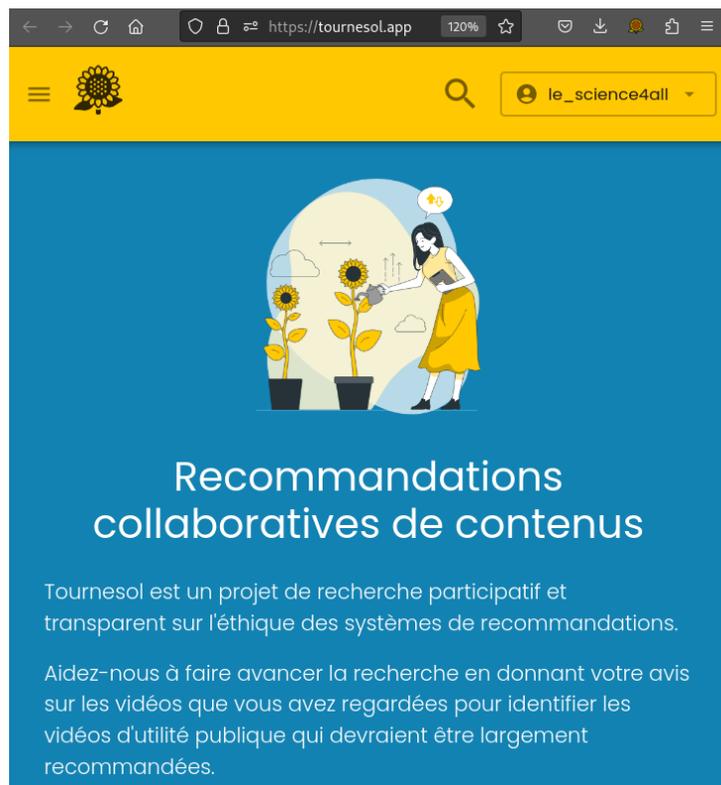
« Les gens rivalisent pour présenter les déclarations les plus nuancées susceptibles de convaincre le plus grand de personnes... »

## Investir dans une démocratie numérique



19 octobre 2023

# Investir dans la recherche en gouvernance sécurisée



Recommandations collaboratives de contenus

Tournesol est un projet de recherche participatif et transparent sur l'éthique des systèmes de recommandations.

Aidez-nous à faire avancer la recherche en donnant votre avis sur les vidéos que vous avez regardées pour identifier les vidéos d'utilité publique qui devraient être largement recommandées.

- Maîtriser l'impact maximal d'un citoyen.
- Démocratie liquide (avec institutions ?).
- Réseau et propagation de confiance.
- Zero-Knowledge Proof of Personhood.
- Généralisation sécurisée des préférences.
- Vote Bayésien sécurisé.
- Diversité et équité des recommandations.
- Apprentissage de la volition.
- ...

# Conclusions

- Présomption de non-recommandabilité.
- Présomption de non-conformité.
- Valoriser le consensus.
- Investir dans une démocratie numérique.