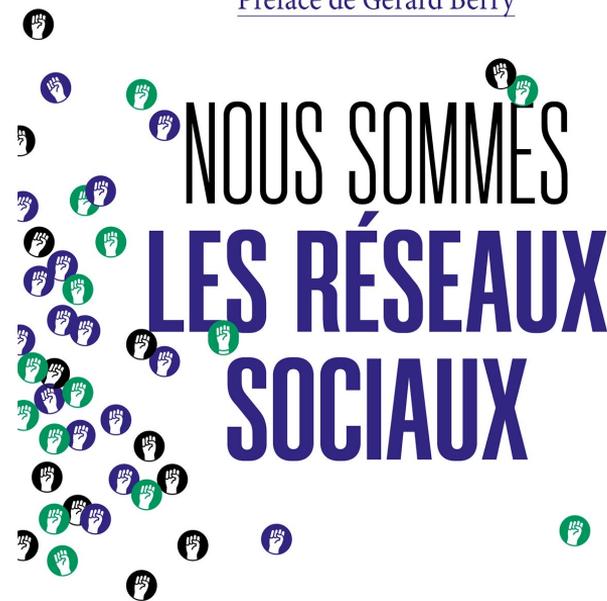


Réseaux sociaux

Serge Abiteboul
Arcep et Inria
Nantes, 2023

SERGE ABITEBOUL
JEAN CATTAN

Préface de Gérard Berry



Organisation

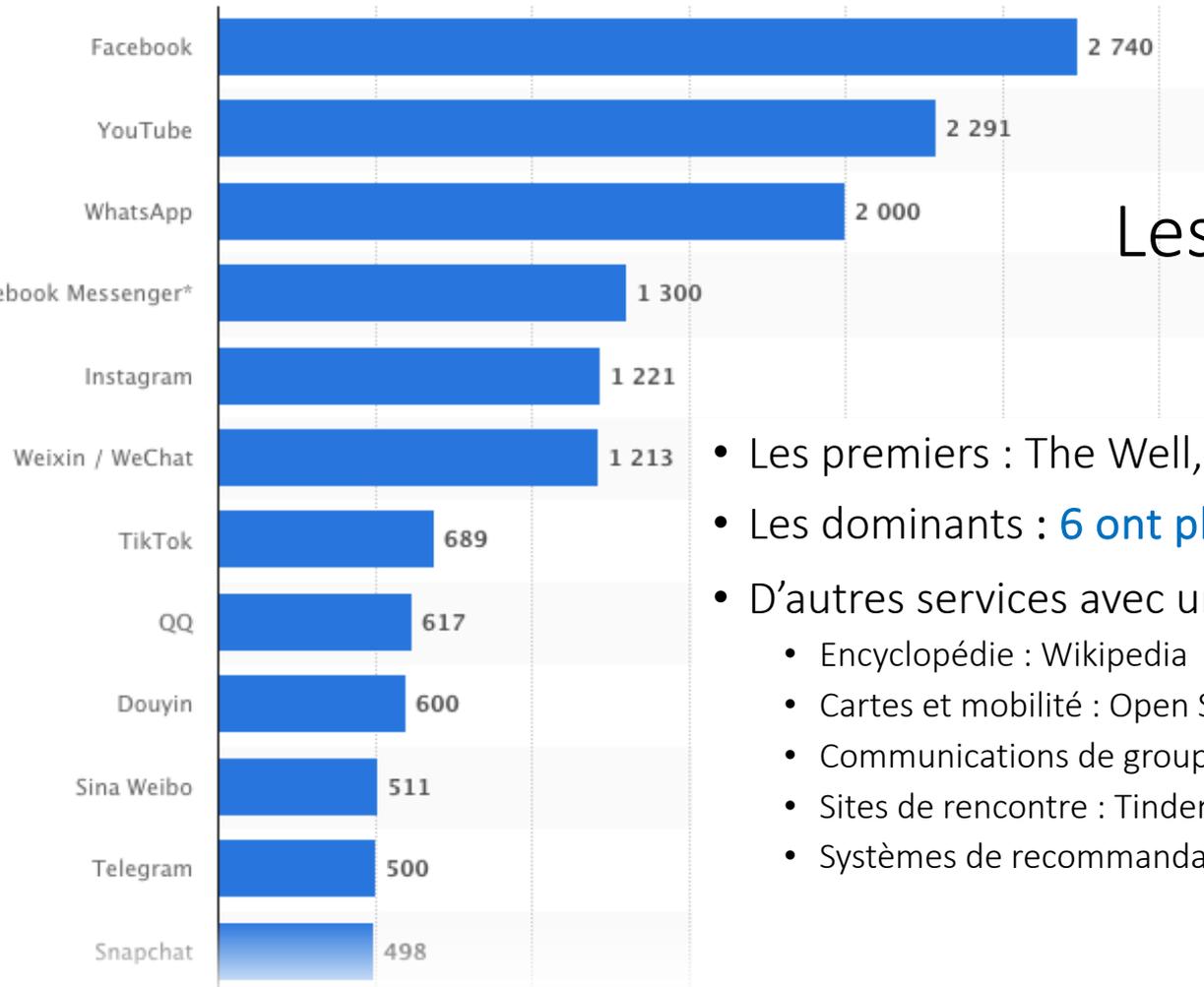
1. Les réseaux sociaux
2. Responsabilité, autorité et confiance
3. L'anatomie des réseaux sociaux
4. La régulation
5. Soft power
6. Conclusion

1. Les réseaux

Point de vue historique

- Un **réseau social** est « *un agencement de liens entre des individus et/ou des organisations, constituant un groupement qui a un sens : la famille, les collègues, un groupe d'amis, une communauté, etc.* » (Wikipédia)
- Un **réseau social (numérique)** est un service numérique offrant la possibilité à des personnes réelles de former un réseau
- Les réseaux sociaux numériques ne font que s'ajouter aux réseaux sociaux existants : la famille, les amis du collège/lycée, l'assocé de pétanque, la copro...
- **Petite et grande conversation, Dominique Cardon**
 - **Réseaux ou médias sociaux**

Les réseaux sociaux numériques



- Les premiers : The Well, forums...
- Les dominants : **6 ont plus d'un milliard d'utilisateurs**
- D'autres services avec une composante réseau social
 - Encyclopédie : Wikipedia
 - Cartes et mobilité : Open Street Map, Google Maps...
 - Communications de groupe : Whatsapp, Snap, Yubo...
 - Sites de rencontre : Tinder, Meetic
 - Systèmes de recommandation : Netflix, Amazon, Youtube...

L'amitié et l'amour

- Permettent des liens entre amis ou amoureux
 - Origine de FB : drague entre étudiants
- Longue tradition de **sites de rencontres**
 - Minitel rose
 - Reste un gros marché

La haine

- Présence massive de message de haine : racisme, sexisme, homophobie...
- **Harcèlement**
- Modifient notre conception de l'amitié
 - Privilégie le chiffre plutôt que la qualité
 - Pression d'échanges permanents

Information

- Les réseaux sociaux facilitent l'accès à l'information et au savoir
 - Tous types d'information, y compris les plus mineures
 - Tous types de savoir, y compris les plus spécialisés, voire exotique
- Rompent l'isolement face aux évènements
- Permet à chacun de participer à la diffusion d'information et de savoir

Désinformation, *fakenews*

- Un phénomène ancien exacerbé avec les réseaux sociaux
 - Les calomnies et les rumeurs
 - Exemple : les campagnes de dénigrement de Marie-Antoinette au XVIII
 - Préfigure la viralité sur les réseaux sociaux
- Ce qui n'a pas changé
 - Comme avant des mensonges et des manipulations (souvent volontaire)
 - Propagées de façon virale (souvent involontairement)
- **Ce qui est nouveau : une viralité bien plus forte**

Pourquoi ils méritent d'être sauvés

**BLACK
LIVES
MATTER**



une

- La « petite conversation »
- La voix de ceux qu'on n'entendait pas
 - Historiquement, seuls les riches et les puissants avaient parole publique
 - Avec les réseaux sociaux, une petite association de quartier peut toucher un grand public
 - Les printemps arabes, Occupy, les gilets jaunes...
 - Twitter-2017 #MeToo & #BalanceTonPorc
- La constitution de communautés d'intérêt
 - L'organisation à l'échelle des quartiers
 - L'économie du partage

2. Leur responsabilité

Doivent-ils contribuer à définir ou répandre le bien ou le mal ? En ont-ils la légitimité ? Comment exercent-ils leur autorité ? Doit-on attendre quelque chose d'eux ? Ont-ils notre confiance ?

Responsabilité : est-ce que les réseaux sociaux doivent participer à l'« amélioration » de la société ?

*With great power comes
great responsibility,
Uncle Ben, Spiderman*



- Rappel : ce sont des entreprises sensées faire du profit
- Au cœur du problème : *business model*
- Gratuit le plus souvent pour les internautes
- Ils vendent l'attention de leurs utilisateurs
 - Publicité
 - Leurs « vrais » clients ce sont les entreprises qui les paient

Une source du problème

la non-responsabilité « légale »

- Section 230 of Communications Decency Act aux US, 1996
 - *No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider*
- Article 12 of the e-Commerce Directive, 2000
 - Même idée
- Pas responsables : ils ne produisent pas les contenus
 - Par opposition avec les médias traditionnels
- Responsables : ils les « poussent », les « accélèrent »
 - Avec des algorithmes secrets, opaques
- Ça change en Europe avec le DSA et le DMA

Le patron de Twitter ne regrette pas d'avoir banni Trump mais reconnaît un précédent "dangereux"



Publié le : 14/01/2021 - 04:17



Le patron de Twitter Jack Dorsey, le 28 octobre 2020 à Washington. © Greg Nash, pool, AFP (archives)

Texte par : [FRANCE 24](#) 4 mn

Jack Dorsey est revenu, dans une série de tweets, sur la décision récente de sa plateforme de bannir le président américain Donald Trump. Il fait notamment part de ses réflexions sur le pouvoir qu'une entreprise comme Twitter détient sur le débat public.

Le peuple s'est exprimé. Trump va être rétabli, Elon Musk le nouveau patron de Twitter



- De quel droit Jack Dorsey coupe-t-il le sifflet d'un président des US ?
 - Ils régulent trop : ils tuent la liberté d'expression
- De quel droit Elon Musk lui rouvre-t-il le sifflet ?
 - Ils ne régulent pas assez : ils portent préjudice à la société

3. L'anatomie des réseaux sociaux

Ouvrons le capot



Interface

GUI
Identification
Personnalisation
PC et surtout
smartphones

Réseau

Communication
Chiffrement
Compression

Centre de
données

Bases de données

- Contenus, contacts, etc.

Recommandation de
contenus

Publicités personnalisées

- enchères

Modération

Algorithmes de recommandation

- **Formule très secrète**
- Un grand nombre de facteurs entrent en jeu
- Différents selon les réseaux sociaux
 - Détermine le « style du réseau »
- Changements dynamiques
- Cause importante de biais

How Does Facebook Choose What To Show In News Feed?

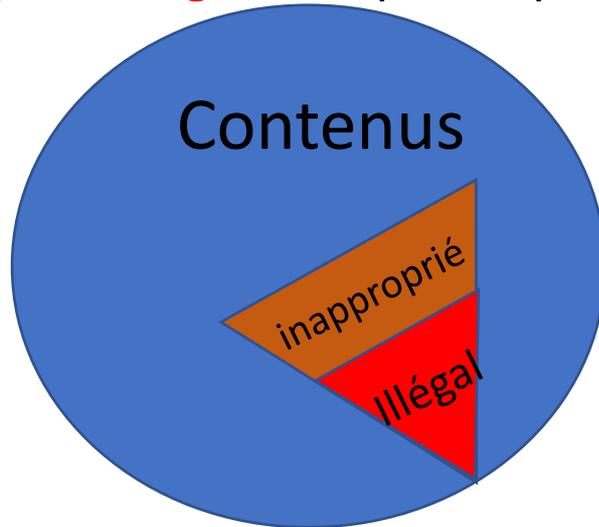
$$\text{News Feed Visibility} = * \underset{\text{Interest}}{I} \times \underset{\text{Post}}{P} \times \underset{\text{Creator}}{C} \times \underset{\text{Type}}{T} \times \underset{\text{Recency}}{R}$$

Interest Interest of the user in the creator	Post This post's performance amongst other users	Creator Performance of past posts by the content creator amongst other users	Type Type of post (status, photo, link) user prefers	Recency How new is the post
--	--	--	--	---------------------------------------

* This is a simplified equation. Facebook also looks at roughly 100,000 other high-personalized factors when determining what's shown.

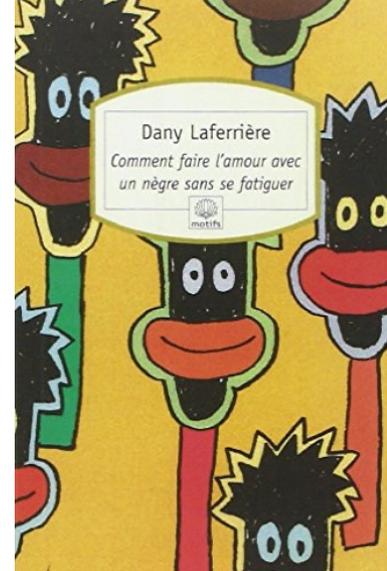
Modération de contenus : sujet central

- La modération de contenus est la fonction par laquelle une plateforme contrôle les contenus en décidant ce qui n'est pas acceptable parce que « **illégal** » ou parce que c'est considéré « **inapproprié** »

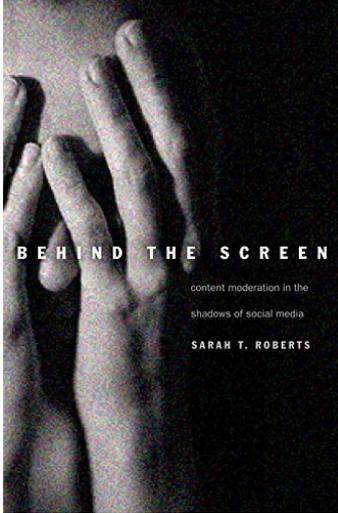


Modération de contenus : pas facile

- 3 V
 - Volume of contenus
 - Variété : fakenews, violence, message de haine, violation de copyright...
 - Vélocité : réaction immédiate attendue
 - Attaque terroriste à Christchurch, Nouvelle Zélande, en 2019 contre deux mosquées
- Langage humain : ambiguïté, jargon, d'argot, figures de style, humour...
- **Absence de contexte**
- Problèmes psychologiques sérieux des humains



Ce n'est pas facile pour des humains Faisons-le faire par des IA



NEWS 01/03/2019 7:25 PM IST

Beheadings, Suicide Attempts, Porn: Why Facebook Moderators In India Are Traumatized

Contract employees monitoring nudity and pornography on Facebook and Instagram view 2,000 posts in an eight-hour shift.

REUTERS

2000

The number of images per hour that content moderators at one Indian center are expected to sort through for Facebook, dating sites, and other clients

Revealed: catastrophic effects of working as a Facebook moderator

Exclusive: Job has left some 'addicted' to extreme material and pushed others to far right

The Guardian

THE TRAUMA FLOOR

The secret lives of Facebook moderators in America

By Casey Newton | @CaseyNewton | Feb 25, 2019

Serge Abiteboul **THE VERGE**

Modération algorithmique avec apprentissage automatique

- L'objectif : étant donné un contenu, on doit décider de le conserver ou non
- Sur quoi peut-on se baser pour prendre cette décision ?
 - Au niveau méta : Une ontologie et des règles pour définir ce qui est inapproprié ; on peut y voir une forme d'éthique appliqué
 - Ensuite, au niveau des données : on a un grand corpus de contenus avec des annotations : *fakenews*, discours de haine, discours violent - les humains ont utilisé les règles pour choisir les annotations.
- La solution : entraîner un algorithme d'apprentissage automatique pour qu'il imite les décisions prises dans le corpus. On peut faire plus avec une analyse plus détaillée basée sur les règles
- Les algos sont déjà meilleurs que les humains pour les textes de discours de haine (pas de preuve)

Où en sont ces IA ? (1)

- Les IA sont meilleures pour : volume et vélocité
 - Quasi seules pendant le confinement
- Elles ont du mal avec la complexité de la langue, des images, des humains
- Imiter les humains
 - Oui, mais énorme zone grise et
 - Les humains ne sont eux-mêmes pas très bons
- Difficulté : Énorme hétérogénéité entre les langues
 - Mis en évidence dans les *Facebook papers* de Frances Hagen en 2022
 - Manque de moyens
 - Modération quasi inexistante dans de nombreuses langues
- Question : absence de contexte ?

Où en sont ces IA ? (2)

- Manque de mesures précises
 - Opacité quasi-totale – le DSA devrait améliorer les choses
- Présence probable de biais - difficile à définir/détecter
- **Le corpus définit les objectifs, mais qui définit le corpus ?**
 - Qui définit le monde de demain ?
- Manque d'accès aux données pour les chercheurs et pour les entreprises moins dominantes
 - Besoins de partager une telle ressource
 - **Données d'intérêt général.** Ne peut pas être une donnée ouverte en raison de problèmes de confidentialité.

Peut-on imaginer d'autres réseaux sociaux ?

- Distribution et hétérogénéité
- Modération par les pairs
- Modèles d'affaires différents
- Comportements éthiques

- Ce sujet concerne la société, les gouvernements, et les développeurs de logiciels/l'industrie

L'interopérabilité

- Le DMA inclut la portabilité
 - Important pour donner plus de liberté à l'utilisateur
 - Comme cela se passe dans les télécoms
- Mais ne suffit pas : Effet de réseau immense dans les réseaux sociaux
 - Exemple encore des télécoms
- Interopérabilité des réseaux sociaux
 - Existe déjà avec le protocole Activity Pub, e.g., Mastodon (*)
 - Pas prévu dans le DMA (seulement des communications interpersonnelles)

(*) : je suis [sergeabiteboul](#) sur Twitter et [@sergeabiteboul@paille.fr](#)

Pour le monde interopérant, notamment toutes les instances de Mastodon

Comment amener les réseaux sociaux à avoir des comportements responsables dans le respect de la démocratie et de la liberté d'expression ?

4. La régulation

L'autorégulation par les médias sociaux

- Les réseaux sociaux **s'autorégulent** pour éviter de perdre des utilisateurs
 - Règles communautaires dépassent les lois nationales comme la nudité
 - Blocage de pages ou même de sites
 - Suppression de comptes d'utilisateurs ou de pages
- Pas facile et la zone grise est très large
- Pas une source de revenus → Peu d'investissements
- Quoi qu'ils fassent, ils sont critiqués
 - **Trop modéré : violation de la liberté d'expression**
 - **Pas assez modéré : ne protège pas les utilisateurs**
- Question clé : Quelle est leur légitimité pour faire de la modération ?
 - Pourquoi leur faire confiance si vous ne savez pas, si vous ne pouvez pas contrôler ce qu'ils font ?

L'exemple de Facebook

- Plus de 15 000 « modérateurs » humains qui suppriment les contenus
- Détection algorithmique des contenus non désirés
- Notification, mécanisme d'appel
- Très critiqué dans la presse
- Les autres grandes plateformes font souvent bien pire

Régulation directe par des États

- États démocratiques
 - Difficulté : masses de contenus et moyens limités des États
 - Difficulté : les réseaux sociaux sont souvent internationaux.
 - Netzwerkdurchsetzungsgesetz (surnom NetzDG) en Allemagne 2017
 - La proposition de loi Avia avec des idées similaires en France
- États moins démocratiques
 - Objectif : **contrôler l'opinion publique**
 - Développer des réseaux sociaux nationaux sous contrôle de l'État
 - Chine : Wechat, Weibo
 - Russie : Vkontakte, Odnoklassniki
 - Peut éventuellement conduire un jour à une partition de l'internet

Régulation supervisée par les États

Une expérience : la mission Facebook

- En 2019
- Suite à une rencontre entre Macron et Zuckerberg
- FB s'ouvre à un groupe d'une dizaine de fonctionnaires français
- Créer un cadre français pour responsabiliser les plateformes de médias sociaux : Agir en France avec une vision européenne, "Mission Facebook", 2019
- Propose une régulation ex ante au niveau européen

The biggest enemy of western people is not war or terrorism, it is their own governments lack of regulation of public health and safety. — Steven Magee

Le Digital Services Act

- Le **Digital Services Act**
 - Règlement européen
 - « Règlement relatif à un marché intérieur des services numériques »
 - Avec le Digital Markets Act pour les aspects concurrentiels
- Protéger l'individu sur le monde numérique
 - En particulier : empêcher la prolifération de contenus indésirables sur la toile
 - Modération de contenus

Le DSA : les plateformes en ligne

- Dépoussière la directive eCommerce 2020 – il était temps
- Passage d'une directive (transpositions nationales et risque de patchwork) à un règlement (applicable immédiatement)
- Idées principales : transparence et sanctions sérieuses
- Régulation des « plateformes en ligne »
 - Réseaux sociaux mais aussi places de marché, plateformes de partage de contenu, magasins d'applications, plateformes de voyage et d'hébergement en ligne

Le DSA : les très grandes plateformes

- Régime dit « des risques systémiques »
 - Plateformes d'au moins 45 millions d'utilisateurs
- Obligations spécifiques
 - flux de contenus non fondé sur la recommandation, plus de transparence sur la publicité, etc.
- Responsabilité d'auto-évaluation des risques
 - En fonction de facteurs : leur interface, leur système de modération, leur régime de publicité...
 - Obligation de moyen pour corriger les dysfonctionnements
 - Limite par analogie : demander à un fabricant de tabac d'évaluer les risques que les cigarettes causent le cancer

Le DSA

- Véritable architecture européenne de régulation
 - Au niveau national, les autorités compétentes récolteront des données, des plaintes
 - Au niveau européen, le « Comité européen des services numériques » coordonnera, concevra les principes généraux de la régulation, et s'efforcera d'homogénéiser au mieux les pratiques
- Le bâton
 - Vrai pouvoir des autorités d'imposer des inspections, d'exiger la fin d'infractions
 - Sanctions importantes : jusqu'à 6% du chiffre d'affaire

Et aussi le DMA

- Digital Market Act qui « vise à rétablir contestabilité et équité sur les marchés du numérique dominés par les très grandes entreprises du numérique »
 - Corriger les comportements les plus problématiques
- Règles ex ante
 - interdit de combiner des données provenant de différents services d'un « conglomérat » de services
 - interdit les méthodes d'auto-préférence,
 - encourager la portabilité ou l'interopérabilité

Opinion personnelle

- Sans doute insuffisant mais
- Premières étapes importantes
- Clé du succès : montée en compétence des autorités pour mieux comprendre le fonctionnement des plateformes
- Cela reste trop un face à face entre les réseaux sociaux et les autorités
- **Nécessité à terme : plus de participation de la société**
 - **Associations, chercheurs, juges, etc.**
- **Le modérateur devient lui-même réseau social**

5. Soft power

Changer les utilisateurs et les médias sociaux, promouvoir le dialogue

Soft power: changer les utilisateurs

- Éduquer les utilisateurs
- Informer les utilisateurs
- Les faire participer - par exemple, les éditeurs de Wikipedia
 - Détecter les mauvais contenus
 - Définir les mauvais contenus

Éducation

- Interrogez les contenus que vous trouvez dans les médias sociaux
- Pensez aux conséquences des contenus que vous publiez
 - Vous êtes ce que vous partagez, C.W. Leadbeater (1854-1934)
- Apprenez à protéger votre vie privée et à vous protéger des personnes nuisibles
- Apprenez la valeur du dialogue (plutôt que de la confrontation)

- Comprenez la technologie sous-jacente, c'est-à-dire l'informatique



Soft power : changer les médias sociaux

- Modération par les pairs, par exemple, WT:Social
 - Jimmy Wales, 2019
 - Modèle économique : dons (pas de publicité)
 - Encourage l'implication de la communauté
- Interaction pacifiée comme dans Kialo (beaucoup moins amusante)
- Réseau P2P, par exemple, Mastodon
 - Médias de type Twitter
 - Logiciel libre
 - Des instances indépendantes peuvent interopérer



Soft power: promouvoir le dialogue

- Moonshot CVE (contre l'extrémisme violent)
- Entreprise sociale britannique établie à Londres, désormais présente dans 39 pays
- Partenariats avec Facebook et Google
- Utilisation de données publiques pour identifier les utilisateurs à l'origine des discours haineux et leurs destinataires
- Méthode de redirection : redirection en ligne vers des structures de dialogue IRL
 - Implication d'anciens haineux (ex. Life after Hate)



6. Conclusion

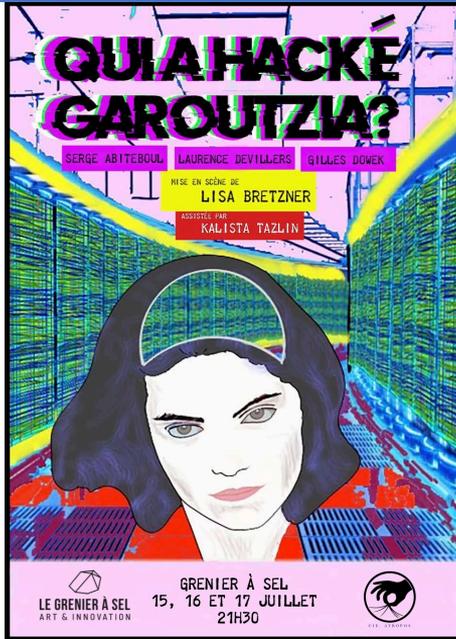
Rien de nouveau sous le soleil

- Comme souvent les nouvelles technologies (et c'est d'abord une techno) apportent le meilleur et le pire
- **Nous devons choisir ce que nous faisons des réseaux sociaux**
 - **Avec l'aide de l'État et des entreprises**
 - **Avec nos propres efforts**
 - **Avec toute la société**

SERGE ABITEBOUL
JEAN CATTAN

Préface de Gérard Berry

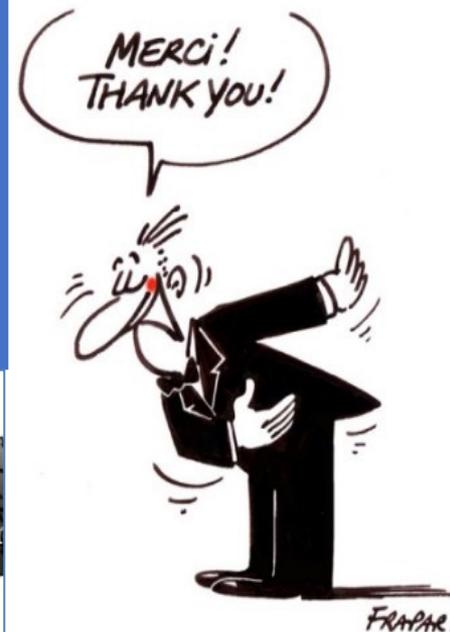
NOUS SOMMES LES RÉSEAUX SOCIAUX



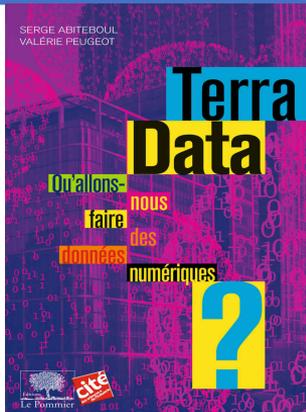
A paraître en février :

Vive les Communs Numériques !

Logiciels libres, wikipédia, le web, la science ouverte, etc., avec François Bancelhon, chez Odile Jacob



le temps des
algorithmes
de serge abiteboul
et gilles dowek



Le bot
qui murmurait
à l'oreille de
la vieille dame
et autres nouvelles numériques



III Blogs
binaire
L'informatique : la science au cœur du numérique
Pourquoi binaire ?
Pour parler d'informatique I

