



L'ETHIQUE DES ALGORITHMES EN SANTE

Christine Balagué

Professeur, IMT-TEM, Titulaire de la Chaire réseaux sociaux et IoT

Vice-présidente du Conseil National du Numérique (2013-2015)

Membre de la CERNA

christine.balague@telecom-em.eu

Pour toute utilisation du contenu de cette présentation, veuillez citer l'auteur, son organisme d'appartenance, le titre et la date du document, ainsi que le volet 3 des ateliers «Ethique, robotique, et applications en santé» de la Plateforme «Ethique et Biosciences» de Toulouse. Merci.



deux écoles filiales et deux partenaires stratégiques

Les algorithmes, un sujet de débat médiatique dans le monde

En Grande-Bretagne, l'opposition plaide pour une plus grande transparence des algorithmes

THE GUARDIAN 23 décembre 2016 6   

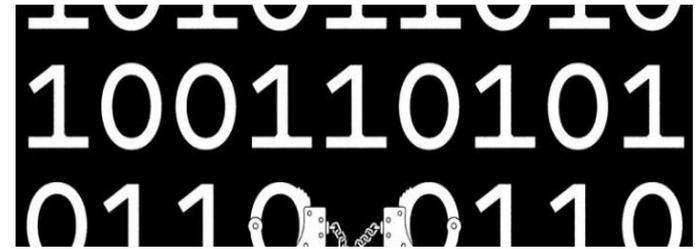


The Opinion Pages | OP-ED CONTRIBUTOR

The New York Times

Make Algorithms Accountable

By JULIA ANGWIN AUG. 1, 2016



france culture

Actualités Art et Création Savoirs Fictions Documentaires Conférences

Dernières publications sur Christine Balagué



16.12.2016

59 min

Pour qui les GAFAs vont-ils nous faire voter ?

La Méthode scientifique

Comment le Big Data peut potentiellement être réutilisé par les politiques ? Comment les réseaux sociaux ont-ils transformé notre rapport à la politique...

04/07/2017



← Les images sont-elles encore produites pour être regardées par les humains ?

L'Intelligence artificielle va-t-elle mettre fin au poker ? →

21 janvier 2017

Peut-on armer la « transparence » des algorithmes ?



 **Recommand** 203 people recommend this.

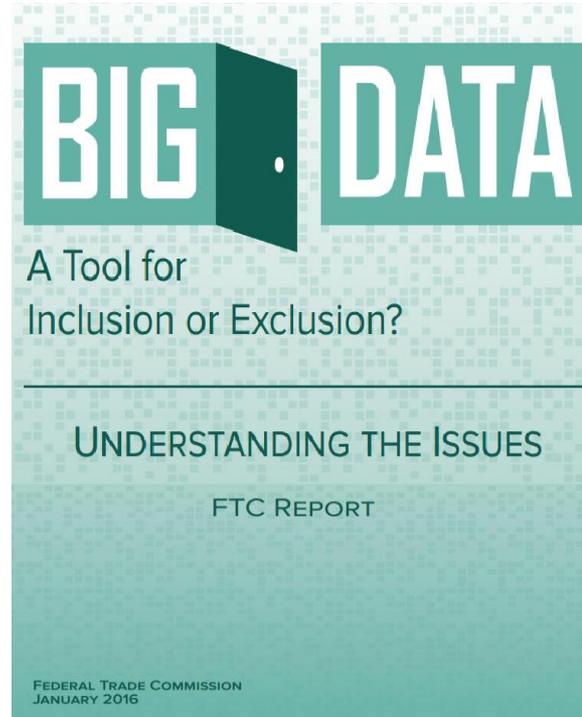
Roger Taylor (@rtayloropendata) est le fondateur du Réseau des services publics ouverts à la Royal Society of Arts. Il est aussi l'auteur, avec Tim Kelsey, le directeur

Rapports et missions gouvernementales

Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights

Executive Office of the President

May 2016



SECRETARIAT D'ETAT CHARGE DU NUMERIQUE

Paris, le

La Secrétaire d'Etat

à **27 OCT. 2016**

Monsieur le Vice-président du Conseil général
de l'Economie, de l'Industrie, de l'Energie et
des Technologies

Mission relative aux modalités de régulation des algorithmes de traitement des
contenus.

Valeurs universelles de l'éthique

Macro-éthique

Principia : « ce qui vient en premier, ce qui est à la source »
ou « ce qui fait autorité »

Principe de justice

Principe d'autonomie

Respect de la vie

Principe de bienfaisance

Utilité
Responsabilité

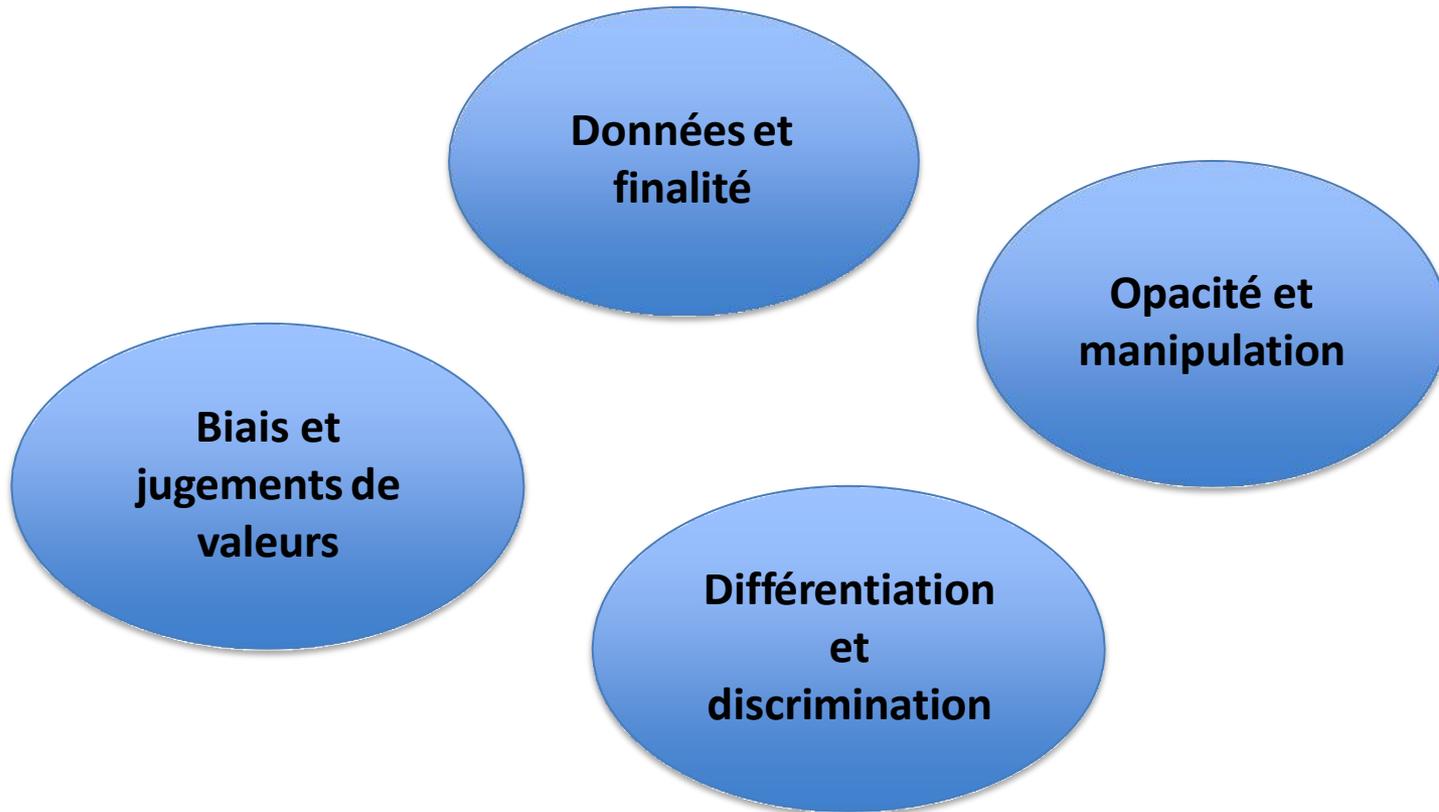
Principe de non-malfaisance

Proportionnalité
Précaution
Incertitude

Micro-éthique

Principles of Biomedical Ethics by Tom L. Beauchamp & James F. Childress (2001)

Quels sont les questions éthiques posées par les algorithmes en santé?





Données et finalité

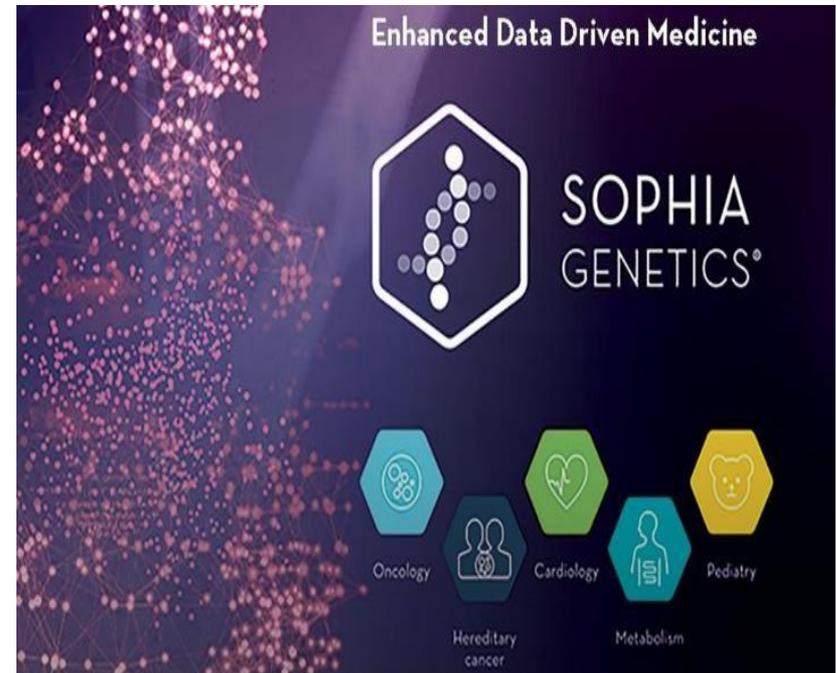


deux écoles filiales et deux partenaires stratégiques

Data driven medicine

Sophia Genetics est une société basée en Suisse qui nourrit de grandes ambitions et revendique être le « *leader global de la médecine basée sur les données (Data Driven Medicine - DDM)* ».

« *Grâce au séquençage à haut débit et aux algorithmes actuels, je suis convaincu que dans cinq à dix ans le cancer sera une maladie maîtrisée* », Jurgi Camblong, PDG Sophia Genetics.

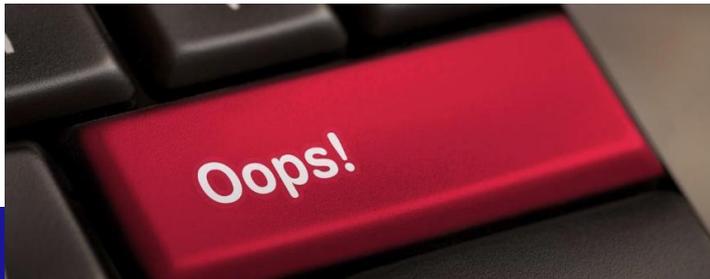


les données de milliers de tests génomiques de patients réalisés par les hôpitaux partenaires de Sophia DDM (ils sont 170 dans 28 pays).

Les problèmes liés aux données, input des algorithmes

- « Objets d'un design humain » (Crawford)
- Qualité
- Biaisées
- Incomplètes
- Vulnérables

Fuite des données de l'OPM (Office of Personnel Management) aux US (cyberattaque): 21,5 millions de personnes concernées/ informations personnelles (dont santé) dans la nature, dont 1,1 million d'empreintes digitales.





Opacité et manipulation



deux écoles filiales et deux partenaires stratégiques

Admission post-bac : le ministère poussé à dévoiler l'algorithme qui oriente les lycéens

LE MONDE | 19.09.2016 à 08h16 • Mis à jour le 20.09.2016 à 09h16 |

Par Séverin Graveleau



➤ Objectif initial APB: rationaliser le système, le rendre plus efficace et moins opaque

➤ Effets inverses:

- Opacité sur l'impact des notes
- Concentration de boursiers (25% district Est de Paris vs 10% en moyenne)
- Sélection aléatoire des universités quand le nbre de candidats est très important pour une formation

Adressée à Ministre de l'Enseignement Supérieur, la Recherche et l'Innovation Frédérique Vidal

Le hasard ne décidera pas de mon avenir !
#TiréAuSort

#TiréAuSort NON France



#TiréAuSort : NON !



L'utilisateur n'a pas conscience des différences de résultats entre utilisateurs à la même requête

The image shows two side-by-side screenshots of Google search results for the query "Egypt".

- Scott's Search:** Shows approximately 300,000,000 results. The top result is "Crisis in Egypt" from humanrightsinst.org. Other results include Wikipedia, The New York Times, and various travel and news sites.
- Daniel's Search:** Shows approximately 321,000,000 results. The top result is the Wikipedia article for Egypt. Other results include travel guides, news sites, and the CIA World Factbook.

The difference in the number of results and the ranking of the top results illustrates how search engines may personalize results based on user location or other factors, even for the same query.



Les fake news en santé



Regenere / Thierry Casasnovas

Home Videos Playlists Channels About



Bienvenue sur la chaine Regenere /Thierry Casasnovas- www.regenere.org

107,011 views 5 months ago

Bienvenue à la découverte de la chaine Regenere!



**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

**Global
Challenges**
www.global-challenges.com

Inoculating the Public against Misinformation about Climate Change

Sander van der Linden, Anthony Leiserowitz, Seth Rosenthal, and Edward Maibach*

Effectively addressing climate change requires significant changes in individual and collective human behavior and decision-making. Yet, in light of the increasing politicization of (climate) science, and the attempts of vested-interest groups to undermine the scientific consensus on climate change through organized “disinformation campaigns,” identifying ways to effectively engage with the public about the issue across the political spectrum has proven difficult. A growing body of research suggests that one promising way to counteract the politicization of science is to convey the high level of normative agreement (“consensus”) among experts about the reality of

Polarization can be amplified when the inherent uncertainty of science itself is used to cast doubt on the existence of a scientific consensus.^[8] For example, ideologically motivated, vested-interest groups known as “Merchants of Doubt” have orchestrated influential “disinformation campaigns” in which they publicly dispute the scientific consensus on various issues, including human-caused climate change.^[9,10] These campaigns have

FULL PAPER

Algorithmes « éthiques » de détection des fake news

Photos?

Daniel Legrand
2 décembre, 01:04

François Fillon et Marion Maréchal-le Pen, ensemble contre ces salopes de gauche qui osent avorter en défiant le plan de Dieu !



LES DÉCODEURS
VENONS-EN AUX FAITS



La Fake News Machine

On peut manipuler une élection avec 400 000 \$!



Figure 1. The Fake News Triangle

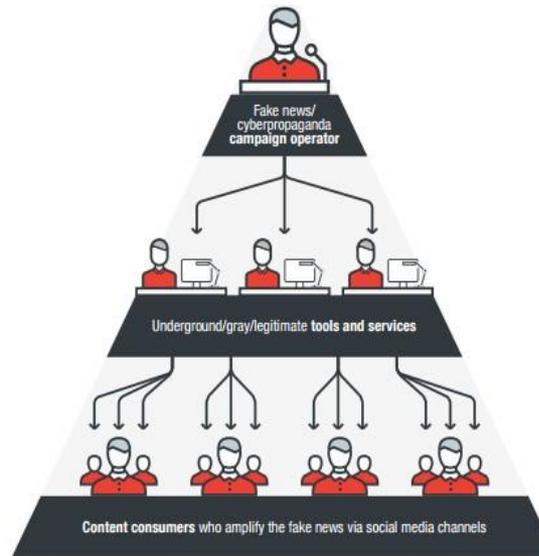


Figure 2. How an operator employs or abuses underground, gray, and legitimate marketplaces to disseminate fake news

为什么选择图创软文?

三大保证为您服务:

写作帮软文代写的六大优势:

软文代写价格:			
优质原创软文/新闻稿			
每篇字数	文章类别	每篇价格	套餐优惠价
500字-800字	软文、新闻稿	100元	400元/3篇
1000字-1500字	软文、新闻稿	200元	800元/3篇

物流配送: 先将将文章发布到一个或多个媒体网站下 (价值60元)

Накрутка голосов.

Победа в конкурсах.

Могу помочь вам победить в интернет конкурсах любой сложности. Опыт в данной сфере более 5 лет и за плечами невообразимое количество тысяч конкурсов.

Гарантирую возможность голоса можно получить как от ботов так и от реальных людей. Для проверки данных пробую накрутку 10-20 голосов до оплаты и далее всякое оплата по частям.

Стоимость услуг определяется индивидуально для каждого конкурса. Средняя стоимость начинаем от 1 рубля за голос для простых конкурсов. Более подробно можно узнать в разделе "Услуги".

Сферы деятельности:

- 1) Накрутка голоса в конкурсах
- 2) Накрутка подписчиков в соц. сетях и на Youtube
- 3) Накрутка лайков, репостов и комментариев
- 4) Заручение аккаунтами
- 5) Накрутка просмотров на Youtube

Связь:

Skype: [contact info]

ICQ: [contact info]

Email: [contact info]

Сайт: [contact info]

WhatsApp: [contact info]

Telegram: [contact info]

Услуги

Lion Gu, Vladimir Kropotov, and Fyodor Yarochkin (2017) The Fake News Machine How Propagandists Abuse the Internet and Manipulate the Public, Trendlabs research paper, Trend Micro

Algorithmes prédictifs en santé: le cas du cancer du sein



Plus précisément, les mutations des gènes BRCA augmentent les risques de cancers de la manière suivante (Antoniou et al – Am. J. Hum. Genet. 2003 ; 72 :117-30) :

Risque de cancer avant 70 ans :

- Sein : 40 à 85 % contre 10 % dans la population générale
- Ovaire : 10 à 63 % contre 1 % dans la population générale

Risque en fonction du gène touché :

- BRCA1 : sein 65 % et ovaire 45 %
- BRCA2 : sein 45 % et ovaire 11 %

Risque de cancer avant 45 ans :

- BRCA1 : sein 25 % et ovaire 10 %
- BRCA2 : sein 7 % et ovaire 1 %

	Tumeurs solides																	Hémopathies malignes												
	Nasopharynx	Tête et cou	Bouche (cavité orale), pharynx, larynx	Œsophage	Adénocarcinome bronchopulmonaire	Estomac	Intestin grêle	Colon rectum	Pancréas	Ampeulle de Vater	Foie	Vessie de la vessie	Rein	Vessie	Sein (avant la ménopause)	Sein (après la ménopause)	Endométrie	Col de l'utérus	Ovaire	Prostate	Testicule	Poumon	Thyroïde	Peau	Lymphome hodgkinien	Lymphome non hodgkinien	Leucémie	Myélome multiple		
Boissons alcoolisées																														
Surcharge pondérale																														
Viandes rouges																														
Charcuteries																														
Sel et aliments salés																														
Compléments alimentaires à base de bêta-carotène																														
Produits laitiers																														
Activité physique																														
Sédentarité																														
Fruits																														
Légumes (non féculents)																														
Fibres alimentaires																														
Allaitement																														

Certainement Probable Suggéré Non concluant Non étudié Suggéré Probable Certainement

Source: site institut national du cancer

Rapport annuel , juin 2017
Institut national du cancer

Algorithmes prédictifs en santé: le cas du cancer du sein

- **Corrélations et non causalité**
- **Comment communiquer au patient les résultats d'un processus complexe ?**
- **Quelle décision doit-il prendre?**

Ethique et algorithmes d'apprentissage machine



Éthique de la recherche en apprentissage machine

Edition provisoire, juin 2017

L. Devillers, M. Dauchet, A. Grinbaum, G. Dowek, D. Bourcier, R. Chatila, S. Abiteboul, N. Boujemaa

Le rapport est disponible sur le site de la CERNA

Applications de l'apprentissage machine

De nombreux systèmes utilisent des modules d'apprentissage automatique :

- moteurs de recherche,
- traduction automatique,
- agents conversationnels (chatbot),
- robots (aspirateurs, sociaux...)

Chatbot médical lancé par la NHS, janvier 2017

- National Health Service aux UK lance un expérimentation pendant 6 mois (1,2 millions de personnes concernées)
- Alternative au 111, numéro dédié aux appels considérés comme peu urgents
- Objectif:
 - désengorger le 999, numéro dédié aux urgences
 - aider les services de santé à mieux gérer les patients (une partie de la prise de contact)
 - ne pas remplacer le médecin mais triage médical en posant des questions
 - réduire les coûts



*Agent conversationnel
de la start up Babylon Health*

Diagnostic automatisé: le cas IBM Watson

A powerful collaborative approach

Quest Diagnostics sequencing & analysis

Quest Diagnostics sequences the tumor's genomic makeup to help clinicians understand the unique cancer-causing mutations specific to that individual patient's cancer

Watson annotation

Watson then compares your patient's results against relevant medical literature, clinical studies, pharmacopeia, and carefully annotated rules created by leading cancer centers

Quest Diagnostics medical director review

A Quest Diagnostics medical director reviews the Watson reports, confirms the results, and prepares a report for your consideration

Watson Genomics from Quest Core report

You receive a report from the Quest Diagnostics medical team that provides therapy and clinical trial information that has been personalized to your patient, including mutations and therapeutic associations that have been graded through clearly-defined levels of evidence



Robots bavards empathiques

2017- : Interaction avec des machines dotées de modules d'apprentissage machine :

- au domicile, dans une boutique ou encore dans un établissement dans l'organisation du soin (assistance, compagnie, surveillance...)
- pour la perception, la décision, l'action et le dialogue (20% foyers US – Alexa/Amazon Echo)

Ces machines mémorisent des données personnelles pour **communiquer, apprendre, dialoguer, agir... et elles pourront même être capable d'empathie**

L'objet peut être perçu comme une personne à travers une illusion empathique !

Les robots assistants

L'adaptabilité à l'environnement que confère la capacité d'apprentissage devrait à l'avenir favoriser l'usage des robots auprès des personnes, notamment comme compagnons ou soignants.

Construire des robots “sociaux” d'assistance aux personnes nécessite d'encadrer leur usage, d'autant plus quand ils sont en contact avec des personnes malades ou âgées.

Les recommandations personnalisées

Les traces que nous laissons à travers nos consultations sur Internet et à travers les objets auxquels nous sommes connectés sont exploitées par des algorithmes d'apprentissage afin de mieux cerner nos préférences de consommation, notre mode de vie et nos opinions.

Par rapport aux simples statistiques, ces algorithmes ont – ou peuvent avoir – la capacité de fournir des prescriptions individuelles

Apprentissage machine

■ Intelligence Artificielle

1^{ère} vague : Systèmes experts

Perception, Apprentissage, Abstraction, Raisonement

2^{ème} vague : Apprentissage statistique

Perception, Apprentissage, Abstraction, Raisonement

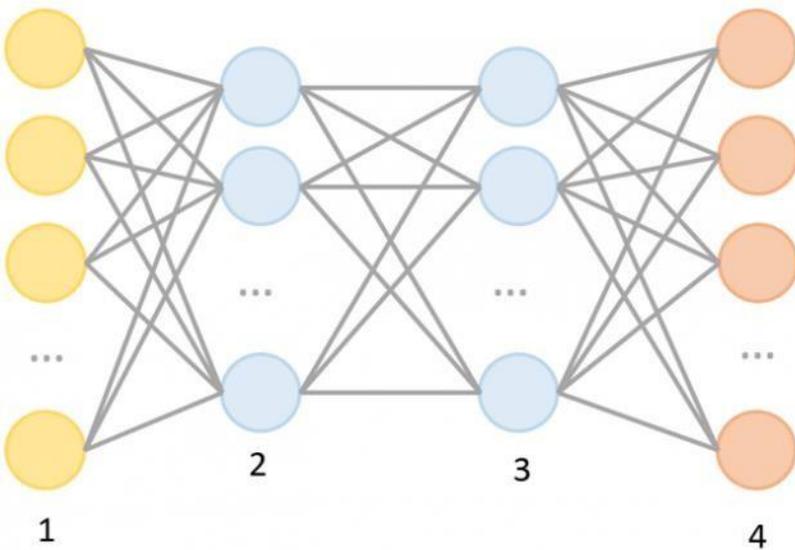
Apprendre sans comprendre

3^{ème} vague : Adaptation contextuelle, sens commun

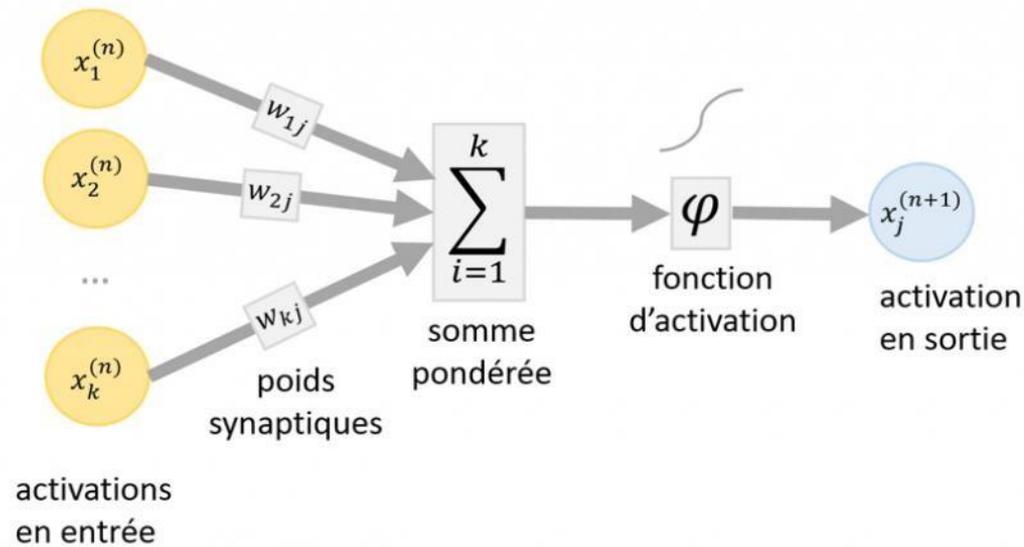
Perception, Apprentissage, Abstraction, Raisonement

Graal -> produire des explications

Réseaux de neurones : principes



(a)



(b)

Opacité des décisions

- Prédire un phénomène à partir d'observations passées présuppose un mécanisme causal.
- Expliquer ce mécanisme n'est pas toujours facile, surtout dans des systèmes complexes
- Approche statistique permettant de découvrir des corrélations significatives dans une masse importante de **données pour construire un modèle prédictif quand il est difficile de construire un modèle explicatif.**

Les questionnements et les préconisations en six thèmes

- Les données des systèmes d'apprentissage
- L'autonomie des systèmes apprenants
- L'explicabilité et l'évaluation des systèmes d'apprentissage
- Les décisions des systèmes d'apprentissage
- Le consentement dans les systèmes d'apprentissage
- La responsabilité dans les relations homme-machine

Les données

[DON-1] Qualité des données d'apprentissage

Le concepteur et l'entraîneur veilleront à la qualité des données d'apprentissage et des conditions de leur captation tout au long du fonctionnement du système. Les entraîneurs du système informatique sont responsables de la présence ou de l'absence de biais dans les données utilisées dans l'apprentissage, en particulier l'apprentissage à en continu, c'est-à-dire en cours d'utilisation du système. Pour vérifier l'absence de biais, ils doivent s'appuyer sur des outils de mesure qui restent encore à développer.

[DON-2] Les données comme miroir de la diversité

Les entraîneurs des systèmes d'apprentissage automatique doivent opérer le choix des données en veillant à ce que celles-ci respectent la diversité des cultures ou des groupes d'utilisateurs de ces systèmes.

[DON-3] Variables dont les données comportent un risque de discrimination

Les entraîneurs (qui peuvent être aussi les concepteurs ou les utilisateurs) doivent se poser la question des variables qui peuvent être socialement discriminantes. Il convient de ne pas mémoriser ni de régénérer par programmation ces variables, par exemple l'ethnie, le sexe ou l'âge. La protection des données à caractère personnel doit également être respectée conformément à la législation en vigueur.

[DON-4] Traces

Le chercheur doit veiller à la traçabilité de l'apprentissage machine et prévoir des protocoles à cet effet. Les traces sont elles-mêmes des données qui doivent à ce titre faire l'objet d'une attention sur le plan éthique.

L'autonomie

[AUT-1] Biais de caractérisation

Le chercheur veillera à ce que les capacités d'apprentissage d'un système informatique n'amènent pas l'utilisateur à croire que le système est dans un certain état lors de son fonctionnement alors qu'il est dans un autre état.

[AUT-2] Vigilance dans la communication

Dans sa communication sur l'autonomie des systèmes apprenants relativement aux humains, le chercheur doit viser à expliquer le comportement du système sans donner prise à des interprétations ou des médiatisations irrationnelles.

L'explicabilité et l'évaluation

[EXP-1] Explicabilité

Le chercheur doit s'interroger sur la non-interprétabilité ou le manque d'explicabilité des actions d'un système informatique apprenant. Le compromis entre performance et explicabilité doit être apprécié en fonction de l'usage et doit être explicité dans la documentation à l'usage de l'entraîneur et de l'utilisateur.

[EXP-2] Les heuristiques d'explication

Dans sa recherche d'une meilleure explicabilité d'un système apprenant, le chercheur veillera à décrire les limitations de ses heuristiques d'explication et à ce que les interprétations de leurs résultats ces soient exemptes de biais.

[EXP-3] Elaboration des normes

Le chercheur veillera à contribuer aux débats de société et à l'élaboration de normes et de protocoles d'évaluation qui accompagnent le déploiement de l'apprentissage machine. Quand les données concernent certains secteurs professionnels spécialisés (médecine, droit, transports, énergie, etc.), les chercheurs de ces domaines devront être sollicités.

Les décisions

[DEC-1] Place de l'humain dans les décisions assistées par des machines apprenantes

Le chercheur, concepteur de machines apprenantes d'aide à la décision, doit veiller à ce qu'aucun biais ne produise un résultat qui devienne automatiquement une décision alors que l'intervention humaine était prévue par la spécification du système, et prendra garde aux risques de dépendance de l'humain aux décisions de la machine.

[DEC-2] Place de l'humain dans l'explication des décisions assistées par des machines apprenantes

Le chercheur doit veiller à ce que les résultats du système soient autant que possible interprétables et explicables pour les personnes concernées, s'impliquer dans les nécessaires évolutions des métiers en faisant usage et dans l'encadrement des pratiques. Des agents experts contribueront à l'explication et à la vérification des comportements du système.

Le consentement

[CON-1] Choix des usagers d'activer ou non les capacités d'apprentissage d'un système

Le chercheur doit créer la possibilité d'utilisation de systèmes avec ou sans leur capacité d'apprentissage. Il doit donner à l'utilisateur au moins un paramètre de contrôle global sur la source des données utilisées pour l'apprentissage.

[CON-2] Consentement dans le cadre de projet

Le chercheur doit délibérer dès la conception de son projet avec les personnes ou les groupes identifiés comme pouvant être influencés.

[CON-3] Consentement à utiliser une machine capable d'apprentissage en continu

Le chercheur doit être conscient que la capacité d'apprentissage et la mise en réseau de telles capacités peut induire des problématiques nouvelles concernant autant le consentement de l'utilisateur que celui de la société.

La responsabilité

[RES-1] Mécanismes de contrôle

Le chercheur doit adapter et inclure dans le système, des mécanismes de contrôle, automatiques ou supervisés (par la machine ou l'être humain), sur les données, sur le fonctionnement au niveau informatique et sur le raisonnement suivi, afin de faciliter l'attribution de responsabilités au regard du bon fonctionnement ou du dysfonctionnement du système.

[RES-2] Déclaration des intentions d'usage

En documentant un système informatique apprenant, le chercheur doit décrire de manière sincère, loyale et complète, les limites qui lui sont connues de l'imputabilité d'une décision ou d'une action du système soit au code-source soit au processus d'apprentissage. Cette documentation servira de déclaration d'intention de la part du concepteur quant aux usages du système informatique qu'il envisage comme normaux. L'absence de telle déclaration ou son caractère tardif peuvent engager la responsabilité supplémentaire du concepteur.



Biais et jugements de valeur



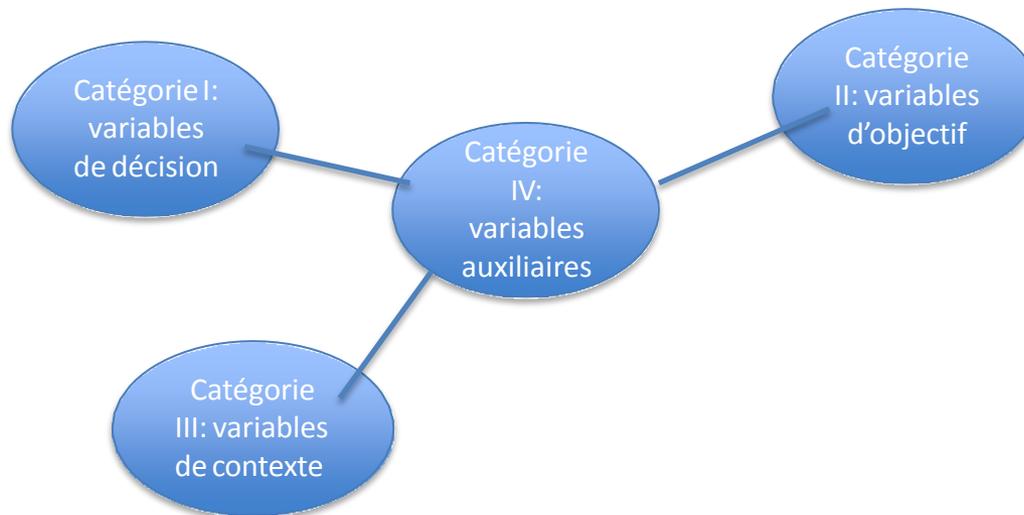
deux écoles filiales et deux partenaires stratégiques

Les algorithmes ont un jugement de valeur

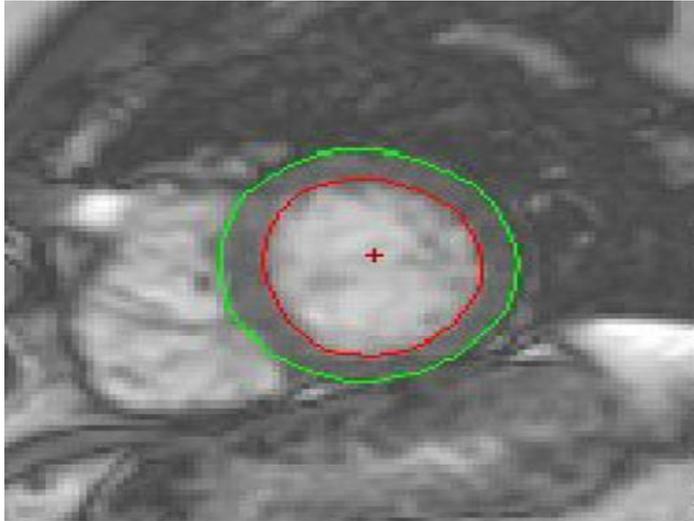
(Kraemer, van Overveld, Peterson, 2011)

■ Design activity:

- Variable de décision (catégorie I): décider qu'une variable devrait avoir une certaine valeur
 - Ex: roue avant du vélo aura 60 cm de diamètre
- Variable objective (Catégorie II): valeur de la variable définissant un objectif atteint
 - Ex: poids de la roue si l'objectif est un vélo léger
- Variables de contexte (catégorie III): densité du carbone
- Variables auxiliaires (catégorie IV): composition de variables (poids est la somme des poids de chaque roue plus du guidon)



Les algorithmes ont un jugement de valeur *(Kraemer, van Overveld, Peterson, 2011)*



■ Exemple: technologies imagerie médicale

- Représentation de structures biologiques par un ordinateur le plus précisément possible
- Objectif: diagnostiquer des maladies
- Risque des algorithmes:
 - Faux Positifs: symptôme de maladie alors que ce n'est pas vrai
 - Faux Négatifs: pas d'identification d'un symptôme de maladie réel



Trade off entre le nombre de faux positifs et le nombre de faux négatifs (Jugement de valeur sur le nombre de faux positifs tolérable vs le nombre de faux négatifs)

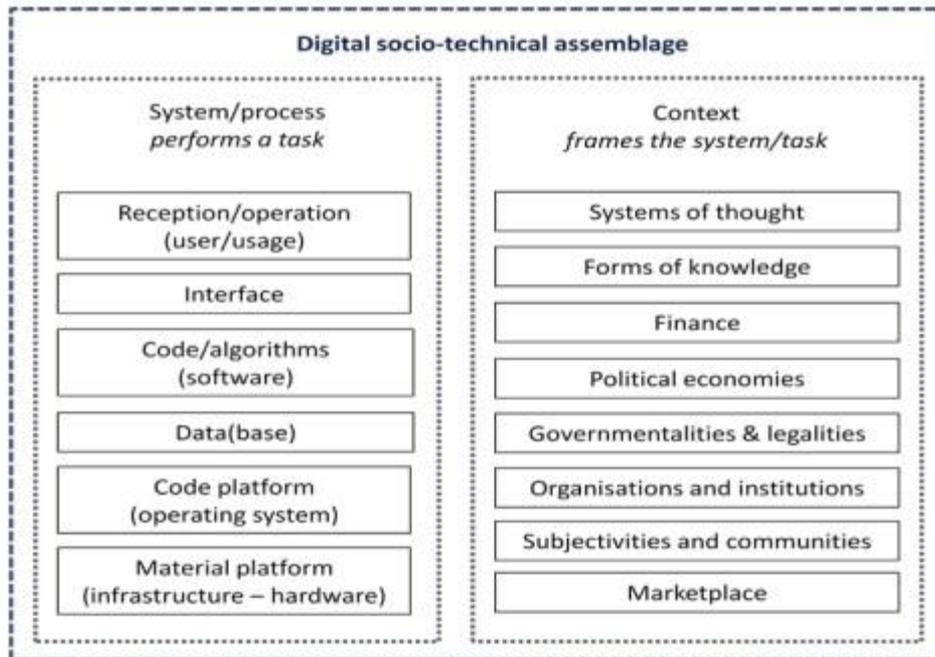
■ Principe de précaution:

- Pour des médecins, il est préférable d'avoir des faux positifs que des faux négatifs,
- Pour les scientifiques, c'est l'inverse, car les connaissances scientifiques sont cumulatives.



Les valeurs des variables de catégorie I impactent les variables de catégories II

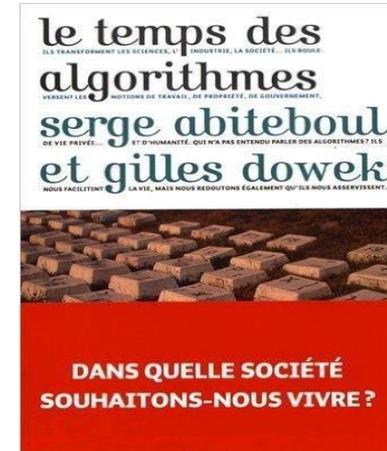
Les algorithmes constituent un élément d'un système plus global



(Kitchin, 2014)

Effets transformatifs, normatifs et non prévisibles des algorithmes

- **Effets transformatifs**
 - Effet sur l'emploi, disruption
 - Les algorithmes formatent nos représentations (D.Cardon)
- **Effets normatifs**
 - Les algorithmes construisent des régimes de puissance et de connaissance (Anderson, 2011)
 - Ils sont utilisés pour séduire, discipliner, décider, réguler et contrôler
 - Technologies du bon comportement
 - Leur grand récit global est de dire que tout est de la faute de l'individu
 - Si on mange mieux, si on conduit de manière responsable etc... tout irait mieux
 - Lorsqu'on fait reposer le problème sur les individus, nous détournons l'attention des solutions systémiques
- **Effets non prévisibles:**
 - Les algorithmes sont un élément d'un système plus global et peuvent donc être impactés par un élément de ce système (Ex: mauvaise qualité des données d'entrée)
 - Ils peuvent avoir des effets inattendus (Ex: algorithmes non supervisés)
 - Les résultats peuvent être erronés (Bugs de codage)



L'IA de Facebook a crée des chatbots qui ont développé leur propre langage....



FAIR, 20 juin 2017



Challenges de la transparence



deux écoles filiales et deux partenaires stratégiques

Challenges



■ Accéder à la black box:

- Public (algorithmes ouverts) vs privé (algorithme propriétaire)

■ Hétérogénéité, intégration

- Complexité
- Plusieurs contributeurs
- Difficulté de comprendre la logique sous-jacente

Big Ball of Mud

■ Ontogénétique et performatif

- Rarement fixé, évolutif, itératif
- Réaction à la situation, à l'input
- Etudier une version sera un snapshot

■ Hors de contrôle

- Résultats varient en fonction du contexte (géolocalisation)
- Impact différent suivant le contexte d'usage (travail vs loisir)

Approches (1)

■ Examiner le code source

- Documentation
- Mapping de l'évolution et des différentes versions
- Programmer une même tâche en différent langages ou plateformes/logiciels (Montfort et al, 2012)

■ Produire une auto réflexion sur le code

- Auto-ethnographie (Ziewitz, 2011; Ullman, 1997)
- Critique du code par le concepteur (chercheur)
- Mesure de l'impact selon le contexte (économique, social, sociétal)
- Tracer et documenter les écarts

■ Reverse engineering

- Examen rigoureux des effets de spécifications (différents scénaris) (Mahnke and Uprichard, 2014; Bucher, 2012)
- Exploration de la manière dont marche le système
- Utilisation de robots
- Etude des communautés qui débattent des algorithmes/ Entretiens semi/non directs

Approches (2)

- **Entretiens de concepteurs, designers/ Ethnographie d'équipe de codeurs**
 - Objectifs des concepteurs, intentions
 - Elaboration de l'algorithme, source d'influences (Diakopoulos, 2013; Mager, 2012)
 - Observer l'élaboration d'un algorithme, immersion dans l'organisation (Rosenberg, 2007; Takhteyev, 2012)
- **Décortiquer l'assemblage socio-technique des algorithmes**
 - Décrire le système global
 - Entretiens, ethnographie de projets (incluant management, institutions, régulation,..)
 - Documentations des acteurs clefs (biographie)
 - Histoire des projets (Montfort et al, 2012; Napoli, 2013)
- **Examiner comment les algorithmes fonctionnent dans le monde**
 - Ethnographies sur la manière dont les acteurs sont conditionnés par des systèmes algorithmiques, et comment ces systèmes sont transformatifs
 - Etudier les résistances, les évitements, les transgressions (utilisation des algorithmes à d'autres fins) (Al-Akkad et al., 2013)
 - Etudier (observation/entretiens/mesure comportementale) sur différents types d'utilisateurs/différents scenarios: conséquences, tactiques d'engagement, sentiments, perceptions

Limites des technologies pour mesurer la transparence des algorithmes

■ Contreproductif:

- Risque de simplification (en ne dévoilant qu'une partie) éliminant des éléments de contextes importants
- Utiliser des algorithmes ou du big data
 - nouvelle forme d'opacité,
 - faible sensibilité aux dynamiques du contexte (McAfee and Brynjolfsson, 2012)
 - connaissance supérieure peu respectueuse des fonctions sociales (sphère publique démocratique) (Birchall, 2012)

■ Etudier comment la transparence est produite

■ Conséquences de la transparence:

- Autocensure, enfermement, anxiété

■ Dans un contexte de néolibéralisme global, l'idéal de transparence sert la vision que le marché est un agent central de la société, générant de la dérégulation et augmentant la self regulation

■ le big data produirait la vérité, ce qui oblitère la connaissance scientifique fondée sur des hypothèses et modèles théoriques pour comprendre les individus

■ De nombreuses activités sociales ne sont pas mesurables par des technologies digitales

Les travaux du Conseil National du Numérique (2013-2015) sur la transparence et la loyauté



Loyauté des plateformes, introduite dans la loi pour une république Numérique du 8 octobre 2016

Article 49 [En savoir plus sur cet article...](#)

I. - Le livre Ier du code de la consommation est ainsi modifié :

1° L'article L. 111-7 est ainsi rédigé :

« Art. L. 111-7. - I. - Est qualifiée d'opérateur de plateforme en ligne toute personne physique ou morale proposant, à titre professionnel, de manière rémunérée ou non, un service de communication au public en ligne reposant sur :

« 1° Le classement ou le référencement, au moyen d'algorithmes informatiques, de contenus, de biens ou de services proposés ou mis en ligne par des tiers ;

« 2° Ou la mise en relation de plusieurs parties en vue de la vente d'un bien, de la fourniture d'un service ou de l'échange ou du partage d'un contenu, d'un bien ou d'un service.

« II. - Tout opérateur de plateforme en ligne est tenu de délivrer au consommateur une information loyale, claire et transparente sur :

« 1° Les conditions générales d'utilisation du service d'intermédiation qu'il propose et sur les modalités de référencement, de classement et de déréférencement de contenus, des biens ou des services auxquels ce service permet d'accéder ;

« Ce décret précise, par ailleurs, pour tout opérateur de plateforme en ligne dont l'activité consiste en la fourniture d'informations permettant la comparaison des prix et des caractéristiques de biens et de services proposés par des professionnels, les informations communiquées aux consommateurs portant sur les éléments de cette comparaison et ce qui relève de la publicité au sens de l'article 20 de la loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique.

« Ce décret fixe également les modalités selon lesquelles, lorsque des professionnels, vendeurs ou prestataires de services sont mis en relation avec des consommateurs, l'opérateur de plateforme en ligne met à leur disposition un espace leur permettant de communiquer aux consommateurs les informations prévues aux articles L. 221-5 et L. 221-6. » ;

Article 4 [En savoir plus sur cet article...](#)

Après l'article L. 311-3 du même code, il est inséré un article L. 311-3-1 ainsi rédigé :

« Art. L. 311-3-1.-Sous réserve de l'application du 2° de l'article L. 311-5, une décision individuelle prise sur le fondement d'un traitement algorithmique comporte une mention explicite en informant l'intéressé. Les règles définissant ce traitement ainsi que les principales caractéristiques de sa mise en œuvre sont communiquées par l'administration à l'intéressé s'il en fait la demande.

« Les conditions d'application du présent article sont fixées par décret en Conseil d'Etat. »



Transalgo



deux écoles filiales et deux partenaires stratégiques

Comment est né Transalgo?

- Axelle Lemaire a confié au Conseil Général de l'Economie (CGE) une **mission** relative aux modalités de régulation des algorithmes de traitement des contenus
- Suite à cette saisine, le CGE a livré le **rapport** « *Modalités de régulation des algorithmes de traitement des contenus* »*
- Parmi les **recommandations** de ce rapport : la mise en place d'une **plateforme scientifique** collaborative destinée à favoriser, d'une part le développement **d'outils logiciels et de méthodes de tests d'algorithmes** et d'autre part la promotion de leur utilisation.

[http://www.economie.gouv.fr/files/files/directions_services/cge/Rapports/2016_05_13_Rapport_Algorithmes\(1\).pdf](http://www.economie.gouv.fr/files/files/directions_services/cge/Rapports/2016_05_13_Rapport_Algorithmes(1).pdf)

La plateforme Transalgo

- Inria s'est vu confié le rôle **d'opérateur** d'une telle plateforme dénommée **TransAlgo**, avec le soutien **du CNum** et de **l'IMT**
- Plateforme scientifique pour le **développement de la transparence et de la responsabilité des systèmes algorithmiques** (dualités algorithmes et des données)
- Le lancement de la plateforme TransAlgo **contribue à ce que puisse s'appliquer l'exigence** de transparence et de responsabilité des acteurs de l'économie numérique, introduit par la **loi pour une République numérique**.
- Cette plateforme sera une **première en Europe**



Merci de votre attention

[@christine.balague@telecom-em.eu](mailto:christine.balague@telecom-em.eu)
[@balague](#)



deux écoles filiales et deux partenaires stratégiques