



Algorithmes & IA dans l'administration publique

Efficacité vs. Droits fondamentaux, démocratie & État de droit ?

ESTELLE PANNATIER | EGOVERNMENT SYMPOSIUM ROMAND | 7 MAI 2024

@AlgorithmWat_CH | @algorithmwatch



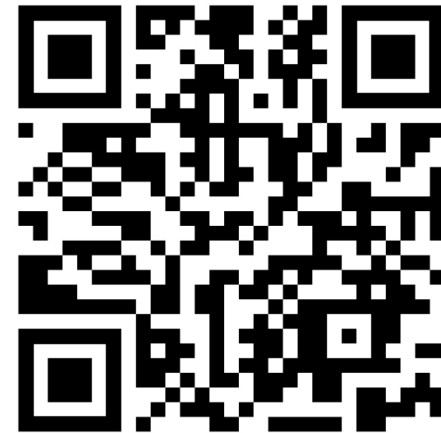
/ AlgorithmWatch

- Organisation de plaidoyer de la société civile (CH & DE)
- Nous nous **focalisons** sur l'impact des **algorithmes** et de **l'IA** sur les **individus** et la **société**.
- Notre **objectif** est de faire en sorte que les systèmes algorithmiques **profitent** réellement aux individus et à la société, au lieu de leur nuire, c'est-à-dire qu'ils renforcent les **droits fondamentaux**, la **démocratie** et la **durabilité** au lieu de les affaiblir.



Que faisons-nous pour atteindre cet objectif ?

- Plaidoyer basé sur des évidences
 - Recherche | Journalisme | Tech → Politique → Plaidoyer & Campagne
- Transparence
 - Fondations
 - Financement de projets
 - Dons privés
 - (Honoraires)





/ Agenda

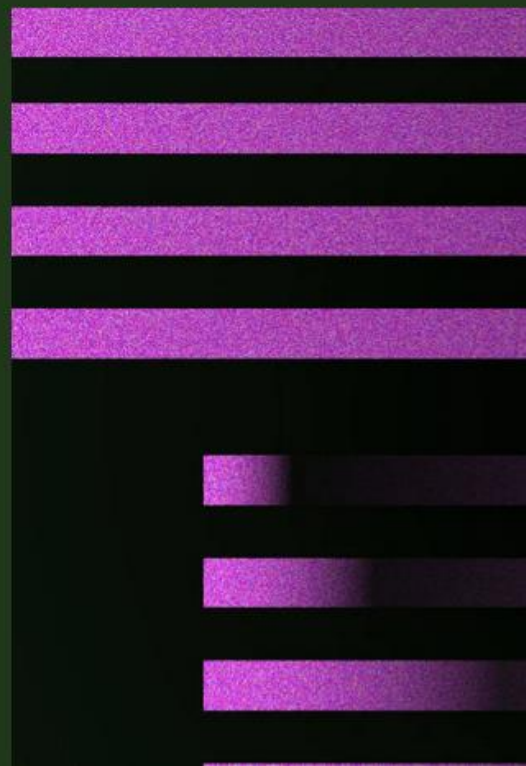
- De quoi parlons-nous ?
 - Où les algorithmes sont-ils utilisés dans/par l'administration publique ?
- Quel est le problème ?
 - Enjeux pour les individus, les droits fondamentaux, la société et la démocratie ?
 - Et quelles sont les chances ?
- Et maintenant ? Que faire – en prenant l'exemple de l'administration publique ?
 - Et que font les autres ?
- Give-Aways



/ De quoi parlons-nous ?

Introducing ChatGPT

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.



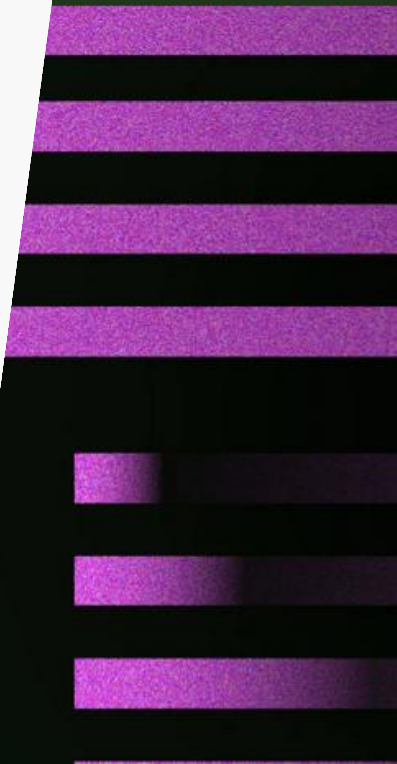
Atlas of Automation Switzerland

AlgorithmWatch CH sheds light on the algorithmic black box



Introd

We've trained
in a conversational way,
possible for ChatGPT to answer followup questions,
admit its mistakes, challenge incorrect premises, and
reject inappropriate requests.





/ De quoi parlons-nous ?

- **Effets des systèmes algorithmiques** sur les individus & la société
 - Pas une définition technologique
- **Systèmes algorithmiques**, qui sont utilisés pour :
 - pour **prévoir, recommander, influencer ou prendre des décisions concernant des personnes**; ou
 - pour **générer des contenus qui influencent les décisions prises par / sur les personnes**.
- vs chaîne d'assemblage, prévisions météorologiques, correction orthographique, ...



/ Algorithmes dans l'administration publique

















/ Et en Suisse ?



Atlas of Automation Switzerland

AlgorithmWatch CH sheds light on the algorithmic black box





Atlas de l'automatisation

Get started now and find out which algorithmic systems are used in your canton or throughout Switzerland!

Search the database

143 results fit the current filter configuration.

Text Search

Hashtag

Policy Area

Employing Entities

Affected Groups

Cantons

Federal

 swiss-wide impact



Better tomorrow

ODARA

Das Ontario Domestic Assault Risk Assessment (ODARA) ist ein Risikobewertungsinstrument, das gemäss den Hersteller*innen das Risikopotenzial eines Mannes berechnet, der seine Partnerin angegriffen hat. Das ODARA besteht aus 13 Ja/Nein-Fragen, die Bereiche wie die Vorgeschichte des Täters in Bezug auf häusliche Gewalt, nicht häusliche Gewalt, Drogenmissbrauch, frühere Freiheitsstrafen und Anzahl Kinder bewertet. Die Gesamtpunktzahl soll das Risikopotenzial angeben.

CANTON Basel-Stadt, Solothurn, Zürich	HASHTAGS #crimeprevention #lawenforcement #predictivepolicing
EMPLOYING ENTITY Police, Public administration	POLICY AREA Public Sector
AFFECTED GROUP Potential criminal offenders, Victims of crime	
PURPOSE Prediction of domestic violence risk	
LINKS https://www.knfp.ch/prognose/odara https://www.apa.org/pubs/books/4318055 https://www.alexandria.unisg.ch/bitstreams/1f96c84a-690f-4729-a48e-8271613b6ea9/download https://www.alexandria.unisg.ch/handle/20.500.14171/111476 https://www.bwl.unibe.ch/ueber_uns/bwl_blog/artikel/2023/eine_kurze_geschichte_der_digitalen_transformation_oder_die_ethischen_probleme_lauern_nicht/index_ger.html	

System «Anyvision» zur
ware habe in der

DyRiAS-Schule

DyRiAS(dynamic system for the analysis of risk)-School assesses the behavior of (former) students regarding their risk potential. The focus lies exclusively on individuals of concern who have drawn attention to themselves by threats of violence or amok fantasies. It is used for risk assessment in cases of threats of amok or death threats by adolescents in a school context and assessment of the risk of suicide in adolescents. DyriAs requires practitioners to answer 30 "yes" or "no" questions about a student. The tool then outputs a score on a scale from one to five, from harmless to dangerous. According to the reference list of the provider, it is used in several cantons in Switzerland.

CANTON Glarus, Luzern, Solothurn, Zürich	HASHTAGS #crimeprevention #predictivepolicing
EMPLOYING ENTITY Police, Public administration, School / University	POLICY AREA Public Sector
AFFECTED GROUP Students	
PURPOSE Prediction of student violence risk	
LINKS https://www.dyriAs.com/en/module/school.html https://www.dyriAs.com/de/referenzen.html	

IPL Algorithm for Refugee Integration

Since 2018, the Swiss State Secretariat for Migration is testing a new, data-driven method for assigning asylum seekers to cantons across the country. The desired goal is to increase the employment rate of refugees. Asylum seekers in the pilot program are assessed by an algorithm designed to maximize their chances of individuals to the canton that y, as under the current Immigration Policy Lab Zurich, and an affiliated ed to give countries hat could easily be existing systems. The land included found work, and their gorithm uses this iloyment at each in a way that gives each

[remove](#) [#migration](#)

y Lab; CH/USA

[test-ai-refugee-integration/algorithmus-verbessert-](#)
[e-technologie-setzen-.html](#)



/ Résultats – et défis



/ Résultats – et défis

- Les systèmes algorithmiques
 - sont de plus en plus utilisés.
 - visent dans la plupart des cas l'augmentation de l'efficacité / l'optimisation des processus.
 - ont une base scientifique qui va de solide à fragile.
 - ne font souvent que des recommandations, mais prennent *de facto* des décisions.



Résultats – et défis

- Les systèmes algorithmiques
 - sont de plus en plus omniprésents
 - visent dans la plupart des cas à améliorer l'efficacité / l'optimisation des processus.
 - ont une base scientifique / mathématique.
 - ne font souvent que des recommandations, mais prennent *de facto* des décisions.

La société automatisée
est une **boîte noire**



/ Résultats – et défis

- Les systèmes algorithmiques
 - sont de plus en plus utilisés.
 - visent dans la plupart des cas l'augmentation de l'efficacité / l'optimisation des processus.
 - ont une base scientifique qui va de solide à fragile.
 - ne font souvent que des recommandations, mais prennent *de facto* des décisions.
- ont des effets sur **les individus & leurs droits fondamentaux**
- ont des effets sur **la société, la démocratie & le bien commun**

Wrongfully Accused by an Algorithm

In what may be the first known case of its kind, a faulty facial recognition match led to a Michigan man's arrest for a crime he did not commit.



How a Discriminatory Algorithm Wrongly Accused Thousands of Families of Fraud

Dutch tax authorities used algorithms to automate an austere and punitive war on low-level fraud—the results were catastrophic.

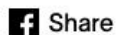


By [Gabriel Geiger](#)



Illustrated By [Cathryn Virginia](#)

March 1, 2021, 3:00pm



Share



Tweet



Snap



MORE LIKE THIS

[World News](#)

**The Dad Who
Years to Ban
'Paedophile
in Europe**

FRANK ANDREWS

01.14.22

[Tech](#)

**Vladimir Putin
About 'Cancel'**

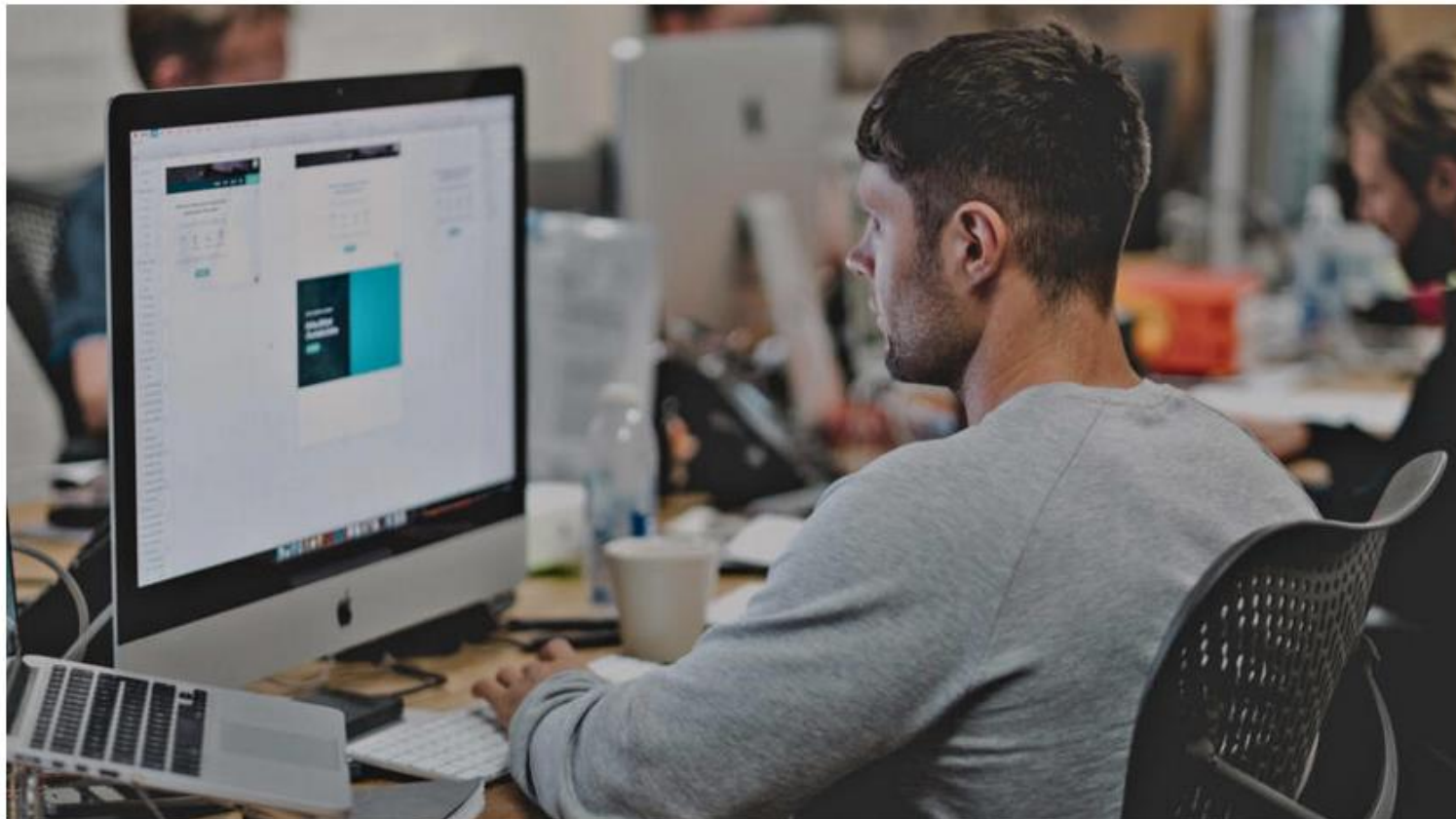
Source : vice.com

Microsoft 365

So überwachen Chefs eure Produktivität am Arbeitsplatz

Der Markt für Kontrollsoftware boomt. So bietet der Riesenkonzern Microsoft eine Produktivitätsbewertung für Unternehmen an, für die Microsoft eine Menge Daten sammelt. Kritiker:innen bezeichnen diesen Service als ein „Instrument zur Arbeitsplatzüberwachung“.

27.09.2021 um 12:46 Uhr - Rahel Lang - in Überwachung - 25 Ergänzungen



FACE IDENTIFIED

STRATFORD CE

BURGER KING



MATCH

Watchlist: Interpol
Match Confidence: 96%
First Name: KIERAN
Last Name: HERSI
Gender: MALE
Height: 1.82 metres
Charges: MURDER

4 OCTOBRE 2023 #ELECTIONS #GENERATIVEAI #IA #PUBLICSPHERE

Nouvelle recherche

Élections 2023 : comment ChatGPT, Bing Chat & Co. pourraient menacer la formation de l'opinion publique

S'informer sur comment voter à l'aide de modèles linguistiques d'IA tel que Bing Chat est une très mauvaise idée. C'est en effet ce que mettent en évidence les résultats d'une analyse menée par AlgorithmWatch et AI Forensics en collaboration avec la RTS et la SRF dans le cadre des élections fédérales en Suisse, ainsi que des élections régionales en Bavière et dans la Hesse. Ces grands modèles de langage apportent à des questions importantes des réponses parfois complètement erronées, en partie mensongères, incomplètes ou obsolètes. Bing Chat, ChatGPT et Co. peuvent ainsi constituer une menace pour la formation de l'opinion publique dans une démocratie.

ASIA

Behind the AI boom, an army of overseas workers in 'digital sweatshops'

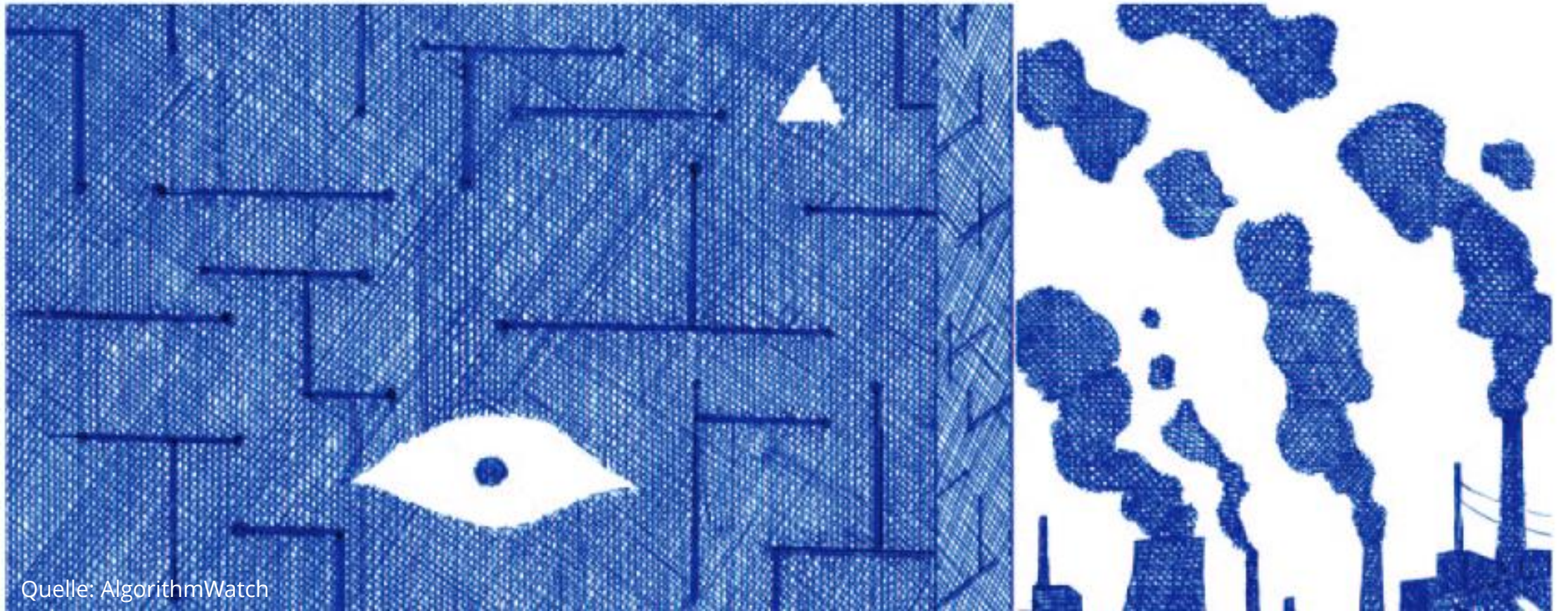
By [Rebecca Tan](#) and [Regine Cabato](#)

August 28, 2023 at 2:00 a.m. EDT



sustain

AI and the Challenge of Sustainability





Impacts sur les prérequis et les infrastructures de la démocratie

- Droits fondamentaux
- Formation de l'opinion publique et coexistence démocratique
- Protection des minorités
- Formation
- Durabilité – écologique | économique | sociale
- Équilibre des pouvoirs & dépendances
 - Justice



/ ... mais où sont les chances pour l'administration ?

- Neutralité et égalité de traitement ?
 - Les algorithmes ne sont pas neutres
 - Ils peuvent identifier des schémas (injustes) > Est-ce un objectif défini comme tel ?
- Facilité d'utilisation & Service Delivery
 - Pour qui est-ce plus simple d'utilisation ? Y a-t-il un compromis sur la traçabilité ?
- Amélioration de l'efficacité, optimisation des processus & réduction des coûts
 - Où s'agit-il d'améliorer l'efficacité / où s'agit-il de prendre des décisions sur des personnes ?
 - Est-ce vraiment un gain d'efficacité ? (Problème – solution)
 - Est-ce vraiment une réduction des coûts? (Snake Oil, solutions Top-down)
- Durabilité
 - Qu'en est-il de la durabilité ?
- **Évaluation au cas par cas !**



/ Et maintenant ? Que faire ?



/ Et maintenant ? Que faire ?

Définir des conditions-cadre (Governance) pour le **développement** et l'**utilisation** des systèmes.

- Transparence → Traçabilité
- Surveillance → Contrôle
- Responsabilité → Obligation de rendre des comptes

- Notamment : conditions-cadre **légales**
 - Appliquer/interpréter les lois existantes & adopter de nouvelles lois



Conditions-cadre dans l'administration publique

- **Transparence** comme premier pas → **Traçabilité** pour les personnes concernées & la société
- Par ...
 - ... **Registres publics** des systèmes de prise de décision algorithmique utilisés dans l'administration publique
 - ...

Konzeptpapier

Ein KI-Transparenzregister
für die öffentliche Verwaltung





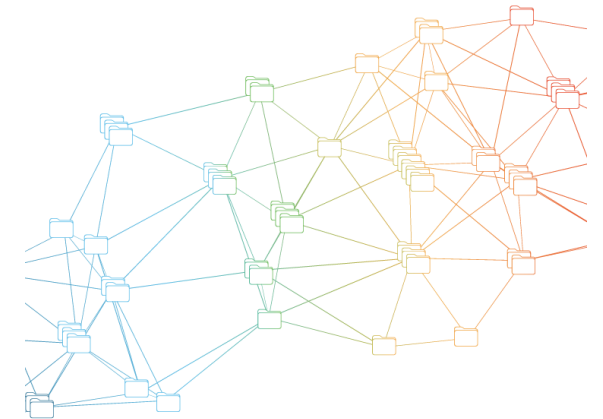
Conditions-cadre dans l'administration publique

- **Surveillance et contrôle individuels & démocratiques** par
 - ... **Analyses d'impact** obligatoires lors de l'utilisation de systèmes de prise de décision algorithmique
 - ... Inclusion des employé·e·s et personnes concerné·e·s



Automatisierte Entscheidungssysteme
im öffentlichen Sektor
**Ein Impact-Assessment-Tool für
die öffentliche Verwaltung**

Von Michele Loi, in Zusammenarbeit mit Anna Mätzener,
Angela Müller und Matthias Spielkamp





Conditions-cadre dans l'administration publique

- Garantir la **responsabilité** et l'**obligation de rendre des comptes**
 - › ... ~~Computer says no~~
- Notamment en
 - ... protégeant contre la discrimination algorithmique.
 - ... garantissant des voies de recours pour les personnes concerné·e·s.
 - ... définissant des lignes rouges.





/ Que font les autres ?



EU :

EU **AI Act** –
Règlement sur
l'IA




Conseil de l'Europe : Convention sur l'IA

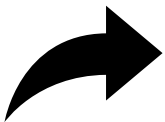


... et Berne ?



/ Berne – État actuel

- Pas d'espace de non-droit : lois existantes (cantonales/nationales, sectorielles, ...)
- Conseil fédéral / administration :
 - Capacity-Building (CNAI, DSCC, BAKOM, EDA, BJ, ...)
 - Négociations Conseil de l'Europe
 - Analyse des besoins de réglementation jusqu'à fin 2024
- Parlement : Interventions 
- Cantons & communes
- De Bruxelles via Strasbourg vers Berne
 - AI Act, Agenda-setting & Standard-setting



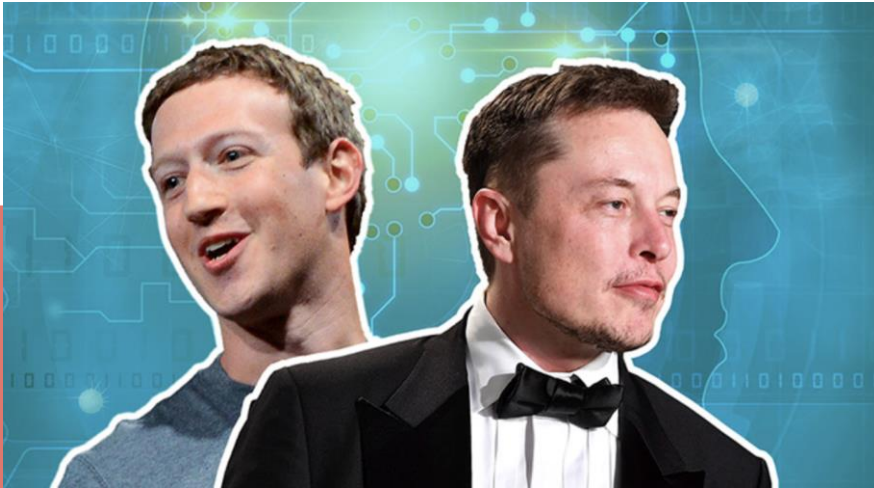


/ Give-Aways

When Your Boss Is an Algorithm

For Uber drivers, the workplace can feel like a world of constant surveillance, automated manipulation and threats of “deactivation.”

Oct. 12, 2018





Les algorithmes et l'IA deviennent-ils une menace pour la société?

- En tout cas, pas parce qu'ils vont anéantir l'humanité.
- Les défis sont multiples
 - C'est une question de justice.
- Les chances pourraient aussi être multiples.



/ Devrions-nous utiliser des algorithmes et de l'IA ?

- Oui.
 - Nous devrions les utiliser là où ils nous profitent et de manière à ce qu'ils nous profitent (à toutes et à tous).
- Mais : les algorithmes et l'IA ...
 - ... sont un outil conçu et utilisé par des individus.
 - ... ne sont pas un but en soi.
 - ... ne sont pas une force de la nature.



Et comment s'assurer que les algorithmes et l'IA nous soient vraiment utiles ?

- En
 - réfléchissant,
 - dans quels domaines, nous voulons utiliser des algorithmes et l'IA.
 - avec quel objectif ils sont utilisés.
 - évaluant de manière réaliste,
 - la contribution des algorithmes et de l'IA à la réalisation de cet objectif.
 - quelles seront les conséquences pour les personnes concernées
 - fixant des conditions-cadre pour le développement et l'utilisation des systèmes
 - étant conscient·e·s de **l'économie politique autour de l'IA**



/ We should talk!

- algorithmwatch.ch | algorithmwatch.org
- algorithmwatch.ch/de/newsletter
- @AlgorithmWat_CH
- info@algorithmwatch.ch

