

PAPIER DE POSITION DU PS SUISSE

RÉGLEMENTATION DES SYSTÈMES D'IA

Adopté lors de la séance en ligne du Conseil de parti
du 6 septembre 2024



Table des matières

Introduction	3
Points clés pour une réglementation sociale-démocrate de l'IA et d'autres systèmes algorithmiques	6
1. Dans l'intérêt des personnes, pas des entreprises	6
2. Les droits fondamentaux doivent être respectés/l'éthique et les analyses d'impact sont obligatoires 6	6
3. Engagement contre la discrimination.....	6
4. Une réglementation neutre sur le plan technologique.....	6
5. Transparence et traçabilité, obligation de déclaration	7
6. Protection contre la manipulation/Promouvoir la littératie médiatique et algorithmique.....	7
7. Devoir de diligence	8
8. L'IA et l'avenir du travail	8
9. Applications indésirables et interdites.....	9
10. Développement de systèmes d'IA ouverts.....	9
11. Utilisation profitable à la société	10
Revendications politiques	11
Glossaire/termes	13
Sources	15
Interventions du PS	16

INTRODUCTION

Le développement de l'intelligence artificielle progresse rapidement et les applications correspondantes sont utilisées dans un nombre croissant de domaines¹. En quelques secondes, des outils accessibles au public comme ChatGPT, Midjourney ou Soundraw créent des textes, des images ou des morceaux de musique. Ils sont utilisés pour l'apprentissage, facilitent parfois des processus de travail, créent de l'art.

Après l'enthousiasme initial suscité par cette nouvelle technologie, de plus en plus de questions sociales et politiques s'imposent : que signifie, pour l'avenir de la formation de l'opinion politique et de la participation, le fait que la propagande puisse être rédigée en quantités illimitées presque sans effort ? Quel est l'effet des algorithmes des plateformes utilisant l'intelligence artificielle sur les informations que les gens consomment ? Quel est l'impact de l'utilisation de grands modèles de langage et d'images comme ChatGPT ou Midjourney sur le monde du travail, par exemple pour les professions du secteur créatif ? Qu'en est-il des droits d'auteur, de la protection des données et de l'obligation de rendre des comptes ? Comment prévenir les discriminations lorsqu'une intelligence artificielle avec des biais est utilisée dans les processus de recrutement ? Comment peut-on s'opposer aux décisions prises par les systèmes d'IA, par exemple lors de l'attribution d'un emploi ou d'un crédit ? Quelles personnes et organisations créent, déterminent et contrôlent les systèmes, algorithmes et métriques utilisés ? Comment peut-on limiter l'énorme consommation d'électricité et de ressources des systèmes d'IA ? En même temps, les systèmes algorithmiques ont aussi un énorme potentiel : comment l'utiliser pour apporter des améliorations concrètes à la société — par exemple pour répondre aux défis écologiques de notre époque ou pour alléger la charge de travail ?

Une lettre ouverte a récemment mis en garde contre la puissance insoupçonnée de la technique. Elle a recueilli à ce jour près de trente mille signatures, dont celles d'Elon Musk et de nombreux chercheurs en IA de renom². Elle appelle à suspendre immédiatement le développement de grandes IA pendant au moins six mois, car des systèmes comme le modèle linguistique ChatGPT-4 sont devenus trop puissants et trop dangereux. Des « risques fondamentaux pour la société et l'humanité » sont menacés par des « IA agissant au niveau humain ». Tant que l'on ne se sera pas mis d'accord sur la manière de régler cela, tous

¹ Selon la définition de l'OCDE, un système d'IA est un « système à base de machines capable de faire des prédictions, de formuler des recommandations ou de prendre des décisions qui influencent l'environnement réel ou virtuel pour des objectifs définis par l'homme ». Le terme "intelligence artificielle" doit toutefois être utilisé avec prudence, car le mot « intelligence » suscite des attentes et des craintes. Dans ce document, il sera question d'une multitude de systèmes qui prennent des décisions automatisées, font des recommandations, émettent des avis ou font des prédictions. Dans son document de synthèse, la *Digital Society Initiative* parle de ce que l'on appelle les « systèmes algorithmiques », *AlgorithmWatch* et la *Digital Society* parlent à leur tour de systèmes décisionnels automatisés (SMA). Un terme large tel que « systèmes algorithmiques » englobe également des applications basées sur d'autres technologies. C'est donc ce terme qui sera majoritairement utilisé dans la suite de ce document.

² <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> [en anglais]. La lettre a également fait l'objet de critiques. L'institut FLI (Future of Life Institute), à l'origine du projet, défend des idéologies controversées, attise un engouement pour l'IA et détourne l'attention des problèmes et des défis réels en se focalisant sur une future superintelligence.

les laboratoires d'IA devraient renoncer à poursuivre leurs recherches³. De nombreux chercheurs de renom, dont certains sont même cités dans la lettre, ont vivement critiqué la lettre et l'institut à l'origine de la lettre^{4,5}.

À juste titre : alors que la crainte que les textes et images générés par l'IA n'inondent les canaux d'information de contre-vérités et de propagande est tout à fait justifiée, la lettre est par ailleurs portée par des fantasmes apocalyptiques sur le remplacement total de l'homme par des machines et « la perte de contrôle de notre civilisation ». Mais ce qui nous attend réellement avec de grands modèles linguistiques comme ChatGPT n'est pas tant une catastrophe *technique* d'ordinateurs malveillants. Les modèles linguistiques menacent bien plus concrètement de devenir un désastre *démocratique* — par la privatisation des technologies linguistiques en tant que futur lieu public politique. C'est précisément là que la politique et la société civile entrent en jeu — et c'est à elles de jouer.

Les développements de ces dernières années l'ont montré : un système d'IA est d'autant plus performant qu'il est alimenté par un grand nombre de données. Mais plus la taille est importante, plus les coûts sont exorbitants. Comme le montre une étude de l'université de Stanford sur les nouvelles IA vocales, la course aux modèles toujours plus volumineux a entre-temps conduit à ce que seule une poignée d'entreprises soient encore en lice — outre OpenAI - le développeur de ChatGPT, citons encore Meta ou Google. Les petites entreprises non commerciales et les universités ne jouent pratiquement plus aucun rôle dans l'atteinte de nouveaux records de taille. Nous sommes ainsi confrontés à un nouvel oligopole qui concentre les technologies linguistiques entre les mains de quelques entreprises privées. Mais la manière dont ces systèmes sont alimentés, avec quelles données et ce qu'ils produisent en fin de compte est une décision éminemment politique. Et cette décision ne doit pas être laissée à quelques entreprises qui ne sont soumises à aucun contrôle démocratique et qui ne rendent de comptes à personne.

Un contrôle et une régulation démocratiques des systèmes d'IA s'imposent au vu de l'évolution récente et des performances des nouveaux systèmes comme GPT4. De nombreuses réglementations sur l'utilisation de l'IA ont été adoptées récemment : aux États-Unis, le gouvernement a présenté en octobre 2022 le « AI Bill of Rights » sur l'utilisation responsable de l'intelligence artificielle⁶ ; le 13 mars 2024, le Parlement européen a donné son feu vert à la loi sur l'intelligence artificielle (« EU AI Act »), qui est entrée en vigueur dans toute l'UE au début du mois de juillet 2024⁷ ; le Conseil de l'Europe a adopté récemment, le 17 mai 2024,

³ Voir aussi : <https://netzpolitik.org/2023/offener-brief-zu-ki-opfer-des-hypes/> [en allemand]

⁴ <https://www.dair-institute.org/blog/letter-statement-March2023> [en anglais]

⁵ <https://www.heise.de/meinung/Welche-Ideologie-hinter-Geoffrey-Hintons-Warnungen-steckt-8988646.html>
[en allemand]

⁶ <https://www.whitehouse.gov/ostp/ai-bill-of-rights/> [en anglais]

⁷ <https://www.consilium.europa.eu/fr/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai/>

une convention sur l'IA visant à garantir une utilisation responsable de l'IA⁸. Dès octobre 2022, le Parlement européen et le Conseil européen ont en outre adopté de nouvelles réglementations de grande envergure pour le secteur de l'information et du numérique avec le « Digital Markets Act » et le « Digital Services Act ».

La Suisse a également fait les premiers pas vers une réglementation. Ainsi, le Conseil fédéral a récemment adopté un rapport et des lignes directrices pour l'administration fédérale⁹. En outre, le Conseil fédéral a annoncé l'élaboration d'un projet de consultation pour la réglementation des plateformes de communication (pour la reprise du « EU Digital Services Act »)¹⁰. Un premier état des lieux de la réglementation des plateformes et de l'IA devrait être disponible fin 2024 ou début 2025. La manière dont le Digital Services Act de l'UE sera repris en Suisse n'est pas encore définie. Les premiers indices laissent entrevoir une reprise affaiblie du DSA – le libre accès aux données pour la recherche, par exemple, n'apparaît pas dans le projet, ce qui pourrait être dû au fort lobbying des entreprises tech – en particulier de Google. La Suisse s'est également montrée réticente jusqu'à présent sur le thème de l'IA. Début 2022 encore, le Conseil fédéral a répondu à une intervention de la conseillère nationale PS Min Li Marti qu'il n'y avait pas besoin de loi sur l'intelligence artificielle. En ce qui concerne les réglementations internationales, la Suisse veut, selon le mot d'ordre du DFAE dans son rapport sur l'IA, participer à l'élaboration des règles internationales en matière d'intelligence artificielle « dans son intérêt ». Jusqu'à présent, cela signifiait surtout une chose : une mise en œuvre *laisser-faire* dans le sens des acteurs de l'économie de marché.

D'un point de vue social-démocrate, la régulation des systèmes algorithmiques est fondamentale. Les systèmes d'intelligence artificielle touchent aux valeurs et aux préoccupations fondamentales de notre parti : l'engagement pour les droits fondamentaux et les droits de l'homme, pour un travail de qualité et pour le maintien et le renforcement d'une sphère publique politique démocratique, la lutte contre la discrimination et la prise de position en faveur des personnes plutôt que des profits des entreprises. C'est la raison pour laquelle le présent document formule des principes et des revendications concrètes concernant l'utilisation des systèmes algorithmiques.

⁸ Voir "Council of Europe Framework Convention on artificial intelligence and human rights, democracy, and the rule of law": https://www.coe.int/fr/web/portal/full-news/-/asset_publisher/y5xQt7QdunzT/content/id/267650696?com_liferay_asset_publisher_web_portlet_AssetPublisherPortlet_INSTANCE_y5xQt7QdunzT_languageId=en_GB [en anglais]

⁹ <https://www.admin.ch/gov/fr/accueil/documentation/communiqués.msg-id-81319.html>

¹⁰ <https://www.admin.ch/gov/fr/accueil/documentation/communiqués.msg-id-94116.html>

POINTS CLÉS POUR UNE RÉGLEMENTATION SOCIALE-DÉMOCRATE DE L'IA ET D'AUTRES SYSTÈMES ALGORITHMIQUES

1. Dans l'intérêt des personnes, pas des entreprises

Les intérêts des personnes et du bien commun doivent être au cœur du développement technologique. Le développement technique n'est pas une fin en soi ni un phénomène naturel, mais un moyen d'atteindre une fin, dans le but d'améliorer la vie des gens et de faciliter l'égalité des chances dans l'accès aux ressources. Il existe et a toujours existé des technologies qui ont été fortement contrôlées ou même limitées dans l'intérêt de l'humanité (p. ex. médecine humaine, technologie nucléaire, etc.). Il est de notre devoir, en tant que société démocratique, de définir quelles technologies nous voulons utiliser et à quelles fins.

2. Les droits fondamentaux doivent être respectés/l'éthique et les analyses d'impact sont obligatoires

Un développement technologique dans l'intérêt des personnes signifie également un développement technologique qui respecte les droits fondamentaux. Ceux-ci doivent être respectés tant dans le développement que dans l'application. Les considérations éthiques doivent être intégrées dès le début du processus, les conséquences doivent être examinées et évaluées.

3. Engagement contre la discrimination

Les systèmes de décision automatisés comportent toujours un risque de discrimination. Dans différents cas, il a déjà été prouvé que des systèmes prétendument neutres et objectifs ont conduit à des comportements discriminatoires¹¹. Les décisions prises par des êtres humains ne sont jamais exemptes de discrimination et de préjugés. Il est donc d'autant plus important que les décisions automatisées soient transparentes, compréhensibles, surveillées et qu'elles puissent faire l'objet d'un recours. Il est en principe judicieux de réglementer la discrimination de manière générale, indépendamment du fait que les décisions soient prises par des hommes ou des machines. Dans ce domaine, l'utilisation de systèmes automatisés peut également présenter des avantages : ils peuvent être étudiés et leur biais n'est pas consolidé par des années de socialisation, mais peut être corrigé par des adaptations du système. Le développement de l'IA explicable et des ensembles de données équilibrés peut aider dans ce domaine.

4. Une réglementation neutre sur le plan technologique

La réglementation doit être neutre sur le plan technologique, car les défis doivent être réglementés indépendamment de la technologie utilisée. En raison de l'évolution rapide, les ré-

¹¹ Voir par exemple <https://fra.europa.eu/en/publication/2022/bias-algorithm> ou <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> [en anglais]

gements ne peuvent être durables que si elles peuvent être appliquées indépendamment de certaines technologies. Néanmoins, des champs d'application spécifiques peuvent être désignés et réglementés individuellement (p. ex. Predictive Policing).

5. Transparence et traçabilité, obligation de déclaration

L'utilisation de systèmes algorithmiques doit être reconnaissable et compréhensible. Cela vaut en particulier dans le cas où une décision automatisée concerne un individu, comme par exemple dans le système judiciaire ou dans les procédures de candidature. Dans ce contexte, la transparence n'est pas le seul critère décisif, la traçabilité des décisions l'est tout autant. Cela implique la publication des données d'apprentissage et des paramètres utilisés pour l'apprentissage dans des répertoires et des registres publics. Les systèmes doivent être mis à disposition pour être vérifiés. Lorsque les autorités utilisent des systèmes algorithmiques, les personnes concernées doivent avoir la possibilité de faire appel et d'obtenir une nouvelle évaluation par un être humain. Pour cela, elles doivent avoir accès aux informations nécessaires. Pour l'utilisation de tels systèmes dans le secteur privé, il faut également une possibilité de recours. Pour toutes les utilisations impliquant des décisions significatives, une évaluation obligatoire des systèmes est nécessaire afin de contrecarrer les éventuels effets discriminatoires de ces systèmes et de leur utilisation. Les instituts de recherche publics ou les acteurs de la société civile doivent mettre à disposition de tels systèmes d'évaluation et les adapter en permanence aux connaissances les plus récentes.

6. Protection contre la manipulation/Promouvoir la littératie médiatique et algorithmique

La liberté d'information et la liberté d'expression sont des droits fondamentaux garantis par la Constitution et des piliers de la formation démocratique de l'opinion et de la volonté. Une restriction de la diffusion de contenus ne doit donc s'accompagner que de la plus grande prudence possible. Parallèlement, il existe des risques de manipulation et de désinformation qui peuvent encore être renforcés par les systèmes d'IA. Pour qu'une société fonctionne, il faut donc aussi savoir à quelles informations on peut se fier, comment elles sont obtenues et pourquoi elles sont affichées à un moment donné. Cela implique de mettre l'accent sur l'éducation et le renforcement de l'éducation aux médias ainsi que sur la promotion des médias afin de maintenir un paysage médiatique large et indépendant.

Les systèmes d'IA jouent également un rôle de plus en plus important dans la diffusion de contenus médiatiques. Les systèmes d'IA déterminent individuellement qui peut voir quelle vidéo ou quel article sur YouTube, TikTok, Instagram ou X. Et ce, dans le but de lier les gens le plus étroitement et le plus longtemps possible à la plateforme concernée et de pouvoir diffuser un maximum de publicité personnalisée. La sélection algorithmique favorise à son tour l'émergence de publics partiels et de pseudo-communautés isolées les unes des autres — un terrain idéal pour la politique ressentimentale, comme le prouvent les nombreuses campagnes de droite de ces dernières années. Alors que les opinions se radicalisent, les contenus contradictoires disparaissent du flux. Dans un cas particulièrement frappant, cette

dynamique — alimentée par l’algorithme de Facebook — a contribué à favoriser le génocide des Rohingyas au Myanmar¹².

Si des systèmes sont utilisés pour filtrer et pondérer les contenus par des algorithmes, cela doit être rendu public. Les utilisateurs doivent avoir la possibilité de définir eux-mêmes les paramètres qui déterminent ce qu’ils voient. Ils doivent notamment pouvoir voir les contributions qui ne correspondent pas au schéma calculé pour eux par le système.

7. Devoir de diligence

Les systèmes algorithmiques peuvent causer des dommages et l’ont déjà fait concrètement¹³. Cela signifie que des questions centrales de devoir de diligence se posent ici. Il est important de noter que les systèmes doivent être réglementés tout au long du cycle de vie du produit. Il faut un devoir de diligence aussi bien pour les développeurs que pour ceux qui les utilisent — sans qu’ils puissent échapper à cette responsabilité, par exemple par une simple exclusion de responsabilité. Il doit également être clair que la sécurité est garantie avant qu’une application ne soit mise sur le marché et que les fabricants et les fournisseurs doivent répondre d’éventuels dommages.

8. L’IA et l’avenir du travail

Les technologies de contrôle automatique du « tracking, tracing, targeting, ranking, scoring et profiling » ont depuis longtemps envahi le monde du travail. Dans les métiers de l’industrie et des services, la composante numérique réside principalement dans le contrôle algorithmique du travail, qui vise à contrôler les processus de travail et à extraire des données¹⁴. Les processus doivent ainsi être simplifiés et automatisés. L’utilisation de systèmes d’IA dans le monde du travail ne doit cependant pas conduire à une plus grande précarisation et déqualification du travail. Au contraire, l’objectif doit être d’utiliser des systèmes algorithmiques pour faciliter le travail et redistribuer les gains de productivité ainsi obtenus aux personnes.

Aujourd’hui encore, de nombreuses fonctions qui semblent automatisées sont effectuées par des êtres humains. Pour modérer le contenu, les grandes entreprises de plateformes et les entreprises de développement d’IA font appel à des clickworkers qui effectuent des tâches souvent traumatisantes dans des conditions de travail abusives, en marquant et en supprimant quotidiennement la violence et les abus de ces plateformes.

¹² <https://www.amnesty.org/en/latest/news/2022/09/myanmar-facebooks-systems-promoted-violence-against-rohingya-meta-owes-reparations-new-report/> [en anglais]

¹³ Voir à ce sujet le scandale des droits aux allocations familiales aux Pays-Bas : <https://netzpolitik.org/2021/kindergeldaffaere-niederlande-zahlen-millionenstrafe-wegen-datendiskriminierung/> [en allemand]

¹⁴ Cf. Simon Schaupp, *Technopolitik von unten. Algorithmische Arbeitssteuerung und kybernetische Proletarisierung*, Berlin 2022. Si le travail consiste de plus en plus à suivre les instructions de systèmes de contrôle algorithmiques, la main-d’œuvre qualifiée devient superflue. Du point de vue des entreprises, cela permet non seulement de supprimer la longue formation initiale, mais aussi de faciliter l’intégration de la main-d’œuvre précaire et migrante.

Avec l'utilisation accrue des systèmes d'IA dans le monde du travail, l'importance de la qualification et de la garantie de l'employabilité augmente. La promotion et le renforcement ciblés de la formation initiale et de la formation continue des travailleur-euses et des demandeur-euses d'emploi permettent de lutter contre la déqualification et la précarisation croissante.

La promotion de la formation continue ne doit ici pas rester l'apanage des entreprises privées : elle doit devenir une responsabilité de la société dans son ensemble. Il faut mettre en œuvre des mesures politiques garantissant que tou-tes les travailleur-euses, soutenu-es par les efforts conjoints des secteurs public et privé, bénéficient d'une formation initiale et d'une formation continue adéquates. C'est là la seule façon d'assurer une transition équitable vers l'ère de l'intelligence artificielle.

9. Applications indésirables et interdites

Il convient d'examiner clairement s'il existe des applications qui ne sont pas souhaitables, voire clairement nuisibles, et qui doivent donc être interdites. Des exemples possibles sont les systèmes de reconnaissance biométrique (comme la reconnaissance faciale) à des fins d'identification dans les lieux publics. L'identification peut se faire non seulement sur la base du visage, mais aussi de la voix ou de la démarche. D'autres systèmes pour lesquels une interdiction (éventuellement assortie d'exceptions) doit être envisagée : la reconnaissance des émotions (éventuellement avec une exception pour les personnes en situation de handicap), la catégorisation biométrique sur la base de données sensibles, certaines applications dans le domaine de la police et de la migration.

10. Développement de systèmes d'IA ouverts

Les systèmes d'IA peuvent contribuer à améliorer la vie des gens. Grâce aux outils de traduction, par exemple, davantage de personnes ont accès à des informations dans une langue qu'elles comprennent ; les algorithmes aident à gérer la production et la consommation d'électricité ; les systèmes de TAL aident à reconnaître et à combattre les *discours haineux* sur Internet. Pour que les systèmes d'IA puissent déployer leur potentiel social et émancipateur, les outils doivent toutefois être aussi simples que possible et, surtout, ouverts et disponibles dans une bonne qualité. Les instituts de recherche publics doivent donc soutenir le développement de technologies ouvertes et assurer la curation et la mise à disposition de données ouvertes.

Pour contrer l'oligopole qui se dessine dans le développement de l'IA, il faut une recherche publique forte. C'est le seul moyen de garantir la transparence et des faits vérifiables. Les universités et autres institutions publiques doivent recevoir des moyens financiers suffisants pour jouer un rôle de premier plan dans l'amélioration de la transparence dans le développement des systèmes d'IA. Actuellement, de nombreux modèles publiés aujourd'hui (comme Gemini de Google, ChatGPT d'OpenAI, etc.) ne permettent pas de mener des enquêtes vérifiables, car les groupes ne rendent pas leurs données publiques et échappent au *processus d'évaluation par les pairs*. Le développement de systèmes d'IA ne doit pas servir les intérêts de profit de certaines entreprises, mais la société dans son ensemble.

11. Utilisation profitable à la société

Les systèmes algorithmiques, en particulier les systèmes d'IA, offrent également de belles possibilités pour l'avènement d'une société plus juste. Les systèmes d'IA peuvent par exemple être utilisés dans le domaine médical pour faciliter les processus de travail et soulager les médecins et le personnel soignant. Un dépistage précoce approprié des maladies, une évaluation des cas d'urgence et une préévaluation des problèmes de santé permettent en outre d'éviter des consultations médicales inutiles et de décharger le système de santé.

REVENDEICATIONS POLITIQUES

- Reprise intégrale de l'« EU AI Act »
- Loi-cadre pour la régulation des systèmes algorithmiques
- Création d'un comité national d'éthique pour l'intelligence artificielle
- Loi générale contre la discrimination, y compris les mesures spécifiques contre la discrimination algorithmique
- Évaluations d'impact obligatoires dans les domaines de la technologie, de la réglementation, de l'égalité des genres, de la protection des données et de la société
- Obligation de transparence et de traçabilité lors de l'utilisation de systèmes algorithmiques, obligation de déclaration, tant dans le secteur public que dans le secteur privé (registre), divulgation des codes sources, des données d'entraînement et des paramètres
- Obligation de prévoir des possibilités de recours (droit à l'expertise humaine)
- Création d'une autorité de surveillance et de contrôle
- Investissements dans la recherche et le développement de technologies d'IA ouvertes, neutres pour le climat et efficaces sur le plan énergétique
- Introduire une réglementation environnementale des centres de traitement de données (chaleur résiduelle, efficacité énergétique, composants à longue durée de vie, économie circulaire)
- Recherche sur les systèmes permettant d'identifier et de corriger les biais dans les systèmes d'IA, tels que l'IA explicable et les ensembles de données représentatifs et curés.
- Réglementation sur la responsabilité du fait des produits et les possibilités de sanctions, notamment en ce qui concerne le devoir de diligence : ni les fabricants ni les fournisseurs ne peuvent céder entièrement leur responsabilité aux utilisateur-trices ou aux client-es.
- Promouvoir l'éducation et la compétence dans l'utilisation des systèmes algorithmiques
- Promotion d'un paysage médiatique large et diversifié
- Éliminer la discrimination dans l'accès aux nouvelles technologies/garantir l'équité en matière d'éducation numérique/garantir les compétences numériques de base de tous les individus, notamment par des offres de formation continue largement accessibles et revisitées en fonction des évolutions technologiques.
- Des règles nationales claires et uniformes sur l'utilisation de l'IA à tous les niveaux d'enseignement

-
- Renforcement de la protection de la jeunesse et des consommateur-trices (p. ex. adaptation de la loi sur la concurrence déloyale, adaptation de la protection de la jeunesse pour les plateformes en ligne entraînant une dépendance), réglementation séparée des plateformes de communication
 - Interdiction de la reconnaissance biométrique dans l'espace public
 - Interdiction du développement et de l'utilisation de systèmes d'armes autonomes, par analogie avec les réglementations relatives aux armes biologiques et chimiques. Engagement international en faveur de l'interdiction de systèmes d'armes de ce type.
 - Les approches réglementaires devraient intervenir dès le stade du développement d'un système d'IA
 - Imposition équitable des entreprises tech et prévention des monopoles grâce à une loi sur la concurrence efficace
 - Répartition équitable des gains de productivité dans l'ensemble de la population : financièrement et/ou sous forme de réduction du temps de travail
 - Interdiction de la surveillance sur le lieu de travail
 - Interdiction de la prévision policière (traduction française plus ou moins officielle de l'anglais « predictive policing »)
 - Développer une stratégie nationale pour contrer les effets négatifs potentiels de l'IA sur le marché du travail : reconversion, imposition en cas de baisse éventuelle du nombre de travailleur-euses, revenu de base inconditionnel
 - La Confédération examine l'impact des systèmes d'intelligence artificielle sur la cybersécurité et la sécurité en général

GLOSSAIRE/TERMES

AI/KI, Intelligence artificielle

Système qui, sur la base d'une entrée, génère une sortie qui semble « intelligente ». Il peut s'agir d'un système « entraîné » par l'apprentissage automatique ou d'un système basé sur des algorithmes classiques.

Apprentissage automatique, ML (« *Machine Learning* »)

Système qui apprend grâce à des données d'exemple. Il existe de nombreuses architectures différentes, dont la complexité varie fortement, par ex. : régression linéaire, filtres bayésiens, réseaux neuronaux, transformateurs, LSTM, diffuseurs, chaînes de Markov, U-Net,...

Apprentissage profond (« *Deep Learning* »)

Grand système de « *Machine Learning* » entraîné avec de très grandes quantités de données. Ce que l'on entend exactement par apprentissage profond évolue dans le temps. Les systèmes qui étaient encore du « *Deep Learning* » il y a quelques années sont aujourd'hui relativement petits.

LLM, Modèles linguistiques étendus (« *Large Language Models* »)

Grands systèmes d'intelligence artificielle entraînés à comprendre et à générer du langage humain. Ils sont « larges » (grands) en termes de quantité de données d'entraînement et de taille du réseau neuronal sous-jacent. Les modèles de la famille GPT, développés par OpenAI, en sont des exemples.

Apprentissage supervisé

Un système ML entraîné avec un ensemble de données contenant à la fois des données d'entrée et de sortie. Une régression linéaire ou un classificateur en sont des exemples.

Apprentissage non supervisé

Un système qui s'entraîne sur des données qui ne possèdent pas d'étiquettes. Des exemples sont par exemple un algorithme de « *clustering* » ou un transformateur de texte.

AGI, Intelligence générale artificielle

Intelligence artificielle capable d'apprendre de nouvelles compétences sans modifier l'architecture. Par exemple, GPT-4 apprend à conduire une voiture.

Biais

Dans ce contexte : les systèmes d'apprentissage automatique sont entraînés avec des données dans lesquelles un biais est déjà présent. Les systèmes adoptent ensuite ce biais. Cela peut être dû au fait que certains points de données sont fortement sous-représentés ou que les étiquettes dans un modèle d'apprentissage supervisé sont faussées par des préjugés humains. Les discriminations systémiques sont ainsi transposées dans les modèles.

Systèmes génératifs

Il y a actuellement un grand engouement pour les systèmes génératifs, comme GPT ou Dall-E, Stable Diffusion, etc. Surtout autour de GPT-4, il y a un battage médiatique non fondé autour de « l'urgence », c'est-à-dire un nouveau comportement qui n'a pas été appris. GPT-4 n'a pas de conscience propre. Cette impression est due au fait que les textes générés donnent l'impression aux gens qu'il y a une conscience.

Le comportement émergent ne peut pas être vérifié, car l'ensemble des données n'est pas public. La plupart des tâches devraient se retrouver d'une manière ou d'une autre dans le jeu de données avec lequel les systèmes ont été entraînés.

Systèmes de recommandation

Systèmes utilisés sur les réseaux sociaux ou les plateformes vidéo pour afficher le contenu en fonction des utilisateur-trices. Ils se basent sur divers signaux, comme les abonné-es ou les interactions d'autres utilisateur-trices similaires. Twitter a publié son algorithme, d'autres plateformes utilisent des systèmes similaires.

Il s'agit généralement une combinaison d'algorithmes classiques et d'apprentissage automatique.

Modèle multimodal

Les modèles multimodaux sont des modèles qui ont été entraînés par l'apprentissage automatique (machine learning) à comprendre et à générer différentes modalités d'informations, par exemple des images, des vidéos et du texte. Exemples : Midjourney, Stable Diffusion ou DALL-E d'OpenAI.

Artificial Neural Network, ANN/Réseaux de neurones artificiels), RNA

Les réseaux de neurones artificiels sont des programmes ou des modèles basés sur l'apprentissage automatique (machine learning) qui s'inspirent de la structure du cerveau humain. Leur structure permet un processus d'identification, d'évaluation et de déduction (à partir) des informations inspiré de la biologie.

SOURCES

<https://www.digitale-gesellschaft.ch/uploads/2022/02/Position-der-Digitalen-Gesellschaft-zur-Regulierung-von-automatisierten-Entscheidungssystemen-1.0.pdf> [en allemand]

<https://algorithmwatch.org/de/regulierung-general-purpose-ai-ki-verordnung/> [en allemand et en anglais]

<https://www.dsi.uzh.ch/dam/jcr:3a0cb402-c3b3-4360-9332-f800895fdc58/dsi-strategy-lab-21-de.pdf> [en allemand]

<https://www.sbf.admin.ch/sbf/fr/home/politique-fri/fri-2021-2024/themes-transversaux/numerisation-fri/intelligence-artificielle.html>

https://chplusplus.org/decisions_automatisees/

INTERVENTIONS DU PS

18.4037 Motion Bendahan

Centre de compétence pour l'intelligence artificielle dans l'administration fédérale

<https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20184037>

20.4338 Postulat Storni

L'intelligence artificielle. Règles de sécurité, transparence et information dans les applications d'apprentissage automatique

<https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20204388>

16.3320 Postulat Marti

Rapport sur les chances et les risques de l'intelligence artificielle et de la robotique

<https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20163320>

19.4462 Interpellation Graf-Litscher

L'éthique numérique. Prévenir la discrimination en matière d'intelligence artificielle

<https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20194462>

22.1051 Question Marti

Suivi de projets d'intelligence artificielle

<https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20221051>

21.4406 Postulat Marti

Rapport sur la réglementation des systèmes décisionnels automatisés

<https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20214406>