



Les nouvelles stratégies de la désinformation :
les défis démocratiques à l'ère de l'IA

Divina Frau-Meigs
Chaire Unesco Savoir Devenir
Association Savoir Devenir



Organisation
des Nations Unies
pour l'éducation,
la science et la culture



Chaire UNESCO Savoir-devenir
à l'ère du développement numérique durable
Université Sorbonne Nouvelle, France

Sommaire

Sommaire:

- 1/ Savoir Devenir et translittératie
- 2/ Désinformation et ses mutations
- 3/ Désinformation augmentée par l'IA

- 4/L'EMI au prisme de l'IA-littératie : compétences
- 5/Quelques points de vigilance dans le design EMI
- 6/Les arènes de formulations et réalisations

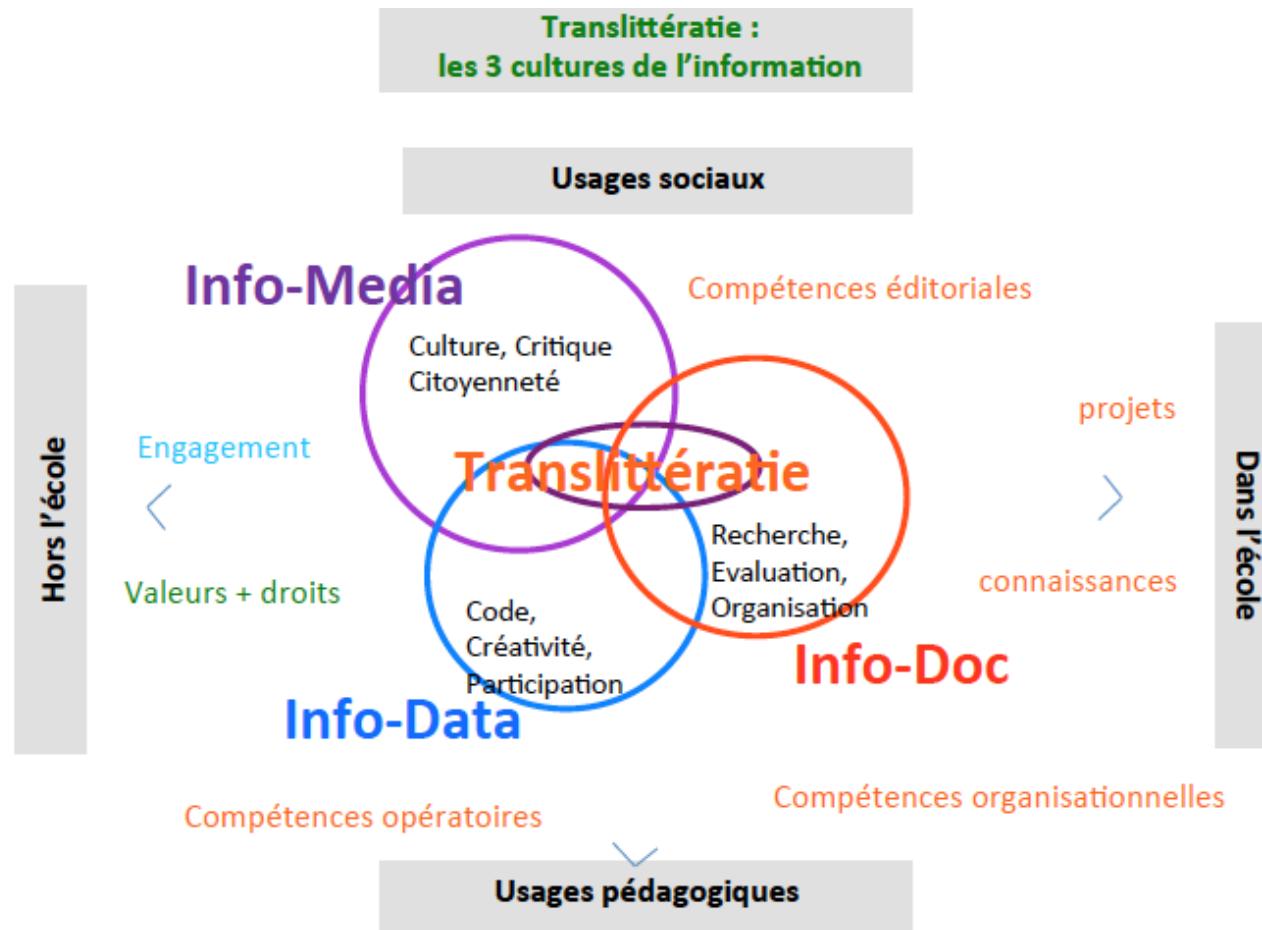
Savoir Devenir

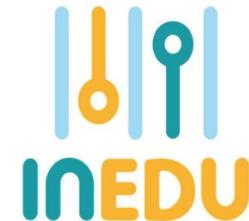
Un pilier de l'éducation

Une philosophie du changement

Une théorie de l'EMI : la translittératie

3 piliers de l'information: communication (médias), documentation (docs) et computation (data)





CROSSOVER



VOICES

Quelques résultats de recherche (à partir de YouCheck! YouVerify et Crossover):

- 1/ Les enseignants en formation initiale et en formation continue expriment un faible niveau de confiance dans leur compréhension des médias sociaux, ainsi que dans leur connaissance du rôle des algorithmes et des données.
- 2/ Les enseignants préféreraient disposer d'exemples et de bases de données faciles à utiliser.
- 3/ Les enseignants ressentent un écart entre l'arène des « réalisations » et l'arène des « formulations » et ont besoin de se sentir soutenus par l'administration pour s'engager dans ces questions.
- 4/ Les enseignants et les jeunes manquent d'un vocabulaire et d'un lexique adéquats pour exprimer leurs représentations et leurs expériences avec l'IA et la désinformation.
- 5/ La recherche manque d'échelles de compétences valables pour concevoir et évaluer des interventions robustes en matière d'éducation aux algorithmes et à l'IA. Ceux qui existent (Digcomp 2.2 par exemple) manquent de granularité.
- 6/ L'opacité des algorithmes (« boîte noire ») eux-mêmes constitue l'un des défis de l'enseignement aux algorithmes et à l'IA.

Nygren, T., Frau-Meigs, D., Corbu, N. and Santoval, S. *Teachers' views on disinformation and media literacy supported by a tool designed for professional fact-checkers: perspectives from France, Romania, Spain and Sweden*, SN Soc Sci (2022) 2:40 <https://doi.org/10.1007/s43545-022-00340-9>

Trois questions de recherche :

- 1/ Quelles sont les **compétences** (connaissances/attitudes/aptitudes/valeurs) dont les enseignants et élèves ont besoin pour évaluer les algorithmes et les systèmes d'IA de manière critique et interagir avec eux de manière efficace?
- 2 / Comment concevoir des interventions dont le **design** favorise une meilleure compréhension des algorithmes et des systèmes d'IA par les citoyens et une meilleure réaction à la désinformation diffusée par ces processus?
- 3/ Comment resserrer l'**écart** entre arène des formalisations et arène des réalisations?

Nygren, T., Frau-Meigs, D., Corbu, N and Santoval, S. *Teachers' views on disinformation and media literacy supported by a tool designed for professional fact-checkers: perspectives from France, Romania, Spain and Sweden*, SN Soc Sci (2022) 2:40 <https://doi.org/10.1007/s43545-022-00340-9>



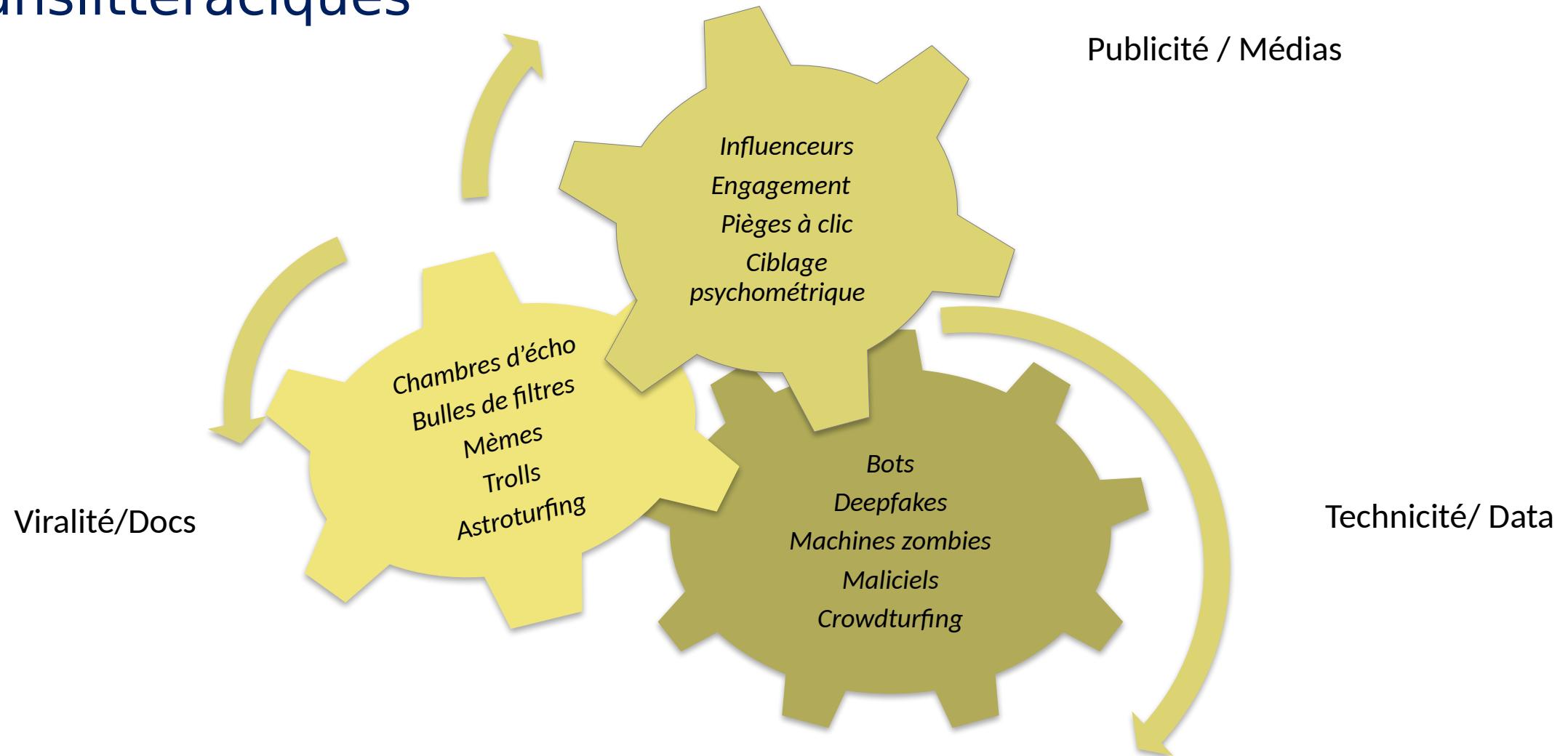
Globalement: se baser sur la recherche fondamentale et la recherche-action
s'adosser à des cadres de compétences granularisées
se donner des principes de design EMI pour les interventions
viser le changement (plus large que connaissance, d'où littératie)
traiter la désinformation comme un désordre de l'information (principe de familiarité)

Mon but: vous donner une vision partagée, mettre des mots sur des actions, faire sens.
Donc, sortir de l'implicite pour aller vers l'explicite (et montrer le 'making of', l'arrière scène).

La désinformation et ses mutations



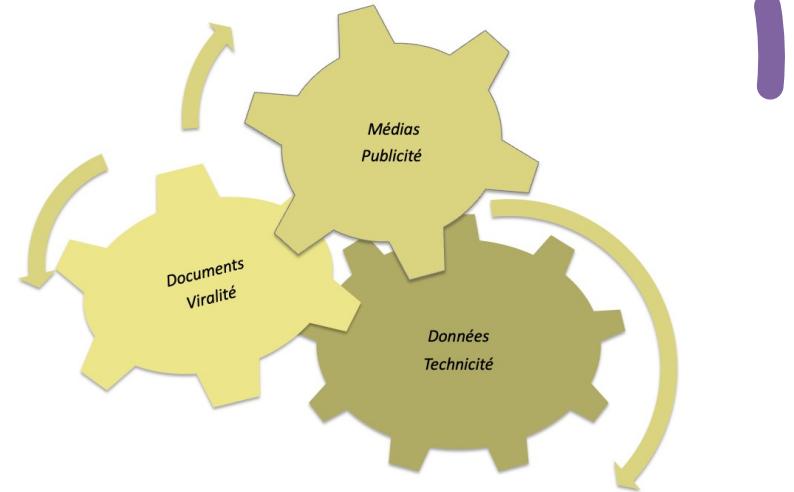
La désinformation fonctionne sur 3 mécanismes translittéraciques



Source: Frau-Meigs Faut-il avoir peur des 'fake news'? La Documentation française 2019

La désinformation et ses mutations: les désordres de l'information en mode translittéracique

Les acteurs malveillants utilisent **les données**
pour modifier les **docs**
et influencer les opinions publiques par les **médias** sociaux (et de masse)



Pour lutter contre la désinformation, il faut comprendre comment les vulnérabilités et faiblesses des systèmes et plateformes sont exploités par les acteurs malveillants. Et il faut développer et adopter des stratégies pour les détourner et en minimiser les effets et les impacts, voire les contrer. C'est un effort collaboratif.

La désinformation et ses mutations: les acteurs malveillants utilisent les données

Découvrir et reconnaître la valeur politique de nos données, c'est une affaire d'économie de l'attention et de politique de l'influence qui affecte le fonctionnement démocratique

Données comme actif : le commerce de données sur les citoyens à partir d'archives gouvernementales et de profils consommateurs est monnayé lors des stratégies de campagne et des élections

Données comme renseignement: les données des élections, en utilisant les algorithmes et en faisant de la veille sur médias sociaux, aident à formuler des messages et actions de sensibilisation auprès des électeurs

Données comme influence politique: des outils pilotés par l'extraction de données (IA), basés sur publicité en ligne et géo-ciblage, sont utilisées pour persuader ou manipuler les opinions

Nouvel or noir de l'économie de l'attention, elles constituent des gisements exploitables et extractibles par le biais des algorithmes et de la puissance de calcul de grands serveurs.



La désinformation et ses mutations : les acteurs malveillants modifient les docs

Découvrir et reconnaître la valeur politique de nos documents, c'est une affaire d'économie de l'attention et de politique de l'influence qui affecte le fonctionnement démocratique

Documents comme **archives** : garante d'authenticité, Internet peut agir comme un dépôt légal mais aussi être copié.

Documents comme **sources** : intégrité de l'information, lieu de vérifiabilité et de fiabilité mais aussi obfuscation

Documents comme **ressources**: base de créativité, référence, création de contenu... y compris fallacieux et factieux

Documents comme **fichiers**: diversité des formats (jpg, docx, ppt, ...), transposables et transfrontières

Pilier de la transformation numérique (du morphing au deepfake), ils sont la base des contenus de désinformation ; ils sont aussi la base des modèles d'apprentissage de l'IA



La désinformation et ses mutations : les acteurs malveillants influencent les opinions publiques par **les médias sociaux** (et de masse)

Découvrir et reconnaître la valeur politique de nos médias sociaux, c'est une affaire d'économie de l'attention et de politique de l'influence qui affecte le fonctionnement démocratique

Médias sociaux comme **chambres d'écho** : renforcer le sentiment d'être un collectif agissant même si minoritaire

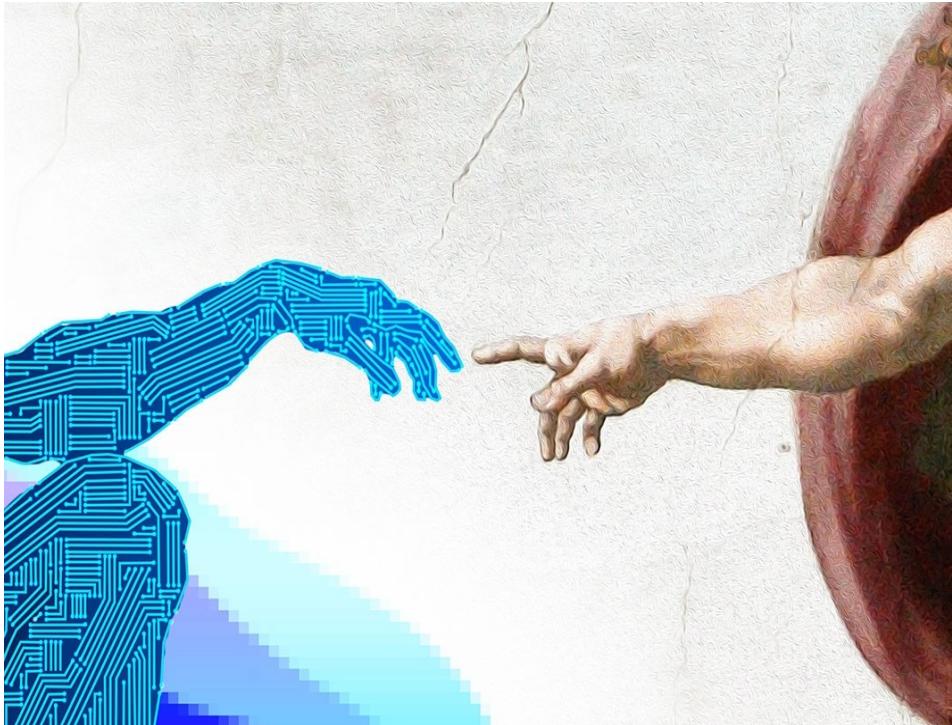
Médias sociaux comme **bulles de filtre** : renforcer les opinions déjà acquises

Médias sociaux comme **relais de désinformation**: viralité, vélocité...

Médias sociaux comme **prescripteurs de l'agenda** des médias de masse : blanchiment d'infox

Acteurs de la transformation numérique (du post au tweet, au reel), ils sont la base de l'amplification des contenus de désinformation ; ils sont aussi utilisés pour rendre les modèles d'apprentissage de l'IA plus proches du langage naturel des usagers (empathie, émotion,...).

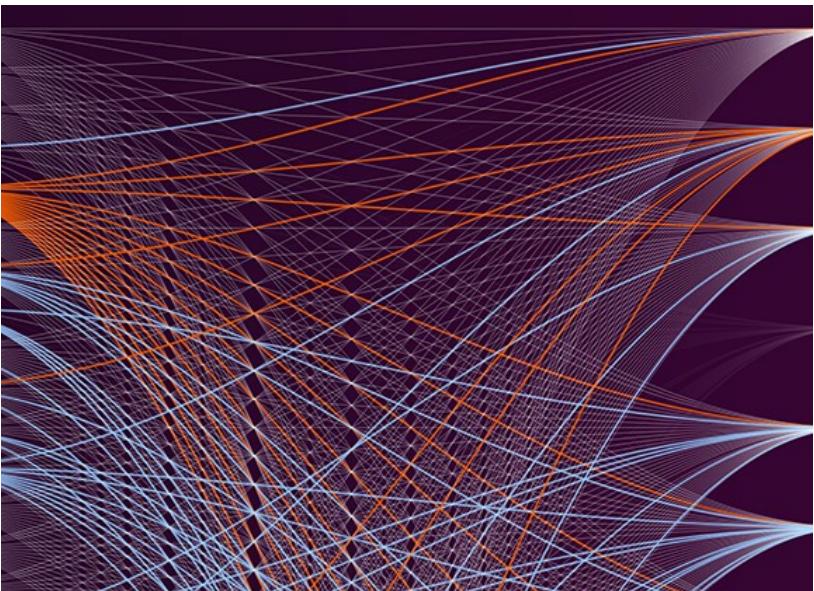
La désinformation et ses mutations... augmentées par l'IA



La science de l'intelligence artificielle (IA), notamment sous sa forme Générative (GIA), a un impact significatif sur l'engagement des personnes avec l'information (documents, médias, data).

Elle permet de produire des systèmes d'AI qui suscitent des inquiétudes quant au contrôle, à l'action humaine sur l'information, à la prise de décision indépendante et à la liberté en général.

L'Intégrité de l'information au risque des données

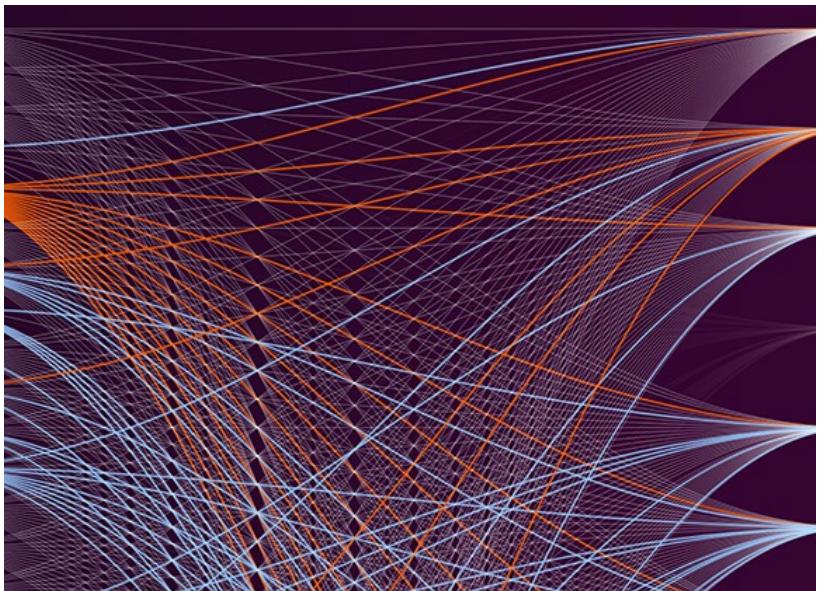


*Les données sont arbitraires mais ont une force de conviction
(Vitesse, Viralité, Volume, ...)*

INTÉGRITÉ DES DONNÉES?

- Sélection contingente : le big data à l'état brut
- Sélection biaisée (principalement dans le monde occidental, anglophone)
- Sélection reflète les biais de recherche et les stéréotypes
- Sélection non basée sur des critères sélectionnés et représentatifs d'un échantillon d'une population donnée
- Absence de modèles sous-jacents et de variables définies
- Propositions statistiques sur la base de ressemblance, association et proximité
- Absence de critères de fausseté ou de vérité

L'Intégrité de l'information au risque des données



La réponse EMI

- ==> Nécessité de connaître la qualité des données à l'entrée (comment elles sont sélectionnées) et les processus de prédiction à la sortie (comment elles sont traitées)
- ==> Nécessité de présenter les informations techniques sous une forme compréhensible pour des catégories spécifiques d'utilisateurs (XAI, évaluation, audits, recours).
- ==> Nécessité d'exiger l'explicabilité et la gouvernance des données (tiers de confiance...)

L'Intégrité de l'information au risque des docs et sources

Les sources sont hybrides ou synthétiques mais ont l'air valides



INTÉGRITÉ DES SOURCES ?

- Erreurs de l'IA ou « hallucinations »
- Références par auto-complétion
- Bibliographies erronées, tronquées
- Réponses proposées précisent peu ou mal les sources primaires et secondaires
- Réponses ne respectent pas les droits d'auteur ou la propriété intellectuelle.
- Absence de marqueurs pour les sources tertiaires, synthétiques, générées par l'IA à partir de rien
- Absence de principe de transparence ou de responsabilité
- Absence de principes tels que l'équité, la diversité et la protection de la vie privée

Source: <https://www.numerama.com/tech/1780810-des-chercheurs-prouvent-que-les-modeles-dia-degenerent-sils-sont-entraines-avec-leurs-propres-resultats.html>

L'Intégrité de l'information au risque des docs et sources

La réponse EMI



==> Attention aux biais d'autorité et de confirmation



==> Nécessité d'éviter l'auto-empoisonnement de l'IA par l'IA afin de garantir un cycle de vie sain de l'information, à l'abri de tout contenu toxique: les IA dégénèrent quand entraînées par leurs propres résultats

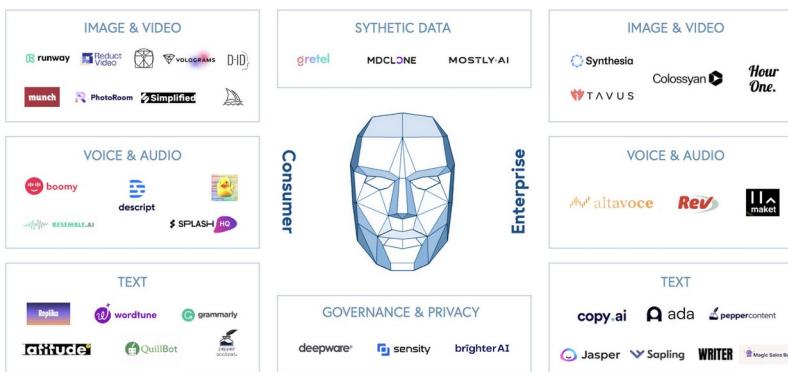


Nécessité d'inventer un système pour les **sources tertiaires**, pour que tous puissent citer et révéler les messages et les traces, utilisation éthique de l'IA

Demander à l'IA de fournir ses sources, vérifier ses sources, faire appel à plusieurs LLMs (pluralisme) et comparer les réponses, utiliser des IA spécialisées ou modèles « RAG » (quand les sujets sont sensibles),

L'Intégrité de l'information au risque des médias

Les médias de masse et sociaux sont hybrides ; les médias synthétiques produisent des contenus réalistes mais artificiels



INTÉGRITÉ DES MEDIAS ?

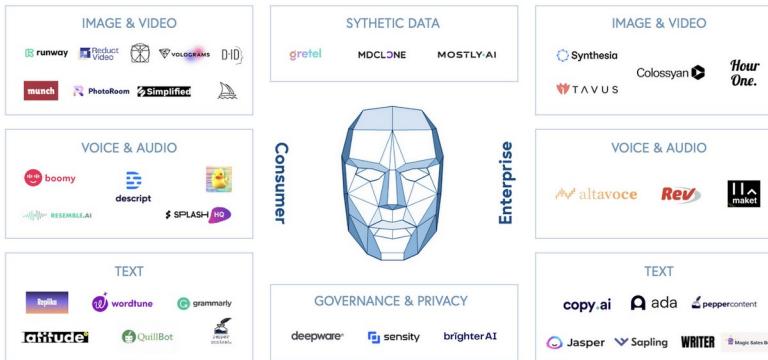
- Deepfakes IA (voix, texte et image) de plus en plus indétectables sur médias sociaux: personnalités générées par IA, auteurs fictifs, faits fabriqués
- Journalistes robots
- Manne publicitaire détournée de la presse (modèle économique basé sur la publicité programmatique – des publicités placées de manière algorithmique par les sociétés d'ad-tech sans se soucier de la nature ou de la qualité des sites)
- Médias de synthèse à usage de désinformation (1254 au 14 mars 2025), ingérence étrangère, déstabilisation et discrédit
- Manque de marqueurs de la citation, de l'expertise
- Absence de règles éditoriales
- Absence de principe de transparence ou de responsabilité

Source:<https://www.bvp.com/atlas/roadmap-the-rise-of-synthetic-media>

L'Intégrité de l'information au risque des médias

La réponse EMI

INTÉGRITÉ DES MEDIAS ?



- Riposter avec IA : Projet Spinoza (pour les journalistes)
- RAG (Génération Augmentée de Récupération) : aider les LLM à donner des réponses plus à jour et plus pertinentes
- Ingénierie inverse (voir Meta Parsing flow model)
- Pluralisme des LLM
- Vérification et fact-checking : ex NewsguardTech...
- Exigence de ligne éditoriale
- Boycott des sponsors et annonceurs

Source <https://www.bvp.com/atlas/roadmap-the-rise-of-synthetic-media>



Organisation
des Nations Unies
pour l'éducation,
la science et la culture



Chair UNESCO Savoir-devenir
à l'ère du développement numérique durable
Université Sorbonne Nouvelle, France

Les systèmes d'IA placent les citoyens dans une négociation de facto avec les géants de la technologie et les fournisseurs d'IA.

Il est essentiel d'exercer son esprit critique pour démonter les principaux mythes sur l'IA ("intelligence", "course aux armements") pour considérer les systèmes d'IA comme des **outils**, sur lesquels il est possible d'exercer un **contrôle et une surveillance**.





Organisation
des Nations Unies
pour l'éducation,
la science et la culture



Chair UNESCO Savoir-devenir
à l'ère du développement numérique durable

Les systèmes d'IA sont une pression de plus sur l'éducation, et notamment la lutte contre la désinformation.

Ces deux facteurs sont en train de reconfigurer l'EMI ... et de la politiser.



L'EMI au prisme de l'IA-littératie : compétences en pratique





EMI et Algo-IA littératie : pas à pas

Attentes des enseignants pour s'engager : formation, ressources validées, soutien de leur hiérarchie

Adopter une approche EMI basée sur expérience et usage

Traiter la question de la « boite noire » et autres idées reçues avec esprit critique

Relever le défi de définir des compétences à partir des scénarios d'usage

Développer des ressources pour le design des interventions en classe

Réponse EMI : le principe de familiarité

Se concentrer sur l'**autonomisation** de l'utilisateur et son appropriation active des opportunités et risques.

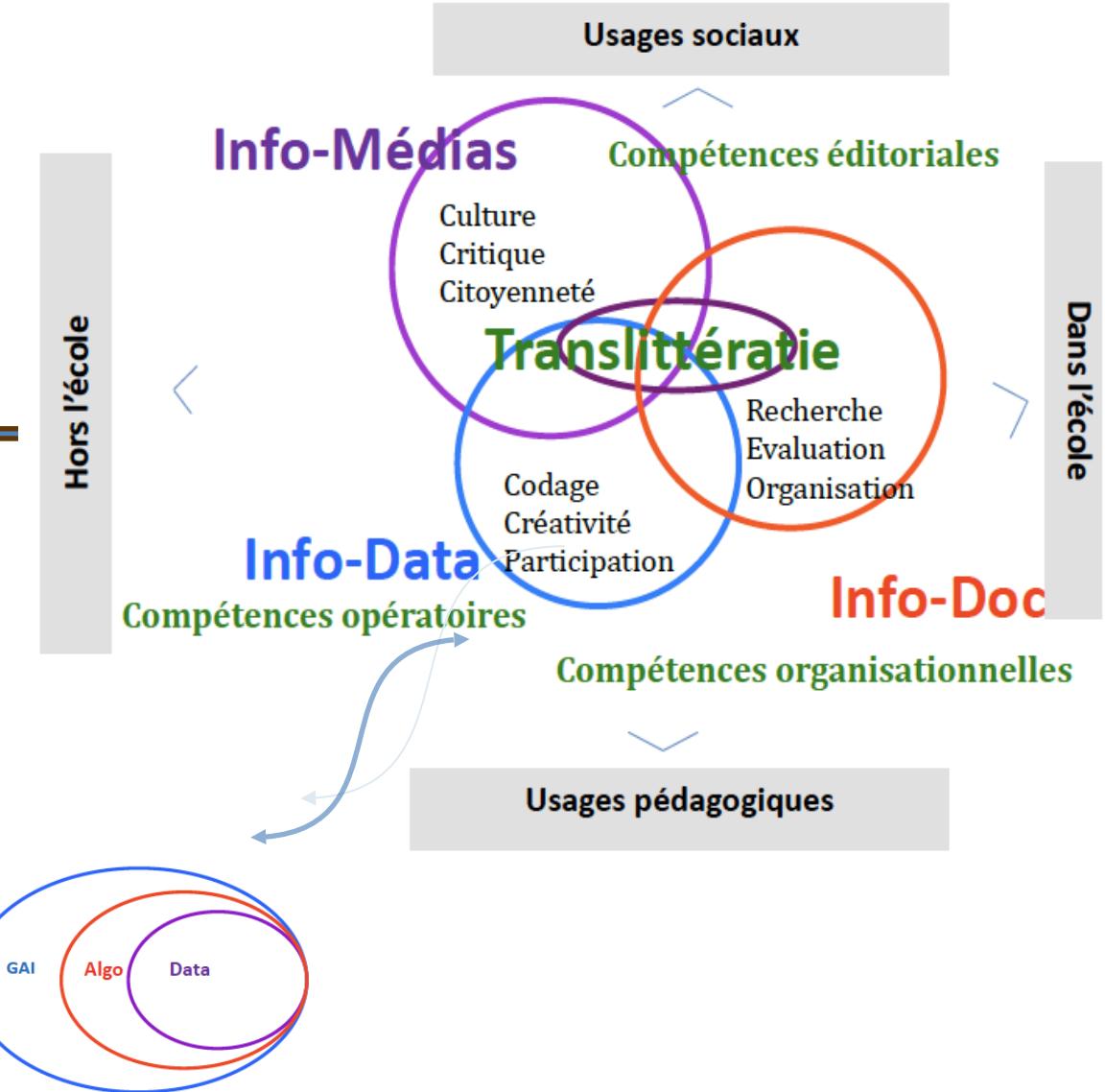
Intégrer les **littératies emboitées de l'IA** (data, algorithme) dans le cadre global de l'EMI comme translittératie (trois cultures de l'information: info-médias, info-doc, info-data).

Partir d'un domaine établi et donc gérable pour les personnes non spécialisées (non-STIM).

Remédier au **déficit de culture générale** sur les questions numériques, y compris l'IA.

Favoriser la **réflexion critique et créative** sur la mise en œuvre des valeurs des droits de l'homme telles que la liberté d'expression, la vie privée, la sécurité, la participation et l'inclusion.

L'EMI est indispensable pour



IA-littératie en EMI : concepts d'usage clé

La littératie est un mode particulier d'apprentissage qui prend en compte, non seulement l'acquisition de connaissances mais aussi et surtout leurs ancrages dans les usages, les expériences et les valeurs.

L'EMI traite l'Algo et IA-littératie en se penchant sur la question du rôle des algorithmes et de l'IA dans la production, la consommation et la diffusion de l'information. Elle aborde les représentations et imaginaires des utilisateurs ainsi que leurs tactiques et scénarios d'usages autour des algorithmes et des systèmes d'IA. Elle vise également à susciter l'esprit critique pour modifier les comportements face aux désordres de l'information (désinformation, discours de haine...).

Glossaire Savoir Devenir

<https://savoirdevenir.net/2024/11/glossaire-emi-data-algorithmes-intelligence-artificielle/>

Une expérimentation à petite échelle (projet pilote)



CROSSOVER

Crossover.social

<https://blog.cellenza.com/data/intelligence-artificielle/les-propositions-de-la-commission-europeenne-pour-une-ia-de-confiance/>

CROSSOVER

Scénarios d'usage (Carol 2000)

Matrice Algo-littératie (qui peut être transférée aux intervention dans les classes)				
Scenario d'usage	Évènement	Focus algorithmique	Compétences EMI	Enjeu de société
1 https://crossover.social/podcast/crossover-podcast-episode-1-the-keyboard-fighters/				
Recherche par mots-clés sur des moteurs de recherche tels que Google keyword: Liberty convoy	Article 1 15/02/2022 Podcast 1 13/07/2022 The guerriers du clavier 'Convoi de la liberté' menace d'envahir Bruxelles	FOCUS sur Algorithmes de RANKING et requêtes <i>Qu'est-ce qu'un mot-clé...</i> 	Analyse des mécanismes de désinformation et des processus de debunking	Contraste entre mobilisations URL et IRL
2 https://crossover.social/podcast/crossover-podcast-episode-2-dangerous-liasons/				
Recherche de communautés d'affinités, de groupes, d'influenceurs, d'acteurs via # sur les réseaux sociaux comme Youtube Hashtag: RT Russia	Article 2 podcast 2 03/11/2022 Algorithmes et propagande : les liaisons dangereuses Bannissement de RT pendant la guerre en Ukraine et remplacement par CGTN Français	FOCUS sur la PARTICIPATION dans les médias sociaux et les tendances <i>Qu'est-ce que l'engagement, comment affecte-t-il le ranking et la dissémination,...</i> 	Comprendre l'économie de l'attention Analyse des mécanismes de cyber-propagande Les processus de recommandation et d'influence	Les « addictions » algorithmiques qui propagent la désinformation

soutenu
par



Financé par
l'Union européenne

Scénarios d'usage

3 https://crossover.social/podcast/crossover-podcast-episode-3-how-algorithms-changed-my-job/				
Recherche de tendances et de comptes influents sur des forums tels qu'Odyssee	Article 3 08/06/22 Podcast 17/01/2023	FOCUS SUR LES ALGORITHMES DE RECOMMANDATION <i>Comment la recommandation se branche sur les usages (navigation...) ? Comment les usages affectent le comportement des algorithmes ?...</i>	Comprendre le rôle des communautés d'influence sur la désinformation	Économie de l'attention / recommandation
4 https://crossover.social/podcast/crossover-podcast-episode-4-algorithm-watchers-digital-fact-checking-prediction-algorithms-disinformation/				
Recherche de désinformation avec un outil intelligent comme Dashboard	Article 4 29/09/2022 Podcast 4 22/02/2023	FOCUS sur la PREDICTION algorithmique Et les biais cognitifs <i>Les algorithmes sont-ils neutres? Comment un outil comme le Dashboard révèle-t-il les dessous de la désinformation ?...</i>	Comprendre les biais algorithmiques Identifier les manipulations possibles S'approprier la vérification des sources	Biais, manipulation

enu
par

CROSSOVER

Quelques résultats de recherche:

- La boite noire et l'opacité : Les quatre scénarios ont permis d'obtenir des informations sur les algorithmes qui vont au-delà de la métaphore de la « boîte noire » (Pasquale 2015 ; Cotter et Reisford 2020). Ils offrent une « compréhension de l'opacité » (Burrell 2016). Les traces laissées lors des enquêtes ont permis d'inférer un certain nombre d'actions des algorithmes qui ont confirmé les hypothèses initiales émises par les développeurs et les journalistes. Ces enquêtes authentiques ont participé à l'empowerment des utilisateurs en révélant et en apportant des « preuves » de l'action des algorithmes dans la vie quotidienne (Le Deuff et Roumanos 2021).
- La validité de la translittératie: ce qui se passe à l'intérieur de la « boîte noire » du système ne détermine pas l'ensemble du processus, en particulier lorsqu'il s'agit de la recherche et de la curation d'informations (Hargittai et al., 2020). Les utilisateurs n'ont pas besoin de connaître l'architecture complète des algorithmes tels qu'ils sont développés par les plateformes pour en comprendre les mécanismes, notamment en termes de résultats et de services dans leur vie quotidienne. Toutefois, les utilisateurs doivent savoir comment certaines fonctions de base (classement, recommandation, prédition) peuvent affecter leurs actions et les conséquences que les algorithmes peuvent avoir dans la vie réelle, pour leur engagement civique et leur consommation de nouvelles par exemple.

Source: Frau-Meigs, Divina. 2024. "Algorithm Literacy as a Subset of Media and Information Literacy: Competences and Design Considerations" *Digital* 4, no. 2: 512-528. <https://doi.org/10.3390/digital4020026>



Quelques résultats de recherche

- La maîtrise des algorithmes peut englober différents types de comportements : détection de la désinformation, vérification du contenu, divulgation/exposition des résultats de la recherche, réponse/réfutation des résultats, demande de transparence et de responsabilité de la part des plateformes et des services.
- La maîtrise des algorithmes implique de soutenir un certain nombre de valeurs, traditionnellement encouragées par l'EMI, telles que l'intégrité de l'information, la qualité des données, la liberté d'expression et le pluralisme et la diversité des médias.
- En matière de changement de comportement, les solutions personnelles (renforcement de la résilience) et collectives (renforcement de la résistance) doivent se mutualiser. Au-delà de l'adaptation, les utilisateurs/citoyens doivent s'engager activement dans leur environnement en ligne, notamment en s'adressant aux développeurs de plateformes et aux décideurs politiques (UE 2018 ; UE 2022)

Frau-Meigs, Divina. 2024. "Algorithm Literacy as a Subset of Media and Information Literacy: Competences and Design Considerations" *Digital* 4, no. 2: 512-528. <https://doi.org/10.3390/digital4020026>

Vers un passage à l'échelle, la question des compétences et ressources



ALGOWATCH

Algowatch.eu

Décliner les compétences en micro-compétences...



Connaissances <i>savoir, comprendre, être conscient de...</i>	
DIGCOMP 2.2	EMI compétences Algo et IA
<p>157. Savoir que les algorithmes, et par conséquent les programmes, sont conçus pour aider à résoudre des problèmes de la vie réelle ; les données d'entrée modélisent les informations connues sur le problème, tandis que les données de sortie fournissent des informations pertinentes pour la solution du problème. Il existe différents algorithmes, et par conséquent différents programmes, pour résoudre le même problème.</p> <p>156. Savoir que le plan d'un programme est basé sur un algorithme, c'est-à-dire une méthode progressive pour produire une sortie à partir d'une entrée.</p> <p>4. Être conscient que les moteurs de recherche, les médias sociaux et les plateformes de contenu utilisent souvent des algorithmes d'IA pour générer des réponses adaptées à l'utilisateur individuel (par exemple, les utilisateurs continuent à voir des résultats ou des contenus similaires). On parle souvent de "personnalisation". (IA)</p> <p>21. Etre conscient que les algorithmes d'IA peuvent ne pas être configurés pour fournir uniquement les informations souhaitées par l'utilisateur ; ils peuvent également véhiculer un message commercial ou politique (par exemple, pour encourager les</p>	<p>Savoir ce qu'est un algorithme (non thématique)</p> <p>Comprendre le fonctionnement des algorithmes de classement des moteurs de recherche ex : Classement par algorithme (thématische)</p> <p>Savoir ce que les algorithmes de recommandation peuvent faire et ne pas faire (non thématique)</p>

Soutenu
par



Financé par
l'Union européenne

Décliner les compétences en micro-compétences...

utilisateurs à rester sur le site, à regarder ou à acheter quelque chose de particulier, à partager des opinions spécifiques). Cela peut également avoir des conséquences négatives (reproduction de stéréotypes, partage d'informations erronées). (IA)	
105. Etre conscient que les systèmes d'IA collectent et traitent plusieurs types de données utilisateur (par exemple, des données personnelles, des données comportementales et des données contextuelles) pour créer des profils d'utilisateur qui sont ensuite utilisés, par exemple, pour prédire ce que l'utilisateur pourrait vouloir, voir ou faire ensuite (par exemple, proposer des publicités, des recommandations, des services). (IA)	Connaître le fonctionnement des algorithmes prédictifs (thématique) Connaître la définition et la fonction de l'intelligence artificielle
NA	Comprendre l'intelligence artificielle générative (IAG) Comprendre l'IAG : Comment les LLM génèrent des réponses Parlez-vous IA ? Savoir traduire des mots clés spécifiques (pour les non-anglophones) : machine learning, big data...
NA	Avoir une culture de base de l'histoire des algorithmes
NA	Comprendre les mécanismes algorithmiques qui peuvent encourager les campagnes de désinformation sur les médias sociaux

Décliner les compétences en micro-compétences...

APTITUDES/ VALEURS <i>Être capable de, savoir faire...</i>	
DIGCOMP 2.2	EMI compétences Algo et IA
119. Savoir que les systèmes d'IA peuvent être utilisés pour créer automatiquement du contenu numérique (par exemple, des textes, des nouvelles, des essais, des tweets, de la musique, des images) en utilisant du contenu numérique existant comme source. Ce contenu peut être difficile à distinguer d'une création humaine	Être capable de reconnaître les informations produites par l'IA
27. Pouvoir reconnaître que certains algorithmes d'IA peuvent renforcer les opinions existantes dans les environnements numériques en créant des « chambres d'écho » ou des « bulles de filtres » (par exemple, si un flux de médias sociaux favorise une idéologie politique particulière, des recommandations supplémentaires peuvent renforcer cette idéologie sans l'exposer à des arguments opposés).	Savoir lutter contre les bulles de filtres et les chambres d'écho

Soutenu
par



Décliner les compétences en micro-compétences...

22. Etre conscient que les données dont dépend l'IA peuvent comporter des biais. Si c'est le cas, ces biais peuvent être automatisés et aggraver l'utilisation de l'IA. Par exemple, les résultats de recherche sur la profession peuvent inclure des stéréotypes sur les emplois masculins ou féminins.	pouvoir juger de la qualité des bases de données (et des biais possibles) sur lesquelles l'IA et les algorithmes travaillent.
113. Savoir comment modifier les configurations de l'utilisateur (par exemple dans les applications, les logiciels, les plates-formes numériques pour permettre, empêcher ou modérer le système d'IA de suivre, rassembler ou analyser des données (par exemple ne pas permettre au téléphone mobile de suivre la localisation de l'utilisateur).	Savoir développer des stratégies pour lutter contre la désinformation algorithmique et le discours de haine
113. Savoir comment modifier les configurations de l'utilisateur (par exemple dans les applications, les logiciels, les plates-formes numériques pour permettre, empêcher ou modérer le système d'IA de suivre, rassembler ou analyser des données (par exemple ne pas permettre au téléphone mobile de suivre la localisation de l'utilisateur). NA NA	Pouvoir limiter l'influence des algorithmes de recommandation sur les médias sociaux Savoir évaluer la fiabilité des sources fournies par l'IA Savoir limiter l'influence des algorithmes de recommandation pour les sites web
8. Savoir comment formuler des requêtes de recherche pour obtenir le résultat souhaité lorsqu'il interagit avec des agents conversationnels ou des haut-parleurs intelligents (par exemple Siri, Alexa, Cortana, Google Assistant), par exemple en reconnaissant que, pour que le système puisse répondre comme il se doit, la requête doit être sans ambiguïté et formulée clairement pour que le système puisse y répondre. (IA)	Pouvoir converser avec l'IA générative

Soutenu
par



Décliner les compétences en micro-compétences...

ATTITUDES et VALEURS <i>Se comporter de sorte que ... (« soft skills »)</i>	
DIGCOMP 2.2 	EMI compétences Algo et IA 
216. Considérer les conséquences éthiques des systèmes d'IA tout au long de leur cycle de vie : elles comprennent à la fois l'impact environnemental (conséquences environnementales de la production d'appareils et de services numériques) et l'impact sociétal, par exemple la plateformisation du travail et la gestion algorithmique qui peuvent porter atteinte à la vie privée ou aux droits des travailleurs ; l'utilisation d'une main-d'œuvre bon marché pour étiqueter les images afin d'entraîner les systèmes d'IA. (IA).	Être sensible aux questions éthiques associées aux algorithmes informationnels
2. Etre conscient que le contenu en ligne mis à la disposition des utilisateurs à titre gratuit est souvent rémunéré par la publicité ou la vente de données de l'utilisateur. est souvent financé par la publicité ou par la vente des données de l'utilisateur.	Avoir une attitude averte et critique à l'égard des modèles économiques sous-jacents aux offres, notamment gratuites
56. Savoir que tous les citoyens de l'UE ont le droit de ne pas être soumis à une prise de décision entièrement automatisée (par exemple, si un système automatisé refuse une demande de crédit, le client a le droit de demander que la décision soit réexaminée par une personne).	Être enclin à lutter contre les manipulations
NA	Être prêt à collaborer avec d'autres pour obtenir des informations plus sûres et de meilleure qualité
NA	Être en mesure de prendre en considération les différents niveaux de risques liés à l'IA dans les informations conformément à la loi sur l'IA

Soutenu
par



Mettre en oeuvre un design EMI: quiz, jeu, exposition

QUIZ
To help you better understand and master AI and algorithms, we suggest you take this series of interactive quizzes. Four topics in which to test your wits. Now it's your turn to play!

Climate change
Health and well-being
Democracy and citizenship
Online Relationships

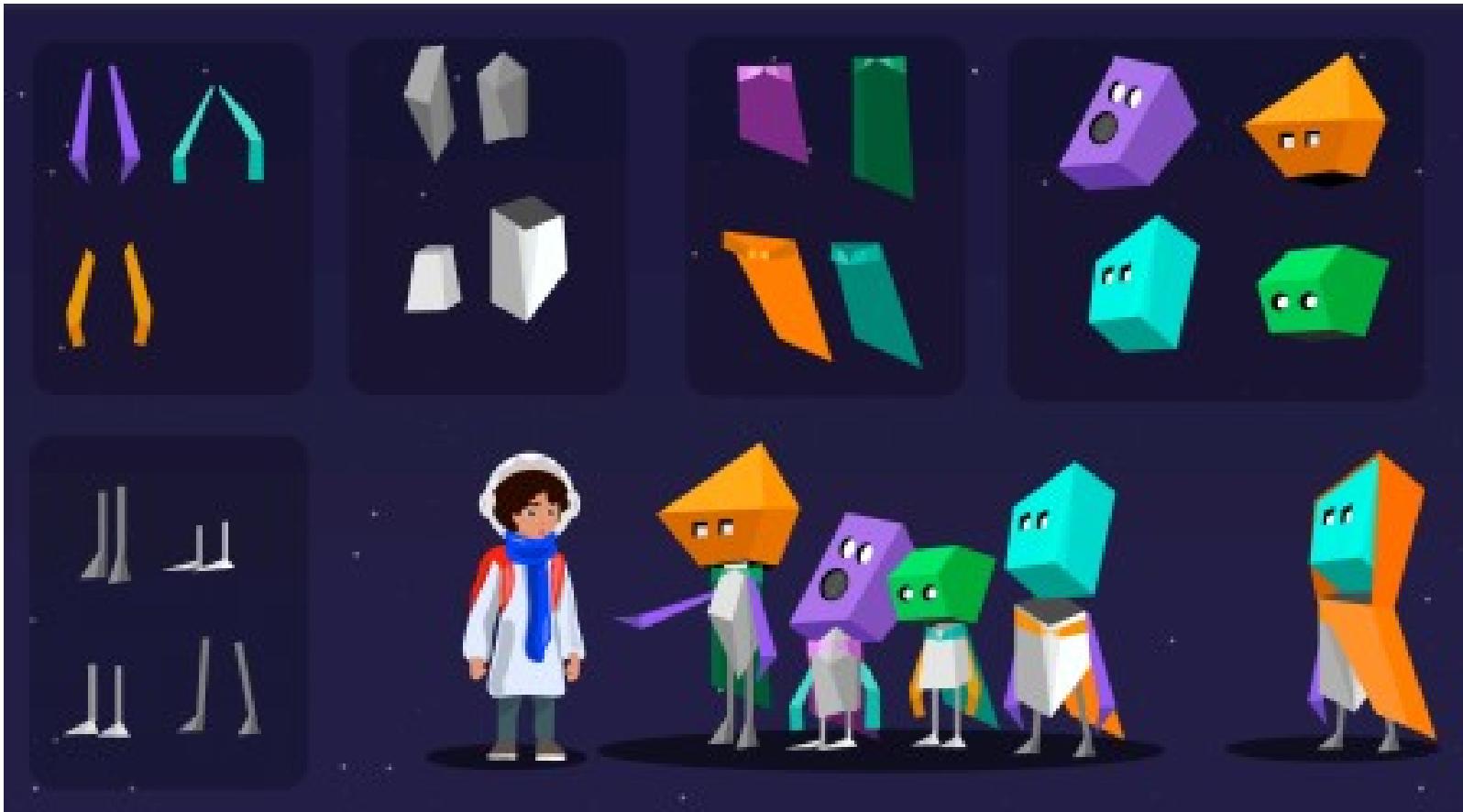
QUIZ
Pour mieux comprendre et maîtriser l'IA et les algorithmes, nous vous proposons de vous lancer cette série de quiz interactifs. Quatre thématiques, autant d'univers où déployer votre sagacité. A vous jouer !

Changement climatique
Santé et bien-être
Démocratie et citoyenneté
Relations en ligne

Soutenu
par



Mettre en oeuvre un design EMI: quiz, jeu, exposition



Soutenu
par

 **MINISTÈRE
DE LA CULTURE**
Liberté
Égalité
Fraternité


Financé par
l'Union européenne

Pour continuer la recherche: 5 écoles dans 4 pays

• Pre- Post test

En termes de questions, quatre dimensions, chacune avec trois questions fermées (4 options chacune) :

- **Définitions (savoir)** - 3 QUESTIONS (identifier la définition de : algorithme, IA ; désinformation) ;
- **Attitudes (savoir être)** - 3 QUESTIONS (dans quelle mesure « VOTRE » vie quotidienne est-elle influencée par les algorithmes ; dans quelle mesure l'IA change-t-elle votre vie ; dans quelle mesure rencontrez-vous de la désinformation dans votre vie quotidienne) ;
- **Aptitudes (savoir faire)** - 3 QUESTIONS (ex. : quelles actions ou outils utilisez-vous pour comprendre l'influence des algorithmes sur votre vie quotidienne ; quelles actions ou outils utilisez-vous pour comprendre l'influence de l'IA dans la vie quotidienne ; quelles actions ou outils utilisez-vous pour évaluer les informations que vous trouvez en ligne).
- **Éthique/Valeurs (savoir devenir)** - 3 QUESTIONS (identifier l'exemple correct de comportement éthique concernant l'algorithme, l'IA ; la désinformation)



Du côté du design EMI : les points de vigilance en lien à l'IA et à la désinformation

Des éléments de design nouveaux



L'explicabilité de l'IA (ou XAI)

3 considérations importantes

- Tirer parti de l'expérience des élèves
- Lever l'angoisse de la « boîte noire »
- (Exiger l'Explicabilité (XAI))
- Former aux interactions avec agents/outils
non humains

-dépasse le mythe de la « boîte noire » et les imaginations anthropomorphiques (« intelligence », « hallucinations », etc.)

-présente les informations techniques sous une forme compréhensible pour des catégories d'utilisateurs spécifiques, afin qu'ils puissent vérifier si elles répondent à leurs besoins

-fournit des explications spécifiques qui peuvent ouvrir la voie à des audits et à des mesures d'évaluation.

- assure que des principes tels que l'équité, l'égalité, la diversité, la sécurité et la vie privée sont appliqués et, si ce n'est pas le cas, permet de prendre des mesures telles que des demandes de recours efficaces (à la fois juridiques et non judiciaires).

-éclaire les principes de transparence et de responsabilité

L'expérience des élèves : imaginaires de l'IA



Deux paniques morales et médiatiques plus récentes :

- la prétendue « intelligence » coévolutive de l'IA (« réseaux neuronaux », « hallucinations », « chats », « conversations »)
- La prétendue menace apocalyptique de l'IA (« course aux armements », « prolifération », « risques existentiels »).



Mèmes Andromo

Sous le capot de la boîte noire: L'explicabilité (XAI) en ressources



Régulation



Transparence des modèles



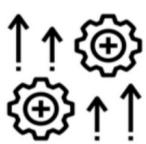
Générer de la confiance



Respecter la réglementation



Comprendre la prise des décisions



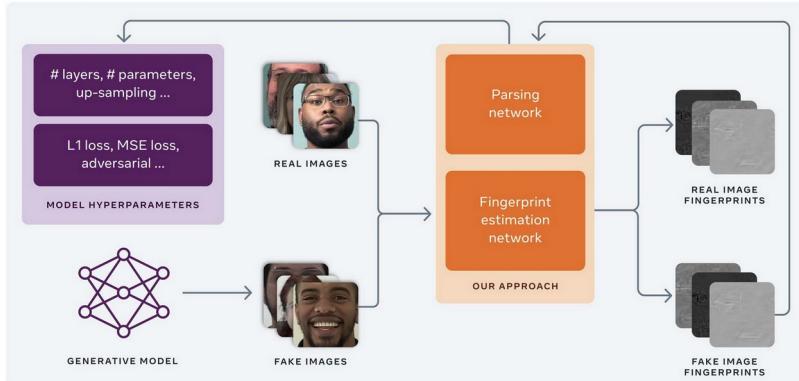
Améliorer la performance des modèles



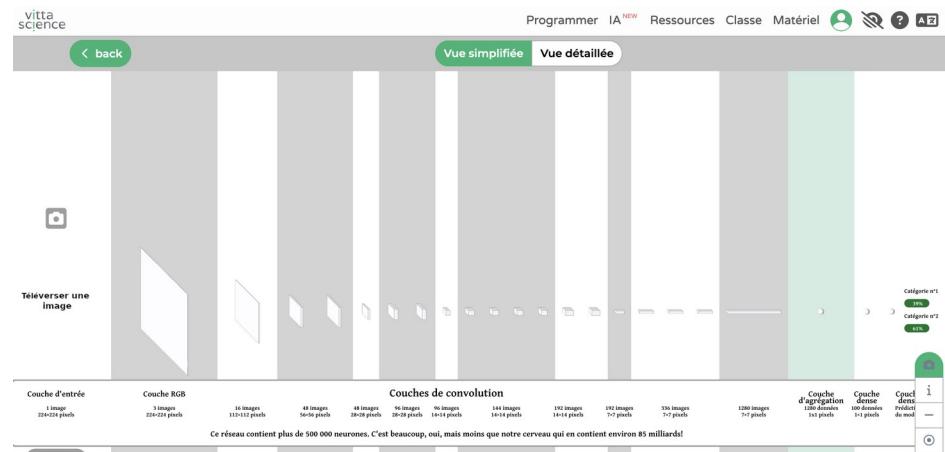
Réduire les biais éthiques et moraux

Rétro-ingénierie

Model parsing flow



Source:
<https://ai.meta.com/blog/reverse-engineering-generative-model-from-a-single-deepfake-image/>



Source: <https://fr.vittascience.com/ia/visualization.php>

<https://www.actuia.com/contribution/jean-cupe/interpretabilite-de-lia-le-nouveau-defi-des-data-scientists/>

Visualisation

Citabilité : Pluralisme des LLM et diversité des sources



2. Diversification des Données d'Entraînement :

- **Ensembles de Données Inclusifs** : Utiliser des ensembles de données diversifiés et représentatifs pour entraîner les modèles d'IA. Cela permet de réduire les biais inhérents aux données historiques ou culturelles ³.
- **Audit et Mise à Jour** : Effectuer des audits réguliers des modèles d'IA pour identifier et corriger les biais. Les modèles doivent être mis à jour avec de nouvelles données pour refléter les évolutions sociétales ³.

3. Éducation et Sensibilisation :

- **Formation des Développeurs** : Former les développeurs et les concepteurs d'IA à reconnaître et à éviter les biais de genre dans leurs travaux. Cela inclut la sensibilisation aux stéréotypes de genre et aux meilleures pratiques pour les éviter ⁴.
- **Campagnes de Sensibilisation** : Utiliser l'IA pour créer des campagnes de sensibilisation qui éduquent le public sur les stéréotypes de genre et leurs impacts négatifs ².

4. Promotion de la Diversité dans les Étapes de Développement :

nder au Chat

de l'IA, il est impératif de rester vigilants et proactifs dans la lutte contre les biais....

4 start.lesechos.fr

Les algorithmes reproduisent les stéréotypes sexistes de notre société | Les Echos Start

17/05/2019 - Comment rectifier le tir ? Féminisation des métiers du numérique, formation et sensibilisation des...

5 news.un.org

Intelligence artificielle : l'UNESCO alerte sur les stéréotypes racistes, sexistes et homophobes | ONU Info

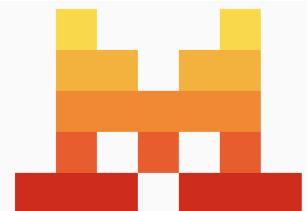
14/03/2024 - A noter qu'en novembre 2021, les États membres ont adopté à l'unanimité la Recommandation sur...

<https://chat.mistral.ai/chat>

6 chut.media

L'IA pour déconstruire les stéréotypes de genre à l'écran | Chut!

04/09/2024 - Pour l'heure, le projet n'est pas encore achevé et aucune date de mise



L'anthropomorphisme: une stratégie de marketing à débusquer = revoir le glossaire!



Expressions manipulatrices!

Intelligence Artificielle= extraction de données

LLM systems = systems de propositions statistiques

Prompt = instruction

Conversation =

manipulation mécanique de symboles

Agent = outil de récupération de données

Hallucination = erreur

Boite noire = manque d'explicabilité en entrée et sortie

...

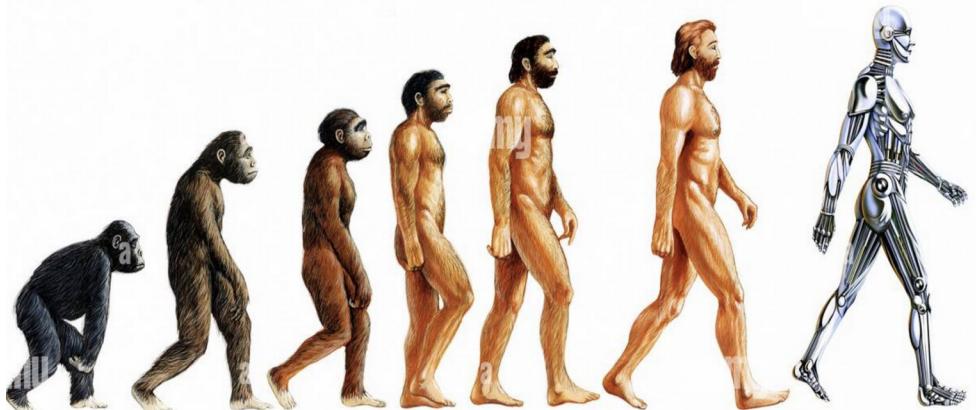


Le marketing de l'IAG s'appuie sur l'anthropomorphisme pour rendre l'IA acceptable et digne de confiance afin que l'IAG puisse pénétrer dans la vie privée des foyers et l'intimité des utilisateurs.

- Les systèmes d'IA sont basés sur des calculs et des extractions statistiques.
- Les systèmes d'IA n'ont pas de capacité
- Les systèmes d'IA manquent de transparence
- Les systèmes d'IA peuvent travailler sur des données périmées (ce qui est mauvais pour l'actualité)
- Les systèmes d'IA n'ont pas d'émotion ou de sentience



Les interactions avec des "agents non-humains" (outils) : se méfier de la coévolution



Les entreprises de l'IAG s'appuient sur l'anthropomorphisme pour rendre les systèmes d'IA acceptables pour créer de l'attachement et être des substituts des relations interhumaines.

- Les systèmes d'IA ne sont pas sensibles
- Les systèmes d'IA n'ont pas de sentiments/émotions et font preuve d'une empathie synthétique préprogrammée.
- Les systèmes d'IA ne s'attachent pas aux humains
- Les systèmes d'IA n'ont pas d'autonomie externe (ils dépendent de l'énergie).
- Les systèmes d'IA n'ont aucune prétention à la vérité et à l'interprétation
- Les systèmes d'IA sont des OUTILS

<https://www.alamy.com/artificial-intelligence-computer-artwork-depicting-the-evolution-of-a-robot-far-right-with-artificial-intelligence-from-a-human-man-homo-sapiens-image335451040.html>

Du côté de l'écart
entre arènes



Arène des formalisations

- Textes officiels: cadre européen; cadre français
- Place de l'EMI dans ce cadre
- Coordination (le guichet unique?)
- RAG éducation nationale
- CRIA (DNE)...

Arène des réalisations

- Ressources: Canopé, Capytale (Vittascience)
- Dispositif Edu up, piloté par le Ministère de l'Éducation Nationale, pour la production de ressources numériques innovantes et adaptées.
- Séminaires de formation
- Collaborations scientifiques (recherche-action)
- ATTENTION A RE-INSERER DE L'EMI ...

Conclusion...

Une résilience effective passe par une approche holiste des cultures de l'information, pour pouvoir rester en surplomb de leurs constantes innovations. S'appuyer sur des outils « intelligents » (statistiques) n'est pas suffisant. Pour amener du changement culturel, il faut faire une évaluation des risques et opportunités, une conscience des vulnérabilités, et s'y mettre de manière collaborative. A la résilience individuelle s'ajoute la résistance collective.

En fin de compte, avec et au-delà de l'IA/GIA et des systèmes d'information complexes, l'objectif de l'éducation, et notamment de l'éducation aux médias et à l'information, est de construire des sociétés de la connaissance viables et durables dotées d'une intelligence collective.

SOURCE: Mireia Freix-Molins, L'autonomisation de l'utilisateur grâce aux réponses

apportées par l'éducation aux médias et à l'information à l'évolution de l'intelligence artificielle générative (IAG), UNESCO 2024

https://unesdoc.unesco.org/ark:/48223/pf0000388547_fra

Merci de votre attention!

Contact :

Divina-frau-meigs@sorbonne-nouvelle.fr

Sites:

<https://savoirdevenir.net/>

<https://Divina-Frau-Meigs.fr>

<https://savoirdevenir-chaireunesco.fr/>



Organisation
des Nations Unies
pour l'éducation,
la science et la culture



Chaire UNESCO Savoir-devenir
à l'ère du développement numérique durable
Université Sorbonne Nouvelle, France

