

Vers une certification pour l'IA et le jeu vidéo

Jonathan Bonneau

Professeur associé à l'École des médias de
l'Université du Québec à Montréal (UQAM)

Laboratoire de recherche en médias
socio-numériques et ludification

Commission spéciale sur les impacts des
écrans et des réseaux sociaux sur la santé
et le développement des jeunes

4 Février 2025

Table des matières

PRÉSENTATION DE L'AUTEUR.....	1
RECOMMANDATION #1.....	2
RECOMMANDATION #2.....	3
RECOMMANDATION #3.....	4
RECOMMANDATION #4.....	5
PISTES VERS CERTIFICATION.....	6
CONCLUSIONS ET CONTACT.....	7

PRÉPARÉ PAR
Jonathan Bonneau

Laboratoire de recherche en médias socio-numériques et ludification

Depuis son lancement à l'été 2016 grâce au soutien de la Fondation Canadienne pour l'Innovation (FCI), le laboratoire constitue un pôle d'expertise unique sur les transformations numériques et leurs effets sur la société contemporaine. Il est membre de l'Alliance des laboratoires universitaires francophones sur le jeu, renforçant ainsi son ancrage dans les études sur le jeu et la communication.

Le laboratoire héberge deux groupes de recherche complémentaires :

- Homo Ludens, consacré à l'étude des rapports entre le jeu et la communication.
- Le Groupe de recherche sur l'information et la surveillance au quotidien (GRISQ), qui analyse l'impact des technologies numériques sur la collecte, l'exploitation et l'utilisation des données personnelles dans les interactions sociales et institutionnelles.

Grâce à l'installation d'infrastructures de pointe, le laboratoire permet l'observation et l'analyse approfondies de phénomènes liés à l'usage des réseaux sociaux, à la production massive de données personnelles et à l'influence du jeu dans les interactions communicationnelles. Ses axes de recherche incluent notamment :

1. Le contexte de mobilité, examinant les pratiques numériques en déplacement.
2. La biométrie et les pratiques, portant sur l'identification et la surveillance des individus à travers les technologies.
3. Les données textuelles et comportementales, étudiant les traces laissées par les utilisateurs sur les plateformes numériques.
4. Les recherches déclaratives sur les pratiques, analysant les discours et perceptions des usagers face aux transformations numériques.

À travers ces travaux, le laboratoire se positionne comme un acteur clé dans la compréhension des dynamiques numériques et l'étude critique de leurs impacts sur les sociétés contemporaines.

Chercheur

Professeur associé à l'École des médias de l'UQAM, Jonathan Bonneau étudie et enseigne l'intelligence artificielle, le jeu vidéo et les univers numériques depuis 2005. Spécialiste de l'immersion et de la création interactive, il analyse l'effet des technologies émergentes sur l'individu et la société.

Directeur des expériences pour divers projets en réalité mixte, il contribue également à la vulgarisation des nouveaux phénomènes technologiques dans la francophonie en tant que chercheur du Laboratoire de recherche en Médias Socio-numériques et Ludification.

Il intervient aussi régulièrement dans les médias et lors de conférences afin de sensibiliser et d'éduquer le public vers une utilisation critique et éclairée des technologies, entre autres par l'entremise d'événements ciblés envers les jeunes et d'organisations telles que celles de l'Université du troisième âge.

Engagé dans la recherche-création, il explore de nouvelles formes de communications humain-machine, axées sur une démarche méthodologique éthique, développant depuis les dernières années des outils d'apprentissage machine et de biométrie servant à appuyer les démarches de diverses thérapies en milieu scolaire et en clinique privée.

Problématique #1

Manque de compréhension des jeunes face à l'IA

L'intelligence artificielle est de plus en plus présente dans la vie quotidienne des jeunes, que ce soit à travers les assistants vocaux, les recommandations de contenu sur les réseaux sociaux, les outils éducatifs automatisés ou encore les jeux vidéo. Cependant, beaucoup d'adolescents et d'enfants ne comprennent pas réellement comment fonctionnent ces systèmes ni les implications de leur utilisation.

Dans les jeux vidéo, par exemple, l'IA est utilisée pour adapter la difficulté, optimiser les interactions avec des personnages non-joueurs ou encore inciter à l'achat via des algorithmes sophistiqués. De même, sur les réseaux sociaux, des algorithmes pilotés par l'IA sélectionnent les publications affichées aux jeunes, influençant leurs opinions, leur bien-être et leur perception du monde.

Ce manque de compréhension engendre plusieurs problèmes : une confiance excessive ou, au contraire, une méfiance injustifiée envers l'IA, une surestimation de ses capacités (en pensant qu'elle est "consciente" ou "neutre") et une vulnérabilité aux manipulations algorithmiques. Sans une éducation adéquate sur le sujet, les jeunes risquent d'adopter des comportements problématiques face à ces technologies, comme une dépendance aux systèmes de recommandation, une exposition prolongée à des contenus biaisés ou une exploitation économique inconsciente dans l'univers du jeu vidéo.

Conséquences

- Perception erronée de l'IA généralisée
- Manipulation des comportements
- Exploitation économique
- Difficulté à réguler l'impact de l'IA

Recommandations

Intégrer l'éducation à l'IA dans les programmes scolaires et sociaux

Expliquant son fonctionnement, ses limites et ses implications éthiques, de manière à mieux comprendre la possible influence de l'IA et développer une utilisation critique de ces outils. (ex. probabilités des "loot boxes", enregistrements, transferts des données, sources de la générativité, mécaniques de type "fear of missing out", arnaques par apprentissage machine, etc.)

Développer des outils de transparence et de contrôle

Permettant aux jeunes d'identifier quand et comment l'IA ou certaines mécaniques interviennent dans leurs expériences numériques et leur donner la possibilité de paramétrer certaines fonctionnalités tout en les informant sur ces ajustements par une vulgarisation des effets impliqués.

Problématique #2

L'innovation au coût de la vie privée

L'intelligence artificielle est omniprésente dans les outils numériques utilisés par les jeunes, que ce soit dans les réseaux sociaux, les plateformes éducatives, les services de streaming ou les jeux vidéo. Ces technologies collectent et analysent une grande quantité de données personnelles, souvent sans que les utilisateurs en aient pleinement conscience.

Dans l'univers des jeux vidéo et des applications en ligne, les comportements des jeunes sont scrutés en temps réel pour adapter leur expérience : ajustement de la difficulté en fonction du niveau de compétence, recommandations personnalisées, ou encore ciblage publicitaire basé sur leurs habitudes de navigation et de consommation. Ce suivi constant peut rapidement basculer dans une forme de surveillance intrusive, limitant leur autonomie et mettant en péril leur vie privée.

Les risques liés à cette collecte massive de données sont multiples : exploitation abusive à des fins commerciales, revente d'informations personnelles à des tiers sans consentement éclairé et mise en place d'algorithmes qui ajustent les offres en fonction de la vulnérabilité économique des utilisateurs. Sans cadre réglementaire strict, ces pratiques peuvent exposer les jeunes à des manipulations économiques et à des atteintes à leurs droits fondamentaux.

Conséquences

- **Surveillance excessive et perte d'anonymat**
- **Exploitation des données à des fins lucratives**
- **Manipulation économique des utilisateurs vulnérables**
- **Diminution du contrôle sur ses propres données**

Recommandations

Limiter la collecte de données au strict nécessaire et renforcer le contrôle des utilisateurs

Interdire la surveillance intrusive des comportements en ligne sans consentement explicite et éclairé tout en permettant d'accéder et modifier nos données personnelles de manière simplifiée.

Protéger les données personnelles des jeunes

Assurer leur anonymisation et interdire leur revente ou toute autre exploitation abusive, profitant des technologies en IA pour prévenir leur collecte, tout en interdisant l'utilisation des données pour des pratiques commerciales agressives, même lorsque transformées.

Problématique #3

Perpétuation des biais et renforcement des inégalités

L'intelligence artificielle est de plus en plus utilisée pour automatiser des processus dans de nombreux domaines influençant la jeunesse, comme l'éducation, les réseaux sociaux, les services en ligne et les jeux vidéo. Pourtant, si elle est mal entraînée ou insuffisamment surveillée, l'IA peut involontairement reproduire et amplifier des biais raciaux, de genre, culturels ou socio-économiques.

Sur les réseaux sociaux, l'IA peut renforcer des bulles de filtre en exposant les jeunes à des contenus qui valident leurs opinions sans leur offrir une diversité de perspectives, imposant même parfois des visions erronées sans justification. Dans les jeux vidéo, les récits procéduraux et les comportements des personnages non-joueurs (PNJ) peuvent involontairement véhiculer des stéréotypes ou représenter certains groupes de manière caricaturale. De plus, les systèmes de modération automatisés peuvent injustement pénaliser des joueurs en raison de modèles biaisés, influençant ainsi leur relation à ces communautés.

Enfin, certaines pratiques d'IA appliquées aux services numériques ciblent des jeunes en fonction de critères problématiques comme leur origine, leur orientation sexuelle ou leur pouvoir d'achat, renforçant ainsi des inégalités et limitant les opportunités de certains groupes.

Conséquences

- **Renforcement des stéréotypes et des discriminations**
- **Traitement inéquitable dans les services automatisés**
- **Arbitrage injuste et censure involontaire**
- **Données de ciblage et de profilage se répercutant dans toute l'expérience numérique**

Recommandations

Favoriser une diversité des équipes de développement et de conception afin de repérer et limiter les biais dès la création.

Par des politiques d'embauche ou des réglementations sur la consultation en lien au contenu, tant du côté de l'intelligence artificielle que du point de vue multimédia.

Modification des approches de nettoyage de bases de données et d'entraînement en vue d'une neutralité encourageant la vérification humaine.

Afin de s'assurer que l'IA ne privilégie pas certaines catégories de jeunes au détriment d'autres, notamment dans les systèmes de recommandation, de sélection ou de "matchmaking".

Problématique #4

Absence de cadre éthique dans le développement

Contrairement à des disciplines comme la médecine, la psychologie ou la recherche, la conception et la programmation informatique ne requièrent ni formation éthique ni certification. Cela entraîne un vide réglementaire laissant place à des pratiques discutables, parfois plutôt par manque d'information et dépriorisation que de manière volontaire.

L'industrie privilégie souvent la rentabilité et l'engagement des utilisateurs au détriment de la responsabilité sociale. L'absence de réglementation stricte et la perception de la neutralité technologique renforcent cette tendance, tandis que la rapidité des innovations et les pressions économiques incitent à des pratiques de monétisation abusives plutôt qu'une réflexion sur le bien-être.

Dans les jeux vidéo, l'IA peut exploiter des données comportementales pour optimiser l'engagement et maximiser les profits grâce à des stratégies telles que les microtransactions dynamiques, la personnalisation de prix en fonction des comportements d'achat ou encore la mise en place de boucles de récompense manipulatrices. La même chose peut être vue dans les nombreux agents conversationnels qui sont offerts aux jeunes et qui misent sur un attachement et une ludification néfaste pour promouvoir leur plateforme ou d'autres logiciels des mêmes développeurs. Seul un encadrement éthique saurait alléger cette tendance.

Conséquences

- **Exploitation des joueurs et des consommateurs**
- **Collecte massive et opaque des données**
- **Renforcement des inégalités et des discriminations**
- **Déséquilibre des rapports de force**

Recommandations

Formation éthique et certification obligatoire pour les développeurs informatiques.

Les chefs d'entreprise, les investisseurs, les employés et les consommateurs doivent être formés à l'importance de prendre des décisions qui favorisent le bien commun et respectent les principes éthiques, puis être tenus responsables des choix faits durant le développement.

Régulation et obligations légales

Obligation de rendre publics les rapports sur le fonctionnement des algorithmes, avant, pendant et suite au développement, ainsi que de se soumettre à des audits éthiques réguliers.

Pistes pour mettre en place une certification éthique des algorithmes

Mise en place d'un organisme indépendant chargé de l'attribution de la certification éthique aux studios de développement, imposant ainsi certaines normes sectorielles basées sur les problématiques mentionnées. Cet organisme aurait pour tâches potentielles :

- 01 **Le test des algorithmes dans un but de prévention des biais ou d'évaluation de la marge d'erreur**

- 02 **La supervision d'obligations légales et d'exigences permettant la certification dans un but d'autorisation de la distribution de logiciels**

- 03 **La mise en place de mentions sur les plateformes et jeux informant les utilisateurs des détails de la conformité aux normes pour chaque produit, ainsi que de la provenance et des échanges des données utilisées par ceux-ci**

- 04 **La coordination des audits réguliers et contrôles indépendants pour assurer le respect des règles sur la durée entière de l'existence du produit numérique**

- 05 **La gestion des incitatifs économiques et fiscaux à la certification, tels que: crédits d'impôt, récompenses, reconnaissances publiques et subventions**

- 06 **La consultation vers la modification des lois et règlements sur la protection des créations numériques afin de favoriser l'ouverture des codes**

- 07 **La promotion de l'instrumentalisation de l'IA pour faciliter une bonne hygiène numérique, tel que par le désengagement des utilisateurs après une période ou un effort donnés, ou encore la régularisation des pratiques de monétisation prédatrices.**

Conclusions

- **Enseigner une pratique critique de l'intelligence artificielle.**
- **Protéger les jeunes contre les pratiques abusives et la collecte de leurs données.**
- **Encourager les studios à concevoir des logiciels plus transparents, neutres et justes.**
- **Responsabiliser l'industrie du jeu vidéo face aux enjeux sociaux et moraux.**

Pour toutes questions:

-  405 Rue Sainte-Catherine Est
Montréal, QC
H2L 2C4
-  bonneau.jonathan@uqam.ca

Bibliographie

- [1] Barraud, B. (2022). *Éthique de l'intelligence artificielle*. L'Harmattan. ISBN : 978-2-14-026517-4.
- [2] Bertail, P., Bounie, D., Cléménçon, S., & Waelbroeck, P. (2019). *Algorithmes : biais, discrimination et équité*. Télécom ParisTech. <https://www.telecom-paris.fr/wp-content/uploads/2019/02/Algorithmes-Biais-discrimination-equite.pdf>
- [3] Bourguine, P. (2019). *Réflexions sur les enjeux de l'IA et les questions d'éthique*. ResearchGate. https://www.researchgate.net/publication/336216393_Reflexions_sur_les_enjeux_de_l%27IA_et_les_questions_d%27ethique
- [4] Burtell, M., & Woodside, T. (2023). *Artificial Influence: An Analysis Of AI-Driven Persuasion*. arXiv preprint arXiv:2303.08721. <https://doi.org/10.48550/arXiv.2303.08721>
- [5] Commission Nationale de l'Informatique et des Libertés (CNIL). (2022, décembre 15). *IA : la CNIL publie ses premières recommandations sur le développement des systèmes d'intelligence artificielle*. CNIL. <https://www.cnil.fr/fr/ia-la-cnil-publie-ses-premieres-recommandations-sur-le-developpement-des-systemes-dintelligence>
- [6] Conrad-Boisnier, C. (2023, 4 septembre). *Sexisme, jeu vidéo et IA : le problème avec le deep learning*. Contrepoints. <https://mastersts.hypotheses.org/7419>
- [7] Demichelis, R. (2024). *L'intelligence artificielle, ses biais et les nôtres*. Éditions du Faubourg. ISBN : 978-2-493594-65-5.
- [8] Farchy, J., & Denis, J. (2020). *La culture des données. Intelligence artificielle et algorithmes dans les industries culturelles*. Presses des Mines, Les Cahiers de l'EMNS.
- [9] Marsh, J. P., Gorayska, B., & Mey, J. L. (1999). *Humane interfaces: Questions of method and practice in cognitive technology*. North Holland.
- [10] Melhart, D., Togelius, J., Mikkelsen, B., Holmgård, C., & Yannakakis, G. N. (2023). *The Ethics of AI in Games*. IEEE Transactions on Affective Computing. <https://doi.org/10.48550/arXiv.2305.07392>
- [11] Ménissier, T. (2023). *Les quatre éthiques de l'intelligence artificielle*. Revue d'anthropologie des connaissances, 17(2). <https://doi.org/10.4000/rac.29961>
- [12] Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche. (n.d.). *Les enjeux éthiques de l'intelligence artificielle*. éducol. <https://eduscol.education.fr/3251/les-enjeux-ethiques-de-l-intelligence-artificielle>
- [13] Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2021). *AI Certification: Advancing Ethical Practice by Reducing Information Asymmetries*. ResearchGate. https://www.researchgate.net/publication/351804294_AI_Certification_Advancing_Ethical_Practice_by_Reducing_Information_Asymmetries
- [14] Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2021). *Towards Fairness Certification in Artificial Intelligence*. ResearchGate. https://www.researchgate.net/publication/352177742_Towards_Fairness_Certification_in_Artificial_Intelligence
- [15] Pérez, J., Castro, M., & López, G. (2023). *Serious Games and AI: Challenges and Opportunities for Computational Social Science*. IEEE Access, 11, 62051-62061. <https://doi.org/10.1109/ACCESS.2023.3286695>
- [16] Riva, G., Anguera, M. T., Wiederhold, B. K., & Mantovani, F. (Eds.). (2006). *From communication to presence: Cognition, emotions and culture towards the ultimate communicative experience – Festschrift in honor of Luigi Anolli (Vol. 9)*. IOS Press.
- [17] Rosenberg, L. (2023). *The Manipulation Problem: Conversational AI as a Threat to Epistemic Agency*. arXiv preprint arXiv:2306.11748. <https://doi.org/10.48550/arXiv.2306.11748>
- [18] Zhu, J., & Mostafavi, B. (2021). *The Social Responsibility of Game AI*. arXiv preprint arXiv:2105.15122. <https://doi.org/10.48550/arXiv.2105.15122>