

---

# RETOUR D'EXPÉRIENCE AVEC UN JEU-DÉBAT POUR SENSIBILISER À L'IA POUR LA SURVEILLANCE SANITAIRE

---

**Carole ADAM<sup>1</sup>**

Univ. Grenoble-Alpes, LIG, Grenoble, France

**Cédric LAURADOUX<sup>2</sup>**

Univ. Grenoble-Alpes, Inria, Grenoble, France

**Résumé.** L'importance de l'Intelligence Artificielle dans la société, et des décisions qui lui sont déléguées (accès aux formations supérieures, détermination de peines de prison, conduite autonome de véhicules, etc) nécessite d'éduquer la population à ses enjeux. Tout le monde ne peut pas avoir une connaissance technique précise sur l'IA, mais comme souligné par l'UNESCO, il est essentiel que la population ait une compréhension basique du fonctionnement de ces algorithmes pour choisir en connaissance de cause de les utiliser ou pas. Pour cela nous avons développé un jeu sérieux à destination des lycées, sous la forme d'un débat citoyen ayant pour but de choisir une solution d'IA pour contrôler une épidémie. Cet article présente un retour d'expérience de l'utilisation de ce jeu en classe. Nous décrivons le fonctionnement d'une séance, les résultats qualitatifs et quantitatifs des premières sessions animées, ainsi que les biais et limites de cette activité.

**Mots-clés.** Jeu sérieux, Intelligence Artificielle, surveillance, vie privée, COVID-19.

**Abstract.** The importance of Artificial Intelligence in society, and of the decisions delegated to it (access to higher education, prison sentences, autonomous vehicles) entails the necessity to educate the population to the associated stakes. Not everybody can learn the technical details behind AI, but as highlighted by UNESCO, it is essential to provide the population with a general understanding of AI algorithms so that they can choose wisely whether to use them or not. To this end, we propose a serious game targeted at high school students, in the form of a civic debate aiming at selecting an AI solution to control a pandemic. This article presents a return of experience of using this game in class. We detail the progression of game sessions, provide qualitative and quantitative results about the first test sessions, and discuss biases and limitations of this activity.

**Keywords.** Serious game, Artificial Intelligence, surveillance, privacy, COVID-19.

## Introduction

L'Intelligence Artificielle est de plus en plus omniprésente dans nos vies. On lui délègue nombre de décisions importantes : accès aux formations supérieures, détermination du risque de récidive et choix des peines de prison, conduite autonome de véhicules, tri automatique de CV avant recrutement... Elle aurait en effet pour avantage de prendre des décisions objectives débarrassées des biais humains, et surtout des décisions plus rapides sur la base de masses de données difficilement analysables par un humain. Cependant, même si les technologies d'IA peuvent offrir de nombreux bénéfices, elles ne sont pas du tout dépourvues de biais. Au contraire elles reproduisent et amplifient les biais humains présents dans les données qui les nourrissent. Elles posent ainsi de nombreux risques : potentiels biais et erreurs, discrimination, ou encore mauvaise

---

<sup>1</sup> [carole.adam@imag.fr](mailto:carole.adam@imag.fr)

<sup>2</sup> [cedric.lauradoux@inria.fr](mailto:cedric.lauradoux@inria.fr)

utilisation. La crise sanitaire a été l'occasion de prendre conscience d'une partie de ces risques avec une surveillance accrue facilitée par l'IA, l'utilisation élargie de la reconnaissance faciale, ou encore de drones.

Il est donc nécessaire pour les personnes soumises aux décisions de ces algorithmes d'en comprendre le fonctionnement, pour choisir en connaissance de cause de les utiliser (ou pas). L'UNESCO (UNESCO, 2021) souligne d'ailleurs cette nécessité. Or les formations à l'IA concernent pour l'instant surtout des ingénieurs ou chercheurs spécialistes de ce domaine. La population générale, elle, a très peu de connaissances sur le sujet et est très sensible aux idées (plus ou moins fausses) véhiculées par les médias : une IA objective, infaillible, qui sauvera le monde ou au contraire le conduira à sa perte. Nous proposons donc un jeu sérieux (Adam et Lauradoux, 2022) à destination des lycéens, sous la forme d'un débat citoyen ayant pour but de choisir une solution d'IA pour contrôler une épidémie. Les participants incarnent les rôles de citoyens avec différents profils dans une ville (pas si) futuriste. Ils doivent débattre au sujet de nouvelles technologies d'IA qui pourraient être adoptées dans leur ville pour lutter contre une épidémie. Ce jeu fait partie d'une série plus large de jeux à débattre, chacun concernant une application sociétale différente (santé, transports...) de technologies d'IA. Cette série de jeux-débats cible les adolescents de niveau collège ou lycée. Le débat est toujours suivi d'un débriefing permettant aux joueurs de partager leur expérience, d'en apprendre plus sur l'IA, et de discuter de ses implications dans leur vie réelle.

Dans cet article, nous discutons d'abord (Section 2) de la nécessité d'éducation du grand public au sujet de l'IA, et de l'apport des jeux sérieux dans ce domaine. La Section 3 présente notre jeu à débattre. La Section 4 se focalise plus en détail sur les solutions d'Intelligence Artificielle que le jeu permet d'introduire. La Section 5 insiste sur l'importance du débriefing pour obtenir un impact pédagogique. Enfin la Section 6 offre des retours d'expérience tirés des premières sessions avec notre jeu, et la Section 7 conclut l'article.

## 2. Nécessité d'éducation

### 2.1. Éducation à l'Intelligence Artificielle

Pendant la pandémie de COVID-19, l'importance d'éduquer la population est devenu encore plus apparente. En effet, le manque d'information, voire la désinformation diffusée sur les réseaux sociaux, peut avoir des conséquences graves. Ainsi, les théories du complot et autres *fake news* ou *infox* se propagent rapidement et peuvent augmenter la réticence envers, voire le rejet des vaccins (Nieves-Cuervo et al., 2021, Lu et Sun, 2022), avec des conséquences néfastes sur la santé. En outre, même si l'Intelligence Artificielle peut aider à lutter contre le COVID-19, il existe également des dangers à éviter : (Naudé, 2020) insiste sur le besoin d'un équilibre prudent entre la confidentialité des données et la santé publique. Tout le monde doit être conscient des avantages et des dangers potentiels de l'Intelligence Artificielle, que ce soit dans le contexte spécifique de la pandémie ou plus généralement.

Cette exigence d'éducation à l'IA est particulièrement mise en avant dans le dernier rapport de l'UNESCO (UNESCO, 2021, notre traduction). Elle ne devrait pas concerner uniquement les technologies sous-jacentes, mais également la valeur et l'utilisation des données personnelles, ainsi que les implications éthiques des technologies d'IA.

*parce que vivre dans des sociétés numériques exige de nouvelles pratiques éducatives, une réflexion éthique, une pensée critique, des pratiques de conception responsables et de nouvelles compétences, au vu des conséquences sur le marché de l'emploi, l'employabilité et la participation citoyenne (p.18)*

*La sensibilisation et la compréhension par le public des technologies d'IA et de la valeur des données devraient être favorisées par une éducation ouverte et accessible, l'engagement citoyen, l'éducation aux compétences numériques et à l'éthique de l'IA, [...] de sorte que tous les membres de la société puissent prendre des décisions éclairées concernant leur utilisation des systèmes d'IA, et soient protégés contre toute influence induite. (p.24)*

*Les États membres doivent promouvoir des programmes généraux de sensibilisation aux avancées de l'IA, notamment concernant les données et les possibilités et défis découlant des technologies d'IA, les répercussions des systèmes d'IA sur les droits de l'homme, [...] et leurs implications. Ces programmes devraient être accessibles aux spécialistes comme aux non-spécialistes. Les États membres doivent développer des programmes d'éthique de l'IA pour tous les niveaux, et promouvoir la collaboration entre l'éducation aux compétences techniques de l'IA et l'éducation aux aspects humanistes, éthiques et sociaux de l'IA. (p.33)*

## **2.2. Jeux sérieux**

Dans le but d'impliquer l'ensemble de la population dans l'apprentissage de l'IA et la compréhension de son utilisation pendant la crise du COVID-19, nous pensons que les jeux sérieux (Crookall, 2010) peuvent être un outil puissant. Les jeux sérieux ont été utilisés avec succès dans différents domaines : éduquer la population aux comportements à adopter en cas de catastrophes naturelles, par exemple les inondations (Rebolledo-Mendez et al., 2009, Taillandier et Adam, 2018) ; sensibiliser au changement climatique et aux actions possibles (Wu et Lee, 2015, Adam et al., 2022) ; former les compétences médicales et chirurgicales (Graafland et al., 2012) ou les compétences communicatives des futurs médecins (Jackson et Back, 2011) ; ou plus récemment expliquer la pandémie et les mesures sanitaires pour améliorer leur acceptabilité (Cottineau et al, 2020, Adam et Arduin, 2022).

Les jeux sérieux n'ont pas besoin d'être informatisés, et nous utilisons ce terme pour désigner tout jeu d'apprentissage, qui utilise des éléments ludiques pour atteindre un objectif pédagogique. Il existe différentes formes de jeux sérieux, des simulations informatiques (Taillandier et Adam, 2018), aux jeux de société ou de cartes (Wu et Lee, 2015), aux récits interactifs, aux jeux de rôles (Jackson et Back, 2011), ou aux débats (Simonneaux, 2001). Compte tenu de notre objectif déclaré d'impliquer le grand public, nous avons conçu un jeu "débranché" auquel tout le monde peut jouer, quelle que soit son éducation, sans ordinateur. Par ailleurs, nous souhaitons développer les capacités d'argumentation des futurs citoyens, pour leur permettre de faire des choix éclairés. Nous proposons donc un jeu sérieux sous la forme d'un jeu de rôle et de débat. La conception de notre jeu est expliquée dans la section suivante.

## **3. Notre jeu sérieux : l'IA contre une épidémie à Wafer city**

### **3.1. Association de médiation scientifique**

Notre jeu fait partie d'une série plus large de jeux-débats sur l'IA dans divers domaines, développés par une association de médiation scientifique (L'Arbre des Connaissances), dédiée à l'ouverture de la science vers les citoyens. Le but de cette série de jeux-débats est d'initier les adolescents au débat citoyen sur des questions d'actualité mêlant sciences et société. Ces jeux se déroulent tous dans la même ville virtuelle (plus ou moins futuriste), appelée Wafer city, une ville de taille moyenne avec un million d'habitants. La moitié de la population vit en centre-ville, l'autre moitié en zone péri-urbaine ou rurale. Les joueurs incarnent le conseil municipal de la ville, dans lequel sont représentés 5 rôles (voir Paragraphe 3.5) : les Alter Wafer, les Central Wafer, les Seniors, les Futuristes, les Lointains. Ces rôles ont été élaborés par l'association dans le but de permettre un débat apaisé (les élèves ne parlant pas en leur nom propre), mais aussi de faire comprendre aux participants la variété des points de vue et la nécessité de considérer

d'autres opinions que la sienne propre et d'argumenter.

L'objectif de ces jeux-débats n'est pas tant d'éduquer les citoyens au fonctionnement précis des technologies d'IA, et il n'y a donc aucun prérequis pour participer. Il s'agit plutôt d'apprendre aux futurs citoyens à prendre des décisions éclairées, en pesant à la fois les bénéfices et les risques avant d'adopter toute nouvelle technologie, même quand elle est présentée comme une solution « miracle » à un problème de société.

Dans chaque jeu de la série, une question différente est posée au conseil municipal. Les deux premiers épisodes de la série étaient consacrés à l'IA dans les transports (octobre 2018) et à l'IA dans la santé (février 2019). Dans le nouvel épisode décrit ici, le conseil municipal doit faire face à une épidémie. Cet épisode a été développé par les deux auteurs avec l'aide de deux membres de l'association et s'inscrit dans le cadre plus générique de la série des jeux débats décrit ci-dessus, La phase de conception s'est étalée sur environ 1 an à partir de février 2021, avec de premières expérimentations en octobre 2021, et la sortie officielle en mars 2022.

### **3.2. Séquence pédagogique d'un jeu-débat**

La séquence pédagogique pour une session de jeu-débat est la suivante :

1. La session commence par une introduction des modérateurs, qui se présentent, interrogent les participants sur leur conception de l'IA, puis donnent quelques définitions, avant d'exposer le scénario du jeu (cf paragraphe 3.3).
2. Les participants sont répartis aléatoirement en cinq groupes, prennent connaissance du rôle qu'ils vont incarner au conseil municipal (cf paragraphe 3.5), et le décrivent aux autres participants.
3. Trois solutions sont présentées l'une après l'autre (cf Section 3.4 et détails Section 4). Pour chaque solution, les participants reçoivent la fiche de description puis discutent dans leur groupe de ses avantages et inconvénients, selon le point de vue du rôle qu'ils représentent. Les animateurs (cf Section 3.6) peuvent leur donner des pistes sous forme de questions, en s'aidant notamment des échelles d'évaluation (cf Paragraphe 5.2) mais restent toujours neutres.
4. Chaque groupe distribue ensuite un total de 5 points entre les solutions, et fournit oralement ses scores et ses arguments, que les modérateurs reportent sur le tableau de jeu (voir un exemple en Tableau 1).
5. La solution gagnante est celle qui cumule le plus de points en additionnant ceux attribués par les 5 groupes. Les résultats sont annoncés, discutés et débattus. Certains groupes peuvent être amenés à changer leurs scores en conséquence, ce qui peut modifier le résultat (voir un exemple de débat en Section 6.2).
6. Après le débat, les modérateurs conduisent une phase de débriefing, qui est essentielle à l'apprentissage (Crookall, 2010). Le contenu de cette phase est présenté en détails en Section 5.
7. L'enseignant de la classe récupère de la documentation et des références pour approfondir le sujet avec ses élèves en classe ultérieurement (voir la liste du matériel de jeu au Paragraphe 3.7)

### **3.3. Scénario du jeu : épidémie à Wafer City**

Dans notre jeu, Wafer city est menacée par une épidémie. Chaque citoyen de la ville dispose d'un dossier médical numérique (contenant les résultats d'analyses, les traitements en cours...) géré par la ville et à la disposition de ses décideurs politiques, qui doivent prendre

les meilleures mesures possibles afin de limiter la propagation de l'épidémie tout en faisant face aux enjeux sociétaux suivants :

- Préservation de la santé mentale. Les règles de confinement sont jugées trop strictes et préjudiciables à la santé mentale, notamment pour les personnes en bonne santé qui ne se sentent pas concernées. Il est essentiel de trouver des mesures qui ne soient pas trop contraignantes pour la population.
- Équité. Dans un double souci d'efficacité et de santé mentale, il est important que tout le monde respecte les règles équitablement, mais sans surcharger les policiers de missions supplémentaires.
- Efficacité. Les personnes asymptomatiques pourraient propager l'épidémie sans le savoir, ce qui rendrait insuffisant le dépistage et la mise en quarantaine des seules personnes symptomatiques. Afin de protéger les personnes les plus vulnérables, il faut prendre des mesures efficaces pour stopper l'épidémie au plus vite.

La description du contexte ne mentionne volontairement jamais le nom du virus responsable de l'épidémie à Wafer city, et évite toute allusion au COVID-19. En effet, le jeu vise à discuter des implications sociétales de l'IA dans un contexte plus large, et devrait rester d'actualité même après la fin de la pandémie actuelle. De même, le contexte juridique est volontairement flou, les mêmes technologies pouvant s'adapter à d'autres objectifs légaux. Par ailleurs, nous voulions également prendre de la distance par rapport à un contexte sanitaire peut-être encore sensible.

### 3.4. Sowana

La société Sowana est spécialisée dans l'Intelligence Artificielle et a déjà travaillé avec la ville sur d'autres sujets, comme les transports en commun ou la santé. Sowana a répondu à l'appel à projets de Wafer city concernant l'épidémie avec trois solutions : *Eye'Wana*, *Wana'Like* et *Wana'Pass*, qui seront détaillées dans la prochaine section. Ces trois solutions font appel à diverses technologies d'Intelligence Artificielle et permettent donc d'éduquer les participants sur leur fonctionnement technique et sur leurs risques.

- *Eye'Wana* utilise la vidéo-surveillance et des algorithmes de reconnaissance faciale pour garantir l'égalité devant la loi : tout le monde est contrôlé et verbalisé de la même manière, sans biais humains.
- *Wana'Like* utilise des algorithmes de recommandation qui analysent le profil des utilisateurs et les données d'affluence, pour recommander des sorties plus sécurisées dans des lieux peu fréquentés.
- *Wana'Pass* est un passeport sanitaire à points, qui analyse les données médicales pour attribuer à chacun des droits d'accès aux lieux à risque adaptés à son état de santé actuel.

Ces solutions sont décrites sur des fiches (voir un exemple en Figure 2) contenant à la fois un descriptif publicitaire de la solution par SOWANA (forcément très positif), et un témoignage d'un utilisateur réel (déjà plus critique et suggérant quelques failles dans l'argumentaire). L'animateur explique aussi rapidement chaque solution avant de lancer les discussions par groupe de rôle.

### 3.5. Rôles représentés au conseil municipal

Les différentes solutions proposées par Sowana ont toutes des avantages et des inconvénients, et seront plus ou moins acceptables pour différents groupes d'utilisateurs. Dans notre jeu-débat, les participants ne jouent pas en leur propre nom, mais assument un rôle en tant que membre d'un groupe spécifique de citoyens. L'idée de faire jouer aux utilisateurs un rôle différent n'est pas

nouvelle et s'est avérée favoriser le changement de perspective (Jarvis et al., 2002, Resnick et Wilensky, 1998). Nous voulions aussi que les participants prennent de la distance avec leur personnage pour garantir un débat plus apaisé, d'autant plus sur un sujet très sensible alors que l'épidémie de COVID-19 n'est toujours pas terminée.

Les participants au jeu sont donc répartis aléatoirement en 5 petits groupes de citoyens, dont le rôle est décrit en détail sur une fiche qu'ils reçoivent et lisent ensemble (cf par exemple Figure 1) :

- Les Central Wafer : les travailleurs, commerçants, artisans du centre-ville, ils tiennent à leur liberté, à leur mobilité et celle de leurs clients. Ils ont des revenus corrects mais peu de temps libre.
- Les Alter Wafer : plutôt jeunes à sensibilité écologiste ou alter-mondialiste, focalisés sur le bien-être de tous et la protection de la nature. Ils vont au travail à pied ou vélo, et demandent des garanties avant d'adopter de nouvelles technologies.
- Les Seniors : des personnes âgées de plus de 60 ans, retraitées, focalisées sur le maintien de leur autonomie. Ils peuvent vivre en ville ou à la campagne.
- Les Futuristes : surtout des jeunes, amateurs de nouvelles technologies, très connectés et actifs en ligne, ils adorent l'innovation mais ont des revenus limités.
- Les Lointains : groupe très hétérogène, vivant en dehors du centre-ville (banlieue ou zone rurale), pour une meilleure qualité de vie ou pour les loyers plus bas, ils font donc de nombreux trajets vers la ville.



**Figure 1** : Extrait des cartes du jeu décrivant 2 des rôles. Illustration © Marie Jamon.

Les membres de chaque groupe discutent entre eux pour trouver des aspects positifs (qu'est-ce qu'on gagne ?) et négatifs (qu'est-ce qu'on perd ?) à chaque solution, du point de vue du profil qu'ils représentent. Cet exercice les force à ne pas accepter ou rejeter « en bloc » une solution, mais à toujours réfléchir à l'équilibre bénéfice-risque. Ils peuvent être guidés dans leur réflexion par le témoignage citoyen figurant sur chaque fiche solution, et par les 3 objectifs sociétaux figurant sur l'ordre du jour. Les animateurs peuvent aussi les faire se questionner sur certains points qu'ils auraient oubliés, tout en restant neutres.

Chaque groupe alloue ensuite un total de 5 points entre les 3 solutions. Chaque groupe-rôle converge ainsi vers des scores (très) différents pour les solutions. Ces scores seront additionnés

pour déterminer la solution gagnante et donc choisie par le conseil municipal. Ces scores devront être argumentés, et donneront aussi lieu à une discussion lors du débriefing.

### **3.6. Rôle de l'animateur**

Le ou les animateurs sont présents pour aiguiller les élèves pendant les débats. Ils s'assurent notamment que chaque groupe a bien compris le rôle incarné, et respecte bien le point de vue de ce rôle pour évaluer les solutions. Ils peuvent aussi aiguiller les élèves en les questionnant et en les faisant s'interroger sur certains aspects des solutions. Les animateurs sont guidés pour cela par les 3 objectifs sociétaux définis par l'ordre du jour (Paragraphe 3.3) et par les échelles d'évaluation figurant dans le guide de l'animateur (Paragraphe 5.2). En revanche, les modérateurs restent neutres, et ne donnent pas de réponses toutes faites ou d'opinion sur les solutions aux élèves. Les animateurs permettent aussi de modérer les débats et d'empêcher les dérives. Enfin, les animateurs conduisent la phase de débriefing final lors de laquelle la classe prend du recul sur toutes les notions abordées lors du jeu (voir Section 5).

### **3.7. Matériel de jeu**

Le matériel de jeu<sup>1</sup> comprend les éléments suivants :

- Cartes décrivant les rôles, une carte différente à distribuer à chaque groupe ;
- Une fiche expliquant l'ordre du jour du conseil municipal, et les problèmes à résoudre ;
- Une fiche par solution, qui fournit l'argumentaire marketing de Sowana, des détails techniques sur l'IA sous-jacente, et le témoignage d'un citoyen destiné à soulever quelques premières interrogations ;
- Un guide du modérateur destiné aux professeurs qui superviseront le jeu. Ce guide contient des informations plus techniques pour approfondir les notions en cours, des définitions utiles, des liens vers de la documentation supplémentaire, des questions pour alimenter le débat, et des informations sur des solutions d'IA similaires déjà réelles à présenter lors du débriefing.

Ces éléments sont fournis lors des séances que nous organisons, mais ils peuvent également être imprimés par tout enseignant désireux de jouer le jeu avec ses élèves, de manière autonome.

## **4. Les solutions d'IA**

Cette section détaille les trois solutions d'IA proposées par Sowana pour combattre l'épidémie à Wafer City, avec leur principe, les notions d'IA qu'elle permet d'introduire aux participants, et les questions éthiques qu'elle soulève. Ces questions éthiques pourront être suggérées pendant la phase de réflexion, puis seront abordées en détails au débriefing.

### **4.1. Eye'Wana**

*Eye'Wana* s'appuie sur la vidéosurveillance pour s'assurer que les règles sanitaires sont respectées de la même manière par tous les citoyens à tout moment. *Eye'Wana* peut détecter diverses infractions aux règles sanitaires (non-port du masque ou non-respect de la distanciation physique ; violation du couvre-feu ou du confinement ; etc.), en identifier l'auteur, et lui notifier l'amende qu'il devra payer en conséquence.

---

<sup>1</sup> Le matériel de jeu et le livret sont disponibles gratuitement en ligne <https://jeudebat.com/jeux/lintelligence-artificielle/la-securite-sanitaire/>

**Technologie.** Cette solution repose sur trois briques : (i) un réseau de vidéosurveillance et une flotte de drones collectent les vidéos et les images des lieux publics, (ii) des algorithmes de vision par ordinateur extraient les visages de ces images, et (iii) des algorithmes de reconnaissance faciale identifient les individus à partir de leur visage.

**Dans la réalité.** Les technologies de reconnaissance faciale sont de plus en plus utilisées (par la police dans certains pays, pour les contrôles d'identité dans certains aéroports, pour déverrouiller un téléphone, etc). Des technologies similaires à Eye'Wana ont été utilisées en Chine (Chun, 2020), où un vaste réseau de caméras de vidéosurveillance avec drones et caméras thermiques, complété par la reconnaissance faciale, a été déployé pour mettre en œuvre un contrôle social pendant la pandémie de COVID-19. Mais des drones ont aussi été expérimentés en France pour surveiller les plages de Nice pendant le confinement<sup>1</sup>.

**Bénéfice-risque.** Eye'Wana promet de décharger les forces de police tout en évitant les préjugés et biais humains : tout le monde aurait la même probabilité d'être contrôlé et verbalisé en cas de non-respect des règles. Cette solution est donc proposée pour répondre au besoin d'égalité devant la loi, mais au prix d'une forte intrusion dans la vie privée des personnes, puisqu'elle traque leurs moindres faits et gestes. Par ailleurs, contrairement à ses promesses d'objectivité, la reconnaissance faciale n'est en fait pas infaillible, et implique même un fort risque de discrimination (Bacchini et Lorusso, 2019, Najibi, 2020). Cela a d'ailleurs conduit à son interdiction récente dans certains états aux États-Unis. Elle pose aussi un certain nombre de questions éthiques (Martinez-Martin, 2019).

## 4.2. Wana'Like

Wana'Like est une application gratuite pour *smartphone* destinée à aider les citoyens à éviter les zones surpeuplées où le risque de contamination est plus élevé. Sowana tient à jour une liste de lieux recommandés, vérifie s'ils respectent le protocole sanitaire, et peut les bannir si des contaminations sont signalées dans leur locaux. L'application analyse les goûts et intérêts de l'utilisateur à partir de son activité en ligne (publications sur les réseaux sociaux, requêtes sur le moteur de recherche Sowana...) pour créer un profil utilisateur. Elle utilise également des données de localisation en temps réel pour déduire la fréquentation actuelle des lieux. En combinant ces données, Wana'Like peut faire des suggestions de loisirs personnalisées mais sûres, et propose des réductions dans les lieux partenaires pour inciter davantage les utilisateurs à suivre les recommandations.



Figure 2 : Extrait de la carte de jeu décrivant Wana'Like - illustration © Marie Jamon.

**Technologie.** Cette solution met en lumière le fonctionnement d'algorithmes très utilisés dans la vie courante, notamment sur les réseaux sociaux, pour la recommandation de contenu, de vidéos

<sup>1</sup> <https://news.trust.org/item/20200320105803-ztaq0>

à regarder, de films à aller voir au cinéma, d'objets à acheter, d'un restaurant, etc. Un algorithme de recommandation permet de filtrer des informations, des objets, afin de trouver automatiquement ce qui pourrait intéresser son utilisateur. Il fonctionne en 3 étapes: (i) recueillir de l'information sur l'utilisateur, soit de manière active (explicite) en lui demandant de classer, noter des éléments, soit de manière passive (implicite) en analysant son comportement en ligne, ses choix antérieurs ; (ii) construire et maintenir un profil de l'utilisateur contenant ses préférences ; (iii) proposer des recommandations, basées sur ses préférences pour les propriétés des éléments disponibles, ou sur les choix d'autres utilisateurs similaires, ou sur une combinaison de ces techniques.

**Dans la réalité.** Wana'Like partage de nombreuses similitudes avec les réseaux sociaux, qui utilisent des algorithmes de recommandation pour faire des suggestions personnalisées à l'utilisateur, et l'orienter vers des contenus (posts, vidéos, photos) susceptibles de l'intéresser. Pendant la pandémie, on a aussi vu Google Maps proposer une mesure de l'affluence dans les lieux publics basée sur les données des smartphones pour recommander d'éviter certains lieux à certaines heures, ou encore le projet NoBis se basant sur le scan de QR codes en entrée et sortie d'un lieu référencé (Avanzi et al., 2021).

**Bénéfice-risque.** Cette solution est beaucoup moins restrictive, mais potentiellement aussi moins efficace, et pose la question de l'accessibilité aux personnes ne possédant pas de *smartphone*. En outre, malgré leur apparence inoffensive (les données sont fournies par l'utilisateur de son plein gré, dans le but d'en tirer un bénéfice), les algorithmes de recommandation posent question. En effet, ils sont très gourmands en données personnelles qui pourraient ensuite être utilisées à mauvais escient, comme pour manipuler les opinions lors d'élections (Aral et Eckles, 2019). Ils ont aussi tendance à orienter vers des contenus de désinformation (Faucon et al., 2021), comme les théories du complot ou anti-vaccination pendant la pandémie de COVID. Des travaux ont ainsi montré une mortalité accrue chez les personnes s'informant principalement sur les réseaux sociaux (Nieves-Cuervo et al., 2021). Ces données personnelles pourraient aussi permettre de déduire des informations sensibles, comme l'orientation sexuelle à partir de photos (Wang et Kosinski, 2018) ou de liens Facebook (Jernigan et Mistree, 2009), ou la fréquentation de lieux sensibles d'après les données de localisation, comme les cliniques d'avortement (Google annonce d'ailleurs récemment<sup>2</sup> supprimer automatiquement les données de visites de tels lieux sensibles). Il existe donc de grands risques liés à l'utilisation commerciale des données personnelles (publicité), la perte de vie privée, ou la manipulation d'opinion. Il y a aussi un conflit d'intérêts entre utilisateur et plateforme, puisque les algorithmes visent à maximiser le temps passé en ligne, au détriment de la santé (mentale) de l'utilisateur (Wells et al., 2021). Enfin, il y a un risque de discrimination des personnes ne possédant pas de *smartphone*.

### 4.3. Wana'Pass

*Wana'Pass* est un passeport sanitaire à points. Il analyse quotidiennement le dossier médical de l'utilisateur et les données de ses capteurs (fréquence cardiaque, température...) pour en déduire son état de santé et son niveau de risque. Il lui attribue ensuite un certain nombre de points, les utilisateurs en meilleure santé recevant plus de points et les personnes infectées n'en recevant aucun. Ces points seront nécessaires pour accéder aux lieux publics (magasins, restaurants, musées, etc.), les endroits les plus risqués nécessitant plus de points d'accès. Chaque lieu devra vérifier et débiter les points d'accès nécessaires ; une fois qu'il n'a plus aucun point (ou tant qu'il est contagieux), l'utilisateur ne peut plus entrer dans ces lieux.

---

<sup>2</sup> Voir par exemple : <https://www.lesnumeriques.com/vie-du-net/google-supprimera-les-donnees-des-personnesvisitant-des-cliniques-pratiquant-l-avortement-n187043.html>

**Technologie.** Wana'pass peut être considéré comme le résultat des progrès de l'IA médicale (Petrooulos, 2020), et montre une utilisation possible de l'IA pour analyser des données médicales. Les premiers algorithmes d'IA médicale utilisaient des systèmes experts pour coder les connaissances médicales et les réutiliser sur de nouveaux scénarios, pour aider par exemple à poser un diagnostic ou proposer un traitement. Les algorithmes d'apprentissage automatique permettent maintenant d'analyser d'énormes volumes de données, qu'un humain ne pourrait traiter, pour apprendre à résoudre une tâche sans être explicitement programmés pour.

**Dans la réalité.** De nombreuses applications d'IA médicale existent déjà, notamment en radiologie, comme la détection de tumeurs sur des mammographies (Sheth et Giger, 2020). Pendant la pandémie de COVID, la France et d'autres pays ont mis en place un passeport sanitaire, mais basé uniquement sur le statut vaccinal, la guérison d'une infection au COVID, ou le résultat d'un test de dépistage. Ce passeport sanitaire manipule donc déjà des données médicales, privées. La mise en place récente du Dossier Médical Partagé en France rend une solution similaire encore plus crédible. Dans un autre contexte, Wana'Pass rappelle aussi le système de crédit social chinois (Shen, 2019, Liu, 2019).

**Bénéfice-risque.** Cette solution vise à répondre au problème d'un traitement équitable plutôt qu'égalitaire : elle protège les personnes à risque et isole les personnes contagieuses, tout en laissant les personnes en bonne santé vivre plus normalement afin de protéger leur santé mentale. Cependant, elle pose la question de discriminer et d'isoler encore plus les personnes âgées sous prétexte de les protéger. Sa mise en place exige aussi des garanties fortes sur la protection des données médicales utilisées. Par ailleurs, l'obligation de présenter un passeport vaccinal pour accéder à un certain nombre de lieux publics, pose le problème des discriminations créées par l'accès inégal aux vaccins (Tanner et Flood, 2021, Gostin et al., 2021, Kofler et Baylis, 2020).

## 5. Débriefing

L'objectif de ce jeu est d'éduquer les participants au sujet de l'Intelligence Artificielle, non seulement d'un point de vue technologique (le fonctionnement général des algorithmes sous-jacents) mais aussi d'un point de vue éthique (les questions soulevées par son utilisation). Le simple fait de jouer et débattre n'est cependant pas suffisant, et il est essentiel de conduire un bon débriefing pour permettre aux joueurs d'assimiler le message.

Le débriefing est en effet une part importante d'un jeu sérieux, et indispensable à l'apprentissage, (Crookall, 2010, Whalen et al., 2018, Dieleman et Huisinigh, 2006). Il poursuit plusieurs buts : clarifier les concepts manipulés pendant la session de jeu, afin d'améliorer l'apprentissage ; relier ces concepts, simplifiés pendant le jeu, avec la complexité du monde réel, pour s'assurer que les participants puissent appliquer leurs connaissances en situation réelle ; et partager les expériences et la réflexion entre les participants. Le débriefing doit être conduit par un modérateur informé, avec à l'esprit un objectif pédagogique clairement défini. Les paragraphes qui suivent décrivent certaines phases du débriefing de notre jeu-débat.

### 5.1. Partage et débat sur les scores

À la fin du débat, chaque groupe donne les notes qu'il a attribuées à chaque solution. L'animateur reporte au tableau les scores de tous les groupes, ainsi que leurs principaux arguments pour justifier ces scores (voir un exemple en Tableau 1). La solution avec le score total le plus élevé "gagne" ce premier tour de scrutin.

L'animateur déclenche alors un débat entre partisans et adversaires des différentes solutions. Par exemple, les membres d'un groupe ayant rejeté une certaine solution sont invités à essayer de

convaincre les membres d'un autre groupe qui eux la soutiennent, et vice versa. S'ils réussissent, les scores au tableau peuvent ensuite être mis à jour, et la solution choisie peut même changer à la suite de ce débat (la section 6 donne un exemple d'un tel débat). Comme recommandé par (Dieleman et Huisingh, 2006), le débat est encadré par un modérateur, qui assure un débat juste, serein et équilibré entre les élèves. Jouer un rôle qui ne correspond pas nécessairement à ses propres opinions aide également les élèves à prendre de la distance avec l'argumentation et à s'exprimer plus librement.

<b>Solution</b>	<b>S1 - Eye'Wana</b>	<b>S2 - Wana'Like</b>	<b>S3 - Wana'Pass</b>
Alter	0 (-) vie privée	3 (+) réductions (+) vie normale (+) contrôle données	2 (-) injuste (-) santé mentale
Centraux	4 (+) commerces ouverts	0 (-) perte clients	1 (-) perte clients
Lointains	1	3 (-) vie privée (+) réductions	1 (-) injuste
Futuristes	2 (+) équitable, emploi	2 (+) liberté (+) réductions	1
Seniors	3 (+) santé (+) vie continue	2 (-) pas de smartphone	0 (-) autonomie (-) vie privée
TOTAL	10	10	5

**Tableau 1** : Transcription d'un tableau de jeu.

## 5.2. Échelles d'évaluation

Afin d'approfondir la comparaison entre les différentes solutions technologiques, le modérateur pose ensuite quelques questions précises. Il guide les participants pour évaluer les solutions d'IA sur les échelles suivantes, importantes à prendre en compte (inspirées de Castelluccia et Métayer, 2020) :

- Leur efficacité (ici contre l'épidémie) : l'équilibre bénéfice-risque est un indicateur clé, employé par les agences de protection de la vie privée avant d'autoriser une technologie utilisant des données personnelles, et par les institutions de santé avant d'autoriser un médicament ;
- Leur impact sur les libertés individuelles : à quel point sont-elles contraignantes ou restrictives ? Combien de liberté est sacrifiée ? Est-ce que cela vaut le coup ? Ont-elles des effets secondaires négatifs, par exemple sur la santé mentale ?
- Leur impact sur la vie privée : quelles sortes de données personnelles collectent-elles ? Ces données sont-elles sensibles ? Qui y a accès ? Que se passe-t-il en cas de fuite ?
- Leur coût économique et écologique : combien cela coûte ? Qui va payer pour le déploiement de ces technologies ? Quelle pollution est générée pour les créer, maintenir, utiliser, recycler ?
- Leur accessibilité pour tous, et les risques de discrimination. Assurent-elles une stricte

égalité (tout le monde est traité de la même façon) ? Sont-elles équitables (les règles sont adaptées à chacun) ?

- Les risques d'erreurs et de mauvaise utilisation : quel risque d'échec de ces technologies ? Quel impact en cas d'erreur ? Quel risque de mauvaise utilisation ? Peuvent-elles être détournées pour servir un autre but que celui initialement prévu ? Si oui, avec quelles conséquences ?

### 5.3. Contextualisation

Pour faire le lien avec la réalité, les modérateurs illustrent enfin comment les différentes solutions virtuelles évoquées pendant le jeu renvoient en fait à des solutions bien réelles déjà mises en place dans certains pays (cf Section 4). Ces exemples illustrent très concrètement les dangers potentiels d'une dérive de l'usage utile à l'usage forcé de l'IA, et aident les participants à comprendre le message qu'ils sont tous concernés par l'IA et son utilisation.

## 6. Retour d'expérience

Ce jeu a été joué pour la première fois avec 6 groupes de 3 classes de lycée lors de la Fête de la Science (pour un total de 103 élèves) en Octobre 2021, comme un premier test grandeur nature. Il a ensuite été testé auprès d'environ 80 étudiants d'une école d'ingénieurs. Pour ces sessions nous avons imprimé le matériel de jeu mentionné plus haut (paragraphe 3.6), et assuré nous-mêmes l'animation et la modération. Dans les paragraphes suivants, nous présentons quelques retours d'expérience issus de ces premières sessions test.

### 6.1. Mise en place

Pendant la Fête de la Science, chaque session a été jouée avec un groupe d'environ 15-17 lycéens (une demi-classe), répartis dans les 5 groupes de rôles, plus un de leurs professeurs présents soit en tant qu'observateur, soit en tant que participant. La séance d'école d'ingénieur s'est déroulée avec les 80 élèves réunis, répartis en 10 groupes (2 par rôle). Chaque séance durait 1h30 à 2h et était encadrée par 2 modérateurs qui animaient les échanges en restant neutres, c'est-à-dire sans les orienter ni faire de suggestions.

### 6.2. Exemple de débat

Nous détaillons ici le débat qui s'est tenu avec l'un des groupes. Nous trouvons cet exemple particulièrement intéressant car il a mené à plusieurs changements de la solution sélectionnée par le conseil municipal.

Après le premier tour de débat (Tableau 2) Wana'Pass gagne avec 10 points, suivie de près par Eye'Wana avec 8 points, puis Wana'Like avec 7 points. Mais nous avons ensuite animé un débat entre les participants des groupes en désaccord sur les différentes solutions, débat qui a conduit à plusieurs revirements. Le Tableau 3 illustre l'évolution des scores totaux de chaque solution au fil des quatre tours du débat, avec le score de la solution gagnante en bleu : on voit ainsi qu'à l'issue du deuxième et troisième tour, c'est Eye'Wana qui l'emporte. Nous détaillons le contenu du débat ci-dessous.

Le groupe des Seniors étant totalement opposé à Wana'Pass (qui ne leur accorderait probablement que très peu de points pour sortir, leur âge étant considéré comme un facteur de risque), il leur a été demandé d'essayer de convaincre d'autres groupes de ne pas choisir cette solution. Ils ont fait valoir que Wana'Pass était inéquitable, et ont convaincu les Centraux (commerçants) de déplacer le seul point qu'ils donnaient à Wana'Pass vers Eye'Wana (pour préserver leur clientèle âgée), et les Futuristes de déplacer 1 point de Wana'Pass vers

Wana'Like. A l'issue de ce deuxième tour, Eye'Wana prend donc la tête avec 9 points (S4-L1-C4-A0-F0), en légère avance sur les 2 autres solutions marquant 8 points chacune (Wana'Like S1-L0-C1-A2-F4, Wana'Pass S0-L4-C0-A3-F1).

	<b>Eye'Wana</b>	<b>Wana'like</b>	<b>Wana'Pass</b>
Seniors	4 (+) pas besoin smartphone	1 (-) besoin smartphone	0 (-) aucun point si âgés
Lointains	0 (-) pas de caméras hors centre-ville (+) force respect des règles	0 (-) risque manipulation	4 (+) simplicité
Centraux	3 (+) bon pour les affaires (+) égalité	1 (-) inégalité entre magasins (+) liberté	1 (-) perte de clients
Alter	0 (-) perte vie privée (-) pollution	2 (-) besoin garanties sur données	3 (-) perte vie privée (+) efficacité
Futuristes	0 (-) perte de liberté	2 (+) coupons réduction, gratuit	2 (+) équitable
<b>Total</b>	<b>8</b>	<b>7</b>	<b>10</b>

**Tableau 2** : tableau après un tour de discussion, scores et arguments pour / contre chaque solution.

Tour	Eye'Wana	Wana'like	Wana'Pass
1	8	7	<u>10</u>
2	9	8	8
3	11	5	9
4	11	1	<u>13</u>

**Tableau 3** : évolution des scores sur 4 tours de débat

Les Lointains ont ensuite été invités à apporter leurs arguments contre Wana'Like, solution à laquelle ils n'attribuaient aucun point. Ils ont fait valoir que cette application pouvait manipuler ses utilisateurs. En conséquence, les Seniors et les Centraux ont retiré 1 point à Wana'Like (leur seul point sur cette solution) et l'ont transféré à Eye'Wana (qui concentre maintenant leurs 5 points), tandis que les Futuristes ont transféré 1 point de Wana'Like à Wana'Pass (inversant ainsi leur changement précédent). A l'issue de ce troisième tour, Wana'Like n'a alors plus que 5 points (S0-L0-C0-A2-F3), Wana'Pass remonte à 9 points (S0-L4-C0-A3-F2), et Eye'Wana prend la tête avec 11 points (S5-L1-C5-A0-F0).

Enfin, les Futuristes et les Alter, farouchement opposés à Eye'Wana (pour ses atteintes aux libertés, et parce que leurs revenus plus faibles les rendaient plus sensibles aux amendes, jugées inéquitables) ont réalisé qu'ils pouvaient changer le résultat : les Futuristes ont repris les 3 points attribués à leur solution favorite mais forcément perdante (Wana'Like) pour les reporter vers leur deuxième option Wana'Pass, afin d'empêcher Eye'Wana de gagner. Les Alter ont aussi reporté 1 point de Wana'Like (sur les 2 qu'ils lui attribuaient) vers Wana'Pass. On obtient donc un total inchangé de 11 points pour Eye'Wana (principalement grâce aux Seniors et Centraux, S5-L1-C5-A0-F0), mais Wana'Like n'a plus qu'un unique point (des Alter : S0-L0-C0-A1-F0), tandis que Wana'Pass rassemble 13 points (4 points des Lointains, 4 des Alter, 5 des Futuristes, et aucun de la part des Seniors et Centraux qui ont plébiscité Eye'Wana : S0-L4-C0-A4-F5).

Il est intéressant de noter que parmi les différentes parties test que nous avons animées (voir les statistiques en section 6.3), c'est une rare occurrence de défaite totale pour Wana'Like (1 point

sur un total possible de 25), et la seule victoire de Wana'Pass.

**Remarque.** On voit bien qu'un tel débat pourrait continuer encore, ce qui n'a pas été le cas faute de temps. Le but du jeu n'est effectivement pas de dégager une solution 'optimale' (il n'y en a précisément aucune) mais de faire ressentir aux participants le fragile équilibre bénéfice-risque des technologies, et comprendre les points de vue différents de personnes avec un profil différent. Nous voulons leur montrer que tout point de vue peut être valable, mais doit être justifié par des arguments pour l'expliquer et convaincre les autres.

### 6.3. Résultats quantitatifs

Nous considérons ici les 6 sessions animées pendant la Fête de la Science en Octobre 2021, avec un total de 103 lycéens. Ces sessions sont en effet plus représentatives du public cible de notre jeu débat. Les résultats quantitatifs (résumés au Tableau 4) sont les suivants :

	Wana'Like	Wana'Pass	Eye'Wana
#Sélections	4	1	2

**Tableau 4 :** Nombre de sélections par solution sur 6 sessions (dont un match nul).

- Wana'Like a été sélectionnée 3 fois, en particulier pour les avantages financiers (coupons de réduction) et pour son aspect moins contraignant par rapport aux autres solutions.
- Lors d'une autre session, Eye'Wana et Wana'Like l'emportent à égalité. Wana'Pass a été rejeté pour sa discrimination envers les plus âgés.
- Eye'Wana est sélectionné seul lors d'une autre session, où il a été considéré plus simple à déployer (pas besoin de smartphone) et parce qu'il force l'égalité entre tous devant les règles. Dans toutes les autres sessions il a été jugé trop intrusif et farouchement rejeté.
- Wana'Pass n'a été sélectionné qu'une seule fois, après le débat résumé ci-dessus, mais plus pour faire barrage à Eye'Wana que par réel choix.

### 6.4. Retours qualitatifs

Nous avons reçu un retour globalement positif de la part de tous les étudiants et de leurs professeurs. Dans cette section, nous discutons de quelques commentaires spécifiques faits par les joueurs après les différentes sessions. Cela souligne l'importance de la phase de débriefing qui permet de tels échanges avec les participants, pendant lesquels ils prennent une certaine distance par rapport aux événements débattus lors du jeu.

**Lien avec le COVID.** Malgré nos efforts pour créer une ville et une épidémie "virtuelles", la plupart des participants ont immédiatement fait référence au "COVID" dans leurs discussions. L'un d'eux a même mentionné être opposé au vaccin COVID, ce qui a prouvé l'importance de garder une certaine distance par rapport au contexte actuel afin d'assurer un débat apaisé et ciblé. Cependant, de telles prises de position n'ont eu lieu que lors des débriefings et n'ont pas influencé les débats, puisque les joueurs devaient alors s'en tenir à leurs rôles. De plus, nous pensons que ce lien s'atténuera avec le temps, à mesure que l'épidémie actuelle s'estompera, tandis que notre jeu restera d'actualité.

**Hypothèse d'infailibilité.** Les étudiants supposaient généralement que les technologies proposées étaient infailibles et, par conséquent, ils ne tenaient pas compte des effets néfastes d'une potentielle erreur. Cela confirme la nécessité d'enseigner les bases de l'IA ainsi que son risque de biais et d'erreurs. Il est nécessaire de savoir comment les algorithmes fonctionnent pour comprendre comment ils pourraient également échouer, et il est essentiel de comprendre l'impact de telles erreurs (discriminations, etc).

**Sécurité vs liberté.** Les étudiants se sont beaucoup plus focalisés sur la protection de leur liberté que sur la sécurité sanitaire et donc la lutte efficace contre le virus. Même si les débats ont évoqué l'efficacité des solutions proposées, celle-ci est passée après des arguments de perte de vie privée ou de liberté de mouvement, et en fait après tout autre argument. Cela n'a rien d'étonnant dans une tranche d'âge qui souvent ne se sent pas directement concernée par le risque sanitaire. Cela montre également un manque d'insistance sur la nécessité d'efficacité des solutions dans nos instructions de jeu. Nous avons depuis amélioré ce point en demandant explicitement aux participants de comparer les solutions sur les différentes échelles listées ci-dessus, y compris leur efficacité. Il s'agit ainsi de discuter s'ils préfèrent accepter une solution très contraignante (et efficace) pour une courte durée, ou une solution moins contraignante (mais moins efficace) qu'il faudra probablement maintenir plus longtemps.

**Surveillance stricte vs diffuse.** Bien que voulant protéger leur liberté et leur vie privée, la plupart des étudiants étaient soit inconscients, soit fatalistes, vis-à-vis de la surveillance déjà exercée sur eux par les algorithmes des réseaux sociaux. La plupart des étudiants ont considéré que les applications les plus intrusives étaient Wana'Pass du fait de l'utilisation de données médicales, ou Eye'Wana car il filme les gens en permanence. Ils sous-estiment complètement les risques liés aux algorithmes de recommandation, puisqu'ils divulguent déjà volontiers de grandes quantités d'informations personnelles sur les réseaux sociaux. Cela montre que l'éducation à l'IA devrait également insister sur la valeur des données personnelles et sur ce qu'il est possible d'en faire. De nombreuses références sont donc fournies dans le guide de l'animateur pour insister sur ce point.

**Public cible.** Ayant animé des sessions à la fois avec des lycéens (15 à 18 ans) et des élèves ingénieurs (22 à 24 ans), nous avons eu des retours assez différents. Les commentaires des participants ont tendance à suggérer que le jeu est mieux ciblé sur les élèves du secondaire, qui ont moins de connaissances techniques sur l'intelligence artificielle et sont heureux d'avoir l'occasion de discuter de la pandémie actuelle et de son impact sur leur vie. Au contraire, certains étudiants ingénieurs ont été déçus de l'accent mis sur les implications éthiques de l'IA alors qu'ils s'attendaient à acquérir plus de connaissances techniques. Étonnamment, certains ont signalé comme point négatif que les solutions semblaient trop irréalistes car "aucune n'était bonne" (alors même que nous leur avons montré des exemples réels de solutions similaires déjà en œuvre dans d'autres pays). Incidemment, c'était précisément le message à retenir : aucune technologie d'IA n'est une solution parfaite, et tout le monde devrait être conscient des pièges avant de les accepter. Il semble donc important de cibler aussi ce public, même s'il est a priori moins ouvert, peut-être en changeant la forme d'intervention.

**Contenu technique et public cible.** Certains participants (élèves ingénieurs) ont regretté l'absence de connaissances techniques sur l'IA dans ce jeu. En fait, le problème vient plutôt du décalage entre leur niveau de connaissances préalable et les connaissances apportées par le jeu. En effet, même si ce jeu-débat semble mettre l'accent sur les aspects sociétaux de l'informatique et de l'IA, il aborde aussi des points techniques. Tout d'abord, la phase d'introduction et le débriefing sont l'occasion de donner des définitions de différents concepts, souvent manipulés mais mal compris, comme l'IA ou les algorithmes. Ensuite, les solutions d'IA proposées sont toutes réalistes et basées sur des briques technologiques existantes, nous permettant d'introduire leur fonctionnement (reconnaissance faciale, fouille de données médicales, algorithmes de recommandation, etc). Enfin, la compréhension de ce fonctionnement permet de prendre conscience des risques associés : des erreurs et biais sont possibles, et les données collectées peuvent être mal utilisées.

**Extension de l'activité.** Le guide de l'animateur fourni (cf Paragraphe 3.7) contient de

nombreuses ressources permettant au professeur d'approfondir les notions souhaitées avec ses élèves en classe ultérieurement. Nous espérons aussi avoir aiguisé leur curiosité et les pousser à s'informer plus sur le sujet. Enfin, pour améliorer encore la transmission de connaissances, nous souhaitons intégrer ce jeu dans une séquence pédagogique à plus long-terme, l'associant avec un cycle de conférences sur l'IA, des groupes de discussion, des ateliers d'élaboration de cahier des charges pour les solutions, etc,

## **7. Discussion des biais et limites de l'activité**

Le but de cette série de jeux débats est de sensibiliser et ouvrir un débat sur les enjeux sociétaux des technologies d'IA utilisées dans notre vie de tous les jours. L'activité s'adresse à des citoyens sans aucune connaissance pré-requise sur l'informatique ou l'intelligence artificielle. Il ne s'agit donc pas d'éduquer les joueurs aux principes exacts de fonctionnement de ces technologies, même si ces deux aspects techniques et sociaux sont profondément intriqués. Un certain nombre de biais découlent de notre conception de l'activité, et nous les discutons ci-dessous.

### **7.1. Représentativité des rôles**

Ce jeu repose sur un débat entre 5 rôles représentés au conseil municipal (seniors, centraux, etc). Ces catégories, imposées par le cadre d'origine de la série de jeux, peuvent sembler stéréotypées. Cependant, le fait que les élèves incarnent des stéréotypes n'est pas gênant, l'important étant de leur faire percevoir les différences de points de vue et d'impact des solutions d'IA selon la perspective. Ainsi par exemple les élèves jouant les Alter (définis comme soucieux de la nature et du bien-être de tous) vont spontanément s'inquiéter de l'impact écologique des batteries des drones de vidéo-surveillance, alors que ceux jouant les Seniors (définis comme des retraités actifs voulant préserver leur autonomie) vont surjouer une moindre familiarité avec les outils numériques, mais aussi réaliser que leur santé plus fragile les expose à être isolés plus que les autres groupes. Les élèves se basent souvent sur leurs grands-parents ou sur des proches pour adopter le profil qu'ils sont supposés incarner dans le jeu ; c'est donc finalement un très bon exercice de prise de recul et de compréhension du point de vue d'autrui. Par ailleurs, les rôles représentés n'ont pas non plus besoin d'être exhaustifs, le but n'étant pas forcément de représenter tous les points de vue mais de faire prendre conscience aux élèves qu'il en existe d'autres différents du leur. Enfin, il est important, pour préserver un débat neutre et éviter les dérapages, de ne pas choisir de catégories trop politisées ou sensibles (comme des partis politiques). Les 5 catégories stéréotypées choisies sont donc un bon compromis pour permettre une compréhension des différences de points de vue tout en préservant un cadre neutre et apaisé.

### **7.2. Techno-solutionnisme**

Le cadre de la série de jeux-débats impose aussi le principe du vote du conseil municipal servant à choisir la solution "gagnante" parmi les 3 propositions de Sowana, sans possibilité de tout refuser. On pourrait accuser ce scénario de techno-solutionnisme puisqu'il force les joueurs à choisir une solution technologique au problème posé. Cependant le jeu exige aussi des joueurs de réfléchir aux points positifs et négatifs de chaque solution avant de prendre une décision, sans jamais les accepter a priori, et d'argumenter leur choix. Ces compétences leur seront très utiles dans leur vie de citoyens. En effet ce scénario est malheureusement assez réaliste : la plupart des élèves seront (et sont déjà) amenés dans leur vie à subir des technologies, et doivent donc d'abord apprendre comment faire la balance bénéfico-risque de technologies déjà existantes, voire éventuellement quels garde-fous ils peuvent exiger (de leurs dirigeants). La question de la construction de ces solutions ne se posera qu'à celles et ceux qui poursuivront des études dans ce domaine, alors que cette activité se veut destinée à tout public.

### **7.3. Concurrence entre les rôles**

On pourrait aussi regretter que ce format de débat force les différents groupes à se mettre dans une forme de concurrence avec les autres groupes, pour influencer le vote vers celle des solutions qui satisfait le mieux leurs propres intérêts. Cependant on observe lors des débats que chaque groupe prend conscience des différents points de vue, notamment grâce à l'obligation qui leur est faite de toujours lister à la fois le positif et le négatif de chaque solution, et de fournir des arguments en faveur de leur point de vue. Cette activité est donc aussi une éducation au débat et à l'argumentation, et en ce sens la mise en concurrence des rôles est utile. Il n'est effectivement pas possible pour les groupes de choisir un consensus incluant une modification de la solution pour convenir à tous les groupes, ou encore une adaptation de son cadre légal. Nous pensons que cela déborderait du cadre de l'activité proposée, tant du point de vue des connaissances préalables requises (fonctionnement des technologies, réglementation en vigueur) que du temps nécessaire (l'activité étant calibrée sur un créneau de cours standard, entre 1h30 et 2h). Cela pourra néanmoins faire l'objet d'ateliers complémentaires (cf paragraphe 7.5).

### **7.4. Cadre législatif flou**

On pourra noter aussi que toutes les technologies proposées n'ont de sens que dans un certain contexte législatif. La même technologie de vidéo-surveillance peut être utilisée pour détecter les personnes rompant le confinement, des agressions dans les transports en commun, un franchissement de feu rouge, ou encore des "mouvements suspects" lors de grands événements sportifs. Les conséquences ne sont bien sûr pas les mêmes sur les libertés individuelles. Or dans le cadre du jeu, le contexte légal est volontairement assez flou, afin de ne pas limiter l'activité à un contexte particulier. Au contraire, les mêmes questions doivent se poser quel que soit le contexte. En effet, une technologie mise en place pour un objectif particulier et dans un contexte légal donné, peut ensuite plus facilement être pérennisée et appliquée à un autre cadre législatif. Il est donc important pour les citoyens de se poser des questions, d'avoir conscience des possibles détournements d'une technologie, et de fixer des garde-fous, comme une durée maximale d'utilisation par exemple (jusqu'à la fin de la pandémie, ou pendant 6 mois). Même sans expliciter le contexte législatif ou la possibilité de le modifier, les élèves se posent souvent ce genre de questions (si besoin, guidés en cela par l'animateur) lors des discussions de groupe. En particulier ils s'intéressent beaucoup aux possibles détournements ou fuites de leurs données, ou encore aux éventuelles erreurs de fonctionnement des algorithmes. Lors du débriefing, les animateurs fournissent donc plus d'informations sur les protections légales qui existent pour les utilisateurs de technologies d'IA. Les citoyens doivent notamment être conscients de leur droit (garanti par le RGPD) à refuser des solutions d'IA prenant automatiquement des décisions les concernant, droit qu'ils ignorent souvent. Ces aspects législatifs sont donc abordés lors du débriefing, et pourraient aussi faire l'objet d'activités complémentaires (cf paragraphe 7.5).

### **7.5. Intégration dans une séquence pédagogique**

Pour répondre à plusieurs des biais ci-dessus, on peut envisager d'intégrer ce jeu-débat dans une séquence pédagogique plus complète. En effet, la durée de l'activité elle-même est contrainte par la durée habituelle d'un créneau de cours, soit 1h30 à 2 heures, ce qui empêche d'aborder certains aspects plus en détail. Mais cette activité peut être prolongée ou précédée de sessions complémentaires. Nous pensons en particulier à démarrer la séquence pédagogique par une conférence sur l'IA, qui aborderait ses grands principes, ses biais possibles, voire ses aspects réglementaires. Après le jeu-débat, d'autres ateliers pourraient être organisés comme une discussion collaborative (avec les mêmes rôles) sur l'élaboration des garde-fous nécessaires avant d'adopter une nouvelle technologie, ou encore la conception de solutions non technologiques au

problème.

## 8. Conclusion

Dans cet article, nous avons présenté un jeu sérieux sous forme de débat municipal visant à choisir des solutions d'IA pour lutter contre une épidémie, et les retours d'expérience des premières sessions de test. Les participants sont répartis en 5 groupes de rôles (personnes âgées, commerçants, écologistes...) : changer ainsi de perspective, devoir comprendre d'autres points de vue et devoir argumenter pour défendre ses idées est un bon exercice pour eux. Les élèves apprennent à débattre, à argumenter leurs positions et à écouter et prendre en compte celles (même très divergentes) des autres. Mais ils ont aussi appris sur l'Intelligence Artificielle et ses dangers, prenant conscience de sa faillibilité et de la valeur de leurs données personnelles, ainsi que de la nécessité de peser à la fois les bénéfices et les risques avant d'adopter une nouvelle technologie.

On peut regretter certains biais, souvent inhérents au cadre imposé par la série des jeux débats dans laquelle nous nous inscrivons, ainsi que par la durée limitée de l'activité. Mais cela ne nous semble pas avoir gêné l'objectif de sensibilisation aux aspects sociétaux de l'IA tel que défendu par cette activité. Ce jeu a bien atteint son objectif de sensibilisation de la population à ces enjeux sociétaux, et la phase de débriefing s'est confirmée comme cruciale dans ce processus. Par ailleurs, les biais de cette activité peuvent être dépassés en l'intégrant dans une séquence pédagogique plus complète permettant d'élaborer plus profondément certains aspects. Nous envisageons effectivement d'étendre son utilisation en créant un cycle pédagogique plus long associant conférences et ateliers participatifs.

Par ailleurs, le débriefing des débats avec les participants nous a également fourni des informations intéressantes sur leur façon d'appréhender l'IA, ses avantages, ses prétendues qualités d'objectivité et d'infailibilité. Cela a conforté notre idée que nous devons éduquer la prochaine génération sur ces technologies. Bien que l'IA puisse certainement aider à lutter contre une crise telle que la pandémie de COVID-19, elle n'est pas infailible et, comme elle devient de plus en plus omniprésente dans nos vies, l'impact potentiel des erreurs devient également plus grave. Il est donc nécessaire que chacun connaisse les bases des algorithmes d'IA (UNESCO, 2021) et puisse faire des choix éclairés quant à leur utilisation.

**Remerciements** – Nous remercions l'Arbre des Connaissances et plus particulièrement Clara Fruchon et Camille Volovitch pour nous avoir aidés à concevoir, tester et diffuser ce jeu.

## Références bibliographiques

Adam, C. et Arduin, H. (2022). Finding and explaining optimal screening strategies with limited tests during the covid-19 epidemics. In *19th International Conference on Information Systems for Crisis Response and Management ISCRAM*.

Adam, C., Jacquier, A., et Taillandier, F. (2022). Un jeu sérieux pour sensibiliser aux enjeux d'une mobilité urbaine durable. *Academic Journal of Civil Engineering (AJCE)*.

Aral, S. et Eckles, D. (2019). Protecting elections from social media manipulation. *Science*, 365(6456):858–861.

Adam, C., et Lauradoux, C. (2022). A Serious Game for Debating about the Use of Artificial Intelligence during the COVID-19 pandemic. *19th International Conference on Information Systems for Crisis Response and Management (ISCRAM), Tarbes, France, pp.*

554–563.

Avanzi, M., Coniglio, R., Cisotto, G., Giordani, M., et Ferro, N. (2021). Nobis: A crowd monitoring service against covid-19. In *IRCDL*, pages 126–137.

Bacchini, F. et Lorusso, L. (2019). Race, again: how face recognition technology reinforces racial discrimination. *Journal of information, communication and ethics in society*.

Castelluccia, C. et Métayer, D. L. (2020). Position Paper: Analyzing the Impacts of Facial Recognition. In *Privacy Technologies and Policy - 8th Annual Privacy Forum, APF 2020, Lecture Notes in Computer Science*, Springer.

Chun, A. (2020). In a time of coronavirus, China’s investment in AI is paying off in a big way. *South China Morning Post*. Last access 20 January 2022.

Cottineau, C. et al. (2020). Understanding the current covid-19 epidemic: one question, one model. *RofASSS (Review of Artificial Societies and Social Simulation)*.

Crookall, D. (2010). Serious games, debriefing, and simulation/gaming as a discipline. *Simulation and gaming*, 41(6):898–920.

Dieleman, H. et Huisingh, D. (2006). Games by which to learn and teach about sustainable development: exploring the relevance of games et experiential learning for sustainability. *Journal of Cleaner Production*, 14(9-11):837–847.

Faucon, L., El-Mhamdi, E.-M., et al. (2021). Recommendation algorithms, a neglected opportunity for public health. *Revue Médecine et Philosophie*, 4(2).

Gostin, L. O., Cohen, I. G., et Shaw, J. (2021). Digital health passes in the age of covid-19: Are “vaccine passports” lawful and ethical? *JAMA*, 325(19):1933–1934.

Graafland, M., Schraagen, J. M., et Schijven, M. P. (2012). Systematic review of serious games for medical education and surgical skills training. *Journal of British Surgery*, 99(10):1322– 1330.

Jackson, V. A. et Back, A. L. (2011). Teaching communication skills using role-play: an experience-based guide for educators. *Journal of palliative medicine*, 14(6):775–780.

Jarvis, L., Odell, K., et Troiano, M. (2002). Role-playing as a teaching strategy. *Strategies for application and presentation, staff development and presentation*.

Jernigan, C. et Mistree, B. F. (2009). Gaydar: Facebook friendships expose sexual orientation. *First Monday*.

Kofler, N. et Baylis, F. (2020). Ten reasons why immunity passports are a bad idea. *Nature Publishing Group*, <https://www.nature.com/articles/d41586-020-01451-0%C2%A0>

Liu, C. (2019). Multiple social credit systems in china. *Economic Sociology: The European Electronic Newsletter*, 21(1):22–32.

Lu, F. et Sun, Y. (2022). Covid-19 vaccine hesitancy: The effects of combining direct and indirect online opinion cues on psychological reactance to health campaigns. *Computers in human behavior*, 127:107057.

Martinez-Martin, N. (2019). What are important ethical implications of using facial recognition technology in health care? *AMA journal of ethics*, 21(2).

Najibi, A. (2020). Racial discrimination in face recognition technology. *Harvard Online: Science Policy and Social Justice*, 24.

- Naudé, W. (2020). Artificial intelligence vs COVID-19: limitations, constraints and pitfalls. *AI and Society*.
- Nieves-Cuervo, G. M., Manrique-Hernández, E. F., Robledo-Colonia, A. F., et Grillo, A. E. K. (2021). Infodemic: fake news and covid-19 mortality trends in six latin american countries. *Pan American Journal of Public Health*.
- Petropoulos, G. (2020). Artificial intelligence in the fight against COVID-19. *Last accessed 20 January 2022*.
- Rebolledo-Mendez, G., Avramides, K., De Freitas, S., et Memarzia, K. (2009). Societal impact of a serious game on raising public awareness: the case of floodsim. In *Proceedings of the 2009 ACM SIGGRAPH symposium on video games*.
- Resnick, M. et Wilensky, U. (1998). Diving into complexity: Developing probabilistic decentralized thinking through role-playing activities. *The Journal of the Learning Sciences*, 7(2):153–172.
- Shen, C. F. (2019). Social credit system in China. Tech. report, City Univ. of Hong Kong.
- Sheth, D. et Giger, M. L. (2020). Artificial intelligence in the interpretation of breast cancer on mri. *Journal of Magnetic Resonance Imaging*, 51(5):1310–1324.
- Simonneaux, L. (2001). Role-play or debate to promote students' argumentation and justification on an issue in animal transgenesis. *International Journal of Science Education*, 23(9):903–927.
- Taillandier, F. et Adam, C. (2018). Games ready to use: A serious game for teaching natural risk management. *Simulation and Gaming*, 49(4):441–470.
- Tanner, R. et Flood, C. M. (2021). Vaccine passports done equitably. *JAMA Health Forum*, 2(4).
- UNESCO (2021). Rapport de la commission sciences sociales et humaines (shs). Technical Report 0000379920fre, UNESCO.
- Wang, Y. et Kosinski, M. (2018). Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of personality and social psychology*, 114(2):246.
- Wells, G., Horwitz, J., et Seetharaman, D. (2021). Facebook knows Instagram is toxic for teen girls, company documents show. *The Wall Street Journal*.
- Whalen, K. A., Berlin, C., Ekberg, J., Barletta, I., et Hammersberg, P. (2018). 'All they do is win': Lessons learned from use of a serious game for circular economy education. *Resources, Conservation and Recycling*, 135:335–345.
- Wu, J. S. et Lee, J. J. (2015). Climate change games as tools for education and engagement. *Nature Climate Change*, 5(5):413–418.