



Aleksej Heinze - Professeur associé en marketing numérique. Il a obtenu son doctorat en 2008 à la Salford Business School (UK).

Adina Dudau - Professeur de gestion publique à l'Adam Smith Business School, Université de Glasgow, et co-rédacteur en chef du European Management Journal.

Joras Ferwerda - Titulaire d'un doctorat en économie, d'une maîtrise en économie et en sciences sociales et d'une licence en économie et en droit de l'école d'économie de l'université d'Utrecht aux Pays-Bas. Il est actuellement professeur associé à l'école d'économie de l'université d'Utrecht aux Pays-Bas.

Corruption et IA : Comprendre les intersec- tions et le nouveau défi pour l'humanité

Un problème imminent pour l'humanité est l'utilisation non éthique de l'Intelligence Artificielle (IA), qui, aux côtés des techniques d'apprentissage automatique, peut favoriser la prolifération de la corruption. Cependant, il existe des avantages notables à l'IA, notamment son utilisation prometteuse dans l'identification et donc la prévention de la corruption.

La corruption est une pratique non éthique et illégale caractérisée par l'abus du pouvoir confié pour un gain personnel ou un avantage injuste. Elle englobe un large éventail de comportements, comme la corruption, la fraude, le détournement de fonds, le népotisme, le copinage et le trafic d'influence. Qu'il s'agisse de l'attribution de contrats de santé publique pendant la pandémie de Covid-19, comme pratiqué par le gouvernement britannique, la fraude à grande échelle en matière d'approvisionnement par le ministère de la Défense ukrainien, ou des cas de fraude comptable individuelle, les conséquences de la corruption nuisent généralement au grand public.

Avec l'augmentation de la transformation numérique des paiements et des enregistrements transactionnels, les opportunités pour les outils et applications d'IA élèvent le jeu de la corruption à un niveau supérieur. Ceux qui cherchent à dissimuler les traces de leurs activités peuvent utiliser des outils d'IA pour fabriquer de fausses informations et des documents. Le public est maintenant familier avec des outils de base tels que ChatGPT par OpenAI, qui ont révolutionné la création de contenu. Cependant, entre de mauvaises mains, le contenu généré par IA ou assisté par IA peut propager de fausses informations et de la désinformation délibérée, favorisant ainsi la corruption parmi d'autres activités préjudiciables.

La qualité de l'intégrité et de l'exactitude des données est un enjeu crucial. Les contenus Deepfake, souvent sous forme de vidéo ou d'images, représentent des individus de manière créative avec peu d'ancrage dans la réalité. Les textes que nous lisons, les vidéos et les images que nous voyons, et les sons que nous entendons peuvent tous être manipulés par des outils d'intelligence artificielle générative. Bien que ces technologies puissent actuellement rencontrer des

difficultés en termes de précision élevée, un œil averti peut repérer des incohérences. Cependant, les mêmes outils et algorithmes apprennent constamment. Avec une formation continue, ils sont susceptibles de devenir encore plus puissants et plus difficiles à détecter à l'avenir. Des histoires de fausses publicités utilisant du contenu généré par IA ont commencé à apparaître.

Le volume d'informations généré par les technologies assistées par IA met la responsabilité sur les individus pour discerner si les informations qu'ils rencontrent sont générées par des humains ou par des IA. Alors que les humains ne sont pas nécessairement plus précis, la connaissance de la source des données devient cruciale. Les outils de détection de l'IA ne sont pas 100% précis à ce stade, avec de nombreux faux positifs. La vitesse et l'imagination des outils d'IA, combinées à des intentions non éthiques, peuvent avoir de nombreuses conséquences négatives pour la société, des e-mails de phishing authentiques en apparence aux réseaux de dissimulation à grande échelle orchestrés à travers du code automatisé.

L'IA peut toutefois jouer un rôle positif en identifiant les tendances et les incohérences dans l'information, aidant ainsi à prévenir et à identifier la corruption. Les transactions numériques étant en hausse et l'information de plus en plus disponible sur les portails numériques, tels que les registres fonciers et les divulgations de propriété des entreprises, les ensembles de données peuvent être évalués avec l'aide d'outils d'IA. Des améliorations similaires peuvent être observées dans la lutte contre le blanchiment d'argent, où les banques peuvent utiliser des systèmes de détection basés sur l'IA pour filtrer les transactions suspectes des transactions normales. Enfin, les organisations peuvent employer des outils de détection de fraude par IA pour toutes leurs activités, signalant immédiatement les activités suspectes et réduisant le temps nécessaire aux «forces négatives» pour couvrir leurs traces.

EST-CE QUE L'IA A AIDÉ DANS VOTRE ÉVALUATION QUOTIDIENNE DES INFORMATIONS ?

Au cours des 20 dernières années, toute une industrie de spécialistes du marketing digital s'est développée, se concentrant sur l'amélioration du contenu pour l'optimisation des moteurs de recherche (SEO). Comme le reconnaît le dernier livre de marketing digital et des médias sociaux de Heinze et ses collègues (Heinze et al., 2024), l'IA générative amène tous les créateurs de contenu et les évaluateurs à réfléchir à la méthode que Google appelle E-E-A-F : expertise, expérience, autorité et fiabilité.

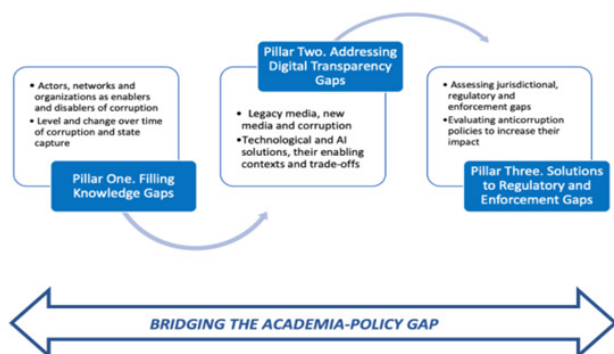
Ces critères d'évaluation E-E-A-F suggèrent la nécessité de plus de transparence pour les créateurs de contenu et l'importance pour les consommateurs d'information de prêter attention à ces éléments que Google et d'autres essaient désormais d'automatiser. Évaluer les producteurs de contenu sur ces quatre éléments aide les algorithmes à juger, mais pas toujours avec succès. Bien sûr, cela pourrait créer des biais algorithmiques futurs qui pourraient avoir des conséquences négatives s'ils ne sont pas supervisés par un humain. Nous avons vu cela dans le marquage automatique et les traductions dans le rapport sur le conflit israélo-palestinien où les messages pro-Palestine ont été censurés par Meta sur Facebook et Instagram à grande échelle, indiquant un algorithme IA injuste qui a été utilisé pour traduire incorrectement le contenu (Human Rights Watch, 2023).

Indépendamment de la manière dont les informations sont produites, les organisations et les individus peuvent appliquer les mêmes principes pour améliorer la transparence et réduire les effets négatifs de la corruption en tant que conséquence de l'IA.

BESOIN DE COMBLER LE FOSSÉ DE LA TRANSPARENCE NUMÉRIQUE PAR LE PROJET BRIDGEGAP

Si les progrès technologiques ont rendu l'information plus accessible, ils ont également créé des possibilités d'utilisation abusive dans différents secteurs et pays.

Dans notre module de travail Bridgégap, nous visons à étudier l'efficacité de divers outils technologiques - tels que l'intelligence artificielle, la blockchain et l'analyse de données cryptées - dans la détection et la prévention des pratiques corrompues. Ces outils ont un grand potentiel, mais leur succès dépend du contexte dans lequel ils sont utilisés. En étudiant des exemples concrets et les meilleures pratiques, le projet vise à identifier ce qui fonctionne et ce qui ne fonctionne pas. (Figure 1).



Bridgégap est un projet de recherche de quatre ans qui réunit 15 organisations travaillant ensemble pour combler le fossé entre la politique et le monde universitaire. Son objectif est de comprendre comment la technologie peut aider à lutter contre la corruption et à promouvoir la responsabilité, en particulier dans l'Union européenne (UE). Un des objectifs principaux est d'investiguer l'efficacité de divers outils technologiques, tels que l'intelligence artificielle, la blockchain et l'analyse de données chiffrées, dans la détection et la prévention des pratiques corrompues. Ces outils ont un grand potentiel, mais leur succès dépend du contexte dans lequel ils sont utilisés. En étudiant des exemples concrets et les meilleures pratiques, le projet vise à identifier ce qui fonctionne et ce qui ne fonctionne pas.

En conclusion, bien que l'IA présente des défis dans la lutte contre la corruption, elle offre également des opportunités d'améliorer la détection, la prévention et d'inciter à plus de transparence sur la manière dont l'IA a été utilisée pour toute création de contenu. Le déploiement responsable et l'examen continu sont cruciaux pour exploiter l'IA comme une force pour l'intégrité et la responsabilité dans la société. ■

Divulgaration :

Les auteurs de cet article travaillent sur un projet de recherche BRIDGEGAP Accord de subvention Numéro de convention : 101132483 [<https://cordis.europa.eu/project/id/101132483>] financé par la Commission européenne. Cette publication [communication] reflète uniquement les opinions de l'auteur, et la Commission ne peut être tenue responsable de toute utilisation qui pourrait être faite des informations contenues dans celle-ci.

