

HAARLEM DECLARATION

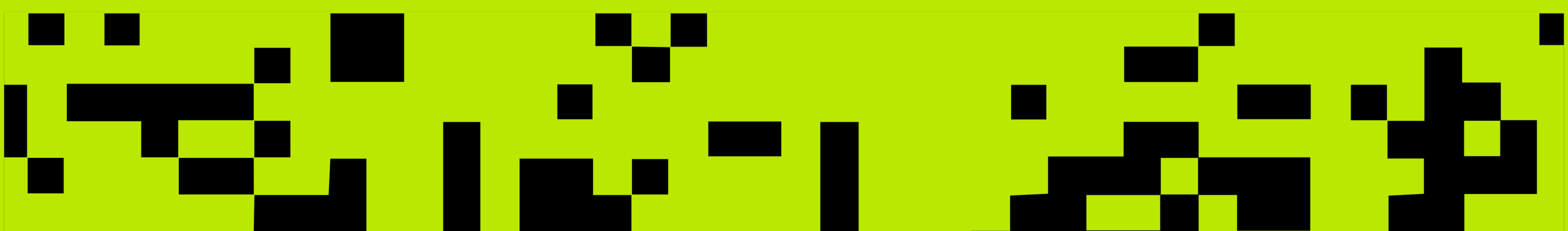
CHAMPIONING AI FOR ADVANCING INCLUSIVE, SAFE, AND RELIABLE DIGITAL MEDIA SPACES!

As a global community of digital media makers – including independent media organizations, journalists, digital content creators and civil society groups using digital media for social impact – we recognize that AI technologies, and generative AI in particular, is actively and rapidly influencing, shaping and transforming the digital media ecosystem. Its influence extends across all aspects of content creation, curation, recommendation, moderation, distribution, promotion, and audience analytics.

While AI technology offers the potential to make routine and labour-intensive digital media processes and functions more efficient, streamlined and cost-effective, its deployment also introduces significant risks. Already visible is the disruption of the digital media ecosystem through the proliferation of misinformation, an erosion of human oversight in editorial decision-making, and the displacement of media workers. As a result, there is an urgent need for a value-driven, responsible and ethics-based approach to integrating AI technologies across digital media work and organisations.

In lieu of the adoption of the Global Digital Compact (GDC) by the United Nations Member States in September 2024, at the Summit of the Future – which outlines as part of its broader key objectives to, "fostering an inclusive, open, safe, and secure digital space that respects, protects and promotes human rights; advancing responsible, equitable and interoperable data governance approaches; and enhancing international governance of artificial intelligence for the benefit of humanity", it is imperative that we, as digital media practitioners, take a proactive stance. We must commit towards using this powerful technology, in its application(s) to digital media, for advancing public good, and in service of the public interest.

We, therefore, under the ambit of Haarlem Declaration 2024, commit to deploying and utilizing all forms of AI-powered tools and technologies, in relation to digital media, in accordance with the following values and principles:



1. Ensuring transparency and explainability

- Human Accountability: Maintain transparency about how AI contributes to tasks and decision-making processes across individual, programmatic, and organizational levels¹.
- Explainable tools: Prioritise the use of AI tools that provide clear, understandable explanations for their recommendations, data privacy settings and data consent management.
- Continuous Education: Educate ourselves and relevant stakeholders about AI functionalities, potential biases, and limitations to enable greater transparency and better inform our decisions

Examples in practice:

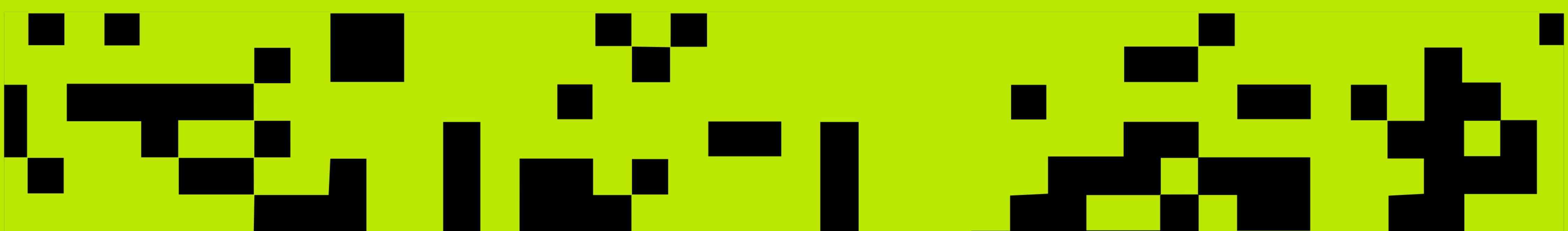
- We commit to conveying to our stakeholders and audiences how AI is utilized in content production, curation, distribution, and audience analytics practices.
- We commit to urge AI developers to prioritise explainability and accessibility in AI systems designs, and we, as digital media practitioners, commit to sharing this information in an engaging and understandable way to promote public good and foster informed dialogue.

2. Promoting ethical data practices

- Protect personal Information: Ensure all AI applications comply with data protection laws and safeguard sensitive data
- Data Minimization: Limit data collection to what is necessary for specific, clearly defined purposes.
- Consent and agency: Obtain informed consent for data use, providing users with transparent, accessible privacy policies that clearly outline how data is collected stored, used, shared, and deleted. Enable users to exercise control and agency over their data with opt-in and opt-out options.

Examples in practice

- We commit to implementing AI tools that have clear terms and processes for data collection and use – including having an opt-in and opt-out options in place
- We commit to calling on AI developers to design tools and systems that prioritize data diversity and inclusivity, guided by frameworks of informed consent and privacy protections.
- We commit to openly communicating to our audiences and stakeholders what types of data we collect, the purpose behind the collection, and how we use this data, whether for content curation, data analytics, or other applications



3. Safeguarding Information Integrity & Content Authenticity

- **Validate AI Outputs:** Implement comprehensive fact-checking protocols for all content generated, curated, or recommended by AI systems to ensure its reliability
- **Avoid and minimize misinformation:** Establish safeguards in our processes to ensure that all AI-supported communications are factual, non-harmful, and support the
- **Quality Control:** Develop and adhere to content standard that prioritise accuracy, credibility, and contextual integrity before dissemination.

Examples in practice:

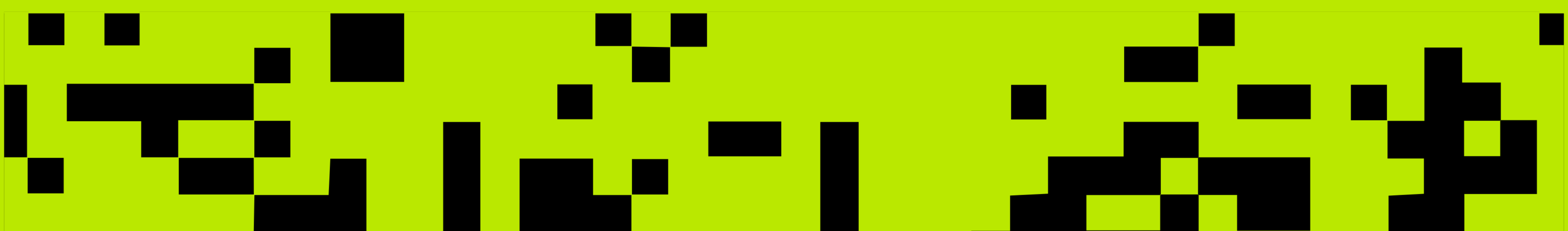
- We commit to upholding journalistic integrity and accuracy, ensuring that AI-generated content adheres to the same standards of truth as human-produced content.
- We commit to respecting the intellectual property rights of independent media and refrain from using our content to train their AI models without explicit permission or financial compensation.
- We commit to full transparency, when AI tools have been used to create or support digital content – be it text, audio, visual, captions, etc.) as well as for fundraising purposes (proposal development, positioning and visibility documents, etc.)
- We commit to upholding necessary measures and knowledge sharing to uphold the content provenance and authenticity;

4. Minimizing Bias, Harm, and Discrimination in use of AI tools

- **Assess AI Applications:** Regularly evaluate AI tools to identify, mitigate, and address potential biases that could result in unjust or discriminatory outcomes.
- **Promote Fairness:** Use diverse datasets and inclusive algorithmic designs to ensure AI systems treat users equitably and prevent the perpetuation of harmful and discriminatory stereotypes.
- **Human Oversight:** Maintain active human supervision to monitor and correct biased AI outputs ensuring alignment with human rights principles

Examples in Practice:

- We consciously implement AI for content moderation in ways that uphold freedom of expression while effectively countering harmful content such as hate speech and disinformation.
- We commit to ensuring that AI-driven content moderation processes we use are transparent, accountable, and free from bias.
- We commit to using AI in content curation to promote a diversity of perspectives, providing relevant and inclusive content to our audiences while actively avoiding the reinforcement of echo chambers or perpetuate inequality.



- We commit to calling on AI developers to design AI systems that prioritize data diversity and inclusivity, ensuring that marginalized and underrepresented groups are meaningfully included in algorithmic decision-making

5. Centring people over technology

- Assess Impact on Roles: Evaluate how AI implementation may affect staff positions and responsibilities and ensure that any transformation is managed with a people-first approach
- Support Staff Transition: Provide training, upskilling, and development opportunities to adapt to new technologies.
- Balance Automation and Human Touch: Use AI to enhance, not replace, human expertise, creativity, and interaction.

Examples in Practice:

- We commit to ensuring that human oversight remains a central feature in all AI processes, ensuring accountability and enabling intervention in cases of error or harm.
- We commit to ensuring that all AI systems are regularly audited and updated to meet evolving ethical standards and societal needs.
- We commit to using AI in content production to enhance creativity and efficiency while ensuring that human oversight remains integral to the editorial process.
- We commit to collaborative knowledge, learning and sharing, including with other media organizations, AI developers, and policymakers to share knowledge and best practices in AI use, including its impact on people who utilize these tools
- We call on governments and regulatory bodies to establish clear guidelines and frameworks that support the ethical use of AI in media, while safeguarding freedom of the press and human rights.

6. Balancing environmental impact of AI use

- Energy Consumption: Recognise the energy demands of AI applications, especially those requiring significant computational power, and actively seek to minimize their carbon footprint.
- Eco-Friendly Options: Opt for AI tools that prioritize reduction of environmental impact efficiency.
- Green Policies: Embed sustainability considerations into procurement and operational policies pertaining to deployment and use of AI tools

Examples in Practice:

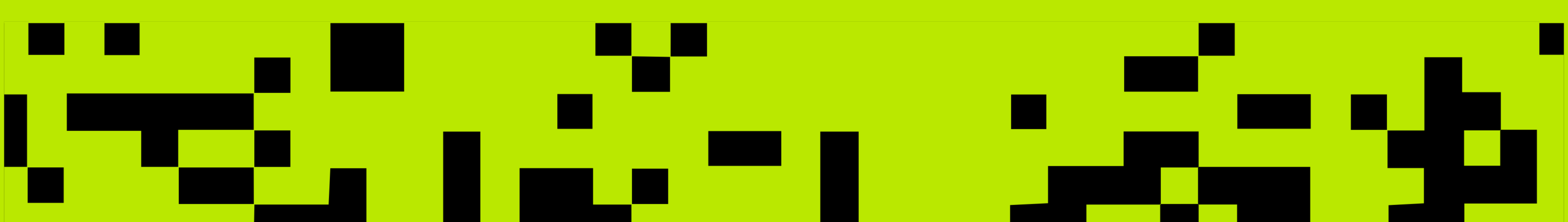
- We commit to integrating a climate and environmental lens into our strategy and implementation processes to promote mindful use of AI technologies and applications in everyday work
- We commit to engaging in collaborative knowledge sharing with other media organizations and our audiences to raise awareness about the environmental impact of AI. This includes supporting and highlighting marginalized narratives and counter-narratives on this pressing and intersecting topic

Practical Commitments

To realise these principles and values in practice and action, we propose to strive towards implementing the following practical commitments:

- **Learn, Share, Deliberate and Empower:** as individuals and a collective, about using AI ethically and responsibly through open and accessible dialogue, learning circles, and lived experiences in relation to using AI in our everyday digital media work
- **Shape & Amplify Diverse Narratives:** Utilize the power of digital media to influence existing, and shape diverse, marginalised, and counter narratives about AI with our respective audiences, online and offline to inform opinions, decisions and actions
- **Co-document our stories of AI practice:** over time, for internal and external mutual reflection on the varying trajectories and outcomes resulting from application of AI in digital media work, including human experience of using AI technology
- **Ethical AI Checklist:** Operationalise ethical and responsible use of AI in digital media work by co-drafting and implementing an ethical AI checklist to regularly monitor, course-correct, promote and ensure ethical use of AI across various digital media tasks, processes and functions
- **Produce Evidence-Backed Research** of both positive and negative implications of AI in digital media – spanning areas of information integrity/disorder, displacement of media workers, propagation of bias resulting from use of AI in a digital media process/function, etc.
- **Collectively Advocate** with/towards our peers and stakeholders in (digital) media ecosystem, where relevant, on minimizing harms and promoting ethical and responsible use of AI in digital media and by digital media actors, including donors and funders

We, the undersigned, believe that these actions – rooted in our shared values and principles -, will ensure that as digital media makers, we remain both alert and informed about existing and forthcoming changes in AI and technology (more broadly). Together, we will continue to co-create, enable, and sustain, inclusive, safe, and reliable digital media spaces for all.



DÉCLARATION DE HAARLEM

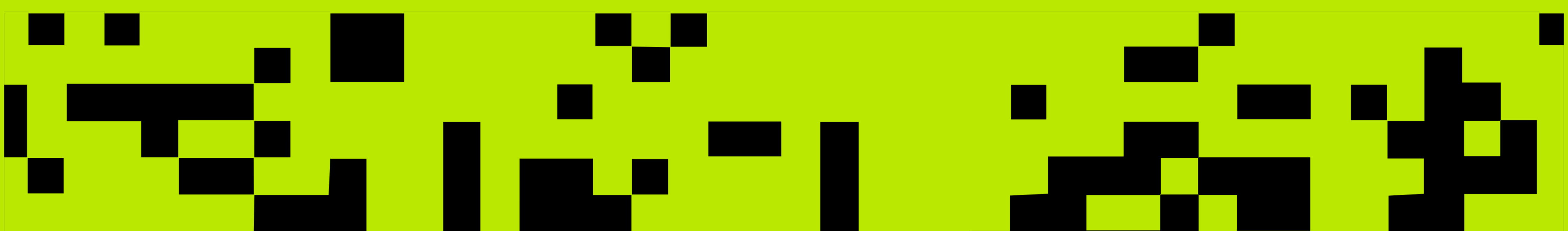
DÉFENDRE L'IA POUR FAIRE PROGRESSER LES ESPACES DE MÉDIAS NUMÉRIQUES INCLUSIFS, SÛRS ET FIABLES !

En tant que communauté mondiale de créateurs de médias numériques - y compris des organisations de médias indépendants, des journalistes, des créateurs de contenu numérique et des groupes de la société civile utilisant les médias numériques pour leur impact social - nous reconnaissons que les technologies de l'IA, et l'IA générative en particulier, influencent, façonnent et transforment activement et rapidement l'écosystème des médias numériques. Son influence s'étend à tous les aspects de la création de contenu, de la curation, de la recommandation, de la modération, de la distribution, de la promotion et de l'analyse de l'audience.

Si la technologie de l'IA offre la possibilité d'améliorer l'efficacité, la rationalisation et la rentabilité des processus et fonctions routiniers et à forte intensité de main-d'œuvre des médias numériques, son déploiement présente également des risques importants. La perturbation de l'écosystème des médias numériques par la prolifération de fausses informations, l'érosion de la supervision humaine dans la prise de décision éditoriale et le déplacement des travailleurs des médias sont déjà visibles. Il est donc urgent d'adopter une approche fondée sur les valeurs, la responsabilité et l'éthique pour intégrer les technologies de l'IA dans le travail et les organisations des médias numériques.

En attendant l'adoption du Pacte mondial pour le numérique (PMN) par les États membres des Nations unies en septembre 2024, lors du Sommet de l'avenir, dont l'un des principaux objectifs est de « favoriser un espace numérique inclusif, ouvert, sûr et sécurisé qui respecte, protège et promeut les droits de l'homme ; promouvoir des approches responsables, équitables et interopérables en matière de gouvernance des données ; et renforcer la gouvernance internationale de l'intelligence artificielle au profit de l'humanité », il est impératif que nous, praticiens des médias numériques, adoptions une position proactive. Nous devons nous engager à utiliser cette puissante technologie, dans ses applications aux médias numériques, pour faire avancer le bien public et au service de l'intérêt général.

Par conséquent, dans le cadre de la Déclaration de Haarlem 2024, nous nous engageons à déployer et à utiliser toutes les formes d'outils et de technologies alimentés par l'IA, en relation avec les médias numériques, conformément aux valeurs et aux principes suivants :



1. Assurer la transparence et l'explicabilité

- **Responsabilité humaine:** Maintenir la transparence sur la manière dont l'IA contribue aux tâches et aux processus décisionnels aux niveaux individuel, programmatique et organisationnel
- **Outils explicables:** Donnez la priorité à l'utilisation d'outils d'IA qui fournissent des explications claires et compréhensibles sur leurs recommandations, les paramètres de confidentialité des données et la gestion du consentement aux données.
- **Formation continue :** Nous éduquer et éduquer les parties prenantes concernées sur les fonctionnalités de l'IA, les biais potentiels et les limites pour permettre une plus grande transparence et mieux informer nos décisions.

Exemples pratiques :

- Nous nous engageons à expliquer à nos parties prenantes et à notre public comment l'IA est utilisée dans la production de contenu, la conservation, la distribution et les pratiques d'analyse de l'audience.
- Nous nous engageons à exhorter les développeurs d'IA à donner la priorité à l'explicabilité et à l'accessibilité dans la conception des systèmes d'IA et, en tant que praticiens des médias numériques, nous nous engageons à partager ces informations d'une manière attrayante et compréhensible afin de promouvoir le bien public et d'encourager un dialogue éclairé.

2. Promouvoir des pratiques éthiques en matière de données

- **Protéger les informations personnelles :** Veiller à ce que toutes les applications d'IA respectent les lois sur la protection des données et protègent les données sensibles.
- **Minimisation des données :** Limitez la collecte de données à ce qui est nécessaire à des fins spécifiques et clairement définies.
- **Consentement et agence:** Obtenez un consentement éclairé pour l'utilisation des données, en fournissant aux utilisateurs des politiques de confidentialité transparentes et accessibles qui décrivent clairement comment les données sont collectées, stockées, utilisées, partagées et supprimées. Permettez aux utilisateurs d'exercer un contrôle et une influence sur leurs données grâce à des options d'acceptation et de refus.

Exemples pratiques

- Nous nous engageons à mettre en œuvre des outils d'IA dont les conditions et les processus de collecte et d'utilisation des données sont clairs, y compris en ce qui concerne les options d'acceptation et de refus.
- Nous nous engageons à demander aux développeurs d'IA de concevoir des outils et des systèmes qui donnent la priorité à la diversité et à l'inclusion des données, en s'appuyant sur des cadres de consentement éclairé et de protection de la vie privée.
- Nous nous engageons à communiquer ouvertement à notre public et à nos parties prenantes les types de données que nous collectons, l'objectif de cette collecte et la manière dont nous utilisons ces données, que ce soit pour la curation de contenu, l'analyse de données ou d'autres applications.

3. Sauvegarder l'intégrité de l'information et l'authenticité du contenu

- **Valider les résultats de l'IA:** Mettez en œuvre des protocoles complets de vérification des faits pour tous les contenus générés, conservés ou recommandés par les systèmes d'IA afin d'en garantir la fiabilité.
- **Éviter et minimiser la désinformation:** Mettre en place des garanties dans nos processus pour s'assurer que toutes les communications soutenues par l'IA sont factuelles, non préjudiciables et qu'elles soutiennent les objectifs de l'UE.
- **Contrôle de la qualité:** Élaborer et respecter des normes de contenu qui privilégient l'exactitude, la crédibilité et l'intégrité contextuelle avant la diffusion.

Exemples pratiques :

- Nous nous engageons à respecter l'intégrité et l'exactitude journalistiques, en veillant à ce que le contenu généré par l'IA respecte les mêmes normes de vérité que le contenu produit par l'homme.
- Nous nous engageons à respecter les droits de propriété intellectuelle des médias indépendants et à ne pas utiliser notre contenu pour entraîner leurs modèles d'IA sans autorisation explicite ou compensation financière.
- Nous nous engageons à faire preuve d'une transparence totale lorsque des outils d'IA ont été utilisés pour créer ou soutenir du contenu numérique (texte, audio, visuel, légendes, etc.) ainsi qu'à des fins de collecte de fonds (élaboration de propositions, documents de positionnement et de visibilité, etc.)
- Nous nous engageons à prendre les mesures nécessaires et à partager les connaissances pour garantir la provenance et l'authenticité du contenu ;

4. Minimiser les préjugés, les préjudices et la discrimination dans l'utilisation des outils d'IA

- **Évaluer les applications de l'IA :** Évaluer régulièrement les outils d'IA afin d'identifier, d'atténuer et de traiter les biais potentiels qui pourraient entraîner des résultats injustes ou discriminatoires.
- **Promouvoir l'équité :** Utilisez des ensembles de données diversifiés et des conceptions algorithmiques inclusives pour garantir que les systèmes d'IA traitent les utilisateurs de manière équitable et empêchent la perpétuation de stéréotypes préjudiciables et discriminatoires.
- **Supervision humaine :** Maintenez une supervision humaine active pour surveiller et corriger les résultats biaisés de l'IA en veillant à l'alignement sur les principes des droits de l'homme.

Exemples pratiques :

- Nous mettons consciemment en œuvre l'IA pour la modération de contenu de manière à respecter la liberté d'expression tout en contrant efficacement les contenus préjudiciables tels que les discours haineux et la désinformation.
- Nous nous engageons à veiller à ce que les processus de modération de contenu pilotés par l'IA que nous utilisons soient transparents, responsables et exempts de préjugés.
- Nous nous engageons à utiliser l'IA dans la curation de contenu pour promouvoir la diversité des perspectives, en fournissant un contenu pertinent et inclusif à nos publics tout en évitant activement de renforcer les chambres d'écho ou de perpétuer l'inégalité.
- Nous nous engageons à demander aux développeurs d'IA de concevoir des systèmes d'IA qui donnent la priorité à la diversité des données et à l'inclusion, en veillant à ce que les groupes marginalisés et sous-représentés soient inclus de manière significative dans la prise de décision algorithmique.

5. Centrer les personnes sur la technologie

- Évaluer l'impact sur les rôles : Évaluez comment la mise en œuvre de l'IA peut affecter les postes et les responsabilités du personnel et veillez à ce que toute transformation soit gérée dans le cadre d'une approche axée sur les personnes.
- Soutenir la transition du personnel : Proposez des formations, un perfectionnement et des opportunités de développement pour vous adapter aux nouvelles technologies.
- Équilibrer l'automatisation et le contact humain: Utilisez l'IA pour améliorer, et non remplacer, l'expertise, la créativité et l'interaction humaines.
-

Exemples pratiques :

- Nous nous engageons à veiller à ce que la supervision humaine reste un élément central de tous les processus d'IA, afin de garantir la responsabilité et de permettre une intervention en cas d'erreur ou de préjudice.
- Nous nous engageons à faire en sorte que tous les systèmes d'IA soient régulièrement contrôlés et mis à jour pour répondre à l'évolution des normes éthiques et des besoins sociétaux.
- Nous nous engageons à utiliser l'IA dans la production de contenu pour améliorer la créativité et l'efficacité, tout en veillant à ce que la supervision humaine reste partie intégrante du processus éditorial.
- Nous nous engageons à collaborer en matière de connaissances, d'apprentissage et de partage, notamment avec d'autres organisations médiatiques, des développeurs d'IA et des décideurs politiques, afin de partager les connaissances et les meilleures pratiques en matière d'utilisation de l'IA, y compris son impact sur les personnes qui utilisent ces outils.
- Nous appelons les gouvernements et les organismes de régulation à établir des lignes directrices et des cadres clairs qui soutiennent l'utilisation éthique de l'IA dans les médias, tout en préservant la liberté de la presse et les droits de l'homme.

6. Équilibrer l'impact environnemental de l'utilisation de l'IA

- Consommation d'énergie : Reconnaître les besoins en énergie des applications d'IA, en particulier celles qui nécessitent une puissance de calcul importante, et chercher activement à minimiser leur empreinte carbone.
- Options écologiques: Optez pour des outils d'IA qui donnent la priorité à la réduction de l'efficacité de l'impact sur l'environnement.
- Politiques vertes: Intégrez des considérations de durabilité dans les politiques d'achat et d'exploitation relatives au déploiement et à l'utilisation des outils d'IA.

Exemples pratiques :

- Nous nous engageons à intégrer une perspective climatique et environnementale dans notre stratégie et nos processus de mise en œuvre afin de promouvoir une utilisation réfléchie des technologies et des applications d'IA dans le travail quotidien.
- Nous nous engageons à partager nos connaissances en collaboration avec d'autres organisations médiatiques et nos publics afin de les sensibiliser à l'impact environnemental de l'IA. Il s'agit notamment de soutenir et de mettre en évidence les récits et contre-récits marginalisés sur ce sujet urgent et intersectoriel.

Engagements pratiques

Pour concrétiser ces principes et valeurs dans la pratique et l'action, nous proposons de nous efforcer de mettre en œuvre les engagements pratiques suivants :

Apprendre, partager, délibérer et responsabiliser : en tant qu'individus et en tant que collectif, sur l'utilisation éthique et responsable de l'IA par le biais d'un dialogue ouvert et accessible, de cercles d'apprentissage et d'expériences vécues en relation avec l'utilisation de l'IA dans nos travaux quotidiens sur les médias numériques.

Façonner et amplifier des récits divers : Utiliser le pouvoir des médias numériques pour influencer et façonner des récits divers, marginalisés et opposés sur l'IA avec nos publics respectifs, en ligne et hors ligne, afin d'informer les opinions, les décisions et les actions.

Co-documenter nos pratiques en matière d'IA : au fil du temps, pour une réflexion mutuelle interne et externe sur les différentes trajectoires et les résultats résultant de l'application de l'IA dans les médias numériques, y compris l'expérience humaine de l'utilisation de la technologie de l'IA.

Liste de contrôle de l'IA éthique : Opérationnaliser l'utilisation éthique et responsable de l'IA dans les médias numériques en élaborant et en mettant en œuvre une liste de contrôle éthique de l'IA pour contrôler, corriger, promouvoir et garantir régulièrement l'utilisation éthique de l'IA dans les différentes tâches, processus et fonctions des médias numériques.

Produire des recherches fondées sur des données probantes concernant les implications positives et négatives de l'IA dans les médias numériques - couvrant les domaines de l'intégrité/du désordre de l'information, du déplacement des travailleurs des médias, de la propagation des préjugés résultant de l'utilisation de l'IA dans un processus/une fonction des médias numériques, etc.

Plaider collectivement auprès de nos pairs et des parties prenantes de l'écosystème des médias (numériques), le cas échéant, pour minimiser les dommages et promouvoir une utilisation éthique et responsable de l'IA dans les médias numériques et par les acteurs des médias numériques, y compris les donateurs et les bailleurs de fonds.

Nous, les soussignés, pensons que ces actions - enracinées dans nos valeurs et principes communs - garantiront qu'en tant que créateurs de médias numériques, nous restons à la fois vigilants et informés des changements existants et à venir dans le domaine de l'IA et de la technologie (de manière plus générale). Ensemble, nous continuerons à co-crée, permettre et soutenir des espaces de médias numériques inclusifs, sûrs et fiables pour tous.

