



**Artificial Intelligence and
Information Warfare**
**L'Intelligence artificielle et la
Guerre de l'information**



Artificial Intelligence and Information Warfare

L'Intelligence artificielle et la Guerre de l'information

The Franco-British Council in partnership with Inria and the Franco-British Data Society hosted a one-day seminar on "Artificial Intelligence and Information Warfare" on Wednesday 10th July 2024 at University College London.

Welcome and Introduction

Michael Butcher, Chair of the Franco-British Data Society, opened the seminar with an overview of the Franco-British Data Society, established in 2018 and outlined its mission to foster connections among data professionals across sectors. The society aims to become a registered charity and is actively involved in educational and innovation initiatives, including a notable project in France engaging students with CO2 measurement data.

He also highlighted upcoming events, including an in-person gathering at the University of Lille and emphasised the society's goal of bridging the gap between the worlds of work and education.

Hervé Mariton, Chair of the Conseil franco-britannique, gave a brief history of the Franco-British Council, established in 1972 to enhance civil society relations between France and the UK. He acknowledged past achievements and ongoing activities, including work on defence, climate change, energy and a new pillar of work on science and society.

Hervé Mariton outlined the agenda for the seminar on "Artificial Intelligence and Information Warfare" and he noted the seminar's focus on practical applications and implications for democracy, media and public debate. The discussion was intended to explore both the threats and opportunities presented by AI, emphasising the need for topical analysis beyond general discussions.

Le mercredi 10 juillet 2024, en collaboration avec inria et la Franco-British Data Society, le Conseil franco-britannique a organisé, à l'University College London, un séminaire d'une journée sur le thème « l'Intelligence artificielle et la Guerre de l'information »

Allocutions de bienvenue et présentation

Michael Butcher, Président de la Franco-British Data Society, a officiellement ouvert le séminaire en donnant un aperçu global de la Franco-British Data Society, établie en 2018, tout en expliquant la mission centrée sur la promotion des rapports entre les professionnels de l'information tous secteurs confondus. La société, qui aspire à devenir un organisme de bienfaisance agréé, s'investit activement dans la mise en œuvre d'initiatives pédagogiques et d'actions innovatrices, dont un grand projet en France qui favorise l'implication d'étudiants dans l'utilisation des données de mesure de CO2.

Il a également mentionné les événements à venir, au nombre desquels une réunion à laquelle il assistera en personne à l'Université de Lille. Il a aussi insisté sur la détermination de la société à combler l'écart entre le monde du travail et le monde de l'éducation.

Hervé Mariton, Président du Conseil franco-britannique, a donné un bref historique du Conseil franco-britannique, établi en 1972, dans un but de resserrer les relations de la société civile entre la France et le R.-U. Il a salué les réalisations accomplies et les activités en cours, dans lesquelles s'inscrivent la défense, le changement climatique, l'énergie et un nouveau projet de travail lié à la science et à la société.

Hervé Mariton a présenté les grandes lignes du programme du séminaire sur « l'Intelligence artificielle



He outlined the key themes which the seminar aimed to address:

- The exploration of AI's role in shaping media landscapes, including the dissemination of disinformation and fake news.
- Ethical considerations and the balance between ensuring factual accuracy and respecting individual freedom of expression.
- The impact on democracy and an understanding of how AI affects democratic processes and public trust in information.
- Technological and organisational Implications and how AI technologies influence media operations and organisational structures

The State of AI in 2024: An Introduction by Dr. Benjamin Guedj

Dr. Benjamin Guedj, Director of Research at Inria, Professor of Machine Learning and Foundational AI at University College London, Turing Fellow at The Alan Turing Institute, director of the Franco-British Data Society and Trustee of the Franco-British Council, acknowledged the significant advancements and transformations in AI and highlighted the ongoing evolution of the field and its profound impact on technology and society.

Dr. Guedj started by addressing the complexity of defining AI. AI encompasses a wide range of technologies and approaches, often driven by contributions from various disciplines including mathematics, computer science, medical sciences, neuroscience and even sociology and psychology. He cites Arthur Samuel's 1959 definition of machine learning, which is a subset of AI focusing on systems that learn from data without explicit programming.

et la Guerre de l'information », et a souligné que ce séminaire portait principalement sur des applications et des implications pratiques dans le cadre de débats démocratiques, médiatiques et publics. L'objectif de la discussion était d'explorer à la fois les menaces et les opportunités que présente l'IA, en insistant sur la nécessité d'avoir recours à des analyses thématiques plutôt que de s'en tenir qu'à des discussions d'ordre général.

Il a présenté les principaux thèmes du séminaire, à savoir :

- L'exploration du rôle de l'IA dans le modelage du paysage médiatique, notamment la diffusion de la désinformation et des fake news (fausses informations).
- Les considérations éthiques et l'équilibre entre s'assurer de l'exactitude factuelle et respecter la liberté d'expression individuelle.
- L'impact sur la démocratie et comprendre comment l'IA influe sur le processus démocratique et la confiance du public dans l'information.
- Les implications technologiques et organisationnelles et comment les technologies de l'IA influencent les opérations médiatiques et les structures organisationnelles.

L'État des lieux de l'IA en 2024 : une présentation par le Dr. Benjamin Guedj

Le Dr Benjamin Guedj, Directeur de Recherche à Inria, Professeur de Machine Learning (apprentissage automatique) et d'IA fondamentale à l'University College London, Turing Fellow à l'Alan Turing Institute, membre du Conseil d'administration de la Franco-British Data Society et membre du conseil d'administration du Conseil franco-britannique, a salué les importantes avancées et transformations dont a fait l'objet l'IA et a mis en évidence l'évolution continue de ce domaine ainsi que ses profondes répercussions sur la technologie et la société.

Le Dr Guedj a débuté sa présentation en soulignant à quel point la définition de l'intelligence artificielle était complexe. L'IA englobe un vaste cortège de technologies et d'approches, bien souvent sous l'impulsion de contributions émanant de disciplines diverses, dont les mathématiques, l'informatique, les sciences médicales, la neuroscience et même jusqu'à la sociologie et la psychologie. Il a cité la définition du machine learning de 1959 d'Arthur Samuel, qui est un sous-ensemble de l'IA

Dr. Guedj emphasised the importance of machine learning, particularly deep learning, which has been a significant driver of recent advancements. Deep learning relies on neural networks with many layers, a concept that gained traction around 2012 due to improvements in computational power. This approach has fueled the recent AI revolution by enabling more complex and powerful models.

In 2024, key areas of AI research include computer vision, speech processing and Natural Language Processing (NLP). These areas represent the forefront of AI research and application, with ongoing developments transforming how we interact with technology.

Dr. Guedj pointed out that despite the impressive capabilities of AI systems like ChatGPT, the term "intelligence" may be misleading. Current AI models, including large language models (LLMs), are fundamentally probabilistic tools predicting sequences of words. These models, while powerful, lack true understanding or consciousness and are essentially sophisticated statistical machines.

AI presents both opportunities and risks. Dr. Guedj discussed the potential for both positive and negative uses of AI technologies. He cautioned that even well-intentioned applications of AI can lead to unintended and possibly harmful outcomes. The field is evolving rapidly and understanding its implications requires ongoing reflection and adaptation.

Looking forward, Dr. Guedj highlighted the goal of achieving more efficient and generalised learning systems, akin to human learning capabilities. The pursuit of "augmented intelligence" rather than "artificial intelligence" may offer a more accurate reflection of the current state of technology.



se concentrant sur des systèmes qui apprennent à partir de données sans programmation explicite.

Le Dr Guedj a insisté sur l'importance que revêt le machine learning, particulièrement le deep learning (apprentissage profond), qui s'est révélé être un moteur majeur des récentes avancées. Ce type d'intelligence artificielle s'appuie sur des réseaux neuronaux composés de nombreuses couches, un concept prisé aux alentours de l'année 2012 suite aux améliorations de la puissance de calcul.

Cette approche a stimulé la récente révolution de l'IA en permettant de créer des modèles plus complexes et plus puissants.

En 2024, la vision par ordinateur, le traitement de la parole et le traitement automatique du langage naturel (NLP - Natural Language Processing) sont des disciplines clés de la recherche et de l'application IA, des développements continus en cours permettant de transformer notre mode d'interaction avec la technologie.

Le Dr Guedj a souligné qu'en dépit des extraordinaires capacités des systèmes d'IA, tels que le ChatGPT, le terme « intelligence » peut porter à confusion. Les modèles d'IA actuels, notamment les grands modèles de langage (LLM - large language models) sont avant tout des outils fondamentalement probabilistes prédisant des séquences de mots. Ces modèles, tout en étant puissants, manquent de véritable compréhension ou conscience et ne sont essentiellement que des machines statistiques sophistiquées.

L'intelligence artificielle entraîne opportunités et risques dans son sillage. Le Dr Guedj s'est exprimé sur les usages à la fois positifs et négatifs des technologies de l'IA, avertissant que même les applications bien intentionnées de l'IA peuvent aboutir à des résultats indésirables et éventuellement dangereux. Ce domaine évolue rapidement et en comprendre les implications exige une réflexion constante et une adaptation de tout instant.

Se projetant dans le futur, le Dr Guedj a parlé de l'intention de créer des systèmes d'apprentissage plus efficaces et plus généralisés, comparables aux capacités d'apprentissage de l'être humain. Il se pourrait que le terme « intelligence augmentée » plutôt que « intelligence artificielle » donne une description plus

Dr. Guedj concluded by acknowledging the transformative potential of AI while emphasizing the importance of addressing its ethical and practical challenges. He invites further discussion and exploration of AI's impact on society and technology.



Keynote speech from David Colon, researcher at the Sciences Po History Centre / teacher of the history of communication, media and propaganda, author of "La guerre de l'information: Les États à la conquête de nos cerveaux" published by Tallandier

The keynote speech by David Colon gave a comprehensive overview of the modern challenges posed by digital information warfare and suggests a multi-faceted approach to countering these threats, emphasizing the need for systemic reforms and international cooperation.

David Colon started by noting that the Davos Forum has highlighted several critical risks with AI including misinformation, disinformation, cyber insecurity, social polarisation and conflict. He emphasises that these risks have been exacerbated by the advent of digital technology and the Internet.

He talked of the evolution of cyber warfare and post-Cold War, the rise of digital technology has transformed intelligence services. Colon discusses how agencies like the NSA have shifted focus from traditional espionage to protecting information systems and data gathering. The concept of cyber war and net warfare emerged which signified ideological conflicts extending into the digital realm.

On early observations of Russian interference, David

précise de l'état actuel de la technologie.

Pour conclure, le Dr Guedj a fait état du potentiel transformateur de l'IA, tout en insistant sur l'importance de faire face aux défis éthiques et pratiques qui l'accompagnent. Il engage à poursuivre la discussion et l'exploration de l'impact de l'IA sur la société et sur la technologie.

Discours d'ouverture prononcé par David Colon, chercheur au Centre d'histoire de Sciences Po, où il enseigne l'histoire de la communication, des médias et de propagande. Auteur de la « La guerre de l'information : Les États à la conquête de nos cerveaux », publié aux éditions Tallandier.

Le discours d'ouverture prononcé par David Colon a donné un tableau complet des enjeux modernes que pose la guerre de l'information numérique et suggère d'adopter une approche multidimensionnelle pour lutter contre ces menaces, en insistant sur le besoin de mettre en place des réformes systémiques et une coopération internationale.

David Colon a débuté sa présentation en signalant qu'à l'occasion du Forum de Davos plusieurs risques majeurs concernant l'IA ont été mis en avant, parmi lesquels la désinformation, la cyber-insécurité, la polarisation sociale et les conflits. Il a ajouté que l'avènement de la technologie numérique et de l'Internet n'ont fait qu'exacerber ces risques.

Il s'est exprimé sur l'évolution de la cyberguerre et de l'après-guerre froide, mentionnant que la montée en puissance de la technologie numérique a transformé les services de renseignements. Il a évoqué comment des agences comme l'Agence nationale de la sécurité (NSA / National Security Agency) est passée de l'espionnage traditionnel à la protection des systèmes d'information et à la collecte de données. L'émergence de la cyberguerre et de la guerre en réseau était signe de conflits idéologiques ayant des ramifications jusque dans la sphère numérique.

S'appuyant sur des précédentes observations d'interférences russes, David Colon en a retracé les origines au milieu des années 1990, évoquant le premier exemple éloquent de ce type d'interférences dans la politique américaine. Il s'est montré critique à l'égard de la conviction initiale que les réseaux numériques allaient

Colon traced the origins of Russian interference to the mid-1990s, highlighting the first significant instance of such interference in U.S. politics. He critiqued the initial belief that digital networks would inevitably promote liberal democratic values and lead to the collapse of autocratic regimes. To give the Russian historical context, David Colon outlined Vladimir Putin's background in the KGB and his connections with the German Stasi, emphasising the long-term strategy of undermining Western societies through disinformation. This strategy has been amplified with the advent of digital technologies.

David Colon went on to explore the use of machine learning and psychological profiling by Russian operatives. The application of predictive analytics, such as those developed by Cambridge Analytica, is discussed in relation to influencing political outcomes by targeting specific personality traits.

The rise of social networks and digital tools has allowed for unprecedented levels of influence by digital technologies on propaganda. He cited the use of social media platforms like Facebook to manipulate public



promouvoir inexorablement les valeurs démocratiques libérales et mèneraient à l'effondrement des régimes autocratiques. Pour établir le contexte historique russe, David Colon a fait état des antécédents de Vladimir Poutine au KGB et de ses liens avec la Stasi en Allemagne de l'Est, mettant l'accent sur la stratégie à long terme consistant à déstabiliser les sociétés occidentales à travers la désinformation. L'avènement des technologies numériques a davantage amplifié cette stratégie.

David Colon a poursuivi en explorant l'utilisation du machine learning et du profilage psychologique par les opérateurs russes. L'application de l'analyse prédictive, comme celles mises au point par Cambridge Analytica, est étudiée en considérant l'influence qu'elle peut avoir sur les résultats politiques en ciblant des traits de personnalités spécifiques.

L'essor grandissant des réseaux sociaux et des outils numériques a permis à la propagande par voies des technologiques numériques d'atteindre une influence sans précédent. Il a cité l'usage qu'il est fait des plateformes de médias sociaux comme Facebook dans un but de manipuler l'opinion publique ainsi que la création de faux profils et de bots pour diffuser la désinformation. Quel meilleur exemple que l'interférence russe dans le cadre de divers élections (aux USA, au R.-U., en France et en Moldavie, p. ex.) et la manière dont ces interférences ont été conçues pour exploiter les vulnérabilités des processus démocratiques.

David Colon s'est également penché sur le rôle des renseignements chinois et de l'influence de plateformes telles que TikTok, expliquant comment le Parti communiste chinois exploite ces plateformes pour conditionner les perceptions et collecter des données, en remarquant que les conséquences de cette stratégie sont plus subtiles et moins évidentes que toute autre forme d'influence.

Il s'est exprimé sur la croissance exponentielle de la mésinformation pilotée par l'IA, les outils de l'IA servant à générer du faux contenu et à manipuler l'opinion publique. Par ailleurs, a-t-il ajouté, la prolifération des faux sites médiatiques et le rôle que joue la publicité dans la promotion de la mésinformation représentent aussi des enjeux majeurs.

David Colon a en outre évoqué la nécessité d'instaurer une réglementation et une coordination à l'échelle

opinion and the creation of fake profiles and bots to spread disinformation. Examples include Russian interference in various elections (e.g., in the U.S., the UK, France and Moldova) and how these operations have been tailored to exploit vulnerabilities in democratic processes.

David Colon also addressed the role of Chinese intelligence and the influence of platforms like TikTok. He explained how the Chinese Communist Party uses these platforms to affect perceptions and collect data, noting that the impact is more subtle and less visible compared to other forms of influence.

He talked of the exponential growth of AI-driven misinformation, with AI tools being used to generate fake content and manipulate public opinion. The proliferation of fake media sites and the role of



advertising in promoting misinformation is also an issue.

David Colon talked of the need for regulation and global coordination. He stressed the importance of international regulation and systemic measures to counter digital misinformation. He suggests that regulations should focus on transparency in AI and digital platforms and global cooperation to address the pervasive threat of misinformation.

He called for urgent action to combat these threats, including regulatory measures for AI and digital platforms, increased transparency and international collaboration. Lessons from countries with strong democratic resilience, such as Finland and the Netherlands, could offer valuable insights, he believed. David Colon concluded by emphasising the critical need for both immediate and long-term strategies to address the evolving landscape of information warfare and its

mondiale. Il a déclaré à quel point il était important de pouvoir s'appuyer sur une réglementation et sur des mesures systémiques internationales pour lutter contre la mésinformation numérique. Selon lui, il serait bon que des réglementations privilégient l'application de la transparence dans l'IA et les plateformes numériques dans une coopération mondiale afin de contrer la menace omniprésente que pose la mésinformation.

Il est temps, a-t-il déclaré, de passer à l'action si l'on veut éradiquer ces trois menaces, à savoir : prendre des dispositions réglementaires en ce qui concerne l'IA et les plateformes numériques, accroître la transparence et la collaboration internationale. À son avis, l'on pourrait tirer des leçons de pays ayant une forte résilience démocratique [?] la Finlande et les Pays-Bas entre autres [?] dont la contribution nous serait précieuse. En conclusion, il a insisté sur l'absolue nécessité d'adopter des stratégies à la fois à effet immédiat et à long terme afin de faire face au paysage évolutif de la guerre de l'information et aux conséquences qu'elle peut avoir sur les institutions démocratiques

La science en action

- **Dr Benjamin Guedj, directeur de Recherche à Inria, Professeur de Machine Learning (apprentissage automatique) et d'IA fondamentale à l'University College London et Turing Fellow à l'Alan Turing Institute.**
- **Garima Gaur, équipe CEDAR Inria France, Docteur en philosophie**

La discussion a porté sur la complexité associée à la lutte contre la mésinformation et sur la nécessité d'adopter une approche multidimensionnelle qui englobe une innovation technologique, des mesures réglementaires et une supervision humaine. Même si des outils comme le Fact-Check Bureau constituent pour les fact-checkers une remarquable avancée, il n'en reste pas moins que le plus grand défi auquel l'on est confronté est d'établir un équilibre entre le rythme rapide selon lequel se développent les technologies et des mesures réglementaires et défensives qui se veulent efficaces.

Un dialogue permanent et une collaboration constante entre les chercheurs, les décideurs politiques et les agences de fact-checking sont également indispensables pour gérer les aspects à la fois techniques et sociétaux de la mésinformation.

impact on democratic institutions.

Science in action

- **Dr. Benjamin Guedj, Director of Research at Inria, Professor of Machine Learning and Foundational AI at University College London and Turing Fellow at The Alan Turing Institute**
- **Garima Gaur, CEDAR team Inria France, Doctor of Philosophy**

The discussion looked at the complexity of tackling misinformation and the need for a multifaceted approach that includes technological innovation, regulatory measures and human oversight. While tools like Fact-Check Bureau represent significant progress in aiding fact-checkers, the broader challenge involves balancing the rapid pace of technological advancement with effective regulatory and defensive measures.

There is also a clear need for ongoing dialogue and collaboration among researchers, policymakers and fact-checking agencies to address both the technical and societal aspects of misinformation.

As Dr. Benjamin Guedj pointed out, technologies like large language models (LLMs) can be used for both beneficial and harmful purposes. The mathematics behind these technologies is consistent, but their applications can lead to vastly different outcomes.

Garima Gaur talked about her work focusing on enhancing fact-checking through technology, particularly by developing systems that help researchers and journalists quickly find relevant fact-check articles. This addresses the overwhelming volume of claims and the speed at which misinformation spreads. She talked about the Fact-Check Bureau – a tool which aims to improve the efficiency of fact-checking by providing a sophisticated query interface and interactive pipeline design for researchers. It does not automate fact-checking but supports human fact-checkers in their work.

On the issue of potential biases in fact-checking agencies and the need to ensure that these agencies' methods are transparent and unbiased, Garima Gaur said that, while her work is focused on the computational aspects, the concern about bias in the fact-checking process itself is valid and highlights the need for continuous evaluation of fact-checking methodologies. It was also

Comme l'a souligné le Dr Benjamin Guedj, rien n'empêche d'utiliser des technologies du type grands modèles de langage (LLM) à des fins aussi bien bénéfiques que nocives. Même si les mathématiques qui sous-tendent ces technologies sont cohérentes, leurs applications peuvent mener à des résultats fort différents.

Garima Gaur s'est exprimée sur son travail dédié à optimiser le fact-checking par le biais de la technologie, notamment en développant des systèmes qui aident les chercheurs et les journalistes à se procurer des articles de fact-checking pertinents sans perdre de temps. Ceci permet de traiter l'immense volume de communiqués reçus et de freiner la vitesse à laquelle se répand la désinformation. Elle a également abordé le sujet du Fact-Check Bureau – un outil dont le but est d'améliorer l'efficacité du fact-checking en fournissant aux chercheurs une interface de vérification sophistiquée et une conception de pipeline de données interactif. Cet outil n'automatise pas le fact-checking, mais il soutient les fact-checkers humains dans leur travail.

Concernant la question des éventuelles partialités chez les agences de fact-checking et de la nécessité de s'assurer que les pratiques desdites agences sont transparentes et impartiales, Garima Gaur a déclaré que ☐ même si son travail est axé sur les aspects informatiques ☐ l'inquiétude concernant la partialité au sein du processus de fact-checking proprement dit existe bien et met en lumière la nécessité de constamment évaluer les méthodologies de fact-checking. La question à savoir si des solutions telles que le Fact-Check Bureau suffisent ou si une approche réglementaire globale et plus centralisée est nécessaire, a aussi été posée.

Au cours de la discussion, il a été pris acte de la tension qui existe entre le rythme accéléré de l'innovation technologique utilisée pour procéder à des attaques (p. ex. : la création de fake news) et le rythme plus lent des mesures défensives (p. ex. : les technologies de fact-checking). La nécessité d'instaurer une réglementation, assurant que les technologies puissent s'appuyer sur des mesures de sécurité adéquates avant d'être largement déployées, a aussi été abordée. Alors que nous devrions éviter de freiner l'innovation, il nous faut nous assurer que les technologies sont mises au point et déployées accompagnées de garde-fous afin d'éviter toute utilisation abusive.

questioned whether solutions like Fact-Check Bureau are sufficient or whether a more centralised, global regulatory approach is necessary.

The discussion acknowledged the tension between the rapid pace of technological innovation used for attacks (e.g., creating fake news) and the slower pace of



defensive measures (e.g., fact-checking technologies). The discussion also touched upon the need for regulation to ensure that technologies have adequate safety measures before they are widely deployed. While we should avoid stifling innovation, we need to ensure that technologies are developed and deployed with safeguards to prevent misuse.

The discussion expressed skepticism about the effectiveness of global governance bodies like the OECD or G20 in addressing the issues posed by misinformation and the potential for regulation to be outpaced by innovation. It suggests that while local solutions like Fact-Check Bureau are useful, they may need to be complemented by broader regulatory frameworks.

Despite technological advancements, Garima Gaur emphasised that human oversight remains crucial. Her tool supports fact-checkers but does not replace them and she acknowledged the ongoing need for human judgment in assessing the veracity of claims.

How do media cope?

- **Laura Ellis, Head of Technology Forecasting at BBC**
- **Laurent Guimier, Directeur général délégué à l'information de CMA Media**
- **Florent Latrive, Directeur adjoint de l'information, en charge de la stratégie numérique, Radio France**

The discussion highlighted the ongoing challenges and solutions in combating disinformation within media

Un certain scepticism a été exprimé au sujet de l'efficacité de certaines agences mondiales de gouvernance [?] comme l'OCDE ou le G20 [?] à résoudre les enjeux que posent la mésinformation et la possibilité que la réglementation soit dépassée par l'innovation. Il a été suggéré, tout en admettant que les solutions prises au niveau local comme le Fact-Check Bureau sont utiles, qu'elles pourraient être renforcées par des cadres réglementaires élargis.

Garima Gaur a insisté sur le fait que, malgré les avancées technologiques, la supervision humaine demeure capitale et que son outil soutient les fact-checkers, mais ne les remplace pas. Elle a confirmé qu'un jugement humain permanent est indispensable quand il s'agit d'évaluer la véracité des informations reçues.

Comment les médias affrontent-ils la désinformation ?

- **Laura Ellis, Head of Technology Forecasting à la BBC**
- **Laurent Guimier, Directeur général délégué à l'information de CMA Media**
- **Florent Latrive, Directeur adjoint de l'information, en charge de la stratégie numérique, Radio France**

La discussion a mis en lumière les défis et solutions perpétuels liés à la lutte contre la désinformation au sein des organisations médiatiques. Elle a insisté sur la nécessité de respecter la transparence, de disposer d'outils de pointe et d'offrir une formation continue aux journalistes pour qu'ils puissent manœuvrer dans un paysage médiatique en pleine évolution.

Laura Ellis a décrit l'approche qu'adopte la BBC pour gérer la désinformation, mettant l'accent sur quatre disciplines clés : collaboration au sein de la salle de presse, détection des deepfakes (enregistrements vidéo/audio), source médiatique et compétences en IA. La source médiatique était son élément central, la BBC ayant travaillé sur la Coalition for Content Provenance and Authenticity (C2PA), [?] une initiative dont l'objectif est d'authentifier le contenu en y ajoutant des métadonnées signées ; permettant ainsi aux utilisateurs de retracer l'origine des actualités et d'en vérifier l'authenticité. Ces métadonnées aident également à suivre le parcours des contenus sur toutes les plateformes. Ellis a évoqué le problème de la confiance dans les médias, compte tenu du fait que les niveaux de confiance ont décliné, et elle a déclaré à quel point il était important de fournir des informations transparentes aux utilisateurs.

organisation and emphasised the need for transparency, advanced tools and continuous journalist training to navigate the evolving media landscape.

Laura Ellis discussed the BBC's approach to tackling disinformation, emphasising four key areas: newsroom collaboration, deepfake detection, media provenance and AI literacy. Her focus was on media provenance, where the BBC has been working on the Coalition for Content Provenance and Authenticity (C2PA). This initiative aims to authenticate content by adding signed metadata, allowing users to trace the origin and verify the authenticity of news. This metadata also helps in tracking the content's journey across platforms. Ellis acknowledged the challenge of trust in media, as trust levels have been declining and the importance of providing transparent information to users.

Laurent Guimier spoke about the revolution brought by AI in journalism and its implications for tackling disinformation. He highlighted the importance of defending journalists and journalism as a democratic pillar amidst rising disinformation. He stressed the need for continuous training for journalists, particularly those who entered the profession before the advent of digital technology. He pointed out the challenges posed by AI and the need for local journalism to ensure reliable information and avoiding reliance solely on AI-generated content.

Florent Latrive discussed the evolution of fact-checking and its role in combating disinformation. He highlighted the increase in misinformation and the need for rapid and effective tools to verify information. He mentioned that fact-checking is no longer sufficient on its own; there's a need to address narratives and the broader context in which information is presented. He emphasised the importance of technological tools like stat checks and collaborations with other media and institutions to combat misinformation effectively.

The discussion raised a number of key issues:

- **Media Provenance and Transparency:** The importance of ensuring content authenticity through signed metadata to build trust and enable verification.
- **Challenges with AI and Disinformation:** The need to address the rapid spread of disinformation facilitated by AI and the need for journalist training to handle new

Laurent Guimier s'est exprimé sur la révolution que l'IA a provoqué dans le monde du journalisme et sur ses implications pour combattre la désinformation. Il a également souligné l'importance de défendre les journalistes et le journalisme en tant que pilier de la démocratie dans un monde où la désinformation s'intensifie, insistant en outre sur la nécessité pour les journalistes de ne cesser de se former, s'adressant en particulier à ceux qui ont débuté dans la profession avant l'avènement de la technologie numérique. Il a attiré l'attention sur les défis que pose l'IA et a déclaré que le journalisme local doit s'assurer de fournir des informations fiables et éviter de s'appuyer exclusivement sur le contenu généré par l'IA.

Florent Latrive a parlé de l'évolution du fact-checking et de son rôle dans la lutte contre la désinformation, mettant en lumière l'essor de la mésinformation et la nécessité de disposer d'outils rapides et efficaces permettant de vérifier l'information. Il a mentionné que le fact-checking n'est plus suffisant en soi et qu'il y a un besoin de revoir les narratifs et le contexte plus élargi dans lesquels est présentée l'information. Il a insisté sur l'importance de disposer d'outils technologiques comme



digital tools effectively.

- **Fact-Checking and Narrative Context:** The need to move beyond fact-checking to consider the broader narrative context and the impact of misinformation on public perception and trust.
- **Impact of Generative AI on Media:** This raises the question of how does the ability to generate more content affect journalism and public trust. The focus should be on improving tools for verification and maintaining human involvement in journalism to preserve authenticity.
- **Role of Context in Journalism:** Understanding the context behind information is crucial. The media must address not only the accuracy of facts but also the narrative and intentions behind them.
- **Disclosure of Content Origin:** There is a need for clear disclosure regarding whether content is AI-generated or human-created to help users navigate and trust information sources.
- **Editorial Priorities and Social Media Influence:** Journalists need to balance social media trends with traditional reporting to ensure accurate and reliable news coverage.

Control and freedom: checks and balances for the industry and government

Point of view of the industry

- **Debbie Widjaja, Chief Product Officer of Zero Gravity, ex Programme Manager of Hate Speech and Misinformation, Meta**
- **Simon Staffell, Director of Corporate Affairs, Microsoft UK**



les statchecks et de collaborer avec d'autres médias et institutions si l'on veut combattre la mésinformation efficacement.

La discussion a soulevé un certain nombre de points clés :

- **Source médiatique et transparence :** L'importance de garantir l'authenticité du contenu par le biais de métadonnées signées pour instaurer la confiance et permettre la vérification.
- **Défis que posent l'IA et la désinformation :** La nécessité de faire face à la diffusion accélérée de la désinformation que facilite l'IA et le besoin de formation des journalistes pour qu'ils soient aptes à utiliser efficacement les nouveaux outils numériques.
- **Fact-Checking et contexte narratif :** La nécessité de dépasser le fact-checking pour prendre en considération le contexte narratif plus élargi et les conséquences de la mésinformation sur la perception et la confiance du public.
- **L'impact de l'IA générative sur les médias :** Ce point soulève la question à savoir de quelle façon la capacité de générer davantage de contenu affecte le journalisme et la confiance du public. La priorité devrait être donnée à l'amélioration d'outils de vérification et au maintien de l'implication humaine dans le journalisme pour préserver l'authenticité de l'information.
- **Rôle du contexte dans le journalisme :** Comprendre le

The discussion focused on the interplay between industry practices and governmental oversight, particularly in the context of technology's impact on society. The first discussion providing the viewpoint of industry featured insights from Debbie Widjaja, Chief Product Officer of Zero Gravity and former Programme Manager of Hate Speech and Misinformation at Meta and Simon Staffell, Director of Corporate Affairs at Microsoft UK. Both speakers provided their perspectives on the roles and responsibilities of technology companies in combating misinformation and protecting democratic processes. The discussion underscored the need for ongoing adaptation and collaboration in addressing the challenges posed by misinformation and the misuse of technology.

Debbie Widjaja, drawing from her experience at Meta, between 2016 and 2019, and her current role at Zero Gravity, discussed the evolution of Meta's approach to handling abuse and misinformation on its platform. When Debbie Widjaja joined Meta in 2016, the company had a relatively small Integrity Team, focusing on abuse and misinformation with minimal resources. The election of Donald Trump in November 2016 marked a turning point, with increased scrutiny on Meta's role in spreading misinformation. This led to a surge in resources allocated to integrity efforts, including a significant increase in the team size and investment, although this reduced later.

By 2018, Meta was dealing with 200 million reports per week, managed by 15,000 human reviewers and numerous automated systems. The complexity of moderating content was highlighted, especially when distinguishing between different types of abuse, such as hate speech and misinformation. Debbie Widjaja detailed how policies operated in practice and the



contexte à l'origine de l'information est capital. Non seulement les médias se doivent de s'assurer de l'exactitude des faits mais aussi du narratif et des intentions sous-jacentes.

- **Divulgarion de l'origine du contenu** : Il est essentiel de savoir clairement si le contenu est généré par l'IA ou créé par l'homme pour aider les utilisateurs à s'y retrouver et à faire confiance aux sources d'information.

- **Priorités éditoriales et influence des médias sociaux** : Il est essentiel que les journalistes fassent la juste part entre les tendances médiatiques sociales et le reportage traditionnel pour garantir l'exactitude et la fiabilité des couvertures de presse.

Contrôle et liberté : freins et contreponds pour les entreprises et les gouvernements

Point de vue des entreprises

- **Debbie Widjaja, directrice des produits de Zero Gravity, ancienne responsable du programme de discours de haine et de désinformation, Meta**
- **Simon Staffell, directeur des affaires générales de Microsoft UK**

La discussion a principalement porté sur l'interaction entre les pratiques industrielles et la surveillance des gouvernements, particulièrement dans le contexte de l'impact de la technologie sur la société. Debbie Widjaja, directrice des produits de Zero Gravity et ancienne responsable du programme de discours de haine et de désinformation à Meta et Simon Staffell, directeur des affaires générales de Microsoft UK ont donné leur avis du point de vue de l'industrie lors de la première discussion. Les deux intervenants ont apporté leur éclairage sur les rôles et responsabilités qui reviennent aux entreprises de technologie dans la lutte contre la désinformation et dans la protection des processus démocratiques. Par ailleurs, la discussion a mis l'accent sur la nécessité de s'adapter et de collaborer en permanence en vue de faire face aux défis que posent la désinformation et l'usage abusif de la technologie.

Debbie Widjaja, s'appuyant sur son expérience acquise à Meta, entre 2016 et 2019, et sur la position qu'elle occupe actuellement à Zero Gravity, a fait part de l'évolution de l'approche adoptée par Meta pour gérer les usages abusifs et la désinformation sur ses plateformes.

challenges faced by reviewers in applying nuanced definitions of violations. Misinformation was not initially classified as an abuse category, leading to a focus on removing fake accounts rather than addressing content directly.

Initially, Meta viewed itself as a neutral platform, similar to a telephone. Over time, it became evident that the platform's algorithm-driven content distribution made it more akin to a curated space, affecting its responsibility for the content shared. Debbie Widjaja noted that the monetisation model, which relies on user engagement, indirectly incentivised the spread of sensational content, complicating efforts to combat misinformation.

Simon Staffell discussed Microsoft's approach to governance and industry collaboration in the context of AI, misinformation and electoral integrity.

Microsoft, along with other entities, tracks extensive nation-state threats targeting democratic processes. These threats involve sophisticated cyber-attacks and misinformation campaigns from countries like Russia, China, North Korea and Iran. The rapid advancement of generative AI presents new challenges. AI's ability to produce high-quality deepfakes and misleading content amplifies existing threats and necessitates advanced detection and countermeasures.

Microsoft was involved in the Tech Accord, which commits technology companies to prevent the misuse of AI, particularly in the context of deepfakes and misinformation. The Accord emphasises content provenance, detection and transparency. Microsoft also launched campaigns to educate the public on deepfakes and misinformation, aiming to build resilience and trust.

During recent UK elections, deepfakes and fake content



Au moment de son intégration à Meta en 2016, la société disposait d'une Équipe d'Intégrité relativement réduite, qui se consacrait à gérer les usages abusifs et la désinformation en ne disposant que d'un minimum de ressources. L'élection de Donald Trump en novembre 2016 a changé la donne, le rôle de Meta faisant alors l'objet d'une surveillance accrue en matière de diffusion de désinformation. Ce tournant s'est traduit par une augmentation des ressources attribuées aux efforts d'intégrité, notamment par un élargissement notable de la taille de l'équipe et par une hausse des investissements alloués, bien que ces derniers aient été réduits par la suite.

À l'horizon 2018, Meta gérait chaque semaine 200 millions de communiqués de presse, administrés par 15 000 réviseurs humains et un nombre considérable de systèmes automatisés. La complexité de la modération du contenu a été mise en évidence, surtout s'agissant de faire une distinction entre différents types d'usages abusifs, tels que les discours de haine et de désinformation. Debbie Widjaja a décrit dans le détail le mode opératoire des politiques dans la pratique ainsi que les défis auxquels sont confrontés les réviseurs au moment d'appliquer des définitions de violations nuancées. Au départ, la désinformation n'était pas classée dans la catégorie des usages abusifs, conduisant à se concentrer à supprimer les fake accounts (faux comptes) plutôt que de s'attaquer directement au contenu.

À l'origine Meta se définissait comme étant une plateforme neutre, comparable à un téléphone. Au fil du temps, il est devenu évident que la distribution du contenu pilotée par algorithme de la plateforme l'apparente davantage à un espace organisé, affectant sa responsabilité en ce qui concerne le partage du contenu. Debbie Widjaja a aussi remarqué que le modèle de monétisation, qui repose sur l'engagement de l'utilisateur, encourage indirectement la diffusion de contenu à sensation, compliquant les efforts à déployer pour lutter contre la désinformation.

Simon Staffell s'est exprimé sur l'approche adoptée par Microsoft en matière de gouvernance et de collaboration au sein de l'industrie dans le contexte de l'IA, de la désinformation et de l'intégrité électorale.

Microsoft, au même titre que d'autres entités, surveille les menaces des États-nations qui visent les processus



did not significantly disrupt the electoral process but did contribute to divisive issues. Simon Staffell highlighted the need for continued vigilance and the importance of a multi-layered defence strategy against misinformation.

The discussion highlighted the evolving role of technology companies in managing misinformation and the impact of emerging technologies like AI on democratic processes and raised a number of key issues:

- **Increased Responsibility:** Both Meta and Microsoft acknowledge the growing responsibility of technology companies to combat misinformation and abuse.
- **Operational Challenges:** The scale of content moderation and the complexity of policy implementation are significant challenges faced by tech companies.
- **Collaborative Efforts:** Effective governance requires collaboration between industry, government and civil society to address the multifaceted issues of misinformation and electoral integrity.

démocratiques, ces menaces englobant les cyberattaques et les campagnes de désinformation menées par des pays comme la Russie, la Chine, la Corée du Nord et l'Iran. Les rapides avancées de l'IA générative soulèvent de nouveaux défis. La capacité de l'IA à produire des deepfakes de grande qualité et un contenu trompeur intensifie les menaces existantes et exige de faire appel à une détection et des contre-mesures de haut niveau.

Microsoft a adhéré au Tech Accord qui prévoit que les entreprises technologiques s'engagent à prévenir tout usage abusif de l'IA, plus précisément dans le contexte des deepfakes et de la désinformation. L'Accord met en exergue la source du contenu, la détection et la transparence. Par ailleurs, Microsoft a lancé des campagnes visant à éduquer le public au sujet des deepfakes et de la désinformation, avec pour objectif de renforcer la résilience et la confiance.

Il a été mentionné qu'au cours des récentes élections, qui ont eu lieu au Royaume-Uni, les deepfakes et le faux contenu n'ont pas particulièrement perturbé le processus électoral mais qu'elles ont contribué à se poser des questions controversées. Simon Staffell a souligné la nécessité de continuer à faire preuve de vigilance et l'importance de disposer d'une stratégie de défense multidimensionnelle contre la désinformation.

La discussion a mis en exergue le rôle de plus en plus présent des entreprises de technologie dans la gestion de la désinformation et l'impact des technologies émergentes, telles que l'IA, sur les processus démocratiques. Plusieurs questions clés ont été soulevées :

- **Une plus grande responsabilité :** Meta et Microsoft ont tous deux fait état de la responsabilité croissante qui repose sur les entreprises de technologie dans la lutte contre la désinformation et l'usage abusif de l'information.
- **Défis opérationnels :** L'envergure de la modération du contenu et la complexité de l'application des politiques sont des défis de taille auxquels sont confrontées les entreprises de technologie.
- **Efforts de collaboration :** Une gouvernance efficace nécessite une collaboration entre les entreprises du secteur, les gouvernements et la société civile pour



Government

• Hervé Letoqueux, Chief operating officer, Viginum

Hervé Letoqueux gave an overview of the function and role of Viginum. Viginum is not an intelligence or legal service. Instead, it focuses on detecting foreign digital interference (FDI) and manipulation of information using automated and massive digital means such as trolls or bots. It employs OpenSource Intelligence (OSINT) and proprietary and other custom tools to identify and investigate digital interference campaigns. It operates within a framework of an Ethical and Scientific Independent Committee, working closely with national defence and other governmental departments.

He talked on the challenges of AI and the increasing use of AI, especially generative AI, which on one hand facilitates the production of misleading or fake content (including deep fakes and fake profiles which blur the line between true and false information) but on the other had represent a good opportunity to enhance the detection of manipulated information.

On election monitoring, Hervé Letoqueux said that, while no major interference was observed in recent French elections, Viginum continues to monitor for potential threats. Notable campaigns such as "Shadowplay" involve synthetic voices and fake profiles promoting foreign interests.

On detection tools, he said that Viginum uses algorithms to detect bots, deepfakes and artificially manipulated content. Research papers and tools are developed internally to aid both civil society and internal operations in identifying and mitigating digital threats.

lutter contre les enjeux multidimensionnels de la désinformation et de l'intégrité électorale.

Les gouvernements

• Hervé Letoqueux, directeur opérationnel de Viginum

Hervé Letoqueux a donné un aperçu d'ensemble de la fonction et du rôle de Viginum. Il ne s'agit pas d'un service de renseignements ou juridique ; en fait, Viginum se dédie à détecter les ingérences numériques étrangères (INE) et la manipulation de l'information en faisant appel à des moyens automatisés et numériques considérables du type trolls ou bots. L'agence utilise OpenSource Intelligence (OSINT) et des outils propriétaires et autres outils personnalisés pour identifier et examiner les campagnes d'interférences numériques. Viginum opère dans le cadre d'un Comité éthique et scientifique indépendant, œuvrant étroitement avec le ministère de la Défense nationale et d'autres ministères.

Hervé Letoqueux s'est aussi exprimé sur les défis que posent l'IA et l'usage de plus en plus répandu qu'il en est fait, insistant en particulier sur l'IA générative qui, d'une part, facilite la production de contenu trompeur ou de contenu faux (notamment les deepfakes et les faux profils qui créent une confusion entre la vraie information et la fausse information) mais qui, d'autre part, a représenté une bonne occasion d'améliorer la détection de l'information manipulée.

Concernant la surveillance des élections, Hervé Letoqueux a précisé que même si en France aucune interférence majeure n'a été observée lors des récentes élections Viginum continue de surveiller les menaces potentielles. Des campagnes importantes comme « Shadowplay » mettent en jeu des voies synthétiques et des faux profils faisant la promotion des intérêts étrangers.

Sur la question des outils de détection, il a déclaré que Viginum se sert d'algorithmes pour détecter les bots, les deepfakes et le contenu artificiellement manipulé. À l'heure actuelle, l'agence travaille en interne sur la mise au point de documents et d'outils de recherche destinés à aider à la fois la société civile et les opérations internes à identifier et réduire les menaces numériques.

Il a également fait état de la résilience du public et des efforts qui sont déployés pour accroître la résilience

He also talked of public resilience and how efforts are also made to enhance societal resilience against misinformation through education and awareness programmes.

Legal overview

• Maître Thaima Samman, Cabinet d'avocats Samman

Thaima Samman gave an overview of the regulatory framework in particular regarding contents and algorithms, whether they are built on AI or not.

Within the European Union, the main legislation addressing platforms' contents is the Digital Services Act (DSA) which has opened the door to algorithm framing. The recently adopted AI Act focuses more specifically on AI algorithms, addressing also, but not only, digital content.

The DSA provides a governance framework on addressing illegal content which imposes to platforms to act (to withdraw or maintain content) once informed of a potential illegal content, according to their legal analysis. It is making them, de facto, the referee for online content. In addition, the DSA established specific rules for very large online platforms (over 45 million unique users in the EU), including transparency requirements and risk assessment and mitigation requirements regarding algorithmic systems to address content issues.



sociétale contre la mésinformation par le biais de programmes éducatifs et de sensibilisation.

Le regard juridique

• Maître Thaima Samman, Cabinet d'avocats Samman

Thaima Samman a donné une vue d'ensemble du cadre réglementaire concernant en particulier les contenus et les algorithmes, qu'ils soient ou non créés en faisant appel à l'IA.

Elle a ajouté qu'au sein de l'Union européenne, la principale législation applicable au contenu des plateformes est le règlement sur les services numérique (DSA \square Digital Services Act) qui a ouvert la voie au cadrage des algorithmes, précisant que le Règlement sur l'IA récemment adopté est davantage centré sur les algorithmes d'IA, en prenant également en compte \square mais non uniquement \square le contenu numérique.

Poursuivant sa présentation, elle a souligné que le DSA constitue un cadre de gouvernance qui gère le contenu illégal obligeant les plateformes à réagir (supprimer ou conserver le contenu, p. ex.) une fois qu'elles sont prévenues de la présence d'un contenu potentiel illégal, selon leur analyse juridique. De facto, cela fait d'elles les arbitres du contenu en ligne. De surcroît, le DSA a établi des règles spécifiques auxquelles sont assujetties les grandes plateformes en ligne (plus de 45 millions d'utilisateurs uniques dans l'UE), dont les prescriptions imposées en matière de transparence, d'évaluation et d'atténuation des risques liés aux systèmes algorithmiques pour gérer les problématiques du contenu.

À ce sujet, le Règlement sur l'IA ne s'arrête pas là dans la réglementation des algorithmes ; en fait, à l'origine, cette législation était basée sur une « approche risque ». Certains cas d'utilisation sont interdits (les systèmes de notation sociale, p. ex.), alors que d'autres cas sont assujettis à des critères spécifiques (documentation algorithmique et transparence, p. ex.) en fonction du degré de risque qu'ils présentent pour les droits fondamentaux. Avec l'émergence de ChatGPT et de systèmes du même genre, l'UE s'est détachée d'une approche strictement basée sur les risques pour ajouter des obligations spécifiques assujettissant également les fournisseurs de modèles d'IA à usage général (documentation technique et transparence, p. ex.), en imposant un niveau supplémentaire de critères pour

The AI Act goes a step further in regulating algorithms. Originally, this legislation was based on a “risk approach”: Some use cases are banned (e.g., social scoring), whereas others are subject to specific requirements (e.g., algorithmic documentation and transparency) based on the level of risk they pose for fundamental rights. With the emergence of ChatGPT and the likes, the EU has moved away from a pure risk-based approach to add specific obligations also on general purpose AI (GPAI) model providers (e.g. technical documentation and transparency), adding an extra layer of requirements for those that pose “systemic risks” (e.g. model evaluations).

The AI legal framework is being complemented by other EU regulations and directives, in particular regarding AI liability. For Thaima Samman, this is likely just the beginning, given the complexity of the challenges that remain. Government, courts and population in general will have to keep learning and adapt to fully harness the potential of AI technology.

On balancing freedom of expression with the need for regulation, Thaima Samman acknowledged this is a key challenge. Talking about content moderation, she reflected that platforms have obligations to address problematic contents with the risk that holding them accountable for keeping or retrieving online contents is giving them a lot of power. For her, if it is not easily avoidable given the contents mass and their circulation speed, they shouldn't be let alone in taking such decisions and democratic countries will need to be creative to improve their regulation system in the digital era.

In any case, judicial and administrative courts remain today the ultimate decision maker on content online, as legal and natural persons keep their right to go to court to challenge a platform's decision. But she considers that there is a strong need for training to ensure informed decision-making at the right level and the right rhythm for the Courts to be able to do their job.

Thaima Samman also spoke of the regulatory approaches which vary between regions around the world, with the EU having more stringent rules compared to the US or China. This includes differing levels of regulation and enforcement regarding AI and digital content.

ceux qui posent des « risques systémiques » (évaluations des modèles, p. ex.).

D'autres réglementations et directives de l'UE viennent s'ajouter au cadre juridique de l'IA, notamment concernant la responsabilité de l'IA. Thaima Samman considère que nous n'en sommes probablement qu'au début, compte tenu de la complexité des défis toujours en présence. Selon elle, les gouvernements, les tribunaux et la population dans son ensemble devront continuer à se documenter, à apprendre et à s'adapter afin d'exploiter au maximum le potentiel de la technologie d'IA.

Thaima Samman pense que la recherche d'un équilibre entre la liberté d'expression et la nécessité d'imposer une réglementation est un enjeu majeur. S'exprimant sur la modération du contenu, elle a affirmé que les plateformes ont l'obligation de gérer les contenus problématiques, tout en sachant que le fait de les tenir responsable du maintien ou de la suppression des contenus en ligne risque de leur donner un pouvoir considérable. À son avis, si ceci ne peut être facilement évité compte tenu du volume des contenus et de leur rapidité de circulation, elles ne devraient pas être les seules à prendre de telles décisions, ce qui revient à dire que les pays démocratiques devront faire preuve de créativité pour améliorer leur système de réglementation dans le domaine du numérique.

Quoi qu'il arrive, de nos jours les tribunaux judiciaires et administratifs restent habilités à prendre les décisions finales concernant le contenu en ligne, car les personnes physiques et morales conservent leur droit de se présenter auprès d'un tribunal pour contester la décision d'une plateforme. Cela dit, Thaima Samman considère qu'une formation s'impose pour que les tribunaux puissent faire leur travail sur la base de décisions éclairées prises au niveau approprié et au bon rythme.

Thaima Samman s'est également exprimée sur les approches réglementaires qui varient d'une région à l'autre du globe, l'UE s'appuyant sur des règles plus rigoureuses que celles en vigueur aux USA ou en Chine. Ceci comprend différents niveaux de réglementation et de mise en application des lois concernant le contenu numérique et généré par l'IA.

Conclusions

The seminar on "Artificial Intelligence and Information Warfare" raised a number of key issues around managing digital misinformation and AI's impact on society including:

- **Disinformation and Fake News:** The seminar emphasised AI's dual role in both perpetuating and combating fake news. AI systems can enhance media literacy but also enable sophisticated disinformation campaigns.
- **Media Integrity:** The importance of maintaining journalistic standards and verifying information sources in an age where AI can manipulate data and create convincing false narratives.
- **Democratic Processes:** The balance between protecting democratic integrity and respecting personal freedoms was a recurring theme, with discussions focusing on AI's impact on public opinion and democratic engagement.
- **Organisational Changes:** AI's influence on media organisations' internal operations was examined, including how AI tools alter news production and distribution processes.
- **Technological Challenges:** The rapid evolution of AI presents new challenges for detecting and managing misinformation.
- **Regulatory Efforts:** The EU's legislative framework, including the DSA and AI Act, aims to balance freedom and control, although challenges remain in fully addressing new technological advancements.
- **Collaboration and Resilience:** Both governmental and industry efforts are essential in developing tools, fostering public resilience and ensuring effective regulation.

Hervé Mariton, Chair of the Conseil franco-britannique, concluded the seminar by thanking all of the participants and attendees. He acknowledged the need for continued dialogue and collaboration between academia, industry, media and government. The goal is to foster a comprehensive understanding of AI's role in information warfare and to develop effective strategies for managing its challenges.

Le séminaire sur « l'Intelligence artificielle et la Guerre de l'information » a soulevé un certain nombre de points majeurs concernant la gestion de la désinformation numérique et les conséquences de l'IA sur la société, notamment :

- **La désinformation et les fake news :** Le séminaire a mis l'accent sur le double rôle que joue l'IA en diffusant les fake news et en les combattant. Les systèmes d'IA ont la capacité d'améliorer la compétence médiatique, tout en menant des campagnes de désinformation sophistiquées.
- **L'intégrité médiatique :** L'importance de maintenir les standards journalistiques et de vérifier les sources des informations dans une époque où l'IA peut manipuler les données et créer de faux narratifs fort convaincants.
- **Les processus démocratiques :** L'équilibre entre la protection de l'intégrité démocratique et le respect de la liberté personnelle a été un thème récurrent, les discussions se concentrant sur les conséquences que peut avoir l'IA sur l'opinion publique et sur l'engagement démocratique.
- **Les changements organisationnels :** L'influence de l'IA sur les opérations internes des organisations médiatiques a été examinée, notamment la façon dont les outils d'IA modifient les processus de production et de diffusion des informations.
- **Les enjeux technologiques :** L'évolution accélérée de l'IA présente de nouveaux défis en matière de détection et de gestion de la désinformation.
- **Les mesures réglementaires :** Le cadre législatif de l'UE, notamment le DSA et le Règlement sur l'IA, a pour objectif d'harmoniser liberté et contrôle, bien que la gestion des nouvelles avancées technologiques pose encore bon nombre de défis.
- **La collaboration et la résilience :** Les efforts que déploient les gouvernements et les entreprises du secteur sont essentiels pour développer des outils, pour soutenir la résilience du public et pour garantir l'instauration d'une réglementation efficace.

Hervé Mariton, Président du Conseil franco-britannique, a déclaré clos le séminaire en remerciant tous les participants et toutes les personnes présentes. Il a évoqué la nécessité d'un maintien de dialogue et d'une collaboration continue entre les académiciens, les entreprises du secteur, les médias et les gouvernements. L'objectif visé est de promouvoir une compréhension élargie du rôle de l'IA dans la guerre de l'information et de formuler des stratégies efficaces pour pouvoir en gérer les divers enjeux.

Franco-British Council

Le Conseil franco-britannique



The Franco-British Council is an independent organisation based in Paris and London. Our mission is to reflect the wide relationships between the two nations and promote constructive dialogue for enhanced future collaboration.

The Council was founded in 1972 on the joint initiative of President Georges Pompidou and Prime Minister Edward Heath, in a context of a developing Europe and of an increasingly globalised community. Since then, it has dedicated itself to the promotion of a better mutual understanding and the development of joint action through bringing together leading representatives of the worlds of culture, politics, defence, science, education and business.

More than 40 years after its creation, in the current geopolitical context and Brexit, the Council's mission resonates even more. Now more than ever, strong ties need to continue to be built between both nations, providing even more opportunities for people to engage at all levels.

The Council is a non-governmental organisation with charitable status. It has a bilateral governance with members and trustees in both countries. It receives funding from the British and French governments and generous partners.

Le Conseil Franco-Britannique est une organisation indépendante basée à Paris et à Londres. Notre mission est de présenter l'étendue des relations existant entre les deux nations et de promouvoir un dialogue constructif en vue d'une coopération future renforcée.

Le Conseil a été créé en 1972 sur une initiative conjointe du Président Georges Pompidou et du Premier ministre Edward Heath au moment de l'adhésion du Royaume-Uni à l'Union européenne et dans un contexte de mondialisation grandissante. Depuis, il s'est consacré à la promotion d'une meilleure compréhension entre les deux pays et au développement d'actions communes en rassemblant des personnalités des secteurs culturels, politiques, économiques, de la défense, de l'éducation, des affaires et de la communauté scientifique.

Plus de 40 ans après sa création, dans le contexte géopolitique actuel et à l'heure du Brexit, la mission du Conseil résonne d'actualité. Aujourd'hui, plus que jamais, il est nécessaire de continuer à forger des liens solides entre les deux nations et de donner à leurs citoyens encore plus d'opportunités d'échanger à tous les niveaux.

Le Conseil Franco-Britannique est une organisation non gouvernementale à vocation caritative. Il a une gouvernance bilatérale avec des membres et des administrateurs dans les deux pays. Il est financé par les gouvernements britannique et français et est généreusement soutenu par de nombreux partenaires privés.