



# Annexe 1



## Glossaire

## **Amplification algorithmique**

Processus par lequel les systèmes de recommandation, les algorithmes de classement ou les outils de curation automatisés des plateformes numériques font remonter et augmentent la visibilité de certains contenus par rapport à d'autres. Cette amplification est généralement fondée sur des métriques d'engagement (par exemple, les clics, les partages et la durée de visionnage), la personnalisation ou les incitations de la plateforme, plutôt que sur l'exactitude ou la fiabilité de l'information.

## **Faux contenu généré par l'IA**

Contenu audio, textuel ou visuel, produit par l'IA générative (voir ci-dessous), qui dépeint des personnes ou des événements d'une manière malveillante ou trompeuse qui diffère de la réalité. Il peut par exemple montrer des personnes faisant des choses qu'elles n'ont pas faites ou disant des choses qu'elles n'ont pas dites, modifier le lieu d'événements réels ou montrer des événements qui n'ont pas eu lieu.

## **« Bouillie d'IA »**

Désigne des médias de faible qualité, notamment des textes, des images ou des vidéos, créés à l'aide d'outils d'intelligence artificielle générative et caractérisés par une production de masse, un minimum d'efforts et l'absence de réelle substance ou d'intégrité artistique. Le terme a une connotation péjorative similaire à « spam », désignant des contenus superficiels ou banals produits à l'aide de l'IA, souvent uniquement à des fins de volume ou d'engagement.

## **Intelligence artificielle (IA)**

Capacité des machines, des systèmes informatiques ou des logiciels à effectuer des tâches qui requièrent généralement l'intelligence humaine. Ces tâches comprennent l'apprentissage, le raisonnement, la résolution de problèmes, la perception, la compréhension du langage naturel et la prise de décision.

## **Cheapfake**

Médias manipulés ou trompeurs créés à l'aide d'outils d'édition simples, peu coûteux ou facilement disponibles.

## **Guerre cognitive**

Utilisation d'informations, ainsi que d'outils psychologiques et technologiques, en vue d'influencer, de perturber ou de contrôler les perceptions, les croyances et la prise de décision d'individus ou de groupes. Le territoire en jeu est l'esprit humain, le but étant d'éroder la confiance, de semer la confusion et de manipuler les comportements à des fins stratégiques ou politiques.

## **Engagement des communautés et redevabilité (activités CEA)**

Méthode de travail qui reconnaît et valorise tous les membres de la communauté en tant que partenaires égaux, dont les divers besoins, priorités et préférences guident l'action de la Fédération internationale, ce qui comprend notamment une réelle participation de la communauté, une communication ouverte et honnête ainsi que des mécanismes permettant d'écouter les retours d'information et d'y donner suite.

## **Biais de confirmation**

Biais cognitif incitant une personne à rechercher, interpréter et mémoriser des informations d'une manière qui confirme ses croyances préexistantes, en écartant ou en négligeant les preuves contradictoires.

## **Théorie du complot**

Croyance visant à expliquer un événement, un phénomène ou une condition sociale en attribuant la cause aux activités secrètes de groupes puissants agissant de concert à des fins malveillantes. Les théories du complot sont généralement résistantes à la falsification (les preuves qui leur sont opposées ne les réfutent pas, au contraire, elles sont souvent absorbées par la théorie), simplifient des réalités complexes pour en faire des complots volontaires et affirment souvent que le monde est sous l'emprise d'acteurs cachés.

## **Comportement inauthentique coordonné**

Terme introduit par Meta (Facebook) pour décrire les activités organisées de groupes de comptes ou de pages s'unissant pour induire les internautes en erreur sur leur identité et leurs activités. Ce comportement n'est pas caractérisé par le contenu des messages, mais par la nature trompeuse, coordonnée et dissimulée des acteurs et de leurs activités. Il est souvent le fait de faux comptes, de réseaux clandestins ou de mandataires cachés, qui visent à manipuler le débat public, à amplifier certains discours ou à influencer des issues politiques.

## **Modération de contenu**

Processus par lequel les plateformes, les organisations et les institutions surveillent, examinent et gèrent le contenu généré par les utilisateurs pour s'assurer qu'il est conforme aux exigences juridiques, aux normes de la communauté et aux lignes directrices en matière de sécurité. La modération peut comprendre la suppression, l'étiquetage ou le filtrage de publications préjudiciables, illégales ou contraire aux politiques en vigueur, ainsi que la promotion d'informations exactes et fiables.

## **Cybercriminels**

Individus ou groupes organisés qui utilisent des ordinateurs, des réseaux ou des technologies numériques pour commettre des infractions. Leurs activités comprennent l'accès non autorisé aux systèmes, le vol de données, l'usurpation d'identité, les escroqueries financières, les attaques par rançongiciel et la diffusion de logiciels malveillants. Les cybercriminels agissent soit de manière indépendante, soit dans le cadre de réseaux en ligne faiblement connectés, soit au sein de groupes criminels transnationaux hautement organisés.

## **Démystification**

Stratégie réactive qui vise à contrer la mésinformation après qu'elle a commencé à se propager. Pour être efficace, elle doit être rapide, claire et diffusée par des messagers en qui le public a confiance. Une démystification efficace ne se limite pas à réfuter les affirmations, elle doit également expliquer pourquoi elle sont fausses. Des techniques de communication telles que le « sandwich de vérité », qui consiste à fournir des informations exactes au début et à la fin, en intégrant entre deux la réfutation des mensonges, contribuent à réduire le risque de renforcer les mythes.

## **Décontextualisation**

Pratique consistant à présenter des informations, des images ou des citations réelles hors de leur contexte d'origine, de manière à en modifier le sens ou les implications. En supprimant des détails cruciaux tels que le moment, le lieu, la source ou l'intention, la décontextualisation peut induire le public en erreur, déformer des événements et renforcer de faux discours, même sans fabriquer de nouveau contenu.

## **Deepfake**

Médias synthétiques – consistant souvent en contenu vidéo ou audio ou en images – créés à l'aide de techniques d'IA, en particulier l'apprentissage profond, pour manipuler ou générer de manière réaliste des contenus qui représentent des événements ou des personnes d'une manière qui n'est pas conforme à la réalité.

## **Processus de démystification**

Processus d'analyse du fonctionnement de la mésinformation. Il consiste à en expliquer les formes et les fonctions, ainsi que les facteurs psychologiques, sociaux et technologiques qui favorisent sa propagation. Contrairement à la vérification des faits, axée sur la vérification de la véracité ou de la fausseté d'affirmations individuelles, le processus de démystification met en évidence les facteurs structurels, émotionnels, algorithmiques et contextuels qui influencent les informations que le public rencontre,

permettant ainsi de mieux comprendre pour quelles raisons la mésinformation est si persuasive et persistante.

### **Désinformation**

Fausse information créée ou diffusée délibérément dans l'intention de tromper ou de nuire.

### **Divulgaration malveillante d'informations personnelles**

Collecte et publication intentionnelles d'informations privées ou identifiables sur une personne sans son consentement, généralement dans le but de la couvrir de honte, de l'embarrasser, de la harceler, de l'intimider, de la menacer ou de lui causer du tort.

### **Vérification des faits**

Processus de vérification de l'exactitude factuelle des informations, affirmations ou déclarations, généralement dans le domaine du journalisme, de la communication publique ou du contenu en ligne. Il consiste à évaluer systématiquement les données probantes, à consulter des sources crédibles et à apporter des corrections transparentes lorsque des affirmations fausses ou trompeuses sont découvertes. La vérification des faits peut être proactive (anticipation) ou réactive (démystification après la propagation des informations).

### **Fausse dichotomie ou faux dilemme**

Sophisme logique affirmant qu'une situation n'a que deux options ou résultats opposés, alors qu'il existe en réalité d'autres possibilités. En réduisant des questions complexes à un choix simpliste de type « ou / ou », les fausses dichotomies faussent le raisonnement, polarisent le débat et limitent la prise en considération d'autres perspectives. La phrase « Vous êtes soit avec nous, soit contre nous. » en est un exemple.

### **Usine**

Dans l'écosystème de la mésinformation, une « usine » désigne une opération organisée qui produit, amplifie ou manipule systématiquement des contenus à grande échelle. Les **usines de contenu** visent généralement à faire du profit en générant de fausses informations alimentées par les clics, tandis que les **usines à trolls** sont motivées par des considérations politiques ou stratégiques et visent à fausser le débat public et à influencer l'opinion.

### **Bulle de filtres**

État d'isolement intellectuel ou informationnel créé par des algorithmes personnalisés qui sélectionnent le contenu en ligne en fonction du comportement antérieur, des intérêts

et des préférences de l'utilisateur. Dans une bulle de filtres, les utilisateurs sont principalement exposés à des informations et à des points de vue qui renforcent leurs croyances existantes, tandis que les autres points de vue sont réduits ou exclus. Ce phénomène risque de limiter la diversité des informations, de renforcer les biais de confirmation et de contribuer à la polarisation.

### **Manipulation de l'information et ingérence étrangères**

Schéma de comportement qui fait peser une menace ou a le potentiel d'avoir un impact négatif sur les valeurs, les procédures et les processus politiques. Cette activité intentionnelle et coordonnée vise à manipuler. Elle peut être menée par des acteurs étatiques ou non étatiques, ou par leurs intermédiaires, à l'intérieur et à l'extérieur de leur propre territoire. (Également connu en anglais sous l'acronyme « FIMI ».)

### **IA générative**

Classe de systèmes d'IA capables de créer de nouveaux contenus, tels que du texte, des images, du contenu audio ou vidéo, ou du code, sur la base de modèles appris à partir de grands ensembles de données. Contrairement aux systèmes d'IA traditionnels conçus principalement pour la classification ou la prédiction, les modèles d'IA générative sont capables de produire des résultats originaux qui ressemblent à des contenus créés par l'être humain, souvent grâce à des architectures d'apprentissage en profondeur telles que les réseaux antagonistes génératifs ou les grands modèles de langage.

### **Hacktivist**

Un hacktivist est un individu ou un collectif qui utilise des techniques de piratage informatique (hacking) comme forme d'activisme politique ou social. L'hacktivism combine le piratage et l'activisme pour poursuivre certains objectifs, par exemple promouvoir l'information libre, protester contre le comportement de gouvernements ou d'entreprises, ou attirer l'attention sur des causes sociales. Les tactiques utilisées peuvent inclure la défiguration de sites web, les attaques par déni de service, les fuites de données et le piratage de comptes en ligne.

### **Informations préjudiciables**

Informations susceptibles de causer, directement ou indirectement, un préjudice à une personne ou à une entité, ou d'y contribuer.

### **Détournement de hashtags**

Utilisation abusive d'un hashtag en vogue ou d'un hashtag de marque sur les médias

sociaux, de manière réactive ou stratégique. Ce phénomène se produit lorsque des utilisateurs adoptent un hashtag destiné à un sujet ou à une campagne spécifique et s'en servent pour promouvoir un message sensiblement différent ou contradictoire. Il implique souvent la diffusion de spam, le trollage, des objectifs politiques ou des théories du complot. Ce détournement peut perturber la conversation initiale et obscurcir ou déformer le contenu prévu.

### **Discours de haine**

Tout type de communication, qu'il s'agisse d'expression orale ou écrite ou de comportement, constituant une atteinte ou utilisant un langage péjoratif ou discriminatoire à l'égard d'une personne ou d'un groupe en raison de leur identité, en d'autres termes, de l'appartenance religieuse, de l'origine ethnique, de la nationalité, de la race, de la couleur de peau, de l'ascendance, du genre ou d'autres facteurs constitutifs de l'identité.

### **Effet de vérité illusoire ou effet de familiarité**

Biais cognitif selon lequel l'exposition répétée à une affirmation augmente la probabilité qu'elle soit perçue comme vraie, indépendamment de son exactitude factuelle. Cet effet est dû au fait que la familiarité améliore la fluidité du traitement, c'est-à-dire la facilité avec laquelle les informations sont traitées mentalement, ce qui est souvent confondu avec l'exactitude.

### **Infodémie**

Surabondance d'informations, qu'elles soient exactes, fausses ou trompeuses, circulant dans les environnements numériques et physiques pendant une crise, ce qui complique la recherche de sources d'informations dignes de confiance et de conseils fiables.

### **Gestion de l'infodémie**

Réponse systématique à l'infodémie, utilisant des outils tels que l'écoute sociale, la démystification et la démystification préventive, en vue de gérer l'environnement de l'information et d'appuyer la prise de décision en matière de santé publique.

### **Intégrité de l'information**

Ce terme se rapporte à la fiabilité, à l'exactitude et à la fiabilité des informations dans un écosystème. Il englobe les efforts visant à protéger les environnements d'information de la manipulation, de la mésinformation, de la désinformation et d'autres contenus préjudiciables, pour que les utilisateurs puissent accéder à des informations factuelles, contextuelles et crédibles et ainsi prendre des décisions en connaissance de cause.

## Opérations d'information

Activités coordonnées qui utilisent l'information, la communication et les tactiques d'influence pour modifier les perceptions, les décisions et les comportements des publics cibles à l'appui d'objectifs politiques, militaires ou organisationnels. Elles combinent souvent des opérations psychologiques, des activités en ligne, et des campagnes de propagande et de désinformation, en vue d'obtenir des effets stratégiques, en ligne et hors ligne.

## Théorie de l'inoculation psychologique

Selon cette théorie, il est possible d'acquérir une résistance à la persuasion de la même manière que par la vaccination

en étant exposé à un argument sous une forme affaiblie, combiné à des réfutations, on produit des « anticorps mentaux » qui aident à résister à des attaques plus violentes par la suite. Dans le contexte des informations préjudiciables, la théorie de l'inoculation sous-tend les stratégies de démythification préventive, c'est-à-dire les interventions éducatives qui exposent des personnes à des exemples de techniques de manipulation (telles que la désignation d'un bouc émissaire, les théories du complot ou l'appel aux émotions) avant qu'elles ne les rencontrent dans la vie réelle. La recherche montre que l'inoculation psychologique peut améliorer la capacité à détecter et à rejeter la mésinformation dans les domaines de la santé publique, de l'action humanitaire et de la politique.

## Loup solitaire

Individu qui planifie et commet des actes violents ou perturbateurs de manière indépendante, sans le soutien opérationnel direct d'une organisation, même s'il en tire son inspiration idéologique. Dans le domaine de l'information, le terme s'étend également aux personnes qui diffusent seules des contenus extrémistes ou conspirationnistes en ligne, amplifiant des discours préjudiciables sans appartenir à un réseau coordonné.

## Malinformation

Informations authentiques diffusées dans l'intention de nuire, par exemple, en sortant une information vraie de son contexte ou en l'utilisant pour discréditer quelqu'un.

## Mésinformation, désinformation et discours de haine

Terme générique désignant différents types d'informations fausses, trompeuses ou préjudiciables. Il couvre la mésinformation (fausses informations partagées sans intention de nuire), la

désinformation (fausses informations créées ou diffusées délibérément en vue de tromper ou de nuire) et les discours de haine. (Également connu en anglais sous l'acronyme « MDH ».)

## Mésinformation

Informations fausses ou inexacts diffusées sans intention de tromper.

## Renseignement de sources ouvertes

Collecte, analyse et utilisation d'informations publiques et légalement accessibles. Les sources comprennent les médias traditionnels, les publications universitaires, les rapports gouvernementaux, les médias sociaux, les images satellite et d'autres données ouvertes. Les gouvernements, les forces de l'ordre, les journalistes, les chercheurs et la société civile se fondent largement sur le renseignement de sources ouvertes pour prendre des décisions, mener des enquêtes et s'informer sur une situation.

## Manipulation politique

Utilisation délibérée de tactiques trompeuses, coercitives ou manipulatoires par des acteurs politiques pour façonner les perceptions, les attitudes et les comportements du public de manière à servir leurs propres intérêts, souvent au détriment d'une prise de décision démocratique éclairée. Elle peut consister à contrôler les discours, à diffuser de fausses informations, à exploiter les émotions, à étouffer les voix dissidentes ou à mener des manœuvres secrètes visant à obtenir le consentement.

## Démythification préventive

Méthode proactive visant à renforcer la résilience du public face aux fausses informations par l'exposition - préalable - à des exemples affaiblis de stratégies trompeuses ou de fausses affirmations, accompagnés de réfutations. S'appuyant sur la théorie de l'inoculation psychologique, elle prépare les individus à mieux reconnaître les contenus manipulateurs et à y résister lorsqu'ils les rencontrent. Il existe deux types de **démythification préventive: l'inoculation psychologique axée sur les idées** vise à contrer les scénarios plus généraux ou les cadres persuasifs susceptibles d'être utilisés pour induire en erreur, tandis que **l'inoculation psychologique axée sur les techniques** se concentre sur les techniques de manipulation plutôt que sur les scénarios, par exemple en enseignant à des personnes comment reconnaître des tactiques telles que la manipulation émotionnelle, la décontextualisation des données ou la désignation de boucs émissaires, afin qu'elles soient mieux équipées pour y résister lorsqu'elles les rencontrent. Ensemble, ces éléments renforcent l'« immunité mentale », c'est-à-dire la résilience cognitive qui aide le public à repérer les fausses

informations, à y résister et à les écarter avant qu'elles ne se fixent.

## Prévention

Stratégies proactives qui anticipent la création d'affirmations trompeuses, de discours ou d'informations préjudiciables avant qu'ils ne se propagent, en vue de réduire leur impact.

## Propagande

Utilisation systématique et délibérée de la communication - symboles, messages ou médias - pour façonner les perceptions, manipuler les processus cognitifs et orienter les comportements afin d'atteindre les objectifs d'un commanditaire. Elle est généralement sélective dans sa présentation des faits et fait souvent appel aux émotions plutôt qu'à l'analyse rationnelle.

## Intermédiaire

Acteur - organisation, média ou individu - utilisé par un État ou un groupe pour mener indirectement des activités pour son profit. Dans les opérations d'information, les intermédiaires amplifient ou propagent des discours tout en dissimulant la véritable origine ou le commanditaire du contenu. Les intermédiaires agissent parfois en connaissance de cause (acteurs alignés) ou à leur insu (« idiots utiles »), permettant aux véritables orchestrateurs de démentir toute implication de manière plausible.

## Résilience

Capacité des individus, des communautés, des organisations ou des pays exposés à des catastrophes, à des crises et aux vulnérabilités qui les sous-tendent, à anticiper les effets, à s'y préparer, à en réduire l'impact, à les surmonter et à se remettre des chocs et des pressions sans compromettre leurs perspectives de développement à long terme.

## Rumeurs

Informations non vérifiées qui se propagent rapidement au sein des communautés. Les rumeurs peuvent s'avérer vraies, fausses ou partiellement vraies par la suite, mais quoi qu'il en soit, elles risquent de semer la confusion ou susciter la méfiance en se propageant.

## Confusion entre satire et réalité

Situations où un contenu parodique ou satirique est pris pour une information authentique. Lorsqu'un contenu satirique (même, article parodique ou sketch comique) circule sans son contexte humoristique ou ironique d'origine, le public risque de l'interpréter littéralement et de le diffuser en pensant qu'il s'agit d'une information réelle.

### **Fraudes et arnaques (en matière d'aide)**

Actes visant délibérément à tromper en vue d'obtenir de l'argent, des ressources ou des informations sensibles, qui ciblent souvent des populations vulnérables en situation de crise. Ils peuvent comprendre de fausses offres d'aide, des collectes de fonds frauduleuses ou des fausses déclarations propagées par des personnes ou des organisations. La désinformation amplifie souvent ces stratagèmes, créant la confusion, un sentiment d'urgence ou la méfiance parmi les communautés et les donateurs.

### **Écoute sociale**

Processus systématique de surveillance, de collecte et d'analyse des conversations, des tendances et des mentions sur les plateformes de médias numériques et sociaux visant à obtenir des informations sur les perceptions, les comportements et les discours qui circulent. Contrairement aux méthodes de base de surveillance des médias sociaux, qui permettent de suivre des métriques (« j'aime », partages

et mentions), l'écoute sociale met l'accent sur l'analyse qualitative du contexte, des sentiments et des questions émergentes.

### **Faux nez**

Fausse identité virtuelle créée et contrôlée par une personne ou une organisation en vue de tromper.

### **Contenu synthétique**

Texte, images, contenus audio ou vidéo, ou autres médias générés, manipulés ou modifiés à l'aide de l'IA ou de systèmes algorithmiques, plutôt que d'être directement créés ou enregistrés par des humains.

### **Troll**

Individu qui perturbe délibérément les conversations, les communautés ou les plateformes en ligne en publiant des contenus incendiaires, hors sujet ou trompeurs, dans l'intention de

provoquer, d'induire en erreur ou de manipuler les autres utilisateurs. Dans le cadre de la désinformation, les trolls opèrent soit seuls, soit dans le cadre de campagnes organisées (« usines à trolls »), amplifiant les discours qui alimentent les divisions ou propageant de fausses informations à des fins politiques, financières ou idéologiques.

### **Manipulation visuelle**

Modification ou utilisation à mauvais escient de photos et de vidéos - au moyen de l'édition, du recadrage ou de l'ajout de fausses légendes - pour tromper ou faire croire quelque chose qui n'est pas vrai.

### **Tatouage numérique**

Technique utilisée pour intégrer une marque (visible ou invisible) lors de la création de contenu numérique, tel que du texte, des images, du contenu audio ou vidéo, ou des médias générés par l'IA, afin d'en indiquer l'origine, l'authenticité ou la propriété.