

**Faculté des sciences économiques,
sociales, politiques et de communication**

Les réactions suscitées par les deepfakes chez les utilisateurs des réseaux socio-numériques

Répercussions sur la réception, transmission, interprétation
et la crédibilité des messages

Auteur : Laura Van de Sande
Promoteur(s) : Suzanne Kieffer

Année académique 2024-2025
Master [120] en communication, à finalité spécialisée : communication
stratégique des organisations

Remerciements

Je tiens tout d'abord à exprimer ma profonde gratitude à ma promotrice, Madame Suzanne Kieffer, pour son expertise, ses précieux conseils et sa disponibilité tout au long de ces deux dernières années. Son accompagnement et son soutien ont fortement contribué à l'élaboration de ce mémoire.

Je remercie également toutes les personnes qui ont participé à cette étude, en particulier celles qui ont accepté de prendre part aux entretiens individuels. Leur implication et leur honnêteté ont été essentielles à l'aboutissement de cette recherche. Sans leur contribution, cette recherche n'aurait pu voir le jour.

Enfin, mes remerciements s'adressent tout particulièrement à mes amis et ma famille pour leur soutien quotidien. Je suis également reconnaissante envers mes collègues étudiants, avec qui l'entraide, les échanges et les partages d'expériences ont grandement enrichi mon travail.

Déclaration sur l'honneur concernant l'IA

Dans le cadre de l'élaboration de l'ensemble de mon mémoire, j'ai eu recours à l'intelligence artificielle tout en respectant les règles de la Faculté Espo. ChatGPT et DeepLWrite m'ont notamment servi d'assistant à la rédaction et à la structuration de mes idées. Ils m'ont permis d'améliorer la qualité et la fluidité de la narration, d'identifier les redondances, de vérifier l'orthographe et la grammaire, ainsi que de traduire quelques passages de texte lorsqu'il était nécessaire.

J'ai également eu recours à des outils comme Image to Prompt et Stable Diffusion 3 Medium pour générer les images constituant mon matériel de recherche, dans le cadre du test d'awareness. Enfin, j'ai utilisé Turbo Scribe AI pour la retranscription des entretiens, tout en effectuant une relecture attentive afin d'assurer l'exactitude des propos rapportés.

La déclaration sur l'honneur concernant l'IA, signée, se retrouve dans les annexes ([cfr annexes 9.1](#)).

Table des matières :

1. Introduction.....	3
1.1. Problématique.....	3
1.2. Enjeux soulevés.....	4
1.3. Objectif et question de recherche.....	5
1.4. Lacunes dans la littérature.....	7
2. Etat de l'art.....	8
2.1. Le deepfake.....	8
2.2. Deepfakes contexte historique et numérique.....	9
2.3. Deepfakes : le pouvoir du visuel.....	12
2.4. Deepfakes : désinformation et manipulation.....	13
2.5. Responsabilité des plateformes et utilisateurs.....	18
3. Méthodologie.....	19
3.1. Objectifs de la méthodologie.....	19
3.2. Méthodes.....	20
3.2.1. Etude de réception via une enquête qualitative.....	20
3.2.1.1. Guide d'entretien.....	22
3.2.2. Test d'awareness sur les contenus.....	26
3.2.2.1. Création du matériel.....	28
3.3. Échantillon de la population et recrutement.....	32
3.4. Script de passation.....	34
4. Analyse et interprétation des résultats :.....	35
4.1. Attitudes cognitives : connaissance du phénomène et de son ampleur.....	36
4.2. Attitudes affectives : de l'émerveillement technologique au doute.....	38
4.3. Attitudes conatives : rapport à la responsabilité, évolution des pratiques et vérification des sources.....	41
4.4. Capacité à discerner le vrai du faux : de la théorie à la pratique.....	44
5. Limites de la recherche :.....	49

6. Discussion et conclusion :	50
7. Bibliographie	54
8. Annexes	59
8.1. Déclaration sur l'honneur concernant l'IA	59
8.2. Première ébauche - Guide d'entretien	60
8.3. Test de la grille de questions	62
8.4. Test du matériel pour le test d'awareness	70
8.5. Premières observations	72
8.6. Formulaire de consentement	73
8.7. Retranscription des entretiens	74

1. Introduction

1.1. Problématique

Dans notre société actuelle, marquée par une transformation rapide et continue, l'ère du numérique a engendré une multitude d'innovations technologiques. Ces innovations technologiques transforment peu à peu notre manière de communiquer, d'interagir avec notre environnement, d'accéder à l'information et même de percevoir le monde qui nous entoure. L'omniprésence des outils numériques a permis l'émergence de nouvelles pratiques médiatiques et sociales, qui redéfinissent continuellement notre rapport à la réalité (Langguth, J. 2021).

Parmi ces avancées numériques, les technologies fondées sur l'intelligence artificielle occupent une place centrale. Les deepfakes, en particulier, représentent à la fois un avenir prometteur en matière de création et de divertissement, mais également un futur inquiétant (De Ruiter, A. 2021).

Grâce à l'apprentissage automatique, il est désormais possible de générer des vidéos, des images ou des enregistrements audio dans lesquels une personne semble dire ou faire quelque chose qu'elle n'a jamais dit ou fait (De Ruiter, A. 2021). Le réalisme de ces contenus rend leur détection particulièrement difficile, même pour des utilisateurs avertis. Il en devient difficile de distinguer le vrai du faux (Bismuth, C. 2023). Les deepfakes et leur réalisme facilitent la diffusion de fausses informations, la production de contenus diffamatoires, ou encore l'usurpation d'identité à des fins malveillantes (Fabre, M. 2022).

La problématique actuelle que posent les deepfakes réside principalement dans le manque de conscience et d'esprit critique des audiences face à l'existence de ces fausses informations. En effet, nombreux sont les utilisateurs des réseaux socio-numériques (RSN) à ne pas être suffisamment informés sur l'existence, le fonctionnement et les risques liés aux deepfakes, ce qui les rend vulnérables à la manipulation et à la désinformation.

De plus, ce manque de conscience et d'esprit critique entraîne des bouleversements majeurs dans le processus d'information et de communication. Il contribue à fragiliser le processus de production et de diffusion de l'information, brouille les repères entre réalité et fiction, et érode la confiance publique envers l'information (De Ruiter, A. 2021).

Dans ce contexte, la lutte contre la désinformation requiert une véritable éducation aux médias (Kazadi, J.D. 2024).

1.2. Enjeux soulevés

Il est ici question de répondre aux inquiétudes du secteur de la communication et d'identifier comment celui-ci peut être affecté par la prolifération des deepfakes (Langguth, J. 2021). Selon mes recherches, de nombreux questionnements sont exprimés par les professionnels de ce secteur, notamment concernant l'efficacité de la transmission d'informations via les réseaux socio-numériques, où les contenus générés par les intelligences artificielles brouillent les repères entre le vrai et le faux. Le réalisme troublant des deepfakes fragilise les mécanismes traditionnels de vérification de l'information, ce qui suscite des inquiétudes quant à la fiabilité des contenus diffusés (Frau-Meigs, D. 2019).

Par l'élaboration de mon mémoire, je vise à apporter des réponses et à analyser, non seulement les réactions des audiences face à ce phénomène, mais également les conséquences des deepfakes sur la transmission, la réception et l'interprétation des informations. Ce travail de recherche a également pour but d'éveiller les consciences auprès des utilisateurs des réseaux socio-numériques, en soulignant l'importance cruciale de vérifier la véracité d'une information, de croiser les sources et de ne pas accorder une confiance aveugle à ce qui circule en ligne.

Ce travail vise ainsi à mettre en lumière les enjeux liés à la confiance dans les contenus numériques à une époque où l'image elle-même, autrefois considérée comme preuve, peut être générée de toutes pièces (Meskin, A. 2008). Il contribue à proposer des pistes de réflexion pour renforcer l'esprit critique des utilisateurs en s'inscrivant dans une volonté de contribuer à une meilleure éducation aux médias et à l'information, en sensibilisant aux risques liés à la désinformation.

1.3. Objectif et question de recherche

Dans le but de répondre aux inquiétudes du secteur, ma question de recherche se compose autour de deux axes.

QR1 : Quelles réactions suscitent les deepfakes chez les utilisateurs des réseaux socio-numériques (RSN) âgés entre 16 et 60 ans ?

QR2 : Quelles en sont les répercussions sur la réception, la transmission, l'interprétation et la crédibilité des messages ?

À travers cette question de recherche, je vise à comprendre quelles réactions suscitent les deepfakes chez les utilisateurs des réseaux socio-numériques, particulièrement en ce qui concerne la manière dont ils perçoivent, partagent et interprètent ces contenus. Cette question de recherche prend en compte l'enjeu majeur du manque de conscience et d'esprit critique des utilisateurs face à ce phénomène, remettant en cause leur capacité à discerner la véracité des informations. Elle explore non seulement l'effet immédiat des deepfakes sur les individus, mais également les effets à plus long terme sur la façon dont l'information est traitée dans l'écosystème des réseaux socio-numériques.

Dans cette démarche, il est d'abord essentiel d'évaluer le niveau de connaissance des utilisateurs par rapport aux deepfakes, pour ensuite analyser leurs réactions. Il ne s'agit pas uniquement d'observer ce que les individus font (réactions visibles/comportements) mais aussi de comprendre ce qu'ils en pensent et ressentent intérieurement, plus communément appelées attitudes. Une attitude se définit comme une prédisposition mentale, une pensée qui ne se manifeste pas nécessairement de manière visible, mais qui oriente les réactions futures (Ifrée, 2023). Une attitude se compose de trois dimensions (Mercator & Dunod, s.d).

Dimension	Signification
Cognitive	Renvoie aux croyances, connaissances et représentations des individus à propos des deepfakes. Notamment leur niveau de connaissance, leur compréhension du sujet et leur perception des conséquences potentielles.
Affective	Concerne les émotions suscitées par l'exposition à ces contenus et qui influencent la manière dont l'information est perçue (méfiance, curiosité, etc).
Conative	Fait référence aux comportements ou intentions d'action, comme la tendance à vérifier ses sources, à partager ou non du contenu, etc.

Afin de répondre à la question de recherche portant sur les réactions suscitées par les deepfakes chez les utilisateurs des réseaux socio-numériques, ainsi que leurs répercussions sur la réception, la transmission, l'interprétation et la crédibilité des messages, cette étude repose sur une méthodologie qualitative articulée autour de deux dispositifs complémentaires. D'une part, je mène une étude de réception au moyen d'entretiens individuels semi-directifs, permettant d'explorer les attitudes (ou réactions non visibles) des utilisateurs, à travers 3 dimensions : cognitive, affective et conative. D'autre part, j'entreprends un test d'awareness à la suite des entretiens, afin de confronter les utilisateurs à des contenus visuels truqués et authentiques pour évaluer leur capacité réelle à identifier des deepfakes. Cette double approche méthodologique permet alors de mettre en lumière les écarts potentiels entre les réactions déclarées et les comportements réels des utilisateurs.

1.4. Lacunes dans la littérature

Les deepfakes représentent une évolution inquiétante des technologies de manipulation numérique, où des contenus audio, vidéo ou image sont modifiés de manière extrêmement réaliste grâce à des algorithmes d'intelligence artificielle (De Ruiter, A. 2021). Cette manipulation visuelle sophistiquée, rendue accessible à un large public grâce aux outils numériques, brouille les frontières entre réalité et fiction, posant un défi majeur à la vérification des faits et à la confiance dans l'information (Bismuth, C. 2023). Leur prolifération a des répercussions importantes sur la désinformation, avec des implications particulièrement préoccupantes dans des contextes sensibles comme les élections ou la politique.

Dans ce climat numérique saturé d'images et de vidéos, les deepfakes alimentent un phénomène d'hyper-trucage qui défie notre capacité à distinguer le vrai du faux. Le rôle des réseaux socio-numériques dans la diffusion rapide de ces contenus, couplé à des algorithmes qui favorisent leur propagation, aggrave le problème. Les utilisateurs, souvent mal équipés pour discerner ces manipulations, deviennent à la fois récepteurs et créateurs de fausses informations, amplifiant ainsi le phénomène de désinformation (Stocker, C. 2020). Les deepfakes exploitent également des biais cognitifs comme le biais de vérité illusoire, où la répétition de fausses informations finit par les rendre crédibles aux yeux des individus (Nelson, J. 2022).

Ce qui manque encore à notre compréhension du phénomène des deepfakes, c'est une évaluation approfondie non seulement de l'état de conscience des utilisateurs des réseaux socio-numériques vis-à-vis de leur existence et de leurs répercussions, mais également de leur aptitude à identifier un contenu truqué. Beaucoup d'entre eux ignorent encore leur fonctionnement, ce qui peut limiter leur aptitude à les détecter. Bien que des solutions techniques, comme les outils de détection basés sur l'IA, émergent, elles ne sont pas suffisantes (Kazadi, J.D. 2024). L'éducation des utilisateurs sur les dangers des deepfakes et le développement de compétences en vérification de l'information restent essentiels pour lutter efficacement contre ceux-ci. Sans une prise de conscience collective, le risque de manipulation et de polarisation des opinions persistera, menaçant la confiance dans les médias et la stabilité démocratique (Simon, FM et al. 2023).

2. Etat de l'art

2.1. Le deepfake

Depuis plusieurs années, notre monde est devenu profondément numérique, transformant nos sociétés, nos habitudes et notre rapport à l'information. Nous évoluons désormais dans un univers où les médias synthétiques dominent, rendant la distinction entre vérité et mensonge, réalité et illusion, plus complexe que jamais (Bismuth, C. 2023).

Mais que sont réellement les deepfakes ?

Les deepfakes sont des contenus numériques ayant subi une manipulation ou une synthétisation, altérant leur apparence d'origine. En d'autres mots, ce sont des contenus visuels ou audios créés de toutes pièces par des algorithmes d'intelligence artificielle (Bismuth, C. 2023). Ces algorithmes s'appuient sur le deep learning, permettant à ces derniers de produire des images et des vidéos extrêmement réalistes, en se basant sur une grande quantité de données existantes. Le deep learning reproduit les mécanismes du cerveau humain permettant de simuler et d'animer les traits du visage d'une personne, ou encore d'animer les éléments d'un décor existant à travers l'utilisation de programmes IA. Cette aptitude rend alors la distinction entre la réalité et la manipulation de plus en plus difficile. Cela remet notamment en question "la vérité dans l'image numérique", la confiance dans l'information et conduit à des hyper-trucages (Fabre, M. 2022, p.2).

Ces hyper-trucages sont des manipulations visuelles si perfectionnées qu'il en devient presque impossible de déceler le vrai du faux. En effet, contrairement aux montages rudimentaires accessibles à tous, que l'on appelle cheap fakes ou encore shallow fakes, les deepfakes utilisent des technologies bien plus avancées et sophistiquées (Robichaud-Durand, S. 2023).

Par ailleurs, c'est à partir des termes "deep learning" et "hyper-trucage" qu'est né le terme de deepfake. C'est en combinant le terme "fake" avec le terme "deep", provenant de deep learning et de trucage, que sont nés les deepfakes. Autrement dit, l'arrivée de fausses nouvelles créées à l'aide de l'utilisation d'outils sophistiqués (Deschamps, C. 2021, p.2).

2.2. Deepfakes contexte historique et numérique

Comme les fake news, les deepfakes sont à l'origine de la désinformation. Ils s'inscrivent dans une lignée de fausses nouvelles. Cette pratique ancienne prend de nouvelles formes avec l'avènement du numérique notamment. En effet, la diffusion de fausses informations n'est pas une pratique nouvelle. Au contraire, ces formes de désinformation ont toujours existé, que ce soit sous la forme de rumeurs, de légendes urbaines ou encore de propagande (Dauphin, F. 2019).

Depuis l'invention de la photographie au 20^{ème} siècle, les médias visuels bénéficient d'une forte crédibilité auprès du grand public et ont largement été utilisés comme preuves ou instruments de propagande. Par conséquent, cette confiance encourage les incitations à créer des éléments visuels falsifiés (Meskin, A. 2008).

En effet, dès les années 1930, on retrouve des exemples de manipulations visuelles à visée politique. L'un des exemples les plus célèbres à cette époque est celui d'Alexander Malchenko (cfr figure 1), effacé des photos officielles après son exécution en 1930 (Dickerman, L. 2000).



Figure 1 : Alexander Malchenko effacé des photos après son exécution en 1930

¹ Figure 1 : Langguth, J., Pogorelov, K., Brenner, S., Filkuková, P., & Schroeder, D. T. (2021). Don't Trust Your Eyes : Image Manipulation in the Age of DeepFakes. *Frontiers In Communication*. <https://www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2021.632317/full#B18>

Jusqu'à récemment, la manipulation photo ou vidéo demandait beaucoup d'effort et était coûteuse, alors réservée à des experts. Cependant, avec le développement des technologies numériques, ces manipulations sont devenues plus simples, plus rapides et surtout plus accessibles (Pierson, 1999). Dès les années 1990, des logiciels comme Adobe Photoshop ont permis au grand public de retoucher des images, donnant naissance à une véritable culture de l'image modifiée. Le terme "photoshoper" est même entré dans le langage courant, démonstration d'une banalisation de cette pratique, devenue commune (Langguth, J. 2021).

Ce qui a marqué un gros changement est l'arrivée des intelligences artificielles. Grâce à des outils performants, il est désormais possible de modifier ou de remplacer le visage d'une personne, dans une image ou vidéo, en quelques clics seulement. Ces modifications grâce à l'intelligence artificielle, souvent imperceptibles à l'œil nu, rendent les manipulations d'autant plus efficaces (Langguth, J. 2021). Or historiquement, la vidéo était considérée comme une preuve difficile à contester, notamment parce que sa falsification demandait des moyens techniques importants (Meskin, A. 2008).

Aujourd'hui, le deepfake donne l'illusion d'être vrai. Il simule un visage, une image vide de sens que l'on peut charger de n'importe quelle histoire ou intention. Pourtant, sans éléments de contexte supplémentaires, il devrait être difficile d'y croire. Cependant, comme le disait le philosophe Peirce : "le vague est réel" (Fabre, M. 2022, p.11). Le sens reste toujours flou. Le deepfake utilise justement cette idée pour manipuler notre compréhension de l'image.

Alors qu'habituellement, l'interprétation se construit progressivement. Le deepfake s'impose comme une certitude : un visage qui semble réel, bloquant alors toute réflexion critique (cfr figure 2). Ce flou n'est pas accidentel, il est tactique. En refusant toute interprétation claire de son intention, le deepfake retient l'information et prend le pouvoir sur son interprète (Fabre, M. 2022).

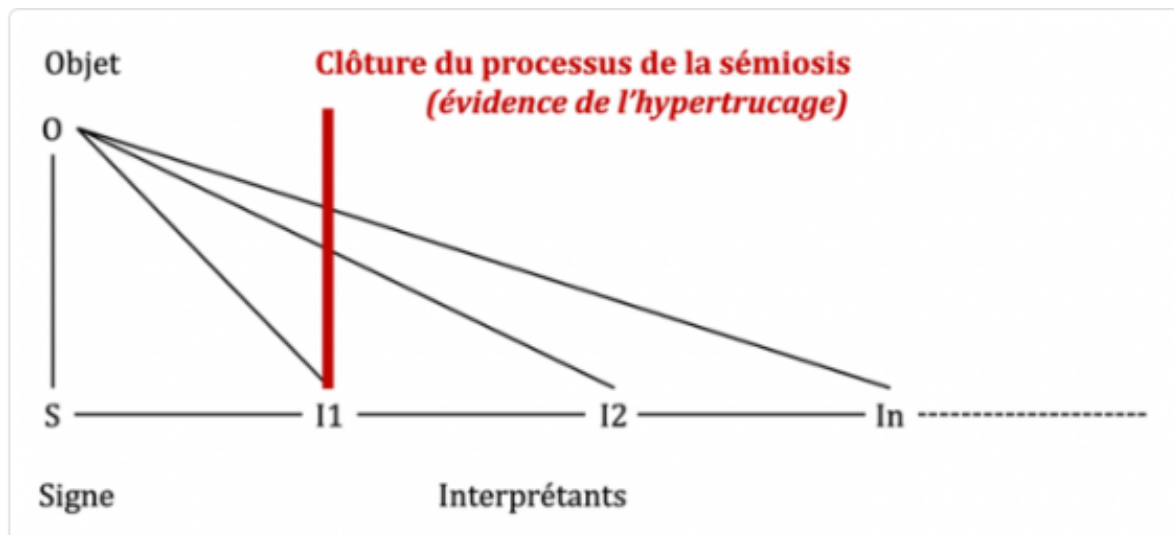


Figure 2 : Clôture du processus de la réflexion critique

Ce qui distingue fondamentalement les deepfakes des manipulations visuelles précédentes, ce n'est pas uniquement leur qualité, mais surtout leur accessibilité. Alors qu'auparavant la production d'images truquées demandait une certaine expertise technique et des outils professionnels, il suffit aujourd'hui d'un ordinateur portable, de quelques tutoriels et d'un ensemble d'images pour produire une vidéo truquée crédible (Beridze, I - Butcher, J. 2019).

La capacité à manipuler le réel n'est donc plus entre les mains des experts ou des régimes autoritaires, mais potentiellement entre celles de n'importe quel individu connecté. Les intelligences artificielles offrent à qui sait les exploiter, une technologie sophistiquée permettant de manipuler les utilisateurs des réseaux socio-numériques. Utilisées par des personnes mal intentionnées, elles offrent un moyen pour des opérations de manipulation d'opinion à grande échelle (Chavalarias, D. 2018). C'est pourquoi leur utilisation préoccupe, notamment dans les contextes électoraux ou géopolitiques sensibles (Langguth, J. 2021).

De plus, le volume de contenus visuels produits chaque jour dans le monde ne cesse d'augmenter. L'enjeu ne réside alors plus seulement dans la création de deepfakes, mais aussi dans leur détection rapide, au milieu d'une masse d'informations visuelles. Cette surcharge informationnelle complique considérablement la vérification de l'authenticité de chaque contenu, facilitant la diffusion et l'ampleur des manipulations (Langguth, J. 2021).

² Figure 2 : Fabre, M. (2022). Urdoxa, le deepfake d'information comme usage tactique du vague. Interfaces numériques, 11(2). <https://doi.org/10.25965/interfaces-numeriques.4863>

Selon une étude menée par ID Crypt Global, environ 1,5 milliard de messages mensongers seraient créés ou partagés chaque jour sur les réseaux socio-numériques. Parmi les plateformes concernées, Facebook en serait la plus touchée, comptabilisant plus d'un milliard de publications de désinformation quotidiennes pour 3 milliards de personnes actives sur la plateforme (Le Calme, S. 2024).

Dans un tel contexte, il devient essentiel de mieux comprendre les mécanismes techniques, sociaux et politiques des deepfakes, afin de mieux anticiper leurs usages potentiels. Leur danger ne réside pas uniquement dans ce qu'ils montrent, mais dans le climat de doute et de confusion qu'ils peuvent instaurer. Si tout peut être fabriqué, alors tout peut être remis en question, y compris la vérité (Langguth, J. 2021).

2.3. Deepfakes : le pouvoir du visuel

““Une image vaut mille mots”, disait Kong Fuzi“ (Bismuth, M. 2023, p.21).

En effet, les deepfakes tirent leur puissance de la nature même de la communication humaine, où le visuel et le sonore jouent un rôle prépondérant. Dès les années 1960, Albert Mehrabian a mis en évidence l'importance du non-verbal dans les échanges, estimant que 93 % d'un message repose sur le ton de la voix et l'image plutôt que sur les mots (Bismuth, M. 2023). En fait, notre cerveau assimile plus facilement une image ou un son qu'un texte, rendant ainsi ces supports particulièrement impactants et influents (cfr figure 3).

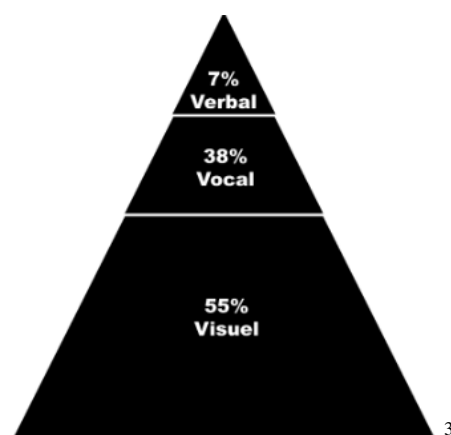


Figure 3 : Importance du non-verbal pour la communication

³ Figure 3 : Bismuth, C (2023) Médias synthétiques. Google Books.
https://books.google.be/books?hl=fr&lr=&id=GJ_KEAAQBAJ&oi=fnd&pg=PP1&dq=les+deepfakes+comme+outils+cr%C3%A9atifs&ots=RMTIpsa4DR&sig=BT7oJqdiyufChm8Ap7B1UstKY1c&redir_esc=y#v=onepage&q&f=false

Le développement des médias numériques et la prolifération des contenus visuels ont accentué cette dynamique. La photographie, le cinéma et ensuite, les réseaux sociaux, ont multiplié notre exposition aux images, qui façonnent nos perceptions et influencent nos décisions. Avec les avancées technologiques, notamment l'intelligence artificielle et les logiciels de retouche, les images de synthèse et les réalités numériques se sont fortement imposées (Bismuth, M. 2023).

Dans ce contexte, les deepfakes exploitent cette dépendance au visuel et au sonore pour manipuler l'opinion, estomper la frontière entre réel et virtuel, et fragiliser voire remettre en question la confiance accordée aux médias. Leur effet est d'autant plus important qu'ils évoluent dans un monde où les images dominent. Images touchant directement à l'affect et aux émotions pouvant alors, plus facilement, nous détourner d'un raisonnement logique (Joffe, H. 2007).

2.4. Deepfakes : désinformation et manipulation

Depuis l'essor du deep learning en 2017, les deepfakes ont envahi les réseaux socio-numériques. Qu'il s'agisse de vidéos, d'images, d'audios ou de textes, ces productions, souvent d'un réalisme troublant, sont conçues pour imiter, manipuler ou tromper, ce qui représente un défi majeur pour la communication et le journalisme (Frau-Meigs, D. 2019). En rendant floue la frontière entre réalité et manipulation, ils compliquent considérablement la vérification des sources, affaiblissent la crédibilité des médias traditionnels et brouillent le lien entre information, vérité et confiance, pourtant essentiel dans une démocratie (Kazadi, J.D, 2024). "Capables d'imiter le "vrai" à la perfection, ils suscitent autant de méfiance que de conviction envers les messages que nous recevons" (Bismuth, M. 2023, p.13).

Ces fausses nouvelles peuvent avoir divers objectifs. Celles-ci peuvent être créées dans le but de faire rire (faire des canulars), de nuire (diffamation), de s'enrichir (chantage) ou encore de manipuler (propagande). Cependant, toutes les formes reposent sur des stratégies de manipulation et de persuasion afin de faire croire en une vérité. Celles-ci sont alors instrumentalisées à travers les différents médias, en usant de trucages et de montages pour amplifier leur effet (Frau-Meigs, D. 2019).

Comme précisé précédemment, les deepfakes ne sont pas de simples montages, mais bien des hyper-trucages permettant à des individus de créer des contenus réalistes. Bien souvent, ces contenus sont utilisés à des fins de désinformation ou de manipulation de l'opinion publique. La démocratisation de ces fausses nouvelles modifie radicalement la dynamique d'information, menaçant par conséquent les valeurs démocratiques et la stabilité de la société (Simon, FM et al. 2023).

Cette banalisation des deepfakes est due à un bouleversement dans l'espace médiatique. Suite à la croissance des réseaux sociaux, n'importe quel utilisateur peut devenir émetteur de ces fausses nouvelles (Naffi, N et al. 2021). Le journaliste n'est plus l'unique médiateur entre les faits et le public. La parole des non-professionnels cohabite désormais avec celle des journalistes, souvent sur un pied d'égalité dans l'univers médiatique. Les citoyens, eux aussi, partagent l'actualité, ce qui accroît la vulnérabilité du public face à la désinformation (Gadras, S, 2022). Dans un contexte de méfiance croissante envers les médias, les utilisateurs se tournent vers des sources alternatives, notamment les réseaux sociaux, où les deepfakes circulent librement. Cette situation affaiblit encore davantage le lien de confiance entre médias et citoyens et remet en question la liberté d'expression, dès lors que l'information fautive prend le pas sur les faits vérifiés (Kazadi, J.D. 2024).

On distingue aujourd'hui deux types de manipulation : celle opérée par les humains, comme les lanceurs d'alerte, qui peuvent contribuer à l'intérêt général, et celle opérée par les machines, comme les deepfakes, qui dissimulent leur intention manipulatrice sous une apparence de véracité. Contrairement aux lanceurs d'alerte qui peuvent parfois dénoncer des injustices, les deepfakes visent principalement à induire en erreur, ce qui soulève de nombreux enjeux éthiques, sociaux, juridiques et politiques (Laville, C. 2022).

Leur potentiel de manipulation est particulièrement préoccupant. D'une part, les deepfakes peuvent nuire à l'identité individuelle d'une personne. Ils peuvent représenter celle-ci de manière non désirée, voire diffamatoire, pouvant nuire à son image. Ce type d'atteinte a des répercussions directes dans la sphère politique, où une vidéo truquée peut fausser la perception du public. La diffusion des deepfakes peut entraver le processus d'information et donc entraver l'opinion publique, la confiance dans les institutions démocratiques et polariser les divisions sociales (De Ruiter, A. 2021).

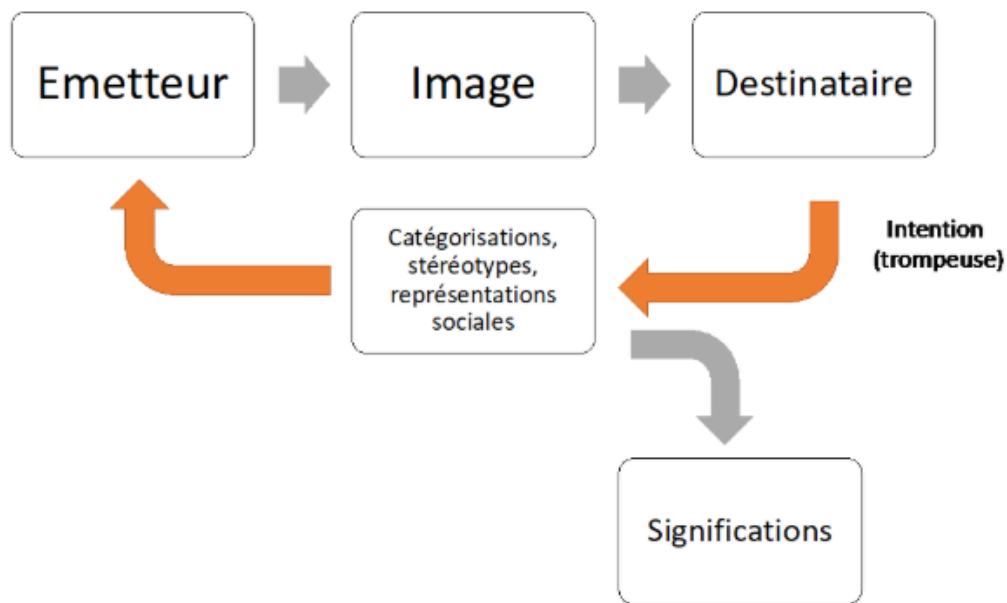
D'autre part, en semant le doute et la confusion, ces contenus trompeurs peuvent grandement influencer les croyances et les comportements des individus (Simon, FM et al. 2023). Les deepfakes sont alors perçus comme un danger pour notre société en certains points : la perte de confiance dans les institutions publiques et la désinformation, notamment politique. Les deepfakes peuvent distordre les discours, manipuler les opinions et fragiliser la démocratie (Robichaud-Durand, S. 2023).

En simulant la voix, les expressions et les gestes d'une personne réelle, les deepfakes parviennent à imiter de façon crédible des figures publiques, allant jusqu'à leur faire prononcer des propos qu'elles n'ont jamais tenus. Une étude de Vaccari et Chadwick montre que près de la moitié des participants peuvent être dupés par un deepfake, même s'ils disposent d'informations contextuelles (Vaccari, C - Chadwick, A. 2020).

Ces contenus exploitent les biais cognitifs déjà mobilisés par les fake news, comme l'effet de vérité illusoire ou le biais de confirmation, tout en y ajoutant une dimension visuelle plus impactante, renforçant leur pouvoir de persuasion.

Le biais de vérité illusoire consiste en la répétition abondante de fausses nouvelles renforçant alors leur crédibilité aux yeux des utilisateurs des réseaux socio-numériques. Il s'agit de percevoir une information répétée comme plus crédible, indépendamment de sa véracité. A force de voir constamment les mêmes informations, ces derniers peuvent finir par se laisser convaincre et y croire (Nelson, J. 2022).

De plus, les utilisateurs des réseaux socio-numériques peuvent être influencés par leurs convictions antérieures, leurs orientations politiques ou encore les stéréotypes et représentations sociales, ce que l'on appelle le biais de confirmation (cfr figure 4). Si ces fausses nouvelles traitent de ces éléments, ils sont alors plus susceptibles de confirmer et de propager, à leur tour, celles-ci. Et ce, même si les informations mises en circulation sont fausses et créées de toutes pièces (Nelson, J. 2022).



4

Figure 4 : Le biais de confirmation

Il est alors pertinent d'évoquer le modèle d'encodage et de décodage développé par Stuart Hall, permettant de comprendre comment les utilisateurs des réseaux socio-numériques interprètent les contenus, y compris ceux issus de la désinformation.

Ce modèle distingue trois moments clés : l'encodage, le message et le décodage. Lors de l'encodage, les producteurs de contenu tels que les médias ou les utilisateurs, fabriquent un message tout en s'appuyant sur des codes culturels influencés par des idéologies et des rapports de pouvoir. Le message ainsi produit ne possède pas un sens unique, il compose un ensemble de significations potentielles. Lors du décodage, le récepteur interprète à son tour, le message à partir de ses propres cadres de références, façonnés par ses expériences, sa culture ou encore sa classe sociale (Hall, S. 1980). De cette manière, le public n'est pas passif : il joue un rôle actif dans la construction du sens et l'interprétation du message, variable d'un individu à l'autre. Cela rejoint le rôle que peut jouer le biais de confirmation sur les réseaux sociaux. Les individus interprètent et relaient les contenus selon leurs croyances ou représentations préalables, contribuant à la circulation de fausses informations (Nelson, J. 2022).

⁴ Figure 4 : Nelson, J. (2022). Fake news et deepfakes : une approche cyberpsychologique. Interfaces Numériques, 11(2). <https://doi.org/10.25965/interfaces-numeriques.4830>

Dans cette logique, leur pouvoir de confusion ne résulte donc pas seulement dans leur caractère mensonger, mais aussi dans leur capacité à éroder la confiance générale dans les photos et vidéos circulant en ligne, en générant une incertitude permanente quant à la véracité de ce que l'on perçoit (Nelson, J. 2022). Par ailleurs, la manipulation ne repose pas uniquement sur le faux. Elle ne s'appuie pas toujours sur des informations totalement erronées. La manipulation peut également prendre racine à partir d'une demi-vérité. Cela permet alors aux fausses nouvelles de paraître plus subtiles et plus difficiles à déconstruire (Deschamps, C. 2021).

Comme conséquence de cela, certains chercheurs parlent d'infocalypse. C'est-à-dire, une abondance de contenus trompeurs compliquant davantage la capacité de discernement des individus. L'exposition grandissante à des contenus trompeurs tend à faciliter le traitement cognitif d'informations erronées. Cela signifie que les fausses informations perçues à travers les réseaux socio-numériques, se fondent dans la masse et n'alertent pas les utilisateurs (Nelson, J. 2022).

Enfin, les fausses nouvelles et plus précisément les deepfakes, se diffusent rapidement et de manière fragmentée. Elles sont capables de s'adapter. Si elles deviennent obsolètes, elles sont susceptibles de muter tout aussi rapidement. La cause principale de l'existence massive de fausses nouvelles est leur facilité et rapidité de diffusion.

Elles peuvent prendre racine n'importe où et être considérées comme véridiques, même de manière totalement involontaire (Gamba, F. s.d).

Face à la montée en puissance des deepfakes, les recherches soulignent que lutter contre les deepfakes ne peut se limiter à des stratégies techniques de vérification ou de détection automatisée. Il est également nécessaire de comprendre les croyances préexistantes, les identités sociales et les mécanismes cognitifs qui rendent certains individus particulièrement vulnérables à la manipulation. Paradoxalement, certaines tentatives de correction peuvent toutefois renforcer les convictions erronées. Ce phénomène est connu sous le nom d'effet boomerang. C'est pourquoi, les réponses à cette crise informationnelle doivent intervenir sur plusieurs plans : informationnel, émotionnel et sociocognitif (Nelson, J. 2022).

2.5. Responsabilité des plateformes et utilisateurs

Premièrement, les réseaux sociaux ou moteurs de recherche, tels que Google, YouTube et Facebook étant les plus grands, ont un rôle dans cette désinformation. En fait, par leur système d'algorithme et de recommandation, ils contribuent à faire apparaître des contenus incendiaires et trompeurs, créés pour manipuler. Au vu du volume de contenus publiés sur ces derniers, il est presque impossible de détecter et de chasser des algorithmes, les fausses nouvelles. De même, certaines plateformes recommandent des fausses nouvelles aux utilisateurs qui pourraient être intéressés, selon leurs centres d'intérêt et leurs préférences. Les plateformes amplifient alors le phénomène (Stocker, C. 2020).

Face à ces défis, il est essentiel de repenser les pratiques, de développer de nouveaux outils de détection et de renforcer l'éducation aux médias. Parce que si l'intelligence artificielle est à l'origine du problème, elle peut aussi être la meilleure arme pour y faire face. Des solutions technologiques émergent notamment, telles que le deepware ou le tatouage numérique, afin de détecter les faux contenus. Toutefois, la solution ne peut pas être uniquement technique. Une éducation aux intelligences artificielles et à leurs dangers est essentielle dès le plus jeune âge (Kazadi, J.D. 2024).

Deuxièmement, les utilisateurs des réseaux socio-numériques jouent également un rôle dans l'existence des deepfakes en 2 points : la création de contenus trompeurs et le manque de discernement.

Les utilisateurs jouent un rôle en partageant de fausses informations. Chaque individu peut interpréter un contenu selon sa propre perception, le transformer en le partageant et ainsi inviter d'autres individus à adopter la même vision erronée. En fait, cela peut être fait de manière volontaire ou non, par le simple partage de celle-ci (Frau-Meigs, D. 2019). Comme l'a évoqué le chercheur François-Bernard Huyghe dès l'émergence du web 2.0 en 2005, "le facteur qui va tout bouleverser est l'équation numérique plus réseaux" (Deschamps, C. 2021, p.2). "Si chacun peut devenir émetteur à son tour et non simple récepteur des mass médias ... il peut informer donc désinformer" (Deschamps, C. 2021,p.2).

Ensuite, les utilisateurs ont un rôle à jouer au niveau du discernement. En tant qu'utilisateur, il est important de vérifier les faits, les sources. C'est ce que l'on appelle le fact-checking (Borel, B. 2017). Cela consiste en la démystification, soit le fait de déconstruire les fausses informations. Toutefois, même si des outils émergent, notre société n'est pas encore assez équipée pour discerner le vrai du faux en matière de tromperie sur les réseaux socio-numériques.

De plus, ce qui donne de l'ampleur aux deepfakes et aux fausses nouvelles, est également le besoin des individus de croire en quelque chose. Peu importe si l'information est démontrable ou non, les individus ont besoin de croire en quelque chose pour se construire une vision de la société qui les entoure. C'est pourquoi, ces derniers ont tendance à croire plus facilement aux fausses informations et à ne pas s'en méfier (Gamba, F. s.d).

3. Méthodologie

3.1. Objectifs de la méthodologie

Dans le but de répondre à ma question de recherche portant sur les réactions des utilisateurs face aux deepfakes et leurs répercussions sur la transmission, la réception, l'interprétation et la crédibilité des messages, il est essentiel de collecter des données précises et approfondies.

Dans un premier temps, il convient d'en apprendre davantage quant à la connaissance des utilisateurs des réseaux socio-numériques face à l'existence des deepfakes et aux risques associés. Il s'agit d'évaluer leur état de conscience ainsi que de comprendre les réactions et attitudes qu'ils pensent avoir lorsqu'ils y sont exposés. Cette approche permet de mieux comprendre leur schéma de pensée et leur posture critique, en identifiant notamment s'ils adoptent une confiance systématique (confiance aveugle) envers les informations diffusées sur ces plateformes, ou s'ils développent, au contraire, une forme de vigilance et de méfiance.

Cette approche se situe dans un contexte où les deepfakes, grâce à leur réalisme visuel et à leur propagation sur les réseaux sociaux, exploitent des biais cognitifs tels que le biais de confirmation et l'effet de vérité illusoire, ce qui rend leur détection et leur rejet plus difficiles.

En effet, l'accessibilité grandissante des outils de création de deepfakes (Beridze, I - Butcher, J. 2019) et leur capacité à brouiller la frontière entre réalité et fiction (Langguth, J, 2021) compliquent encore davantage la tâche des utilisateurs, qui jouent un rôle actif dans la diffusion de ces contenus, qu'ils en aient conscience ou non (Gadras, S, 2022).

Dans un second temps, il est impératif d'explorer les comportements réels des utilisateurs lorsqu'ils interagissent avec des contenus trompeurs, créés de toutes pièces. Cette approche vise à confronter leurs déclarations, ce qu'ils pensent ou affirment faire, à leurs réactions réelles face à ces contenus. Ces informations me permettent de vérifier si ces derniers sont conscients de leurs propres comportements et s'ils réagissent effectivement comme ils le prétendent.

Il s'agit notamment de mettre à l'épreuve les résultats avancés par Vaccari et Chadwick, qui estiment que près de la moitié des individus peuvent se laisser tromper par un deepfake (Vaccari, C - Chadwick, A, 2020). Cette vérification me permet de confirmer ou de nuancer leurs conclusions dans un contexte différent d'analyse.

3.2. Méthodes

3.2.1. Etude de réception via une enquête qualitative

Ma recherche repose principalement sur une étude de réception réalisée au moyen d'une **enquête qualitative** conduite par la réalisation d'entretiens individuels approfondis de type semi-directif.

Le choix de cette méthode s'appuie sur les différents objectifs menés. Contrairement à une approche quantitative visant à la récolte de données mesurables, l'approche qualitative a pour but principal de comprendre ou d'expliquer un comportement, une motivation ou encore une perception. Cette méthode met l'accent sur des données verbales détaillées, plutôt que sur des données mesurables. C'est pourquoi, cherchant à approfondir le point de vue d'un groupe cible restreint, cette méthode est adaptée. Dans le cadre de ma recherche, il s'agit de

comprendre de quelle manière les utilisateurs des réseaux socio-numériques pensent percevoir, interpréter et réagir aux deepfakes.

Pourquoi opter pour des entretiens individuels de type semi-directif ?

Les entretiens individuels permettent de creuser en profondeur un sujet avec des répondants. Il ne s'agit pas simplement de répondre par une réponse fermée, mais bien d'argumenter ses réponses. Le choix du format semi-directif est également stratégique : il me permet de guider l'entretien tout en laissant une marge de spontanéité aux répondants, leur offrant ainsi la possibilité d'aborder des aspects inédits qui pourraient enrichir mon analyse. Aspects auxquels je n'ai peut-être pas songé.

Je mène entre dix et quinze entretiens individuels, plus précisément onze entretiens. Ce choix repose sur la volonté d'atteindre une certaine saturation des données durant la période de recherche qui m'est attribuée. En effet : “Dans le domaine de la recherche qualitative, la saturation des données joue un rôle crucial pour garantir la validité et la fiabilité des résultats. Il s'agit d'un concept que les chercheurs utilisent pour déterminer le moment où la collecte de données supplémentaires ne permet plus d'obtenir de nouvelles idées ou informations” (De Abreu, G. 2024).

Ce nombre prend également en compte les contraintes de temps liées à mon statut étudiant, tout en restant réaliste. Cependant, cela me permet tout de même une collecte d'informations diversifiée afin d'obtenir une vision nuancée des réactions, attitudes et comportements concrets des utilisateurs des réseaux socio-numériques, sans compromettre la qualité de la recherche.

3.2.1.1. Guide d'entretien

Afin d'établir mon guide d'entretien, je m'appuie non seulement sur les objectifs menés, mais également sur mes recherches théoriques et lectures scientifiques préalables (cfr tableau 1).

Tableau 1 : Concepts clés utilisés pour l'élaboration de la grille de questions

Auteurs	Concepts clés
Frau-Meigs, D. 2019	Les utilisateurs ont un rôle à jouer lorsqu'ils partagent de fausses informations, de manière volontaire ou non.
Borel, B. 2017	Les utilisateurs ont un rôle à jouer au niveau du discernement, notamment en ce que l'on appelle le fact-checking.
Langguth, J. 2021	L'enjeu des deepfakes réside moins dans leur création que dans leur détection rapide, rendue difficile par la surcharge d'informations visuelles dans l'espace numérique.
Bismuth, C. 2023	Les deepfakes transforment nos habitudes dans l'univers numérique et notre rapport à l'information.
Langguth, J.2021	L'utilisation des deepfakes à grande échelle préoccupe, notamment dans les contextes électoraux ou géopolotiques sensibles.
Simon, FM et al. 2023	En semant le doute et la confusion, ces contenus trompeurs peuvent grandement influencer les croyances et comportements des individus.
Fabre, M. 2022	L'aptitude des deepfakes à imiter les mécanismes cognitifs humains rend la distinction entre réel et manipulation plus difficile.
Langguth, J. 2021	Si tout peut être fabriqué, tout peut être remis en question, y compris la vérité.
Kazadi, J.D. 2024	Les deepfakes brouillent la frontière entre réel et manipulation, fragilisant la vérification des sources et la crédibilité des médias traditionnels.
Meskin, A. 2008	Depuis le 20ème siècle, souvent utilisés comme preuves, les médias visuels bénéficient d'une forte crédibilité auprès du grand public.
Bismuth, C. 2023	Les deepfakes tirent leur efficacité dans un univers où le visuel et le sonore jouent un rôle prépondérant.
Nelson, J. 2022	L'exposition grandissante à des contenus trompeurs tend à faciliter le traitement cognitif d'informations erronées.

Le Calme, S. 2024	Environ 1,5 milliard de messages mensongers seraient créés ou partagés chaque jour sur les réseaux socio-numériques.
Nelson, J. 2022	Les fausses nouvelles exploitent des biais cognitifs comme l'effet de vérité illusoire ou le biais de confirmation, tout en y ajoutant une dimension visuelle impactante, renforçant leur pouvoir de persuasion.

De surcroît, j'ai notamment réalisé un premier test de ma grille de questions en date du 7 juin 2024 ([cfr annexe 8.2](#)). Pour ce faire, j'ai pris l'initiative d'administrer celui-ci à 2 individus : le premier âgé de 23 ans, le second âgé de 47 ans. Ce choix n'est pas anodin, il visait à évaluer la clarté et la compréhension des questions ainsi que du sujet auprès d'individus aux profils différents. Cette étape m'a permis d'ajuster le guide d'entretien en fonction des résultats obtenus afin d'en optimiser la pertinence et l'efficacité ([cfr annexe 8.3](#)) ([cfr annexe 8.5](#)).

A la suite du test de mon dispositif méthodologique, j'ai pu m'apercevoir que certains ajustements étaient nécessaires. L'une des premières modifications a concerné l'adaptation du vocabulaire employé. Bien que familiers au concept de contenus numériques truqués, générés par l'intelligence artificielle, le terme deepfake ne résonne pas de la même manière chez tous les répondants. Le répondant plus âgé connaissait le concept. Toutefois, il ne connaissait pas le terme officiel, anglophone : deepfake. Pour éviter toute confusion et assurer une bonne compréhension des questions, j'ai alors dû introduire brièvement celui-ci en début d'entretien. Par ailleurs, certaines questions jugées redondantes dans les réponses obtenues, ont été fusionnées, reformulées ou encore supprimées. Ces modifications ont également été guidées par les concepts clés identifiés dans l'état de l'art, permettant ainsi d'aligner le guide d'entretien avec les objectifs menés.

Selon moi, après ajustement, le guide d'entretien semi-directif explore en profondeur différents aspects des deepfakes : les connaissances, les perceptions des participants, les comportements, ainsi que les répercussions que peuvent avoir ces deepfakes sur le processus d'information (cfr tableau 2). Ce guide d'entretien me sert alors de base pour ma recherche tout en laissant une certaine flexibilité, essentielle afin de suivre les pistes intéressantes qui pourraient émerger durant la discussion.

Tableau 2 : Grille de questions

Thème 1 : Introduction - Connaissance générale

(Objectif : évaluer la connaissance générale du sujet et de son ampleur par les répondants)

Aviez-vous déjà entendu parler des deepfakes auparavant ? Si oui, que savez-vous à leur sujet ? Comment en avez-vous entendu parler ?

Avez-vous déjà été confronté à des avertissements ou des campagnes de sensibilisation sur les deepfakes ?

Pensez-vous que les deepfakes occupent une place importante sur les réseaux sociaux ? A combien estimeriez-vous leur présence, en pourcentage ?

Pensez-vous qu'en tant qu'utilisateur des réseaux sociaux, vous avez un rôle à jouer avec les deepfakes ? (partage, méfiance, esprit critique, discernement, etc.)

Thème 2 : Perceptions et réactions face aux deepfakes

(Objectif : comprendre la réaction cognitive et affective des utilisateurs)

Quelle a été votre première réaction/pensée lorsque vous avez pris connaissance de l'existence des deepfakes ?

Avez-vous déjà eu le sentiment qu'une image ou une vidéo sur les réseaux sociaux pouvait être truquée ? Pouvez-vous me raconter ?

Avez-vous déjà été influencé par une information qui s'est avérée fausse ? Quelles en ont été les conséquences ? (Croyance, comportement)

Selon vous, qu'est-ce qui rend un deepfake difficile à détecter à première vue ?

Selon vous, comment distinguer un deepfake d'une vraie vidéo ou photo ?

Sur une échelle de 1 à 10, à quel point vous sentez-vous capable de reconnaître un deepfake ?

Pourquoi ce nombre ?

Thème 3 : Comportements sur les réseaux socio-numériques à l'ère des deepfakes
(Objectif : comprendre la réaction conative et les comportements des utilisateurs)

Depuis que vous avez connaissance des *deepfakes*, cela a-t-il modifié votre manière d'utiliser les réseaux sociaux ?

Doutez-vous à présent de toutes les informations que vous rencontrez sur les réseaux sociaux ? Ou seulement quand quelque chose vous paraît étrange ?

Avez-vous pris des mesures spécifiques pour vous protéger contre les fausses informations ?

Lorsque vous voyez une nouvelle vidéo ou image en ligne, prenez-vous systématiquement le temps d'en vérifier la source ? Comment procédez-vous ?

Aujourd'hui, êtes-vous plus vigilant(e) lorsque vous partagez du contenu sur vos réseaux ? Si oui, qu'est-ce qui a motivé ce changement ?

Thème 4 : Répercussions sur la réception, la transmission et l'interprétation des messages.
(Objectif : analyser et comprendre les effets des deepfakes sur la confiance envers les médias, et la perception du réel)

Considérez-vous les réseaux sociaux comme une source fiable ? Pourquoi ?
Et les médias traditionnels/journaux présents en ligne ? Pourquoi ?

Selon vos réponses, quelles sont alors les sources d'informations auxquelles vous faites le plus confiance ? (Pas uniquement en ligne, en général.)

Avez-vous tendance à croire plus facilement une information lorsqu'elle est accompagnée d'une image ou d'une vidéo, plutôt qu'un simple texte ? Pourquoi ?

Si une même information est répétée plusieurs fois par différentes personnes, par différents canaux, cela renforce-t-il votre confiance en sa véracité ?

Thème 5 : Conclusion

Y a-t-il quelque chose que vous aimeriez ajouter à ce sujet ?

3.2.2. Test d'awareness sur les contenus

En complément de cette étude qualitative, je mets en place un test d'awareness auprès des participants aux entretiens individuels. Cette expérimentation consiste à leur demander de différencier une publication authentique d'un deepfake afin d'évaluer leur capacité de discernement, en d'autres termes, leur capacité à identifier les contenus trompeurs. L'objectif ici est de recueillir des réponses honnêtes et vérifiables concernant la conscience et l'esprit critique des audiences vis-à-vis des deepfakes sur les réseaux socio-numériques.

Si les entretiens permettent d'explorer les représentations et perceptions des répondants quant aux deepfakes, ils ne garantissent pas que les comportements déclarés correspondent à leurs comportements réels. En effet, je vais pouvoir recueillir des informations sur le comportement et la conscience que pensent avoir les répondants vis-à-vis des contenus trompeurs. Toutefois, il existe souvent un écart : ce qu'ils pensent faire n'est pas toujours véridique. C'est pourquoi, il est crucial de confronter ces déclarations à une évaluation pratique, afin de vérifier la cohérence des propos par une méthode telle que celle-ci.

Lors d'un premier test réalisé le 7 juin 2024, j'ai présenté simultanément aux répondants une vidéo authentique de Mark Zuckerberg ainsi qu'un deepfake le mettant en scène, afin d'évaluer leur capacité à déceler le vrai du faux. Toutefois, le plus jeune des répondants connaissait déjà l'existence de ce deepfake, ce qui lui a permis de l'identifier immédiatement, sans procéder à une réelle analyse. Aucun critère de différenciation n'a été exprimé, compromettant ainsi la validité de l'évaluation. Par ailleurs, la méthode utilisée pour ce test s'est révélée peu appropriée : la présentation simultanée des deux vidéos a nettement facilité la distinction. Placées côte à côte, les différences sont devenues évidentes, ce qui remet en question la capacité de discernement réelle des répondants à détecter une deepfake ([cfr annexes 8.4](#)).

Afin de garantir la validité des résultats et d'éliminer tout biais lié à une connaissance préalable des répondants avec les contenus, j'ai conçu un matériel de recherche inédit, spécifiquement pour cette étude. Pour ce faire, je génère deux deepfakes au format photo et

utilise une image authentique issue des réseaux sociaux. Cette démarche permet de m'assurer qu'aucun répondant n'a été exposé à ces contenus auparavant, ce qui renforce l'objectivité de leur évaluation.

Chaque participant est confronté à trois images : une image authentique et deux deepfakes, présentés dans un ordre aléatoire. Cette répartition permet d'observer les variations dans la capacité de discerner le vrai du faux selon l'ordre de présentation et d'éviter toute logique de déduction. Comme par exemple : si la première image est fausse, la seconde est forcément vraie.

Contrairement au test initial, les contenus ne sont plus présentés simultanément. Chaque image est désormais montrée de manière individuelle pendant 15 secondes. Ce choix n'est pas anodin, il vise à reproduire de manière réaliste l'expérience de navigation sur les réseaux sociaux où les utilisateurs, en particulier les plus jeunes, font défiler les contenus rapidement, sans leur accorder une grande attention. Au cours des 20 dernières années, la durée moyenne d'attention a connu une nette diminution, passant d'environ 12 secondes à 8 secondes. Cette baisse est encore plus importante pour la génération Z, dont l'attention moyenne serait inférieure à 8 secondes (Blavity Inc, 2024). Accorder 15 secondes permet alors de concilier deux objectifs essentiels : simuler une expérience proche des usages habituels des réseaux sociaux, tout en laissant un laps de temps suffisant au participant pour observer l'image. Le participant partage alors sa première impression en indiquant s'il pense que l'image est un deepfake ou une image authentique. Ce processus est ensuite répété pour les images suivantes.

A la suite de l'observation de chacune des images, celles-ci sont présentées simultanément afin de permettre au répondant de confirmer ou d'ajuster ses réponses initiales. Durant cette étape, il a l'occasion de réaliser une analyse comparative plus approfondie en s'appuyant sur des détails, non nécessairement visibles au premier coup d'œil.

L'ensemble des précautions énoncées ci-dessus est essentiel pour maintenir l'objectivité de l'analyse et éviter tout biais lié à une déduction ou reconnaissance préalable des images utilisées. De cette manière, cette nouvelle approche garantit une fiabilité des données accrue et une évaluation plus fiable de leurs perceptions et critères de différenciation.

3.2.2.1. Création du matériel

Le choix du contenu authentique et la création du matériel à savoir : deux deepfakes, se reposent sur les concepts clés identifiés au cours de mes lectures scientifiques (cfr tableau 3).

Tableau 3 : Concepts clés utilisés pour la création du matériel

Auteurs	Concepts clés
Meskin, A. 200	Depuis le 20ème siècle, souvent utilisés comme preuves ou instruments de propagande, les médias visuels bénéficient d'une forte crédibilité auprès du public.
Langguth, J. 2021	Il est désormais possible de modifier ou de remplacer le visage d'une personne dans une image ou une vidéo rendant les manipulations d'autant plus efficaces.
Fabre, M. 2022	Le deepfake s'impose comme une certitude. Un visage qui semble réel bloque alors toute réflexion critique.
Langguth, J. 2022	L'utilisation des deepfakes préoccupe, notamment dans les contextes électoraux ou géopolitiques sensibles.
Langguth, J. 2022	"Si tout peut être fabriqué, tout peut être remis en question, y compris la vérité."
Bismuth, C. 2023	Notre cerveau assimile plus facilement une image ou un son qu'un texte, rendant ces supports particulièrement influents. Les deepfakes utilisent alors cette dépendance pour estomper la frontière entre réel et virtuel.
Joffe, H. 2007	Les images touchent directement à l'affect et aux émotions, pouvant alors facilement nous détourner d'un raisonnement logique.
Kazadi, J.D, 2023	Les deepfakes compliquent considérablement la vérification des sources en rendant floue la frontière entre réalité et manipulation.
Hendrik, B. 2018	Le message ainsi produit ne possède pas un sens unique, il compose un ensemble de significations potentielles.
De Ruiter, 2021	Les deepfakes peuvent nuire à l'identité individuelle d'une personne, pouvant nuire à son image et, par conséquent, avoir des répercussions directes dans la sphère politique notamment.
Nelson, J. 2022	Les fausses nouvelles exploitent le biais de confirmation, tout en y ajoutant une dimension visuelle impactante.

Plusieurs travaux issus de la littérature scientifique ont orienté mes choix dans la conception du matériel.

Meskin (2008) met notamment en évidence l'efficacité du visuel pour les deepfakes en raison du statut historique accordé à la vidéo et à la photo, en tant que preuves visuelles authentiques et difficilement falsifiables. Joffe (2007) confirme cette tendance pour le visuel, en insistant sur le fait que les vidéos et photos manipulées agissent directement sur les émotions, ce qui facilite non seulement l'adhésion aux messages qu'elles véhiculent, mais également leur mémorisation.

Dans le prolongement de cette analyse, Langguth (2021), Bismuth (2023) et Fabre (2022) affirment que la modification d'un élément visuel, en particulier celle d'un visage, accentue le réalisme des deepfakes. L'apparition d'un visage constitue une forme d'évidence trompeuse, qui tend à refléter authenticité et confiance. Dès lors, la manipulation devient plus difficile à identifier, laissant peu de place au doute (Langguth, J. 2021) (Bismuth, C. 2023) (Fabre, M. 2022). En outre, d'autres auteurs mettent en lumière les implications politiques et géopolitiques de ces technologies, qui, utilisées à grande échelle, peuvent influencer l'opinion publique et déstabiliser les institutions, notamment lorsqu'elles ciblent des figures politiques (De Ruiter, A. 2021).

C'est à partir de cet ensemble d'observations théoriques que j'ai conçu mon support de recherche. J'ai tout d'abord sélectionné une image authentique d'Emmanuel Macron et de Donald Trump au cœur d'une rigolade, un contexte volontairement informel (cfr figure 5). Ce choix s'appuie sur le principe du biais de confirmation. Biais par lequel les individus ont tendance à interpréter les informations en fonction de leurs croyances ou représentations préexistantes (Nelson, J. 2022). De cette manière, un contexte inattendu ou inhabituel, comme deux figures politiques partageant un moment de complicité, peut susciter le doute et rendre la détection plus challengeante.



Figure 5 : Image authentique de Donald Trump et Emmanuel Macron

Sur la base d'une photo originale représentant Emmanuel Macron et Donald Trump, j'ai généré deux deepfakes en utilisant une combinaison d'outils d'intelligence artificielle. Pour ce faire, j'ai tout d'abord inséré l'image authentique dans l'outil **Image to Prompt**, un logiciel en ligne capable d'analyser une image et d'en extraire un prompt. Ce prompt généré reprend l'ensemble des éléments visuels de l'image (décor, visage, expressions, ambiance, etc.) et constitue une base textuelle modifiable pour la création de contenus truqués. En ajustant ce prompt, il est possible de transformer l'image initiale afin de l'insérer dans un contexte différent.

Une fois les prompts retravaillés, je les ai insérés dans l'outil **Stable Diffusion 3 Medium**, un générateur d'images synthétiques basé sur l'intelligence artificielle. Par le biais de cette méthode, j'ai pu produire deux deepfakes : l'un très réaliste (cfr figure 6), l'autre de qualité approximative (cfr figure 7). Cette différence de qualité a été pensée pour observer si le réalisme visuel influence la capacité des participants à détecter un contenu truqué et à quel point la qualité du visuel peut agir sur leur jugement.

⁵ Watson, J (2025). Macron et Trump ravivent leur « bromance » au milieu de tensions stratégiques sur l'Ukraine. *France 24*. <https://www.france24.com/fr/am%C3%A9riques/20250225-macron-et-trump-ravivent-leur-bromance-en-pleine-tensions-strat%C3%A9giques-sur-l-ukraine>

Voici notamment le prompt qui a servi à créer la figure 6 : "Donald Trump et Emmanuel Macron sont assis face à face dans un salon officiel à la décoration élégante, avec un mobilier doré et une cheminée en arrière-plan. Trump, en costume noir avec une cravate rouge, est penché légèrement en avant, les mains jointes, l'air attentif et sérieux. Macron, en costume noir avec une chemise blanche, une cravate noire et une épinglette sur le revers, adopte une posture expressive, parlant avec des gestes marqués de la main droite, illustrant un échange. L'ambiance reflète un moment de discussion formel entre deux chefs d'État."



Figure 6 : Deepfake représentant Donald Trump et Emmanuel Macron

Voici également le prompt ayant servi à la création de la figure 7 : "Donald Trump et Emmanuel Macron sont assis côte à côte dans un salon luxueux de style européen, orné de moulures dorées, de meubles anciens et d'un décor raffiné. Ils se serrent la main fermement, chacun affirmant une expression sérieuse. Toutefois, leurs regards ne se croisent pas. Ils regardent tous deux droit devant eux. Trump porte un costume noir avec une cravate rouge, Macron un costume noir avec une cravate noire. L'image semble nette, avec des contours précis et sans défauts apparents."



Figure 7 : Deepfake représentant Donald Trump et Emmanuel Macron

3.3. Échantillon de la population et recrutement

Pour répondre à ma question de recherche, j'ai décidé de me concentrer sur une population belge âgée de 16 à 60 ans, femmes et hommes, utilisateurs réguliers des réseaux socio-numériques. Bien que les plus jeunes soient davantage susceptibles d'utiliser ces derniers, l'écart générationnel tend à se réduire. Aujourd'hui, les données révèlent que près de 80 % des personnes âgées de 30 à 49 ans utilisent les réseaux sociaux, et environ 65 % des 50-64 ans sont également actifs sur ces plateformes (Walden University, 2022). Ce constat justifie alors pleinement le choix de cet échantillon permettant de refléter de manière représentative la diversité des usages des réseaux socio-numériques, au sein d'une population belge. En tenant compte de ces diversités, je peux alors mieux comprendre les réactions des utilisateurs et les répercussions des deepfakes sur les utilisateurs et la manière dont ils perçoivent, partagent et interprètent les messages.

Ensuite, étudier les utilisateurs réguliers des réseaux socio-numériques âgés de 16 à 60 ans permet de se concentrer sur un groupe démographique délimité, tout en gardant une étude gérable en termes de taille et de complexité. De plus, ces utilisateurs sont en mesure de me fournir un consentement éclairé. Cela signifie qu'ils sont en mesure de comprendre l'objectif de cette enquête, mais également de donner leur accord pour l'enregistrement des données dans le cadre d'entretiens individuels, par exemple. Si j'élargis ma cible vers une audience plus jeune, il est nécessaire que je demande l'autorisation des parents pour interroger leurs enfants, ce qui ajoute de la complexité à cette étude.

Dans le cadre de cette étude, le recrutement des participants s'effectue par convenance, c'est-à-dire en sollicitant des personnes facilement accessibles. Pour ce faire, j'utilise exclusivement les réseaux sociaux comme canal de diffusion. Une annonce est publiée sur mon profil Facebook personnel, ainsi que dans certains groupes Facebook ciblés, dans le but de maximiser la visibilité de l'appel à participation. Dans ces publications, je présente brièvement le sujet de mon étude, en précisant les objectifs menés et les critères de sélection des participants, à savoir : être utilisateur régulier des RSN, être âgé entre 16 et 60 ans et résider en Belgique. J'indique également que la participation demande environ 30 minutes d'entretien.

Une fois les participants sélectionnés, et avant chaque entretien, je prends le temps de leur faire lire et signer un formulaire de consentement. Ce document détaille les objectifs de la recherche, le déroulement de l'entretien, la nature des données collectées, leur utilisation ainsi que les droits des participants, notamment celui de se retirer à tout moment sans justification. Cette étape est essentielle, car elle garantit le respect des principes éthiques de la recherche et assure la transparence entre la chercheuse et les répondants. De cette manière, chaque participant s'engage en toute connaissance de cause et en comprenant les modalités de protection de ses données personnelles.

Le formulaire de consentement se retrouve dans les annexes ([cfr annexes 8.6](#)). Les versions de ce document signées par l'ensemble des répondants sont conservées avec soin afin de pouvoir être présentées en cas de besoin.

3.4. Script de passation

Pour toute personne souhaitant reproduire cette recherche, ce script de passation constitue une trame à suivre systématiquement lors de chaque entretien individuel. Ce script sert à mener l'entretien de manière uniforme auprès de chacun des répondants. Il garantit ainsi une approche claire et homogène pour l'ensemble des entretiens, contribuant à la rigueur méthodologique de la recherche.

a) Formulaire de consentement

“Avant de commencer, veuillez prendre connaissance de ce document et le compléter ([cfr Annexe 8.6](#)).”

b) Introduction à l'entretien

“Bonjour et merci d'avoir accepté de participer à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux socio-numériques.

Pour information, les deepfakes sont des contenus photos ou vidéos, truqués, créés par des intelligences artificielles. Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage d'une personne ou encore les éléments d'un décor, à partir d'une simple photo ou description. Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes. Bien entendu, vous êtes libre, à tout moment, de ne pas répondre à certaines questions ou encore de mettre fin à l'entretien si vous en ressentez le besoin. Avez-vous des questions avant de commencer ? ”

c) Grille de questions

A présent, l'entretien peut débuter en suivant l'ordre établi dans la grille de questions (cfr tableau 2) tout en laissant place aux interactions spontanées.

d) Test d'awareness :

A l'issue de la grille de questions se déroule le test d'awareness. Celui-ci débute par une brève explication de son déroulement. Je précise notamment aux participants que les images présentées peuvent être toutes vraies, toutes fausses ou une combinaison vraies-fausse. De plus, j'invite les participants à justifier chacun de leurs choix.

4. Analyse et interprétation des résultats :

Dans le but de répondre à ma question de recherche, j'ai mené des entretiens individuels semi-directifs auprès de onze individus. Ensemble, ils constituent un échantillon de la population belge âgée de 16 à 60 ans, répartie de manière équilibrée selon leur genre et leur âge (cfr tableau 4).

Tableau 4 : Répartition des répondants selon le genre et l'âge

	Homme	Femme	Total
Moins de 18 ans	1	0	1
Entre 18 et 25 ans	1	4	5
Entre 26 et 45 ans	0	1	1
Entre 46 et 60 ans	2	2	4
Total	4	7	11

L'ensemble des données recueillies lors de ces entretiens se retrouve dans les annexes ([cfr annexes 8.7](#)). Ces données reprennent l'ensemble des réponses obtenues à partir de la grille de questions ainsi qu'à travers le test d'awareness mené auprès des participants. La conduite de ces entretiens a permis d'atteindre un point de saturation des données. Au fur et à mesure des échanges, les informations récoltées devenaient récurrentes et aucune nouvelle perspective de recherche n'émergeait.

Pour répondre à la question de recherche, il est essentiel de commencer par une analyse du niveau de connaissance des répondants à propos du phénomène des deepfakes. En effet, cette connaissance ou méconnaissance pourrait avoir des répercussions sur les réactions des utilisateurs, mais également sur la manière dont l'information est partagée, interprétée et perçue comme crédible par ces derniers sur les réseaux socio-numériques.

4.1. Attitudes cognitives : connaissance du phénomène et de son ampleur

Connaissance du phénomène des deepfakes :

L'analyse des entretiens administrés révèle une connaissance générale, mais incomplète du phénomène des deepfakes. Si la majorité des participants, notamment les plus jeunes (16-26 ans), en ont entendu parler ou y ont été fréquemment exposés sur les réseaux socio-numériques, cette connaissance reste rudimentaire. Beaucoup comprennent le concept de contenus visuels modifiés par l'intelligence artificielle, sans pour autant en connaître le terme exact, ni distinguer clairement les deepfakes d'autres formes de manipulations numériques comme les piratages ou les arnaques en ligne. Cette confusion est particulièrement présente chez les répondants plus âgés ou moins actifs sur les réseaux socio-numériques, ce qui témoigne d'un manque de compréhension du sujet. La reconnaissance du phénomène dépend donc largement du niveau d'exposition. Pour certains, ce sont les réseaux sociaux qui jouent un rôle informatif, notamment lorsqu'ils croisent ces contenus dans leur fil d'actualité. La médiatisation de certains cas, comme celui d'une fausse vidéo de Brad Pitt, ou des séquences dans les journaux télévisés ou les séries, peut également contribuer à éveiller les consciences.

Ce manque de conscience peut également provenir d'une absence de mises en garde claires. Les pratiques de prévention mentionnées durant les entretiens reposent essentiellement sur des dynamiques communautaires. Les participants mentionnent la présence de "*warnings*" sous certaines vidéos, de commentaires appelant à la vigilance, ou encore des contenus explicatifs publiés par des utilisateurs des réseaux socio-numériques cherchant à démystifier les deepfakes. Cependant, ces pratiques sont ponctuelles, souvent improvisées et non formelles. De nombreux participants disent n'avoir jamais été avertis de manière claire, mais en déduisent par eux-mêmes l'importance de se méfier de certains contenus.

Toutefois, les deepfakes sont, pour la majorité des répondants, immédiatement associés à des usages trompeurs. Ils sont perçus comme des outils de manipulation, de discrédit, de recherche de buzz ou de profit, notamment dans le domaine politique. En effet, plusieurs d'entre eux en mobilisent spontanément des exemples pour illustrer leurs propos et insistent

sur les dangers de ces contenus lorsqu'ils sont utilisés dans le but de déstabiliser une figure politique ou d'entacher sa réputation, comme en témoignent ces propos : *“j'imagine bien un truc genre Trump, qui aurait été le sujet d'un deepfake et qui, je ne sais pas, pète les boulons suite à quelque chose, que les gens y croient et que cela ait des répercussions sur sa carrière politique.”* Cette perception rejoint alors les inquiétudes exprimées par Langguth (2021), qui souligne que l'utilisation des deepfakes à grande échelle soulève des risques majeurs, notamment dans les contextes électoraux et géopolitiques sensibles.

Connaissance de l'ampleur du phénomène des deepfakes :

Lorsqu'il s'agit d'évaluer l'ampleur du phénomène sur les réseaux socio-numériques, les réponses recueillies montrent une forte hésitation, avec des estimations très variables, allant de 2 à 50 % de contenus trompeurs circulant en ligne. Cette variabilité des réponses souligne une difficulté à appréhender concrètement la place des deepfakes dans l'écosystème numérique. Cette difficulté peut être comprise par la surcharge informationnelle qui caractérise les médias sociaux. La quantité de photos et vidéos publiée chaque jour rend l'identification des deepfakes presque impossible pour les utilisateurs non avertis (Langguth, J. 2021). Prenons l'exemple de Facebook, plus d'un milliard de messages mensongers sont créés par jour, pour 3 milliards d'utilisateurs actifs. Cela reviendrait à dire que plus d'un tiers du contenu circulant sur Facebook, soit 33,33 %, est mensonger (Le Calme, S. 2024). Dans un tel contexte, les deepfakes peuvent se fondre dans la masse, ce qui explique cette confusion dans l'évaluation de l'ampleur du phénomène.

Cependant, même s'ils reconnaissent ne pas toujours être en mesure de les identifier avec certitude, la majorité des répondants s'accordent à dire que ces contenus semblent prendre de l'ampleur, comme en témoignent ces propos : *“J'en croise tous les jours, pas beaucoup, mais quand même. J'en croise tous les jours maintenant, alors qu'avant pas. Donc ça prend de l'ampleur.”* Toutefois, tandis que certains affirment en croiser régulièrement, d'autres semblent n'en avoir jamais vraiment vu. Ce contraste peut notamment s'expliquer par une méconnaissance du phénomène, mais aussi par le fonctionnement des algorithmes propres aux médias sociaux. En effet, les contenus proposés aux utilisateurs sont souvent façonnés par leurs préférences, ce qui peut limiter leur exposition aux deepfakes (Stocker, C. 2020). Dans cette logique, les participants consommant majoritairement des contenus humoristiques ou liés à leurs loisirs peuvent être moins susceptibles de rencontrer des deepfakes. A l'inverse,

les participants intéressés par les sujets politiques ou d'actualité sont potentiellement plus exposés à ce genre de manipulation.

4.2. Attitudes affectives : de l'émerveillement technologique au doute

La découverte des deepfakes provoque chez les répondants de multiples réactions, allant de l'émerveillement technologique à la peur, le doute ou encore l'indifférence (cfr figure 8). Ces réactions contrastées se répercutent directement sur la manière dont les messages sont perçus, partagés et interprétés dans l'espace numérique.



Figure 8 : Nuage de mots - réactions des répondants face à l'existence des deepfakes

De l'émerveillement technologique à l'inquiétude :

Pour plusieurs répondants, les deepfakes suscitent tout d'abord un certain enthousiasme, exprimé par des termes comme *“dinguerie”*, *“stylé”*, *“c'est bien fait”*, *“c'est fou”*, *“c'est super réaliste”*, ou encore *“waouh tout ce qu'on peut faire avec la technologie”*. Cet émerveillement met en lumière la capacité des intelligences artificielles à imiter le réel à la perfection.

Toutefois, cette phase d'émerveillement laisse rapidement place à une prise de conscience des dangers associés à cette technologie. Cette phrase : *“parce que si tu sais l'utiliser dans le bon, tu sais l'utiliser dans le mauvais”*, exprimée par un répondant, en est un bon exemple.

De même, des mots tels que : *“une belle arnaque”* sont particulièrement lourds de sens, montrant une vision péjorative du phénomène.

Cette inquiétude se renforce auprès des utilisateurs à l'idée que de telles manipulations se retrouvent *“entre de mauvaises mains”*, pour reprendre les paroles d'un des répondants. En effet, la capacité à manipuler le réel n'est plus uniquement réservée aux experts, mais à n'importe quel individu connecté qui sait utiliser cette technologie, comme le déclarent les répondants (Chavalarias, D. 2018). Cela fait donc émerger une certaine forme de méfiance exprimée de la sorte : *“On ne connaît pas la source, on ne connaît pas la personne et ses idées.”* Il est vrai que les réseaux socio-numériques permettent à tout le monde de devenir producteur et diffuseur d'informations. La parole des professionnels et des non-professionnels cohabite sur un même pied d'égalité (Naffi, N et al. 2021). Cette évolution contribue à affaiblir la crédibilité des messages.

Le doute suscité par les deepfakes :

Une des réactions les plus importantes est le doute. L'expression *“on est un peu perdu, on ne sait plus s'il faut y croire ou pas.”* ou encore *“tu ne sais plus dire si c'est vrai ou si c'est faux, même en pensant que oui, c'est vrai, j'ai vu l'info ailleurs, l'autre information que j'ai vue, est-elle vraie ?”* formulée par des répondants, traduit clairement un sentiment de confusion, qui rejoint directement l'accessibilité de ces technologies.

Ce doute prédominant peut drastiquement affecter la réception, le partage et l'interprétation des messages. Pour Simon et al. (2023) ce n'est pas uniquement le contenu du message qui influence les croyances, mais le doute permanent qu'il installe, qui modifie en profondeur le traitement de l'information (Simon, FM et al. 2023). Les deepfakes rendent floue la frontière entre information et manipulation, ce qui fragiliserait le lien de confiance entre les citoyens et les institutions médiatiques (Kazadi, J.D. 2024).

Cette observation se confirme pleinement dans les discours recueillis lors des entretiens, concernant les réseaux sociaux. Plusieurs répondants expriment une méfiance marquée à l'égard de ces plateformes : *“ Fiabiles non, c'est du divertissement ”* ou *“ Tout le monde met tout et n'importe quoi. Tu ne sais pas dire ce qui est vrai.”* Ces propos montrent que la

capacité de n'importe quel utilisateur à diffuser des contenus, qu'ils soient authentiques ou truqués, alimente ce climat de doute et d'incertitude. Ainsi, on peut confirmer l'idée de Kazadi (2024), selon laquelle les réseaux socio-numériques brouillent la relation entre vérité et confiance. En revanche, ce constat mérite d'être nuancé lorsqu'on s'intéresse aux médias traditionnels, notamment à ceux présents en ligne. Contrairement à ce que la théorie suggère, le lien de confiance entre citoyens et médias professionnels ne semble pas rompu. Plusieurs répondants affirment faire confiance à ces médias, tout en questionnant leur objectivité. Quand je leur demande si les médias professionnels présents en ligne sont fiables, voici un aperçu de leurs réponses : “ *Oui car à la base ils sont censés être neutres, même si parfois ça ne l'est pas trop.*” ou encore “*Oui car c'est quelque chose d'officiel, mais parfois je me doute qu'il n'y a pas tout de vrai. Enfin ... pas que cela soit faux, mais qu'ils effacent ou cachent des infos.*” Ces propos montrent alors que la confiance envers les médias traditionnels en ligne est davantage liée à leur manque de neutralité, influencée par leurs opinions ou leurs intérêts, plutôt qu'à la possible présence de fausses informations dans leurs contenus. En ce sens, la thèse selon laquelle : si tout peut être fabriqué, tout peut être remis en question, y compris la vérité, trouve une signification (Langguth, J. 2021). Les deepfakes installent bien un climat de confusion, mais dans le cas des journalistes, la vérité n'est pas remise en question, simplement la manière dont elle est racontée.

L'indifférence face au phénomène :

Enfin, certains répondants adoptent une posture de distance ou d'indifférence face au phénomène des deepfakes. Soit parce qu'ils estiment ne pas être directement concernés par ceux-ci ou parce qu'ils les perçoivent comme une évolution technologique inévitable. Ce type de réaction traduit une certaine banalisation du phénomène liée aux évolutions numériques constantes. Ces évolutions numériques, déjà à l'œuvre dès les années 1990, où la retouche photo s'est imposée comme une banalité, jusqu'à en devenir un verbe courant (Langguth, J. 2021).

Cette indifférence ou cette banalisation peut toutefois avoir des conséquences importantes sur la diffusion de deepfakes. Comme le rappellent Frau-Meigs (2019) et Borel (2017), les utilisateurs jouent un rôle central dans la propagation des fausses informations : soit en les partageant involontairement, soit en manquant de discernement ou d'action. Cette passivité,

combinée à une surcharge informationnelle, rend alors la détection des deepfakes encore plus difficile et accroît la désinformation (Frau-Meigs, D. 2019) (Borel, B. 2017).

4.3. Attitudes conatives : rapport à la responsabilité, évolution des pratiques et vérification des sources

Le rapport à la responsabilité des utilisateurs :

Si certains répondants affirment clairement avoir un rôle à jouer face aux deepfakes, d'autres expriment au contraire une forme d'impuissance face à l'ampleur du phénomène.

D'une part, quelques répondants pensent ne pas avoir de rôle à jouer face aux deepfakes, n'ayant ni l'influence ni les outils nécessaires pour agir. Ils rejettent cette responsabilité, estimant que ce n'est pas aux utilisateurs d'agir, mais plutôt aux créateurs de contenus ou aux plateformes. Certains affirment même : *“je ne partage rien”, “je ne like rien”* ou encore *“je ne sais pas si ce sont les utilisateurs qui ont un rôle à jouer ou plutôt les personnes qui créent les deepfakes.”* Ce sentiment d'impuissance peut alors s'expliquer par la méconnaissance de l'espace numérique, du phénomène et des gestes à adopter pour limiter la diffusion de deepfakes. Pourtant, la solution ne peut pas être uniquement technique. Comme l'affirme Kazadi (2024), l'éducation aux médias a sa part de responsabilité face à cette indifférence. L'éducation aux médias et aux risques associés est nécessaire dès le plus jeune âge (Kazadi, J.D. 2024). Faute de cela, certains utilisateurs s'en remettent uniquement aux plateformes et aux créateurs de contenus. Cependant, même les utilisateurs lambda jouent un rôle clé dans l'écosystème numérique. L'action ou l'absence d'action contribue à faire circuler ou au contraire freiner la propagation d'une information trompeuse. En partageant une vidéo sans en vérifier la source, ou en laissant un deepfake se diffuser sans réagir, ils participent malgré eux, à la désinformation (Frau-Meigs, D. 2019).

D'autre part, la majorité des entretiens révèlent un sentiment de responsabilité, mais encore largement guidé par le niveau de connaissance du phénomène et le degré d'engagement de chacun sur les plateformes sociales. Plus les répondants semblent sensibilisés aux deepfakes et aux risques associés, plus ils semblent s'impliquer dans leur propagation. Ils estiment que des gestes simples peuvent déjà contribuer à freiner leur propagation, notamment : en évitant de relayer ces contenus, en les signalant, en vérifiant leurs sources, en alertant leur entourage,

ou encore en apprenant à les identifier dans leur fil d'actualité. Ces petites actions qu'ils réalisent renvoient directement à ce que Borel (2017) désigne comme la responsabilité individuelle des utilisateurs à faire preuve de vigilance, de discernement et à adopter des pratiques de vérification de l'information, appelées également fact-checking (Borel, B. 2017).

Évolutions des pratiques des utilisateurs sur les réseaux socio-numériques :

Même si la plupart des répondants se disent conscients du phénomène des deepfakes et affirment avoir un rôle à jouer dans la lutte contre ceux-ci. Cette prise de conscience ne se traduit pas, pour autant, en un changement concret des habitudes sur les réseaux socio-numériques. Si certains déclarent faire preuve de plus de prudence, de vigilance et de moins de naïveté envers les messages qu'ils reçoivent, cela reste intuitif et dépend fortement du contenu rencontré. Un répondant illustre parfaitement ce léger changement en expliquant : *“Je m'en méfie plus. Quand il y a un discours un peu fou, je me pose la question de savoir si c'est un vrai truc ou non. Alors qu'avant on prenait simplement l'information, sans réfléchir.”* En revanche, d'autres déclarent ne pas avoir modifié leurs comportements en ligne. Pour certains, cela s'explique par le fait qu'ils ne se sentent pas concernés par les deepfakes, estimant ne pas en croiser au quotidien. Pour d'autres, l'absence de changement se justifie par la nature de leur utilisation des réseaux socio-numériques, qu'ils utilisent pour échanger avec leurs proches ou se divertir, et non comme une source d'information. Dans ces cas-là, la présence éventuelle de deepfakes ne suscite pas d'inquiétude, car elle ne touche pas à des sujets jugés sérieux ou d'actualité. De plus, certains répondants affirment ne pas avoir modifié leurs habitudes car ils estiment encore pouvoir faire la part des choses. Comme l'exprime l'un d'eux : *“... la photo, la vidéo, je ne vérifierai pas parce que je suis sûre de mon jugement, pour l'instant. C'est vraiment le contenu. Là, parfois, j'aurai un doute, et là, j'irai voir sur d'autres sites, j'irai me renseigner ailleurs, oui.”*

Malgré l'absence de changement pour la réception de l'information, tous les répondants affirment néanmoins prêter plus d'attention à ce qu'ils relaient. Beaucoup déclarent ne rien partager du tout ou se limiter à des contenus humoristiques, perçus comme inoffensifs, même s'ils étaient générés par l'intelligence artificielle. Cela s'explique notamment par la motivation de ne pas faire circuler de fausses informations, mais surtout par la peur de *“paraître bête”* aux yeux de leurs proches ou de leur communauté en ligne.

La vérification systématique des sources :

Par ailleurs, bien qu'ils soient conscients de l'existence des deepfakes, les répondants continuent à accorder une forte confiance aux visuels, pour autant que ceux-ci soient de bonne qualité : *“s'il y a une image ou une vidéo qui est réaliste, je crois que tout le monde serait plus influencé et plus à même d'y croire”*. Comme le souligne la théorie de Meskin (2008), l'image et la vidéo restent perçues comme une preuve, un fait incontestable. Les répondants sont en accord avec cette théorie en déclarant : *“l'air de rien, je pense que ça influence quand même parce que tu as une espèce de preuve à l'appui”, “on a plus facile à faire confiance à un visuel”*. De plus, ce propos : *“une image ou du son, c'est plus percutant”*, rejoint la théorie de Joffe (2007) soulignant le fait que les images et vidéos touchent directement aux émotions, influençant nos jugements en dépit de la logique ou du raisonnement critique.

Enfin, la plupart des répondants disent douter des informations reçues uniquement lorsque quelque chose attire leur attention, ou ne correspond pas à ce qu'ils pensent. Cela renvoie directement au biais de confirmation (Nelson, J. 2022). Le doute ne surgit auprès des répondants que lorsque quelque chose cloche : un visage est étrange, une voix ne correspond pas, ou des propos sont jugés incohérents selon la connaissance du sujet ou de la personne. En l'absence de ces éléments, les contenus sont le plus souvent perçus comme authentiques, renforçant la vulnérabilité face aux manipulations.

Si beaucoup reconnaissent ne vérifier la source que lorsqu'un contenu paraît bizarre. Certains choisissent simplement d'ignorer l'information en cas de doute : *“Quand j'ai un doute, je passe, je scrolle.”* Mais lorsqu'il y a vérification des sources, les méthodes évoquées sont : une recherche Google, la comparaison avec d'autres articles, ou la consultation de comptes perçus comme fiables tels que Brut, RTBF et Hugo Décrypte. Le recours à l'entourage et à la sphère sociale joue également un rôle important, comme l'explique un des répondants : *“J'essaie de voir aussi l'avis des gens dans les commentaires, et d'en parler autour de moi, à des gens que je crois plus capables de pouvoir m'informer.”* Cependant, cette stratégie n'est pas fiable. La popularité d'un contenu, le volume de commentaires ou la fréquence à laquelle l'information revient, ne doivent pas être perçus comme preuves de fiabilité. Cette stratégie illustre le biais de vérité illusoire, selon lequel une information fréquemment rencontrée

paraît plus crédible, indépendamment de sa véracité. La validation de l'information repose alors sur des repères biaisés pouvant renforcer, une fois de plus, la désinformation.

4.4. Capacité à discerner le vrai du faux : de la théorie à la pratique

Comment détecter un deepfake ?

Dans l'ensemble des entretiens menés, une idée revient souvent : la difficulté à identifier un deepfake repose principalement sur son réalisme visuel et sonore. Pour plusieurs répondants, la qualité de la vidéo ou de l'image joue un rôle majeur : *“Quand c'est de bonne qualité, cela donne l'impression que c'est sérieux, alors qu'en fait, pas forcément.”* D'autres insistent sur le niveau de détail et de vraisemblance du contenu en déclarant : *“Quand on a une image avec beaucoup d'informations vraies et qu'il y a juste un petit détail faux ... on se dit, s'il y a autant de choses qui sont réelles, c'est que tout est vrai.”* En effet, comme le soulignent Nelson (2022), Fabre (2022) ou Deschamps (2021), les deepfakes ne se construisent pas nécessairement sur un mensonge total. Ils exploitent l'ambiguïté et la demi-vérité, ce qui rend la distinction entre réel et manipulation extrêmement difficile, portant alors à confusion. Un visage qui semble réel ou une voix familière suffisent à bloquer la réflexion critique.

Pour d'autres répondants, la difficulté à identifier un deepfake d'un contenu authentique, réside dans leur méconnaissance du sujet. Pour ceux peu familiarisés avec l'intelligence artificielle, la détection semble presque impossible : *“Je ne connais pas assez l'informatique pour voir la différence ... Même en zoomant, je ne verrais rien.”*

Malgré les difficultés à identifier un deepfake au premier coup d'œil, les répondants mobilisent néanmoins différentes stratégies pour tenter d'y parvenir. Nombre d'entre eux se fient à des indices visuels précis, notamment pour les vidéos, qu'ils jugent plus faciles à reconnaître que les photos, comme en témoigne cet extrait : *“...les photos c'est compliqué, mais genre, si c'est une vidéo, il y a des choses que l'IA ne sait pas forcément refaire, surtout les mains, elle ne sait pas faire comme si c'était naturel.”* Le mouvement, les expressions faciales, la synchronisation entre la voix et les lèvres ou encore les anomalies sur les parties du corps, sont perçus comme des éléments d'analyse. En revanche, les photos sont perçues comme plus trompeuses, car elles sont statiques et souvent d'une qualité supérieure. L'absence de mouvement réduit la possibilité de repérer certains détails, comme le déclare un répondant : *“Une photo, si on ne s'y attarde pas, on peut penser que c'est vrai ... surtout si on n'a jamais vu de photo de référence.”*

Détection d'un deepfake, de la théorie à la pratique :

Lorsque je demande aux personnes interrogées d'évaluer, sur une échelle de 1 à 10, leur capacité à reconnaître un deepfake, leurs réponses varient de 1 à 8 (cfr tableau 5). Cela se traduit notamment par la connaissance préalable et la fréquence d'exposition des répondants à ces contenus truqués.

Certains répondants affichent une confiance élevée, autour de 8/10, s'appuyant sur leur expérience. Des répondants déclarent notamment : *“Je dirais 7 parce que j'en vois beaucoup”*, ou encore *“Je les chope tous, enfin je crois”*. Toutefois, pour certains, cette confiance est modérée, entre 4 et 6/10, exprimant une forme d'hésitation constante. Cette hésitation est généralement liée au contexte ou à la qualité du contenu rencontré : *“Ça dépend des moments, de mon niveau de fatigue ...”* ou *“Si le truc est bien fait, ça va être compliqué de le repérer.”* Pour d'autres, la maîtrise du sujet abordé par le deepfake influence aussi la capacité de détection : *“Si c'est un domaine que je connais bien, je peux dire 6.”* Toutefois, un des répondants affiche une grande prudence et se note entre 0 et 1/10, estimant ne jamais avoir été confronté à des deepfakes et n'étant donc pas capable de les identifier.

Tableau 5 : Capacité à discerner un contenu authentique d'un deepfake (/10)

Répondant n°	1	2	3	4	5	6	7	8	9	10	11
Capacité à discerner /10 selon les répondants	6	6	4	7	4,5	8	7	0-1	5	5	7
Capacité à discerner /10 selon le test d'awareness	3,33	3,33	6,67	6,67	6,67	10	3,33	6,67	10	3,33	6,67

Dans le cadre de ma recherche, je cherchais à évaluer la capacité réelle des répondants à distinguer un contenu authentique d'un deepfake. L'objectif était de confronter leurs perceptions à leurs performances réelles. Ils m'ont notamment expliqué les méthodes qu'ils pensent utiliser pour repérer un deepfake, mais il s'agit aussi de comprendre comment ils y parviennent concrètement.

Le tableau 5 compare alors l'autoévaluation des répondants concernant leur capacité à identifier un deepfake (notée sur 10) avec leur performance réelle mesurée par le biais du test d'awareness, noté selon la même échelle. On observe que plusieurs répondants tendent à surestimer leur capacité de détection. Par exemple, les répondants 1 et 7 se donnent une note de 6 ou 7/10, alors que leur score réel tourne autour de 3,33/10. A l'inverse, les répondants 3 et 4 se donnent une note plutôt basse, mais réussissent bien au test avec une note de 6,67/10. Ces écarts suggèrent alors un décalage entre perception et réalité. Les individus peuvent être inconscients de leurs compétences réelles, dans le bon sens comme dans le mauvais.

Tableau 6 : Résultats du test d'awareness

	Image 1 : Vraie	Image 2 : Deepfake	Image 3 : Deepfake	Total de bonnes réponses
				
Répondant 1	Deepfake	Vraie	Deepfake	1/3
Répondant 2	Deepfake	Vraie	Deepfake	1/3
Répondant 3	Deepfake	Deepfake	Deepfake	2/3
Répondant 4	Vraie	Vraie	Deepfake	2/3
Répondant 5	Deepfake	Deepfake	Deepfake	2/3
Répondant 6	Vraie	Deepfake	Deepfake	3/3
Répondant 7	Deepfake	Vraie	Deepfake	1/3
Répondant 8	Deepfake	Deepfake	Deepfake	2/3
Répondant 9	Vraie	Deepfake	Deepfake	3/3
Répondant 10	Deepfake	Vraie	Deepfake	1/3
Répondant 11	Deepfake	Deepfake	Deepfake	2/3
Total	3/11	6/11	11/11	

De plus, je souhaitais vérifier si les résultats de mon étude allaient dans le sens de ceux exprimés par Vaccari et Chadwick (2020), démontrant que près de la moitié des participants pouvaient être trompés par un deepfake, même s'ils disposent d'informations contextuelles. Plusieurs observations issues du test d'awareness viennent alors confirmer ces résultats, mais également les nuancer.

Premièrement, lorsque les participants ont été confrontés à l'image 2 (cfr tableau 6), étant un deepfake, 6 répondants sur 11 sont parvenus à l'identifier correctement. Tandis que 5 d'entre eux l'ont identifiée comme une image authentique. Cette répartition presque équitable rejoint alors les résultats de Vaccari et Chadwick. Cependant, cette tendance ne s'est pas répétée pour l'image 3 (cfr tableau 6), étant elle aussi un deepfake. Cette fois, l'ensemble des répondants, sans exception, l'a identifiée correctement. Cela peut notamment s'expliquer par la qualité inférieure de l'image, perçue comme moins réaliste. Certains répondants déclarent : *“Je trouve que les gens ont l'air moins réel”* ou encore *“La tête d'Emmanuel Macron est disproportionnée et même l'alignement avec son corps n'est pas correct.”* En effet, des détails visuels moins convaincants ont facilité la détection, ce qui montre que la qualité du deepfake joue un rôle crucial pour sa crédibilité. C'est pourquoi, cela conduit à nuancer l'hypothèse initiale. Si une part importante de la population étudiée peut être dupée, cela dépend fortement du niveau de réalisme du deepfake.

Par ailleurs, un autre élément marquant qui ressort du test d'awareness est la manière dont les participants ont interprété l'image 1 (cfr tableau 6). Celle-ci, bien qu'authentique, a semé le doute auprès des participants. En effet, 8 participants sur 11 l'ont identifiée comme étant un deepfake. Ce résultat suggère alors que l'exposition à de potentielles images truquées tend à générer un climat de doute constant, ce qui pourrait se répéter sur les réseaux socio-numériques. Toutefois, cela pourrait être accentué par le cadre du test, incitant les participants à douter systématiquement de ce qui leur est présenté.

Enfin, les justifications fournies par les répondants pour justifier leurs choix démontrent de manière évidente l'influence de leurs connaissances préalables et de leurs représentations personnelles sur leur jugement. L'image 1 (cfr tableau 6) en est un bon exemple. Tous les participants ont mobilisé des stéréotypes ou opinions concernant Emmanuel Macron et Donald Trump. Certains ont estimé que les voir rire ensemble ne correspondait pas à l'image

sérieuse qu'ils se faisaient des présidents, comme en témoignent ces extraits : *“Ça ne me semble pas très professionnel, donc je dirais que c'est faux”, ou encore “ils ne se permettraient pas ça à un rendez-vous.”* A l'inverse, d'autres répondants, en s'appuyant sur les stéréotypes associés à ces figures politiques, ont estimé qu'une telle scène pouvait parfaitement se produire. L'un d'entre eux affirme notamment : *“ce sont deux bouffons, donc ça pourrait être vrai.”*

Cette manière dont les convictions antérieures influencent l'évaluation d'une image renvoie directement au biais de confirmation, mais aussi au modèle de codage/décodage proposé par Stuart Hall (Nelson, J. 2022) (Hall, S. 1980). Ces théories suggèrent que les individus interprètent les messages, y compris visuels, à travers leurs expériences, leurs croyances et leurs représentations sociales. De cette manière, il est essentiel de souligner que la reconnaissance d'un deepfake ne dépend alors pas uniquement de son réalisme visuel, mais aussi du bagage idéologique de celui qui le perçoit.

5. Limites de la recherche :

Comme toute recherche scientifique, cette étude présente certaines limites méthodologiques qu'il convient de souligner, pouvant potentiellement nuancer les résultats et leurs interprétations.

Tout d'abord, des contraintes techniques et financières ont influencé la création de mon matériel de recherche utilisé dans le cadre du test d'awareness. Bien que la méthodologie initiale prévoyait l'utilisation de deepfakes au format photo et vidéo, cela n'a pas pu être mis en œuvre. En effet, en raison de ressources matérielles limitées et d'un manque de compétences techniques nécessaires à la création de vidéos par l'intelligence artificielle, le matériel de recherche est composé uniquement de photos. De même, dû à ce manque de compétences techniques, la qualité des photos est parfois inférieure à celle des deepfakes rencontrées sur les réseaux socio-numériques. De plus, plusieurs participants ont mentionné le fait que les photos truquées semblaient plus difficiles à détecter que les vidéos. Il aurait alors été particulièrement intéressant de tester les deux types de formats afin de confronter cette perception. D'autant plus que Meskin (2008) souligne la confiance historique accordée à la vidéo et à la photo.

Ensuite, le recrutement des participants par convenance représente une autre limite importante dans cette recherche. Les personnes interrogées appartiennent majoritairement à un environnement social, culturel et géographique relativement semblable, provenant de mon réseau personnel Facebook. Cela ne garantit pas une réelle diversité des profils. Les résultats obtenus peuvent donc ne pas représenter avec précision la population belge étudiée. Par ailleurs, il est également possible que certains participants, du fait de cette proximité avec la chercheuse, aient inconsciemment cherché à bien faire ou à donner des réponses qu'ils estiment souhaitables dans le cadre de la recherche. En conséquence, cela peut altérer l'authenticité de leurs réponses.

Enfin, bien que la saturation des données qualitatives ait été atteinte, la composition de l'échantillon présente un certain déséquilibre. Les catégories d'âge les plus représentées sont les 18-25 ans et les 46-60 ans. C'est pourquoi, nous ne pouvons pas écarter le fait qu'une présence plus marquée des répondants âgés de moins de 18 ans ou de 26-45 ans aurait potentiellement permis de faire émerger de nouvelles perspectives.

6. Discussion et conclusion :

Pour rappel, la méthodologie adoptée combinant un entretien individuel et un test d'awareness avait pour objectif de confronter les réactions "théoriques" des utilisateurs face aux deepfakes à leurs réactions réelles. Cette approche visait à apporter des éléments de réponse à ma question de recherche. Une question de recherche articulée autour de deux axes.

***QR1** : Quelles réactions suscitent les deepfakes chez les utilisateurs des réseaux socio-numériques (RSN) âgés entre 16 et 60 ans ?*

Les deepfakes suscitent des réactions cognitives, affectives et conatives chez l'ensemble des utilisateurs. Tout d'abord, les résultats de cette recherche démontrent que les utilisateurs des réseaux socio-numériques ont une connaissance partielle du phénomène et de son ampleur. S'ils sont capables d'en identifier les grands principes, ils en ignorent les mécanismes techniques et confondent parfois les deepfakes à d'autres formes de manipulations numériques. Cette incompréhension est davantage marquée chez les utilisateurs plus âgés ou moins actifs sur ces plateformes. Toutefois, cette méconnaissance n'empêche pas une prise de conscience des dangers, en particulier en matière de manipulation de l'opinion publique. Par exemple, la création de faux propos à l'encontre d'une personnalité politique, comme le souligne Langguth (2021). Cette conscience des risques provoque alors chez les utilisateurs un sentiment d'inquiétude, de méfiance, de doute, mais parfois également de résignation, face à un phénomène qui prend de l'ampleur et qu'ils perçoivent comme inévitable.

Cependant, malgré cette méfiance, peu d'actions concrètes sont mises en place pour s'en protéger. La vigilance évoquée par certains des utilisateurs reste intuitive et n'est pas accompagnée systématiquement de véritables pratiques de vérification. De plus, les utilisateurs continuent d'accorder une forte confiance aux visuels, les percevant comme des preuves irréfutables, ce qui est problématique à l'ère des deepfakes (Meskin, A. 2008).

En revanche, une forme de prudence se présente quant au partage d'informations. Beaucoup affirment ne plus relayer de contenus sur leurs réseaux sociaux ou auprès de leur entourage. Cela par peur de propager involontairement de fausses informations.

Enfin, cette recherche révèle que les utilisateurs ne sont pas pleinement conscients de leur propre capacité à détecter un deepfake. Si certains se sous-estiment, générant du doute excessif. D'autres, à l'inverse, se surestiment, ce qui représente un risque encore plus préoccupant. Ceux qui croient pouvoir facilement repérer un contenu truqué peuvent être d'autant plus vulnérables. Ils risquent de croire à de fausses informations en pensant qu'ils auraient forcément détecté celles-ci.

QR2 : Quelles en sont les répercussions sur la réception, la transmission, l'interprétation et la crédibilité des messages ?

Cette incertitude causée par la difficulté à appréhender le phénomène et son ampleur, ainsi que les difficultés à évaluer sa propre capacité de détection, peut avoir des répercussions importantes sur l'ensemble du processus de communication.

Premièrement, l'incertitude affecte directement la réception des messages. D'une part, le doute permanent qu'instaurent les deepfakes, plus que le contenu lui-même, bouleverse la manière dont l'information est reçue. Cela valide pleinement le constat de Simon et al. (2023) suggérant que ce n'est pas seulement le contenu qui influence la réception du message, mais le climat de doute qu'il installe. Certains utilisateurs se mettent alors à douter de tout, même de contenus authentiques, menant à une interprétation suspicieuse des messages et freinant leur efficacité.

D'autre part, chez les utilisateurs qui surestiment leur capacité de détection, le danger réside dans l'absence de doute. Ces derniers peuvent involontairement accorder leur confiance à des contenus truqués. Cela est particulièrement dangereux lorsque l'on sait que les deepfakes tirent une partie de leur efficacité du biais de vérité illusoire (Nelson, J. 2022). Uniquement parce qu'un contenu est visuellement convaincant ou massivement diffusé, les utilisateurs peuvent le percevoir comme crédible.

Concernant la transmission de l'information, une tendance claire se dégage. Par crainte de relayer de fausses informations, les utilisateurs préfèrent ne plus rien partager. Cette volonté, bien que responsable, limite également la diffusion de contenus authentiques et pertinents, notamment ceux issus des journalistes ou des professionnels de la communication. Dû à cela,

la visibilité de certains messages est alors réduite, affectant alors les principes de communication sur les réseaux socio-numériques.

Pour ce qui est de l'interprétation des messages, cette étape est également fortement influencée par les deepfakes. Les utilisateurs interprètent les messages, non pas de manière neutre ou objective, mais à travers leurs croyances personnelles antérieures, comme l'illustre le biais de confirmation (Nelson, J. 2022). Un contenu conforme à leurs convictions a donc plus de chance d'être cru, même s'il est faux. De plus, les réactions observées chez les utilisateurs des réseaux socio-numériques confirment la pertinence du modèle de codage/décodage développé par Stuart Hall (1980). Les individus ne reçoivent pas nécessairement les messages comme ils ont été codés/pensés, mais les décodent en fonction de leur expérience, de leur bagage idéologique et de leur environnement social. Cette dimension idéologique accentue alors le risque que les deepfakes ciblent des stéréotypes ou des croyances, les renforcent et augmentent la diffusion de fausses informations au sein des réseaux socio-numériques.

Enfin, en matière de crédibilité des messages, les réactions obtenues confirment les travaux de Kazadi (2024). Comme il l'avait prédit, la confiance envers les réseaux sociaux est fragilisée et remise en question, notamment en raison de la cohabitation entre sources fiables et non fiables sur ces derniers (Naffi, N et al. 2021). Les utilisateurs en deviennent méfiants envers les contenus qui circulent, car ils le savent, ils peuvent être partagés par n'importe qui (Chavalarias, D. 2018). Cela brouille alors les repères. Cependant, cette fragilisation de la relation de confiance envers les sources d'information ne s'applique pas totalement aux médias traditionnels (également présents en ligne). La crédibilité des messages des médias traditionnels n'est pas totalement remise en cause. Les utilisateurs continuent à leur accorder une certaine confiance, tout en soulignant parfois leur manque de neutralité. Cela suggère alors que, pour les utilisateurs, la confiance envers les médias traditionnels en ligne est davantage liée à leur manque de neutralité, influencée par leurs opinions ou leurs intérêts, plutôt qu'à la possible présence de fausses informations dans leurs contenus. En ce sens, la thèse selon laquelle si tout peut être fabriqué, tout peut être remis en question, y compris la vérité, trouve une signification (Langguth, J. 2021).

Les deepfakes installent bien un climat de confusion, mais dans le cas des journalistes, la vérité n'est pas remise en question, simplement la manière dont elle est racontée.

Conclusion générale :

En conclusion, les deepfakes suscitent des réactions diverses chez les utilisateurs, bouleversant le processus de communication. De la prise de conscience au doute, ou de l'indifférence à l'action limitée, ces réactions ont des conséquences directes sur la manière dont l'information est reçue, partagée, interprétée et jugée crédible.

Si certains concepts théoriques sont vérifiés, comme le biais de confirmation, l'effet de vérité illusoire, le modèle de codage/décodage et la confiance aux visuels. Certains sont à nuancer, comme la rupture de confiance envers les médias traditionnels. Les utilisateurs des réseaux socio-numériques ne sont ni totalement dupes, ni totalement vigilants. Cela est d'autant plus flagrant pour les audiences plus âgées, moins averties, ou moins actives sur les réseaux sociaux.

C'est pourquoi, les résultats confirment l'urgence d'une éducation aux médias dès le plus jeune âge (Kazadi, J.D. 2024), incluant la sensibilisation aux biais cognitifs et l'incitation au fact-checking (Borel, B. 2017). De cette manière, les utilisateurs des réseaux socio-numériques seraient armés au mieux pour lutter contre la propagation des deepfakes et surtout contre la désinformation et la manipulation de l'opinion publique.

Pour aller plus loin dans cette réflexion, il serait judicieux de s'intéresser aux pratiques mises en place par les professionnels de la communication et des journalistes afin d'assurer l'efficacité et la crédibilité de leurs messages, à l'ère des deepfakes. En effet, dans un environnement numérique fortement bouleversé par les deepfakes, marqué par la rupture de confiance et un climat de doute généralisé, ils sont contraints d'adapter leurs pratiques et stratégies de communication.

7. Bibliographie

Beridze, I., Butcher, J. (2019). When seeing is no longer believing. *Nature Machine Intelligence* 1, 332-334 <https://doi.org/10.1038/s42256-019-0085-5>

Bodker, H. (2018). Stuart Hall's encoding/decoding model and the circulation of journalism in the digital landscape. *Routledge eBooks* (p-57-71). <https://doi.org/10.4324/9781315158549-7>

Borel, B. (2017). Fact-Checking won't save us from fake news. https://online225.psych.wisc.edu/wp-content/uploads/225-Master/225-UnitPages/Unit-02/Borel_FiveThirtyEight_2017.pdf

Bismuth, C. (2023). Médias synthétiques. *Google Books*. https://books.google.be/books?hl=fr&lr=&id=GJ_KEAAAQBAJ&oi=fnd&pg=PP1&dq=les+deepfakes+comme+outils+cr%C3%A9atifs&ots=RMTIpsa4DR&sig=BT7oJqdiyufChm8Ap7B1UstKY1c&redir_esc=y#v=onepage&q&f=false

Blavity Inc. (2024). The power of micro-moments seizin Gen Z's limited attention for maximum impact <https://blavityinc.com/the-power-of-micro-moments-seizing-gen-zs-limited-attention-for-maximum-impact/#:~:text=Instant%20Gratification%20in%20Micro%2DMoments,expectations%20for%20rapid%2C%20impactful%20interactions.>

Cathala, H-P. (s.d.). Le temps de la désinformation. *Google Books*. https://books.google.be/books?hl=fr&lr=&id=mOtXDwAAQBAJ&oi=fnd&pg=PT3&dq=d%C3%A9sinformation&ots=FvVjgl9zq&sig=OUJYjQtjzcWDmI0ymgPzaundlGY&redir_esc=y#v=onepage&q=d%C3%A9sinformation&f=false

Chavalarias, D. (2018). Au-delà des « fake news » : à l'ère numérique, nos démocraties doivent évoluer pour ne pas mourir. *Hal*. <https://hal.science/hal-01924436/>

Dauphin, F. (2019). Les fake news au prisme des théories sur les rumeurs et la propagande. Open Editions Journals, 53, 15-32. <https://doi.org/10.4000/edc.9132>

De Abreu, G. (2024). Qu'est-ce que la saturation des données ? Comprendre son utilisation dans la recherche qualitative. Mind the Graph Blog. <https://mindthegraph.com/blog/fr/quest-ce-que-la-saturation-des-donnees/>

Deschamps, C. (2021). Les technologies du faux : un état des lieux. Academia. [https://d1wqtxs1xzle7.cloudfront.net/82058710/Les technologies du faux-libre.pdf](https://d1wqtxs1xzle7.cloudfront.net/82058710/Les_technologies_du_faux-libre.pdf)

De Ruiter, A. (2021). The distinct wrong of Deepfakes. Philosophy & Technology, 34, 1311-1332. <https://doi.org/10.1007/s13347-021-00459-2>

Dickerman, L. (2000). Camera Obscura : Socialist Realism in the Shadow of Photography. JStor, 139-153 <https://www.jstor.org/stable/779160>

Fabre, M. (2022). Urdoxa, le deepfake d'information comme usage tactique du vague. Interfaces numériques, 11(2). <https://doi.org/10.25965/interfaces-numeriques.4863>

Frau-Meigs, D. (2019). Faut-il avoir peur des fake news ? Google Books. https://books.google.be/books?hl=fr&lr=&id=jiLyEAAQBAJ&oi=fnd&pg=PA3&dq=cons%C3%A9quences+des+fake+news+sur+le+processus+de+communication&ots=qf4SZyId4V&sig=wSj4dx_Hj_upPWdM6C4DqVkVFdA&redir_esc=y#v=onepage&q&f=false

Gadras, S. (2022). 2000-2020 : âge critique du journalisme ? Les transformations contemporaines de la profession. Le manuel de journalisme. <https://hal.science/hal-04773976/>

Gamba, F. (s. d.). Fake news : un univers complexe et déroutant. CFR : TANGRAM 45. <https://www.ekr.admin.ch/publications/f850.html>

Hall, S. (1980). Encoding—Decoding. Routledge eBooks (p.44-55)
<https://doi.org/10.4324/9780367809195-6>

Ifrée. (2023) Fiche ressource “Attitude, comportements et dissonance cognitive.”
https://publication.ifree.asso.fr/wp-content/uploads/2023/06/2_FicheIfree_Attitude_comportement_dissonance_v.pdf

Joffe, H. (2007), Le pouvoir de l’image : persuasion, émotion et identification. Cairn (p.102-115) https://shs.cairn.info/article/DIO_217_0102?tab=texte-integral

Kazadi, J. D. (2024). Les sources d’information journalistiques face aux deepfakes. Communication Technologies et Développement, 16. <https://doi.org/10.4000/12nfh>

Langguth, J., Pogorelov, K., Brenner, S., Filkuková, P., & Schroeder, D. T. (2021). Don’t Trust Your Eyes : Image Manipulation in the Age of DeepFakes. Frontiers In Communication <https://www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2021.632317>

Laville, C. (2022) Le journaliste, un professionnel de la transparence ? Cairn (p.29-36) <https://shs.cairn.info/revue-sens-dessous-2017-2-page-29>

Le Calme, S. (2024) La pandémie de la désinformation : 1,5 milliard de fake news sont publiées chaque jour sur les médias sociaux, selon ID Crypt Global. <https://www.developpez.com/actu/356788/La-pandemie-de-la-desinformation-1-5-milliard-de-fake-news-sont-publiees-chaque-jour-sur-les-medias-sociaux-selon-ID-Crypt-Global-Facebook-et-X-hebergent-le-plus-grand-nombre-de-contenus-de-desinformation/>

Mark Zuckerberg Harvard Speech (s.d.). [Vidéo]. Facebook. <https://www.facebook.com/Benzinga/videos/456953483417290>

Mercator - Dunod (s.d) Ressources de formation marketing à l’ère de la data et du digital [https://www.mercator.fr/lexique-marketing-definition-attitude#:~:text=On%20distingue%20trois%20composantes%20d,conative%20\(tendance%20%C3%A0%20agir](https://www.mercator.fr/lexique-marketing-definition-attitude#:~:text=On%20distingue%20trois%20composantes%20d,conative%20(tendance%20%C3%A0%20agir)

Meskin, A., Cohen, J (2008) Photography and philosophy : essays on the pencil of nature. Google Scholar <https://aardvark.ucsd.edu/perception/agnosticism.pdf>

Missika, J-L (s.d) Le business de la haine. Google Books. https://books.google.be/books?hl=fr&lr=&id=Z3pZEAAQBAJ&oi=fnd&pg=PT2&dq=com bien+de+deepfakes+sur+les+r%C3%A9seaux+sociaux&ots=zGeNyUd4yU&sig=1k9Gftzf08piJsVX5m-J8q2aNjk&redir_esc=y#v=onepage&q&f=false

Naffi, N., Davidson, A.-L., & Berger, F. (2021). Infocalypse : la propagation des hypertrucages menace la société. Research Gate. https://www.researchgate.net/profile/Nadia-Naffi/publication/362542334_Infocalypse_la_propagation_des_hypertrucages_menace_la_societe/links/62f7ed88c6f6732999ca1feb/Infocalypse-la-propagation-des-hypertrucages-menace-la-societe.pdf

Nelson, J. (2022). Fake news et deepfakes : une approche cyberpsychologique. Interfaces Numériques, 11(2). <https://doi.org/10.25965/interfaces-numeriques.4830>

Pouillet, Y. (2021). La régulation des réseaux sociaux. Cairn. https://www.cairn.info/revue-etudes-2021-6-page-19.htm?WT.tsrc=cairnEmailAlert&WT.mc_id=ETU_4283

Robichaud-Durand, S. (2023). L'hypertrucage : analyse du phénomène des « deepfakes » et recommandations. Lex Electronica,. <https://doi.org/10.7202/1108807ar>

Saleh, I., Bouhaï, N., & Leleu-Merviel, S. (2021). H2PTM'21 : Information : enjeux et nouveaux défis. Google Books https://books.google.be/books?hl=fr&lr=&id=2-FMEAAQBAJ&oi=fnd&pg=PA81&dq=manipulation+par+l%27utilisation+des+fake+news+ou+deep+fakes&ots=go5SB7BCu-&sig=XBOJ86oMEPoyIt8fSq_aK7IuCp8&redir_esc=y#v=onepage&q&f=false

Simon, F. M., Altay, S., & Mercier, H. (2023). Misinformation reloaded ? Fears about the impact of generative AI on misinformation are overblown. *Misinformation Review*. <https://doi.org/10.37016/mr-2020-127>

Stöcker, C. (2020). How Facebook and Google accidentally created a perfect ecosystem for targeted disinformation. Dans *Lecture Notes in Computer Science* (p. 129-149). https://doi.org/10.1007/978-3-030-39627-5_11

Stromback, J. (2022), Knowledge resistance in high-choice information environments. Routledge eBooks. <https://doi.org/10.4324/9781003111474>

Vaccari, C., Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society*, 6 (1) <https://doi.org/10.1177/2056305120903408>

Walden University. (2022). How age influences social media preferences. Walden University. <https://www.waldenu.edu/programs/business/resource/how-age-influences-social-media-preferences>

Watson, J (2025). Macron et Trump ravivent leur « bromance » au milieu de tensions stratégiques sur l'Ukraine. *France 24*. <https://www.france24.com/fr/am%C3%A9riques/20250225-macron-et-trump-ravivent-leur-bromance-en-pleine-tensions-strat%C3%A9giques-sur-l-ukraine>

We're increasing transparency on ads. (s. d.). [Vidéo]. Instagram. <https://www.instagram.com/p/ByaVigGFP2U/>

8. Annexes

8.1. Déclaration sur l'honneur concernant l'IA

Je déclare qu'il s'agit d'un travail original et personnel, que toutes les sources référencées ont été indiquées dans leur totalité et ce, quelle que soit leur provenance, et que l'utilisation des outils d'intelligence artificielle est conforme aux consignes d'utilisation publiées par la Faculté ESPO et disponibles à cette adresse : <https://intranet.uclouvain.be/fr/myucl/facultes/espo/intelligence-artificielle-consignes-d-utilisation-pour-les-etudiant-es.html>

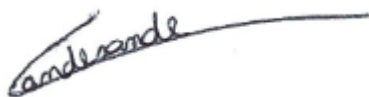
Je suis consciente que le fait de ne pas citer une source, de ne pas la citer clairement et complètement, de ne pas annoncer dans une partie spécifique en début du mémoire l'utilisation que j'ai faite des outils d'intelligence artificielle ou de les utiliser de manière contraire à ce qui est prévu dans la note facultaire constituent des irrégularités graves au regard du *Règlement des études et des examens de l'Université catholique de Louvain* (Chapitre 4, section 7, article 107 à 114) au sein de l'Université. J'ai notamment pris connaissance des risques de sanctions administratives et disciplinaires encourues.

Au vu de ce qui précède, je déclare sur l'honneur ne pas avoir commis de plagiat ou toute autre forme de fraude et de n'avoir pas utilisé les outils d'intelligences artificielles en dehors de ce qui est accepté à l'UCLouvain.

Nom, Prénom : Van de Sande Laura

Date : 30/05/2025

Signature de l'étudiant-e :



8.2. Première ébauche - Guide d'entretien

Introduction :

Bonjour, merci d'avoir participé à cet entretien. Je m'appelle Laura Van de Sande et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux sociaux numériques.

Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes et confidentielles. Bien entendu, vous êtes libre, à tout moment, de ne pas répondre à certaines questions ou encore de mettre fin à l'entretien si vous en ressentez le besoin.

Avez-vous des questions avant de commencer ?

Consentement :

Avant de commencer, puis-je enregistrer notre conversation à des fins d'analyse ?

Grille de questions :

Thème 1 : Introduction - Connaissance générale

Avez-vous déjà entendu parler des deepfakes ? Si oui, pouvez-vous expliquer ce que vous en savez ?

Quand et comment avez-vous découvert les deepfakes pour la première fois ?

Où avez-vous principalement entendu parler des deepfakes ?

Avez-vous déjà été confrontée à des avertissements, des campagnes de sensibilisation sur les deepfakes ?

Thème 2 : Perceptions et réactions face aux deepfakes

Quelle a été votre première réaction lorsque vous avez pris connaissance de l'existence des deepfakes ?

Avez-vous déjà rencontré une image ou une vidéo sur les réseaux socio-numériques que vous soupçonniez d'être créée de toutes pièces ? Si oui, comment avez-vous réagi ?

Pouvez-vous décrire la situation ?

A l'inverse, avez-vous déjà cru à une image/vidéo créée de toutes pièces ? Si oui, comment avez-vous réagi face à cette découverte ?

Thème 3 : Comportements face aux deepfakes

Comment faites-vous pour essayer de déterminer si une vidéo ou une image est un deepfake ?

Avez-vous des techniques spécifiques que vous utilisez pour vérifier l'authenticité des vidéos et des images en ligne ?

Lorsque vous voyez une nouvelle vidéo ou image en ligne, vérifiez-vous systématiquement la source ? Si oui, comment procédez-vous ?

Avez-vous changé vos habitudes de consommation d'informations en raison des deepfakes ?

Quels critères utilisez-vous pour évaluer la crédibilité d'une publication ou d'une information en ligne ?

Quelle est votre confiance dans votre capacité à identifier un deepfake ?

Quelles sources considérez-vous comme fiables pour vérifier les informations ?

Considérez-vous les réseaux socio-numériques comme une source fiable ?

Avez-vous déjà, sans le savoir, partagé une fausse information ? Si oui, comment cela est-il arrivé, et comment avez-vous réagi ?

Thème 4 : Impact des deepfakes sur les audiences

Etes-vous préoccupé par l'impact potentiel des deepfakes sur la société ? Pourquoi ?

Quels sont, selon vous, les dangers potentiels des deepfakes ?

Avez-vous changé vos habitudes de consommation d'informations en raison des deepfakes ?

Avez-vous pris des mesures spécifiques pour vous protéger contre les fausses informations (deepfakes) ?

Les deepfakes ont-ils affecté votre confiance dans les médias numériques et les institutions ? Comment ?

Les deepfakes ont-elles influencé votre opinion sur les technologies numériques en général ?

Pensez-vous que les gens sont suffisamment informés sur les deepfakes et leurs dangers ?

Thème 5 : Conclusion

Y a-t-il quelque chose que vous aimeriez ajouter à ce sujet ?

8.3. Test de la grille de questions

Répondant	Âge	Genre
Numéro 1	23 ans	Femme

I : Avez-vous déjà entendu parler des deepfakes ? Si oui, pouvez-vous expliquer ce que vous en savez ?

R : Oui, la première fois que j'ai vu, je pense, c'était sur TikTok. C'était une image de Tom Cruise qui avait été modifiée, il disait n'importe quoi. J'ai aussi parlé de ça dans un cours. C'était un cours de technologie et communication, je crois.

I : Quand et comment avez-vous découvert les deepfakes pour la première fois ?

R : Je pense que j'y ai déjà répondu.

I : Où avez-vous principalement entendu parler des deepfakes ? (réseaux sociaux numériques, actualités, bouche-à-oreille, etc.)

I : J'ai appris l'existence des deepfakes par les réseaux sociaux et le bouche-à-oreille. Mais on m'a expliqué vraiment ce que c'était en cours, quoi. Je ne connaissais pas le terme précis avant.

I : Avez-vous déjà été confrontée à des avertissements, des campagnes de sensibilisation sur les deepfakes ?

R : Non, je ne pense pas, peut-être les petites mentions en dessous d'une publication. Exemples : les avertissements sur TikTok qui mettent en garde contre les deepfakes, comme

les alertes “contenus sensibles” par exemple. Ou encore dans la description de certains contenus. Mais je ne pense pas qu’il y ait beaucoup de sensibilisation là-dessus, en général.

I : Quelle a été votre première réaction lorsque vous avez pris connaissance de l’existence des deepfakes ?

R : C’est dingue ce qu’on peut faire avec la technologie ! Vraiment, le fait de pouvoir imiter des gens, etc., on se croirait dans un film de science-fiction.

I : Avez-vous déjà rencontré une image ou une vidéo sur les réseaux sociaux numériques, que vous soupçonniez d’être créée de toutes pièces ? Si oui, comment avez-vous réagi ? Pouvez-vous décrire la situation ?

R : Oui, oui, parfois, quand je vois des célébrités parler, par exemple sur TikTok. Je me dis que c’est bizarre. En fait, je me demande si c’est vraiment eux, ou pas.

I : Pourquoi tu te poses cette question ?

R : Parce que déjà, il n’y a pas de compte certifié/vérifié. Donc ça laisse un gros doute. Ou alors, parfois, ce sont des tout nouveaux comptes, ce qui paraît vraiment bizarre. Et puis le contenu de leurs paroles n’est pas cohérent avec ce qu’ils ont l’habitude de dire ou avec leur image, leur personne.

I : Comment réagis-tu alors, quand tu penses être face à un deepfake ?

R : Je regarde les commentaires, les contenus du compte. pour voir si ça semble vrai ou pas, et je fais le tri dans ma tête. Je me fais mon avis par rapport à ça en général.

I : A l’inverse, avez-vous déjà cru à une image/vidéo créée de toutes pièces ? Si oui, comment avez-vous réagi face à cette découverte ?

R : A mon avis, oui. C’est sûr !

I : En as-tu une précise en tête ?

R : Si j’y ai cru, je ne le sais pas. Je ne sais pas qu’en fait elle était fausse. Mais c’est sûr que ça a dû arriver. Si on fait pas attention, ou si on ne se renseigne pas un minimum. Tu tombes dans le piège, quoi. Parce que parfois c’est hyper ressemblant et très réaliste. Il y a des trucs, ça se voit à des kilomètres c’est faux. Et d’autres images sont hyper bien faites.

I : Comment faites-vous pour essayer de déterminer si une vidéo ou une image est un deepfake ?

R : J’essaie surtout de regarder le contexte. Je pense par exemple à une vidéo de Macron en pleine période de guerre. Ça n’a aucun sens. Même si la vidéo paraît réaliste et on dirait pas que c’est faux, si on le met dans le contexte, jamais Macron n’aurait ce discours.

I : Tu regardes alors plus en profondeur ce qui est montré dans la vidéo ?

R : Oui, c'est plus par rapport au contexte et à la réalité. C'est grâce à ça que je fais le lien, pour voir si c'est vrai.

I : Avez-vous des techniques spécifiques que vous utilisez pour vérifier l'authenticité des vidéos et des images en ligne ?

R :

I : Par exemple une méthode opératoire que tu fais à chaque fois ?

R : Vérifier le compte de la personne qui a publié ça. Voir si c'est "Le Soir". Là c'est déjà plus plausible que si c'est un compte où il est marqué user 292 aha.

I : Lorsque vous voyez une nouvelle vidéo ou image en ligne, vérifiez-vous systématiquement la source ? Si oui, comment procédez-vous ?

R : Non du tout. Je vais plus vérifier si ce sont des personnalités importantes que si ce sont des personnes lambdas.

I : Avez-vous changé vos habitudes de consommation d'informations en raison des deepfakes ?

R : Non, comme avant. Au pire, je crois juste à quelque chose et c'est pas mal. Mais ça va pas je pense changer ma vie, aha.

I : Quels critères utilisez-vous pour évaluer la crédibilité d'une publication ou d'une information en ligne ?

R : Je regarde l'origine de la vidéo ou de l'image. Si c'est un repost ou non. Je me demande qui a posté ça ? Je regarde ce qui est dit dedans, les commentaires. Je fais une petite analyse si j'ai un gros doute.

I : Quelle est votre confiance dans votre capacité à identifier un deepfake ?

R : Oui, il y a des deepfakes qu'on sait reconnaître plutôt facilement si on a le contexte, etc. On sait que sur internet, il y a plein de deepfakes et de fake news, et on sait qu'il ne faut pas croire à tout sur les réseaux sociaux. Donc, à partir de ce moment-là, on est déjà plus méfiant. Je me mettrai donc 7 ou 8 sur 10 peut-être.

I : Quelles sources considérez-vous comme fiables pour vérifier les informations ? Considérez-vous les réseaux sociaux numériques comme une source fiable.

R : En premier, les journaux traditionnels qui sont présents en ligne. Par exemple, Le Soir, L'avenir, etc. Les autres, je ne sais pas vraiment s'ils disent la vérité. Tous les utilisateurs sont comme nous. Ils peuvent dire ce qu'ils veulent en faisant croire que c'est vrai.

I : Et les réseaux sociaux, sont-ils fiables selon toi ?

R : Oui et non. Cela dépend de qui l'info vient.

I : Avez-vous déjà, sans le savoir, partagé une fausse information ? Si oui, comment cela est-il arrivé et comment avez-vous réagi ?

R : Non, car je ne repartage jamais rien sur les réseaux sociaux.

I : Même en parler autour de toi ? Par exemple, que tu aurais parlé avec ton entourage de l'information, et puis, plus tard, tu as appris que c'était totalement faux ?

R : C'est possible, mais je ne saurais pas dire. Si oui, je ne m'en suis pas rendu compte. Mais c'est totalement possible que ce soit arrivé. Par exemple, je crois très facilement aux poissons d'avril et à ce genre de choses. Des choses comme ça je peux vite tomber dedans si c'est pas gros comme une maison.

I : Êtes-vous préoccupé par l'impact potentiel des deepfakes sur la société ? Pourquoi ?

R : A petite échelle, oui. Je me dis que les gens vont croire tout ce qu'on dit sur les réseaux sociaux. Et après, ils vont le dire aux autres, cela va se propager. Et ça va faire effet boule de neige. Par exemple, si on parle un peu politique. Ils vont croire des choses qui sont dites, alors que ce n'est pas vrai. Et donc, ils vont peut-être changer d'avis ou d'opinion. Et ça peut avoir un grand impact sur les résultats. Je pense par exemple à Donald Trump qui a été victime de deepfakes.

I : Quels sont, selon vous, les dangers potentiels des deepfakes ?

R : C'est une question difficile ... De croire à des choses qui ne sont pas vraies et qui influencent la vie réelle. Sur les choix de la vie, les opinions, les comportements, etc. Par exemple, encore une fois, pour les élections. Cela a plus d'impact, je pense, dès que cela touche à des célébrités ou à des politiques. Des gens connus, en fait. Car je ne vois passer des deepfakes qu'à leur sujet. Je n'en vois pas sur des gens lambdas, on va dire.

I : Avez-vous ...

R : Des fois, aussi le danger, c'est, pardon, haha, de faire peur aux gens. Il y a parfois de la manipulation.

I : Avez-vous changé vos habitudes de consommation d'informations en raison des deepfakes ?

R : Non pas du tout.

I : Avez-vous pris des mesures spécifiques pour vous protéger contre les fausses informations (deepfakes) ?

R : Je me dis juste qu'il ne faut pas croire tout ce qu'on dit. Tant que cela n'est pas dit par une source vraiment fiable, je ne vais pas y croire. Si c'est sur Le soir, oui, par exemple. Mais si c'est une personne lambda... Je vais quand même vérifier. Car parfois, juste les gens crient au loup et, en fait, c'est faux.

I : Les deepfakes ont-ils affecté votre confiance dans les médias numériques et les institutions ? Comment ?

R : Non du tout. Car je me dis que, par exemple, Le soir, ce sont des journalistes. Ils ont un diplôme de journaliste, ils ont les compétences pour vérifier les informations un minimum avant de les reposter. Donc non, je fais confiance aux journaux etc. en tout cas.

I : Les deepfakes ont-elles influencé votre opinion sur les technologies numériques en général ?

R : Oui, car les technologies ça peut être super, mais il y a aussi des mauvais côtés. Il faut juste que ça soit utilisé à bon escient et pas n'importe comment. Sinon, tout le monde peut croire à n'importe quoi.

I : Pensez-vous que les gens sont suffisamment informés sur les deepfakes et leurs dangers ?

R : Non, clairement pas. Je pense même que certaines personnes ne savent pas que ça existe ou alors oublient que cela existe. Il y a certains articles, etc. qui en parlent, mais le problème, c'est que les gens doivent les lire. Je trouve qu'on n'en entend pas assez parler à la radio ou à la télé, par exemple. Ce ne sont pas des infos qui ressortent beaucoup. Il faudrait peut-être en parler même, dès le plus jeune âge. Dans les écoles secondaires ou quoi. Les personnes âgées ont aussi besoin qu'on leur en parle, je pense.

I : Y a-t-il quelque chose que vous aimeriez ajouter sur ce sujet ?

R : Il faut faire attention à tout ce qu'on voit sur les réseaux sociaux. Il y a du bon et du mauvais. Donc il faut réussir à peser le pour, ou le contre.

Répondant	Âge	Genre
Numéro 2	47 ans	Homme

I : Avez-vous déjà entendu parler des deepfakes ? Si oui, pouvez-vous expliquer ce que vous en savez ?

R : Oui, cependant je ne connaissais pas le terme deepfake, mais je sais qu'il existe des vidéos de personnalités publiques qui sont déformées, par exemple. Certaines personnes changent leurs paroles ou encore les mettent dans des situations délicates.

I : Quand et comment avez-vous découvert les deepfakes pour la première fois ?

R : Mes enfants m'en ont parlé. J'ai également entendu parler quelquefois de cette pratique à la télévision, au journal parlé.

I : Avez-vous un exemple d'une deepfake dont vous vous souvenez ?

R : Non, je n'ai pas d'exemple précis, mais je sais que cela est fréquemment utilisé sur les personnes politiques ou des célébrités.

I : Avez-vous déjà été confronté à des avertissements, des campagnes de sensibilisation sur les deepfakes ?

R : Non, pas vraiment. Mes enfants m'ont averti à ce sujet et j'en ai entendu parler au journal parlé, mais cela s'arrête là.

I : Quelle a été votre première réaction lorsque vous avez pris connaissance de l'existence des deepfakes ?

R : Je ne me suis pas tout de suite inquiété de ce que cela pouvait avoir comme effet sur les gens. Je me rappelle juste m'être dit que ce qu'il était possible de faire dans cette génération avec la technologie était incroyable.

I : Vous souvenez-vous de la première deepfake que vous avez vue sur les RSN ?

R : Non pas précisément. Mais cela devait avoir un rapport avec une célébrité je suppose.

I : Avez-vous déjà rencontré une image ou une vidéo sur les réseaux sociaux numériques que vous soupçonniez d'être créée de toutes pièces ? Si oui, comment avez-vous réagi ? Pouvez-vous décrire la situation ?

R : Je soupçonne rarement des contenus d'être truqués. Je sais que cela existe, mais il est difficile, je trouve, de les reconnaître. J'ai déjà été confronté à des faux contenus et j'ai su les reconnaître car ils étaient mal réalisés. Sinon, je trouve cela délicat.

I : A l'inverse, avez-vous déjà cru à une image/vidéo créée de toutes pièces ? Si oui, comment avez-vous réagi face à cette découverte ?

R : J'ai déjà cru à de fausses nouvelles, je pense, oui. Comme je n'arrive pas toujours à les reconnaître, je pense que j'ai dû me faire avoir. Mais je ne saurai pas dire à quelles informations j'ai bien pu croire, à tort.

I : Comment faites-vous pour essayer de déterminer si une vidéo ou une image est un deepfake ?

R : Je ne me pose pas souvent la question de si c'est vrai ou faux, mis à part si ce que je vois me paraît trop gros. C'est vrai que lorsque je regarde mon téléphone, je ne fais pas forcément attention à ce qui défile. Mais je pense que si j'ai un doute, je demanderai à mes enfants qui sont bien plus compétents pour cela. Après tout, ils savent tout de la technologie.

I : Qu'entendez-vous par "gros" ?

R : Je parle par exemple, d'un discours improbable, d'une qualité image ou vidéo médiocre alors qu'il s'agit d'un président qui est représenté, etc. Ou encore, parfois, il est noté clairement dans les commentaires de la vidéo que c'est une fausse.

I : Avez-vous des techniques spécifiques que vous utilisez pour vérifier l'authenticité des vidéos et des images en ligne ?

R : Non, à part mon instinct, je dirais. Si je trouve la vidéo bizarre, je me poserai plus de questions. Sinon, je ne vais pas plus chercher que ça.

I : Lorsque vous voyez une nouvelle vidéo ou image en ligne, vérifiez-vous systématiquement la source ? Si oui, comment procédez-vous ?

R : Non, je ne vérifie pas toujours la source. Après, je sais qu'il ne faut pas tout croire sur les réseaux sociaux. Et donc, je lis ce que je vois sur les réseaux sociaux, mais je sais que je dois faire le tri, et me renseigner sur ces sujets, par exemple au journal qui est plus sûr.

I : Avez-vous changé vos habitudes de consommation d'informations en raison des deepfakes ?

R : Non, pas forcément, je me méfie peut-être un peu plus. Mais il n'y a pas de gros changements. Je prends peut-être plus le temps de lire les commentaires quand une photo ou une vidéo me semble bizarre, pour voir ce qu'en disent les autres.

I : Quelle est votre confiance dans votre capacité à identifier un deepfake ?

R : Faible, aha. Comme je l'ai dit, je ne me méfie pas plus que ça. Surtout, je pense que les technologies ne permettent pas toujours d'identifier si c'est vrai ou faux. Les technologies font ça bien et sont très fortes.

I : Qu'est-ce qui rend les technologies "fortes", selon vous ?

R : Je pense que ça va de pire en pire. Elles s'améliorent de jour en jour. Elles sont capables de faire de plus en plus de choses. Elles sont même capables aujourd'hui de faire le travail des hommes.

I : Quelles sources considérez-vous comme fiables pour vérifier les informations ? Considérez-vous les réseaux sociaux numériques comme une source fiable ?

R : Le journal, les journaux/revues locales, etc. Tout ce qui vient d'un organisme public, je dirais.

I : Avez-vous déjà, sans le savoir, partagé une fausse information ? Si oui, comment cela est-il arrivé et comment avez-vous réagi ?

R : Je ne sais pas. Je pense que c'est possible, oui. Je partage et repartage toutes sortes de choses sur les réseaux sociaux. Mais je n'ai jamais volontairement partagé de fausses infos.

I : Êtes-vous préoccupé par l'impact potentiel des deepfakes sur la société ? Pourquoi ?

R : Oui, même si au début j'étais plus fasciné par ce que pouvait faire la technologie, je sais que cela peut avoir des conséquences. Personnellement, je redoute le fait que plus personne ne croit en rien. Aussi, par exemple, pour mes enfants, j'aurai peur qu'ils soient victimes de ces sortes de montage, ce qui pourrait leur créer une mauvaise réputation, du harcèlement, etc.

I : Quels sont, selon vous, les dangers potentiels des deepfakes ?

R : Croire en de mauvaises informations, entacher la réputation de personnes connues ou non, etc. Je ne sais pas trop, mais je suis certain que cela a de mauvais impacts.

I : Avez-vous changé vos habitudes de consommation d'informations en raison des deepfakes ?

R : Comme j'ai dit, je pense que je suis un peu plus méfiant, mais c'est pas maladif. Je me méfie de certaines informations si je m'en méfie ou je doute de la vérité, je n'y prête pas plus attention et continue à faire défiler les contenus sur mon téléphone. Et je garde en tête qu'il ne faut pas tout croire sur Facebook, surtout quand je sais que des contenus peuvent être modifiés.

I : Les deepfakes ont-ils affecté votre confiance dans les médias numériques et les institutions ? Comment ?

R : Cela n'a pas enlevé ma confiance au journal parlé, ou aux journaux papiers, ou encore aux chaînes infos sur internet. Je sais qu'ils ne relayeront jamais de fausses informations, car c'est leur devoir de nous informer.

I : Les deepfakes ont-elles influencé votre opinion sur les technologies numériques en général ?

R : Non, depuis que je sais que certaines machines peuvent remplacer les humains, je sais qu'il y a du bon et du mauvais dans les technologies.

I : Pensez-vous que les gens sont suffisamment informés sur les deepfakes et leurs dangers ?

R : Non, par exemple, moi qui suis déjà plus âgé, je ne suis pas toujours capable de différencier une vraie vidéo d'une fausse, par exemple. Alors j'imagine que, pour les personnes âgées c'est encore plus difficile. Ils ne savent parfois pas bien utiliser leur téléphone, donc cela me paraît périlleux pour eux. Il faudrait peut-être nous apprendre à différencier une vraie vidéo d'une fausse, plutôt que juste nous dire que cela existe.

I : Y a-t-il quelque chose que vous aimeriez ajouter sur ce sujet ?

R : Non rien de spécial

8.4. Test du matériel pour le test d'awareness

J'avais notamment pris pour premier test une vidéo authentique de Mark Zuckerberg : <https://www.facebook.com/Benzinga/videos/456953483417290>, publiée sur sa page Facebook.

Ainsi qu'un deepfake dont il a été victime : <https://www.instagram.com/p/ByaVigGFP2U/>

Répondant	Âge	Genre
Numéro 1	23 ans	Femme

I : Savez-vous reconnaître parmi les deux vidéos, la vidéo qui est truquée ? Savez-vous m'identifier le deepfake ? Je vous laisse quelques minutes pour y regarder à votre aise.

R : Je pense que c'est la deuxième.

I : Vous dites que le deepfake est la deuxième vidéo, c'est bien ça ?

R : Oui, je pense qu'il s'agit de la deuxième vidéo.

I : C'est la bonne réponse. Comment êtes-vous parvenu à identifier le deepfake ? Quels éléments avez-vous pris en compte ?

R : Honnêtement, je pense avoir déjà vu cette vidéo sur mes réseaux. J'ai choisi uniquement comme cela. Je pense l'avoir vue sur TikTok. Et en voyant les vidéos, je trouve que les détails sur le visage de Zuckerberg quand il parle sont étranges comparés à la première vidéo.

Répondant	Âge	Genre
Numéro 2	47 ans	Homme

I : Savez-vous reconnaître parmi les deux vidéos, la vidéo qui est truquée ? Savez-vous m'identifier le deepfake ? Je vous laisse quelques minutes pour y regarder à votre aise.

R : Je pense qu'il s'agit de la deuxième, mais je ne suis pas sûr du tout. Je dis cela un peu au hasard.

I : C'est la bonne réponse. Comment êtes-vous parvenu à identifier le deepfake ? Quels éléments avez-vous pris en compte, même si cela est plus du hasard d'après vous ?

R : Je ne sais pas trop ... La deuxième vidéo me semblait plus fausse que la première. Je ne sais pas trop quoi dire comme éléments qui m'ont aidé. C'est sans doute le fait que la vidéo soit filmée de face et non de profil.

I : Que voulez-vous dire par là ?

R : Et bien, je pense que si cela avait été une vidéo de profil, cela n'aurait pas été possible de la trafiquer. Ou alors beaucoup plus difficile, et du coup cela aurait été moins bien fait peut-être. Je ne sais pas trop.

8.5. Premières observations

Les répondants, bien que familiers avec le concept de contenus truqués grâce aux intelligences artificielles, ne connaissent pas toujours le terme anglais exact des deepfakes. En effet, ces derniers ont découvert le concept par le biais du bouche-à-oreille ou encore les réseaux sociaux. Au départ, intrigués par les capacités des technologies, ils ne se sont pas inquiétés des dangers potentiels que ceux-ci pouvaient avoir.

Avec le temps, même s'ils reconnaissent que ces derniers peuvent avoir des conséquences graves comme la manipulation, la désinformation, la mise à mal d'une réputation ou encore l'influence des opinions. Les répondants n'ont pas changé, pour autant, leurs habitudes de consommation de l'information sur les réseaux socio-numériques. Ils affirment savoir que les deepfakes existent et qu'il ne faut pas croire à tout, mais qu'ils s'informent toujours via ces canaux. Néanmoins, ils affirment s'informer par le biais de canaux officiels comme les radios ou les chaînes d'informations présentes en ligne. En effet, ils ne perdent pas confiance en ces institutions. Selon eux, ce sont des professionnels qui se doivent de dire la vérité et de vérifier leurs sources. A l'inverse, ils savent qu'il ne faut pas croire à ce que disent les utilisateurs lambdas car ils peuvent simplement partager leurs opinions.

Toutefois, ils pensent tous les deux avoir déjà cru à une fausse information et même peut-être l'avoir relayée, sans s'en rendre compte. Ils expliquent cela par le fait qu'ils ne se méfient pas beaucoup et ne vérifient pas toujours leurs sources. De plus, ils trouvent cela extrêmement difficile de différencier les deepfakes des vrais contenus.

Selon eux, les intelligences artificielles sont très compétentes et, malheureusement, les utilisateurs ne sont pas assez sensibilisés à ce sujet. Toutefois, s'ils doutent qu'un contenu soit créé de toutes pièces, ils essayent de regarder notamment ce que disent les autres utilisateurs en commentaires, la qualité du visuel, les propos tenus, le contexte par rapport à l'actualité, ou encore la personne qui a publié l'information. Mais ils affirment que même avec cela, si le deepfake est réalisé parfaitement, ils ne se sentent pas capables de l'identifier et peuvent facilement y croire, sans s'en méfier.

8.6. Formulaire de consentement

Formulaire de consentement - Enregistrement de données

Objet de l'étude :

Vous êtes invité(e) à participer à une collecte de données orales dans le cadre d'une étude menée par Laura Van de Sande, étudiante de l'Université Catholique de Louvain.

Cette étude vise à comprendre les réactions suscitées par les deepfakes auprès des utilisateurs des réseaux socio-numériques, en particulier en ce qui concerne leur perception, leur interprétation et leur comportement face aux informations transmises.

Votre droit de vous retirer à tout moment :

Votre participation à cette étude est volontaire. Vous êtes entièrement libre de participer ou non, et de vous retirer de l'étude à tout moment sans préjudice.

Poser des questions à tout moment :

Vous pouvez poser des questions sur la recherche à tout moment, soit durant l'entretien, soit par courrier électronique à l'adresse suivante laura.vandesande@student.uclouvain.be

Ce qui est attendu de vous :

Si vous acceptez de participer, il vous sera demandé de :

- Lire et signer ce formulaire de consentement (5 min).
- Répondre à une série de questions portant sur les deepfakes (20 min).
- Participer à un test d'awareness portant sur votre capacité à distinguer un deepfake d'un contenu authentique (10 min).

Confidentialité et protection de la vie privée :

Votre participation est entièrement anonyme et confidentielle. La confidentialité des données collectées dans les enregistrements sera assurée conformément au règlement général sur la protection des données (RGPD). Ni les données collectées, ni les résultats de la recherche ne pourront en aucun cas conduire à votre identification.

Risques :

Aucun risque identifié avec cette étude.

Consentement à la participation :

En remplissant ce formulaire de consentement, vous certifiez que vous avez lu et compris les informations ci-dessus. Vous êtes libre de retirer votre consentement ou de vous retirer de cette étude à tout moment et sans préjudice. |

- ☐ J'ai lu et compris les informations ci-dessus et j'accepte de participer à cette étude.
- ☐ Je refuse de participer à cette étude.

Signature :

8.7. Retranscription des entretiens

Répondant	Âge	Genre
Numéro 1	23 ans	Femme

I : Bonjour et merci d'avoir accepté de participer à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux socio-numériques. Pour information, les deepfakes sont des contenus photos ou vidéos, truqués, créés par des intelligences artificielles. Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage d'une personne ou encore les éléments d'un décor, à partir d'une simple photo ou description. Cet entretien durera environ 40 minutes et toutes vos réponses seront anonymes et confidentielles. Bien entendu, vous êtes libre, à tout moment, de ne pas répondre à certaines questions ou encore de mettre fin à l'entretien si vous en ressentez le besoin. Avez-vous des questions avant de commencer ?

R : Non non, pas pour l'instant.

I : Super. Avant de commencer, veuillez prendre connaissance de ce document, et le compléter.

R : Oui, pas de soucis.

I : Nous pouvons maintenant commencer par la première question. Avez-vous déjà entendu parler des deepfakes auparavant ? Si oui, que savez-vous à leur sujet ?

R : Oui, euh, c'est le fait d'utiliser des personnes connues afin de leur faire dire des choses qu'elles n'auraient pas dites. Par exemple, les gens de la politique qu'on utilise et auxquels on fait dire n'importe quoi, alors que pas du tout, ils n'ont jamais dit ça.

I : Ok, je vois. Et comment en avez-vous entendu parler ?

R : TikTok haha. Malheureusement, j'y passe beaucoup de temps, et sur ce genre d'applications, il y en a quand même beaucoup.

I : Avez-vous déjà été confronté à des avertissements ou des campagnes de sensibilisation sur les deepfakes ?

R : Oui.

I : Pouvez-vous m'expliquer sous quelle forme cela était présenté ?

R : C'était un petit warning en dessous d'une publication du style " attention ceci est un contenu créé avec l'IA." Voilà, c'est tout.

I : Et pensez-vous que les deepfakes occupent une place importante sur les réseaux sociaux ?

R : Oui.

I : Pouvez-vous m'en dire plus ? M'expliquer pourquoi ?

R : J'en vois beaucoup quand même, donc je pense que c'est parce qu'il y en a quand même un nombre important. Et cela peut poser problème quand on ne sait pas que c'est un deepfake, car c'est tellement bien fait que cela peut porter à confusion.

I : Et donc, quand vous naviguez sur les réseaux sociaux, vous diriez qu'il y en a combien, en pourcentages ?

R : Je ne sais pas ... ,je dirais peut-être 30 %.

I : Pensez-vous qu'en tant qu'utilisateur des réseaux sociaux, vous avez un rôle à jouer avec les deepfakes ?

R : Euh ouais.

I : Comment ?

R : Bah, je pense qu'en fait ... Je ne sais pas si ce sont les utilisateurs qui ont un rôle à jouer ou plutôt les personnes qui créent les deepfakes qui devraient être obligées de mettre un message à chaque fois que c'est un deepfake.... Ou alors, peut-être, les programmes à la base pour créer le deepfake qui laissent une bande en dessous de l'image ou de la vidéo. A chaque fois du genre : "Ceci est de l'IA."

I : Oui, d'accord, je comprends ... Ensuite, quelle a été votre première réaction ou pensée lorsque vous avez pris connaissance de l'existence des deepfakes ?

R : J'ai trouvé ça fou. J'ai pensé "Waouh quelle dinguerie, ahah."

I : Et pourquoi, selon vous, c'est une "dinguerie" pour reprendre vos mots ?

R : Et bien, parce que cela peut être utilisé dans plein de domaines. Bah pour le marketing, par exemple, c'est trop bien. Plus besoin de payer les acteurs, les deepfakes font le travail.

I : Et avez-vous déjà eu le sentiment qu'une image ou une vidéo sur les réseaux sociaux pouvait être truquée ? Et pouvez-vous me raconter le contexte ?

R : Je n'ai pas compris la question, je crois haha

I : Par exemple, si vous étiez face à une image ou une vidéo mais que vous n'étiez pas sûre de sa véracité.

R : Ah oui, ok. Et bien, en général, les trucs de politique, je m'en méfie toujours un peu. Je ne m'y connais pas. Et en général, si ce n'est pas sur une chaîne euh ... connue, on ne sait pas trop si c'est une vraie info ou si c'est un deepfake. Et surtout, on ne sait pas ce que la personne pense.

I : Vous parlez de la personne qui publie l'information ?

R : Oui, c'est ça. On ne sait pas si c'est quelqu'un qui n'apprécie pas une personne et qui le fait exprès pour rire de la tête de la personne ou pas.

I : Ok .. Et..

R : Dans le sens où, imaginez, il y a quelqu'un qui est hyper extrémiste et qui fait dire des paroles what the fuck à quelqu'un, qui, à la base, ne pense pas cela du tout. Ça n'a aucun sens et pourtant cela peut arriver. Mais comme ce n'est pas habituel, on est un peu perdu et on ne sait pas s'il faut y croire ou pas. Et c'est ça le problème, en fonction de nos idées, je pense que certains y croient plus que d'autres.

I : Oui, je vois. Et avez-vous déjà été influencé par une information qui s'est avérée fausse par la suite ?

R : Je ne pense pas.

I : Donc, par exemple, cela n'a jamais influencé une de vos pensées concernant une personnalité publique ou même influencé un de vos comportements, etc.

R : Ah oui ! Mais deepfakes, je ne pense pas. Juste des fausses infos, mais pas des deepfakes.

I : Selon vous, qu'est-ce qui rend un deepfake difficile à détecter à première vue ?

R : Euh ...

I : Quand vous le voyez sur votre fil d'actualité, par exemple.

R : La qualité de la vidéo ... parce que quand c'est de bonne qualité, cela donne l'impression que c'est sérieux, alors qu'en fait, pas forcément. Et euh ... le fait de ne pas connaître forcément la personne qui publie la vidéo. Par exemple, sur Tik Tok, on voit plein de choses, de pleins de comptes que l'on ne connaît pas. On ne connaît pas le contexte, en fait.

L : Et donc, selon toi, comment distinguer une deepfake d'une vraie image ou vidéo ?

R : Si c'est sur un compte officiel, par exemple.

I : Donc simplement vérifier s'il s'agit d'un compte officiel qui publie l'information. Pas d'autres choses en particulier ?

R : Pas forcément, non. J'avoue que je n'ai pas trop d'idées.

I : Sur une échelle de 1 à 10, à quel point vous sentez-vous capable de reconnaître un deepfake ?

R : Je ne sais pas ... Euh ça dépend dans quel contexte, parce que si c'est un domaine où je ne connais rien, cela va être compliqué de le reconnaître. Mais si c'est un domaine qu'on connaît bien, ouais, 6/10.

I : Et pourquoi ce nombre ?

R : Parce que ça dépend beaucoup du domaine et de mes connaissances dans ce domaine.

I : Ensuite, depuis que vous avez connaissance des *deepfakes*, cela a-t-il modifié votre manière d'utiliser les réseaux sociaux ?

R : Beh, je pense que je m'en méfie plus. Quand il y a des discours un peu fous, je me méfie et je me pose la question de savoir si c'est un vrai truc ou non. Alors qu'avant on prenait simplement l'information et c'est tout quoi. Sans forcément réfléchir si c'était faux ou vrai.

I : OK.

R : On voyait une personne connue et on se disait oui, c'est vrai ce qui est dit. On ne soupçonnait pas.

I : Donc, quand même plus de méfiance ?

R : Oui totalement.

I : Doutez-vous à présent de toutes les informations que vous rencontrez sur les réseaux sociaux ? Ou alors seulement quand quelque chose vous paraît étrange ?

R : Je pense que c'est surtout quand c'est un sujet qui me dérange. Mais je pense que c'est comme ça pour tout. Quand cela nous arrange, c'est ok, sinon c'est un problème et on ne pose pas la question plus rapidement.

I : Avez-vous pris des mesures spécifiques pour vous protéger contre les fausses informations ?

R : Pas spécialement.

I : Cela peut être dans le comportement, les habitudes, etc. Si vous avez changé une habitude pour vous protéger, par exemple.

R : Je ne pense pas.

I : Et lorsque vous voyez une nouvelle vidéo ou image en ligne, prenez-vous systématiquement le temps d'en vérifier la source ?

R : Pas du tout.

I : Mais quand vous le faites, comment procédez-vous ?

R : Je ne sais pas trop. En fait, je crois que je ne vérifie juste pas, haha.

I : D'accord. Et donc, aujourd'hui, vous faites plus attention à ce que vous voyez sur les réseaux sociaux, mais êtes-vous plus vigilante lorsque vous partagez du contenu sur vos réseaux ?

R : Ah oui , oui, oui, je n'ai pas envie de paraître bête. Comme certaines personnes qui partagent l'avis de décès d'une célébrité alors que ce n'est pas vrai. Je ne veux pas publier n'importe quoi non plus. Donc je fais un minimum attention.

I : Et qu'est-ce qui vous motive à faire attention à présent.

R : J'ai juste pas envie de décrédibiliser mon image, je pense. Je n'ai pas envie de passer pour une grosse débile, quoi.

I : Oui, ok, je comprends bien haha. Considérez-vous les réseaux sociaux comme une source fiable ? Et pourquoi ?

R : Fiabiles, non. Fin c'est du divertissant. C'est pas toujours très fiable, cela dépend quel compte. Par exemple, Hugo Décrypte c'est fiable, car son but est de partager des infos sur l'actualité, mais les autres comptes, c'est pas trop fiable.

I : Et les médias traditionnels ou journaux présents en ligne, pensez-vous qu'ils sont fiables ?

R : Oui, car à la base, ils sont censés être neutres. Même si, parfois, ça ne l'est pas trop. C'est quand même censé être neutre. Si ce sont des chaînes officielles, ils n'ont pas le choix. Il n'y a pas le ressenti personnel de la personne qui en parle en donnant son avis. Non, la personne donne juste l'info et c'est tout.

I : Selon vos réponses, quelles sont alors les sources d'informations auxquelles vous faites le plus confiance ?

R : Je n'ai pas compris.

I : A quoi vous fiez-vous pour trouver de l'information qui, selon vous, est juste ?

R : Ah beh, les comptes certifiés du genre Hugo Décrypte ou des chaînes officielles du style RTL, etc. C'est à ça que je fais confiance, parce que c'est un truc reconnu comme source d'info sérieuse et fiable.

I : OK, avez-vous tendance à croire plus facilement une information lorsqu'elle est accompagnée d'une image ou d'une ...

R : Oui.

I : ... vidéo, plutôt qu'un simple texte ?

R : Ah grave.

I : Et pourquoi ?

R : Bah déjà, je pense que je le retiens juste plus facilement. Un truc que je dois lire, je ne le lis juste pas. Alors qu'une image avec du son ou quoi, c'est plus percutant quoi.

I : Oui, je vois. Pour être sûre que je comprenne bien, vous voulez dire que cela passe moins inaperçu dans votre fil d'actualité ?

R : Ah oui, oui, oui. Déjà en fait, tout ce qui est à lire, je ne le lis pas. Je lis rarement un truc de plus de 3 lignes, donc bon.

I : Ok. Et si l'information est répétée plusieurs fois, cela renforce-t-il votre confiance en sa véracité ?

R : Ah oui. Même si c'est une bêtise, comme par exemple Iced Tea. Ce n'était pas un deepfake, simplement une publication qui disait qu'ils arrêtaient le goût pêche. Alors qu'en fait, c'était un poisson d'avril. Mais j'y ai cru, car tout le monde en parlait et le repostait sur Twitter, etc.

I : D'accord. Je n'ai plus d'autres questions à vous poser, simplement s'il y a quelque chose que vous souhaitez ajouter, dire à ce sujet.

R : Bah, je pense que ce serait cool, et ça devrait être obligatoire que les logiciels créateurs de deepfakes mettent un petit message en dessous du contenu et qu'on ne puisse pas le télécharger sans ça.

I : Vous parlez par exemple d'un bandeau de prévention sur les contenus, comme il y a pour l'alcool notamment ?

R : Oui, voilà, c'est ça.

I : Parfait, ici s'achèvent les questions. A présent, je vais vous montrer 3 photos, vous aurez 15 secondes pour regarder celles-ci, une à la suite de l'autre, mais pas en même temps. Les trois images peuvent être vraies, fausses ou vraies et fausses. Vous avez donc 15 secondes par image pour l'analyser et me dire si vous pensez qu'il s'agit d'un deepfake ou non.

R : Ok.

I : Voici la première image :



R : (Observe pendant moins de 30 secondes l'image)

Ok, c'est faux !

I : Pour quelles raisons ?

R : Ils rigolent trop pour que ce soit réel. Ils sont censés être sérieux. Le contexte me paraît bizarre.

I : OK. Voici la seconde image :



R : (observe l'image pendant 15 secondes)

I : Parvenez-vous à identifier s'il s'agit d'un deepfake ou d'une image authentique ?

R : Ici, je trouve que les gens ont l'air moins réel, même s'il y a quand même une bonne qualité d'image. Je dis que c'est faux.

I : D'accord, voici la dernière image.



R : (Observe l'image pendant 15 secondes)

Ah, celle-ci, je pense qu'elle est vraie.

I : Pourquoi diriez-vous cela ?

R : A cause des expressions faciales. On voit que Macron a le front ridé, donc je me dis que cela ne peut pas être faux. Ce n'est pas comme l'autre photo, où on dirait un dessin. Ici, le contexte et les personnes me paraissent vrais.

I : A présent, en regardant les 3 côte à côte, confirmez-vous vos réponses ? C'est-à-dire : la première fausse, la seconde fausse et la troisième vraie.

R : Oui, je confirme, les deux premières sont fausses. Cela me semble être le plus logique.

I : Sur quels éléments avez-vous fondé votre réponse ?

R : Pour la première image, la qualité m'a l'air normale, mais le fait qu'ils rigolent à ce point-là, je trouve cela suspect. Pour les deux autres, je me suis surtout basée sur la qualité et sur le contexte mais aussi sur les expressions du visage.

I : Ok donc je vais vous donner les réponses. La première était vraie.

R : Ah mince, j'ai hésité.

I : La deuxième fausse et la troisième vraie.

R : Ok, la deuxième je me doutais, c'était évident, mais la troisième je ne m'y attendais pas. Pareil pour la première, le fait qu'ils rigolent m'a trompé.

I : Voilà, ici s'achève l'entretien. Merci d'avoir répondu à toutes mes questions.

R : Avec plaisir.

Répondant	Âge	Genre
Numéro 2	23 ans	Femme

I : Bonjour et merci d'avoir participé à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux sociaux. Pour information, les deepfakes sont des contenus photos ou vidéos truqués créés par des intelligences artificielles. Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage d'une personne ou encore les éléments d'un décor à partir d'une simple photo ou description. Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes et confidentielles. Bien entendu, vous êtes libre à tout moment de ne pas répondre à certaines questions ou même d'interrompre l'entretien si besoin. Avez-vous des questions avant de commencer ?

R : Non.

I : Parfait, alors nous pouvons passer à la première question. Aviez-vous déjà entendu parler des deepfakes auparavant ? Et si oui, que savez-vous à leur sujet ?

R : Alors oui, c'était dans un cours à l'UNIF où on devait faire un travail sur les deepfakes. Et donc, pour moi, c'est quand l'IA crée des images qui sont fausses et qui pourraient paraître vraies parce qu'elles sont bien faites. Donc, par exemple, une connue, c'était Emmanuel Macron, je crois, qui était en plein milieu d'une guerre à Paris ou d'une manifestation. Je ne sais plus ce que c'était. Ou alors, par exemple, utiliser les deepfakes pour parler au nom de quelqu'un. Par exemple, j'avais vu sur TikTok un gars qui avait fait une deepfake au sujet de Tom Cruise. Il disait des choses bizarres, et en fait, ce n'était pas du tout lui.

I : Ok. Et avez-vous déjà été confrontée à des avertissements ou des campagnes de sensibilisation à ce sujet ?

R : Non, je ne crois pas. Peut-être au journal télévisé, qui disent : « attention, il est de plus en plus présent, machin ». Mais c'est tout, quoi.

I : Et sur les réseaux sociaux, vous n'y avez jamais été confronté ?

R : Non, je ne pense pas.

I : Et pensez-vous que les deepfakes occupent une place importante sur les réseaux sociaux ?

R : Euh ... importante, je ne sais pas. Il y en a, mais je dirais qu'il n'y a pas que ça. Après, ça dépend aussi des gens, parce que, comme on le sait, les réseaux sociaux, c'est beaucoup les algorithmes qui remettent ce que tu as regardé dans ton fil d'actualité. Donc, par exemple, si on prend TikTok, je pense que si tu regardes la deepfake jusqu'au bout, t'en auras tout le temps. Mais si tu n'as pas l'habitude d'en regarder, je ne sais pas, je dirais peut-être quoi, qu'il y a 20-30% maximum de deepfakes sur les réseaux.

I : Et pensez-vous qu'en tant qu'utilisateur des réseaux sociaux, vous avez un rôle à jouer avec les deepfakes ?

R : On peut les signaler. Mais c'est tout ce qu'on peut faire, je pense. Et ne pas y croire, parce que sinon... Et surtout ne pas partager. Ça, c'est le pire à faire, je pense, parce que hop, les fausses rumeurs, les fake news partent super rapidement.

I : Et quelle a été votre première réaction ou pensée lorsque vous avez pris connaissance de l'existence des deepfakes ? Qu'est-ce que tu as ressenti ? Qu'est-ce que tu en as pensé ?

R : Je me suis dit, waouh, tout ce qu'on peut faire avec la technologie. Maintenant, on peut inventer des trucs et créer des images qui paraissent vraiment vraies et qui ne le sont pas. Donc, c'est dingue, mais parfois, ça peut même faire un peu peur, parce que si tout le monde commence à y croire, ça peut un peu foutre le bordel, je pense, dans la société. Mais oui, il faut être conscient et ne pas croire tout ce qu'on voit sur les réseaux sociaux, je pense.

I : Quand vous dites “foutre le bordel dans la société”, vous pensez à quoi ?

R : Je ne sais pas, j'imagine bien un truc genre Trump qui aurait été le sujet d'un deepfake et qui, je ne sais pas, pète les boulons suite à quelque chose, que les gens y croient et que cela ait des répercussions sur sa carrière politique, etc. Oui, des choses comme ça.

I : Ok, intéressant. Et avez-vous déjà eu le sentiment, en étant sur votre fil d'actualité, d'être face à une image ou vidéo qui pouvait être truquée ?

R : Euh...

I : Donc que vous aviez le sentiment que l'image à laquelle vous étiez confrontée, cela pouvait être faux. Si oui, pouvez-vous me raconter dans quel contexte ?

R : Oui, comme j'ai dit tout à l'heure, du coup, l'image, la vidéo avec Tom Cruise, où il parlait et tout, mais ça se voyait un peu que ce n'était pas lui. Enfin, il faisait un genre de parcours où je ne sais plus quoi. Ça date, je ne me souviens pas parfaitement, mais c'était dans son jardin. Et je ne sais pas, ça me paraissait bizarre. Je me suis dit que ce n'est pas le genre de type qui est sur les réseaux sociaux et qui commence à faire des choses comme ça. Donc, ça me paraissait suspect ... Après, on aurait vraiment dit que c'était lui. Donc, voilà.

I : Ok. Et avez-vous déjà été influencée par une information qui s'est avérée fausse par la suite ?

R : Pas que je me souviene.

I : Par exemple, quand je parle d'influencer, ça peut être influencer une croyance ou un comportement.

R : Ça ne me dit rien. Mais je n'ai pas une bonne mémoire, donc bon, peut-être que je ne m'en rappelle juste pas.

I : Pas de soucis. Et selon vous, qu'est-ce qui rend un deepfake difficile à détecter à première vue ?

R : Que ça ressemble vraiment fort à la réalité. Parfois, je crois qu'il faut bien regarder le contexte, etc., se renseigner sur tout ça, pour être sûre qu'il s'agisse d'un mensonge et pas d'une réalité.

I : Oui. Et donc, comment le distinguer, selon vous ?

R : Je pense que tu ne peux pas le distinguer à l'œil nu, ou alors tu regardes vraiment précisément les détails de l'image. Mais plutôt se renseigner, je pense. Si tu vois une image un peu bizarre, tu te renseignes sur les réseaux, mais plus sur Internet, pour voir qu'est-ce qu'on en dit, si le contexte va dans ce sens-là, etc.

I : Et sur une échelle de 1 à 10, à quel point vous sentez-vous capable de reconnaître un deepfake sur les réseaux ?

R : Dix.

I : Combien, je n'ai pas compris ?

R : Six.

I : Et pourquoi ce nombre ?

R : Parce que je pense être capable de reconnaître peut-être plus que la moitié des deepfakes, mais après, je ne dirais pas que je suis ... À tout moment oui, il y en a qui passent, et j'y crois, alors que c'est complètement faux. Donc... Oui, après, j'ai un peu aussi ... Enfin, je ne sais

pas, ça dépend aussi des deepfakes, mais maintenant, l'image générée est par l'IA, et ça, ça se voit plus que c'est en mode dessin que réalité, donc ...

I : Ok, oui.

R : Ça se voit facilement, et d'autres, oui, c'est plus subtil, donc ... Peut-être que je ne le verrai pas.

I : Oui, je comprends. Du coup, depuis que vous avez connaissance des deepfakes, cela a-t-il modifié votre manière d'utiliser les réseaux sociaux ?

R : Non, je ne pense pas, parce que j'ai toujours fait un peu gaffe à vérifier un peu les infos que je recevais, pour voir si c'était vrai ou non. Donc ... Pas forcément.

I : Et à présent, doutez-vous de toutes les informations que vous rencontrez sur les réseaux sociaux, ou seulement quand il y a quelque chose qui vous paraît étrange, par exemple ?

R : Seulement si quelque chose me paraît étrange, sinon j'y crois. Après les infos que je trouve souvent, celles que je crois, ce sont celles des sites, enfin des comptes vérifiés du style : Brut, Hugo Décrypte, RTBF, RTL. C'est là-dessus que je m'informe, le reste s'il y a un truc qui me paraît un peu spécial, je vais vérifier ...

I : Oh mince, l'enregistrement déconne, ... Donc, vous disiez, quand il s'agit de sites officiels, par exemple les journaux en ligne, etc. ?

R : Oui, c'est ça. C'est surtout là-dessus que je m'informe. Donc, en soi, je sais que je peux faire confiance. Ce sont des principes déontologiques qui sont respectés. Donc ... Il n'y a pas de souci.

I : Pour être sûre que je comprenne bien. Dans votre fil d'actualité, même s'il y a d'autres comptes qui passent avec des informations, vous n'allez pas avoir tendance à les lire, et lire simplement celles qui sont ...

R : Oui. Et même si je lis et que je vois que ça paraît bizarre, je vais m'informer sur Internet.

I : Vous allez vérifier ?

R : Oui.

I : OK. Donc, avez-vous pris des mesures spécifiques pour vous protéger contre les deepfakes ?

R : Ben, je vérifie sur Internet ce qui est dit.

I : Et comment vérifiez-vous la source d'une image ou d'une vidéo, par exemple ?

R : Ben, déjà, je regarde c'est quoi le compte qui raconte ça. Si ça a l'air de débutants, qui disent n'importe quoi, ou si ça a l'air d'un truc sérieux. Et puis, oui, je vais taper sur Internet, je vais voir si c'est vrai ou non. Je vais refaire des recherches de mon côté, quoi.

I : Donc vous avez dit que vous étiez plus vigilante avec ce que vous voyiez, mais est-ce que vous êtes également plus vigilante avec ce que vous partagez sur les réseaux sociaux ?

R : Je ne partage rien, du coup, il n'y a pas de souci là-dessus.

I : Ok, parfait. Et considérez-vous que les réseaux sociaux sont une source fiable ? Et pourquoi ?

R : Attendez, je rebondis à ma dernière réponse. Je ne repartage rien, mais par contre, quand je vois souvent, par exemple, sur les réseaux sociaux, sur Instagram. Il y a beaucoup de gens avec la politique, etc., qui repartagent le compte et des choses qui sont dites. Mais les gens repartagent les infos de manière exagérée. Et parfois, je suis là, "mais ça sort d'où, quoi ?". C'est juste une personne lambda qui partage des infos, quoi. Donc, là-dessus, je suis beaucoup plus critique.

I : Ok. Et donc, considérez-vous que les réseaux sociaux sont une source fiable ? Et pourquoi ?

R : Oui et non. Oui, si tu suis ... Si tu suis des comptes journalistiques comme Brut, Hugo Décrypte, RTL, RTBF, tout ça. Ça, d'office, c'est vrai. Ils ne peuvent pas mentir, enfin c'est vérifié, etc. Mais je pense qu'il y a tellement de fausses informations que le reste, je ne regarde même pas. Et comme je dis, si je vois passer ça, des repartages ou des reposts sur Instagram, souvent, je me dis, mais ça sort d'où, cette info ? Y a aucune preuve d'où il a été chercher ça. C'est juste un mec lambda qui a exagéré, extrapolé un truc. Et voilà.

I : D'accord. Et donc, les médias en ligne, ça, c'est fiable, selon vous ?

R : Ouais.

I : Et donc, à l'heure actuelle, quelles sont les sources d'informations auxquelles vous faites le plus confiance ? Pas forcément en ligne, cela peut être en général.

R : Oui, la télé, les trucs journalistiques, les sites éditoriaux, que ce soit la télé, la radio, ou les réseaux sociaux.

I : Ok, donc tant qu'il s'agit d'un média de confiance ...

R : Oui, exactement.

I : Et avez-vous tendance à croire plus facilement une information si elle est accompagnée d'un visuel ou d'un son, par exemple, genre une image, une vidéo, plutôt qu'un texte ?

R : Ouais, quand même, je pense. L'air de rien, je pense que ça influence quand même, parce que t'as une espèce de preuve à l'appui. Et c'est pour ça que les deepfakes sont dangereux. Parce que, oui, ce sont des fausses preuves. Donc... Oui, je pense, inconsciemment, même ça influence.

I : Et si, par exemple, une même information est répétée plusieurs fois, par plusieurs comptes et différents canaux, est-ce que cela va renforcer votre impression que l'information est vraie ?

R : Ça dépend de l'information en question. Si c'est gros comme un bateau, je vais me dire que non. Si elle est très subtile, ouais, peut-être. Et je vais peut-être me renseigner alors, parce que si j'en vois, je me renseigne en tout cas.

I : Ok. Super, c'est la fin de mes questions. Je ne sais pas si vous avez quelque chose à ajouter sur le sujet.

R : Non, pas vraiment.

I : Non ? Donc, à présent, je vais vous montrer trois images. Trois images, une à la suite de l'autre. Vous aurez 15 secondes pour observer chaque image.

R : Ok. 15 secondes par image.

I : A chaque fois, après 15 secondes, vous pourrez me donner votre avis. Me dire si vous pensez qu'il s'agit d'une vraie ou d'une fausse image. Ensuite, vous pourrez les voir, en même temps, et me confirmer ou ajuster vos réponses.

R : OK. Et c'est la même image alors ?

I : Ce sont trois images différentes.

R : OK.

I : Voici la première.



R : (Observe l'image pendant 15 secondes)

Je pense que j'ai déjà vu cette image, donc je crois que c'est vrai.

I : Vous pensez que c'est vrai ?

R : Ouais, enfin cette image me dit ... Après, je ne l'analyse pas dans les détails, tous les vêtements et tout ça, mais ... oui.

I : OK, à présent, voici la deuxième image.



R : (observe très rapidement) Ok, non.

I : Pour quelles raisons ?

R : Bah, déjà, je trouve que l'image a l'air un peu fausse comme ça. Et même, à quel moment, il se marre avec Donald Trump comme ça.

I : Et du coup, quand vous dites que cela a l'air un peu faux, que voulez-vous dire par là ?

R : La qualité est trop nette comme ça. Vous voyez ce que je veux dire ?

I : Oui, parfaitement.

R : Top.

I : Ok, et maintenant la troisième. Celle-ci.



R : (Observe l'image pendant 15 secondes)

Ah non, ils n'ont pas l'air réel, quoi.

I : Ok c'est cela qui vous fait dire qu'elle est fausse ?

R : Oui

I : Et à présent, en regardant les trois photos, est-ce que vous confirmez vos choix ou est-ce que vous voulez les adapter ?

R : J'hésite pour la première ... Non, allez, je garde mon avis.

I : OK. Je vous donne alors les réponses. Donc, la première réponse, c'était une fausse image.

R : Ah bon ?

I : Oui, et la deuxième, c'était une vraie image.

R : Là où il se tape une barre ?

I : Oui, c'est ça.

R : Ah ouais ?

I : Et la troisième, c'est bien évidemment une fausse.

R : D'accord ça va, il y en avait une flagrante et les autres, je ne m'y attendais pas.

I : Sur quels éléments vous êtes-vous appuyée pour faire vos choix ?

R : Ben, le contexte ... Le contexte et la qualité de l'image. Parce que même la troisième, ben oui, la tête a l'air plus grosse. Même quand on voit le décor, c'est un peu un décor dessiné comme ça. Et même maintenant que je regarde un peu plus longtemps, que j'analyse un peu les mains, on voit que ce sont des mains dessinées aussi, pas en mode photo. La première, maintenant que j'analyse un peu plus, on voit qu'ils ont été détournés d'une image. Mais je pensais avoir déjà vu l'image, donc ça a dû jouer aussi. Et même, je trouvais qu'elle paraissait vraiment réelle niveau couleur, etc. Et celle où il se tape une barre, déjà le contexte, il est ... bizarre. Enfin, j'y croyais à moitié et puis je trouvais que ... En fait, ça m'étonne un peu qu'elle soit vraie parce que c'est fort net. Le décor est trop parfait par rapport à la première image.

I : Oui, je comprends. Et, pour voir si je comprends bien, j'ai l'impression que vous avez peut-être été influencée par le contexte de la deuxième image parce que vous pensez qu'en tant que politiciens ils doivent être sérieux.

R : Si ... Ouais, j'ai du mal à les imaginer se taper des barres ensemble et se tenir la main. Je ne sais pas, c'est un peu bizarre connaissant les personnes.

I : Oui, je comprends. Et bien, voilà, l'entretien est terminé, merci pour vos réponses, cela va beaucoup m'aider.

R : J'espère que j'ai bien répondu à toutes les questions.

I : Oui, c'est super. De toute manière, il n'y a pas de bonne ou de mauvaise réponse.

R : Super alors, parce que parfois c'est difficile de se souvenir de tout.

I : C'est normal haha, mais merci beaucoup !

Répondant	Âge	Genre
Numéro 3	19 ans	Femme

I : Bonjour et merci d'avoir participé à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux sociaux. Pour information, les deepfakes sont des contenus photos ou vidéos truqués créés par des intelligences artificielles. Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage d'une personne ou encore les éléments d'un décor à partir d'une simple photo ou description. Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes

et confidentielles. Bien entendu, vous êtes libre à tout moment de ne pas répondre à certaines questions ou même d'interrompre l'entretien si besoin. Avez-vous des questions avant de commencer ?

R : Non.

I : Nous pouvons alors commencer par la première question, qui est : aviez-vous déjà entendu parler des deepfakes auparavant ? Et si oui, que savez-vous à leur sujet ?

R : Je n'en avais jamais entendu parler avant aujourd'hui.

I : Ok, et donc avez-vous déjà été confronté à des avertissements ou des campagnes de sensibilisation sur les deepfakes ?

R : Sur les réseaux sociaux comme TikTok, il y a des gens qui préviennent de ne pas se fier à tout. Il y a des gens qui disent oui, ne vous fiez pas à toutes ... toutes les informations que vous pouvez voir, que les gens peuvent inventer certaines choses ou modifier la réalité. Mais donc, j'imagine que c'est un peu la même chose, mais pour d'autres informations.

I : Ok, oui, donc si je comprends bien, vous connaissiez le concept mais pas le mot.

R : Oui, en soi.

I : Et donc, pensez-vous que les deepfakes occupent une place importante sur les réseaux sociaux ?

R : Oui.

I : Pouvez-vous expliquer pourquoi vous pensez cela ?

R : Oui, car les gens, ils aiment bien toujours avoir quelque chose à dire et donc ils aiment bien inventer des choses et choquer les gens et avoir le nouveau petit potin à dire, et donc ils sont prêts à inventer toutes sortes de fausses images ou de fausses informations pour pouvoir attirer les autres et avoir de la visibilité.

I : Quand vous dites que ça occupe quand même une place importante sur les réseaux sociaux, si vous deviez donner un pourcentage par jour, par exemple, de faux contenus que vous croisez dans votre fil d'actualité ?

R : 35-40 % ?

I : Vous diriez qu'il y en a 35-40 % dont vous êtes consciente ou en général ?

R : 35-40 % auxquels je pense être affichée. Consciente, moins.

I : Et pensez-vous qu'en tant qu'utilisateur des réseaux sociaux, vous avez un rôle à jouer avec les deepfakes ?

R : Oui.

I : Comment ?

R : Je dois déjà pas en répandre pour éviter de répandre des mauvaises choses dans la tête de toute ma communauté. Et aussi essayer de me sensibiliser pour ne pas être touchée et pas être impactée. Pour éviter que ça change ma manière de réfléchir et de voir les choses, à cause de gens qui veulent me faire croire des choses qui sont fausses.

I : Et quelle a été votre première réaction ou pensée lorsque vous avez pris connaissance de l'existence des deepfakes ?

R : Je suis un peu choquée parce que c'est quand même assez réaliste, donc je suis choquée qu'on puisse faire une telle chose.

I : Choquée dans le bon sens ou dans le mauvais sens ?

R : Pour moi, tout ce qui est bon est quand même mauvais, parce que si tu sais l'utiliser dans le bon, tu sais l'utiliser dans le mauvais. Et donc, le fait de savoir faire des choses comme ça, ça peut vraiment être dangereux. Par exemple, si ça tombe entre de mauvaises mains et s'il y a des grands moyens pour faire répandre des choses dans toute la population, ça peut très vite circuler et les gens peuvent très vite y croire. Donc plus de mauvais, bon ce serait ... Je ne sais pas trop comment l'intelligence artificielle, en créant des images peut faire de bonnes choses, je ne vois pas forcément.

I : Et avez-vous déjà eu le sentiment qu'une image ou une vidéo sur les réseaux sociaux pouvait être truquée ?

R : Euh ...

I : Donc, par exemple, que vous étiez sur votre fil d'actualité et que vous vous posiez la question : “ est-ce une vraie ou fausse image/vidéo que je regarde” ?

R : Alors oui, mais ça ne m'arrive pas souvent, je crois quand même assez bêtement ce que je vois je crois.

I : Et donc, si ça vous est déjà arrivé, vous pouvez me raconter peut-être le contexte, l'image ou la vidéo que vous avez vue ?

R : Souvent, je ne me pose pas trop de questions et j'y crois bêtement, et c'est en allant voir les commentaires et en voyant l'avis des gens que je me rends compte que je me suis fait avoir. Même les images de l'actualité ou des paroles reprises par d'autres personnes, comme je ne suis pas assez informée, j'y crois et donc je sais que ça peut être dangereux. Cela désinforme un peu tout le monde, dont moi, haha.

I : Avez-vous déjà été influencée par une information vidéo ou photo qui s'est avérée fausse par la suite ?

R : Que j'ai été influencée ?

I : Oui, en fait, pour être sûre que vous compreniez bien. Par exemple, une information qui a pu vous influencer à croire quelque chose qui était faux sur une personne, sur une situation, ou même une information qui vous a influencée dans un comportement ?

R : Oui, quand je vois l'apparence d'une célébrité, par exemple, et que je me demande comment cela se fait qu'il ait maigri si vite, etc. Je me pose la question, mais au final j'y crois. Ou les retouches photos.

I : D'accord. Je pensais notamment, par exemple, à des discours de personnes connues qui ont été modifiés, sortis de leur contexte.

R : Il n'y a rien qui me revienne en tête pour le moment, mais souvent les extraits qui sont mis sur TikTok, par exemple, je les ai déjà vus donc ça ne m'influence pas, car j'ai déjà vu le discours dans le contexte.

I : Et selon vous, qu'est-ce qui rend un deepfake difficile à détecter à première vue ?

R : On ne sait pas que c'est faux. Oui, c'est ça. Quand on a une image, avec beaucoup d'informations vraies et qu'il y a juste un petit détail faux. Parce que donc on se dit, s'il y a autant de choses qui sont réelles, c'est que tout est vrai.

I : Et donc, selon vous, comment distinguer un deepfake d'une vraie photo ou d'une vraie vidéo ?

R : On ne sait pas, ou alors il faut être très informé sur le sujet, ou encore faire des recherches et pouvoir aller vérifier les sources des images. Mais maintenant, il y a beaucoup de sources différentes et il y en a beaucoup qui peuvent être truquées.

I : Donc, si vous avez une vidéo ou une image qui circule sur les réseaux sociaux à laquelle vous êtes confrontée, vous pensez ne jamais savoir si celle-ci est vraie ou non ?

R : Oui, c'est ça, je n'aurai jamais vraiment la réponse.

I : Et donc, sur une échelle de 1 à 10, à quel point vous sentez-vous capable de reconnaître un deepfake ?

R : 4 haha.

I : Et pourquoi ce nombre ?

R : Car je ne me sens pas particulièrement douée pour ça. J'ai quand même un avis critique, mais je crois que je pourrais être assez vite bernée par des photos et croire bêtement, parce que je vais considérer les gens plus informés que moi, donc je leur ferai confiance.

I : Et depuis que vous avez connaissance des deepfake, cela a-t-il modifié votre manière d'utiliser les réseaux sociaux ? Du coup, depuis que vous connaissez le concept de fausses images et fausses vidéos.

R : Ben oui.

I : Et comment ça ?

R : Maintenant, j'essaie de moins croire directement et d'essayer de voir aussi l'avis des gens dans les commentaires, ceux qui essaient de raisonner, et d'en parler autour de moi à des gens que je crois plus capables de pouvoir m'informer sur la vraie situation, avec qui ont des vraies infos.

I : Vous voulez dire que, quand vous voyez quelque chose en ligne, pour être sûre de ce que vous voyez, vous allez peut-être demander à votre famille, à vos amis, etc. s'ils ont entendu parler de ça et leur demander ce qu'ils en pensent ?

R : Oui. S'ils ont entendu parler de l'info par d'autres moyens qui seront peut-être plus crédibles que les réseaux sociaux, par exemple à la TV et tout.

I : D'accord, et à présent que vous savez tout ce qu'on peut faire avec les deepfakes, est-ce que vous doutez de toutes les informations que vous rencontrez sur les réseaux, ou alors seulement quand il y a quelque chose qui vous dérange, quelque chose qui attire votre attention ?

R : Je crois que je ne doute pas de tout, mais de quelque chose qui va vraiment porter mon attention. Je me dis : "Est-ce que c'est vraiment possible ?" Ou alors, si ce sont des sujets qui sont plus délicats, à même à discussion, des sujets pour lesquels les deepfakes, ça pourrait être dangereux. Sinon le reste, je vais en avoir un peu plus rien à faire, si ce n'est pas vrai.

I : Et donc, avez-vous pris des mesures spécifiques pour vous protéger contre les deepfakes, contre les fausses infos ?

R : Non.

I : Donc vous dites quand même avoir changé votre manière de consommer l'info, mais vous n'avez rien changé pour autant ?

R : Je demande à mon entourage, mais sinon je continue quand même à aller sur des sources pas très fiables, où mes habitudes sur les réseaux sont pareilles.

I : Et lorsque vous voyez une nouvelle vidéo ou une image en ligne, prenez-vous systématiquement le temps d'en vérifier la source ?

R : Non.

I : Et si vous deviez vérifier la source, comment feriez-vous ?

R : Je n'en ai aucune idée, à part demander à mon entourage et voir ce que les autres en disent.

I : Et aujourd'hui, vous êtes plus vigilante à ce que vous voyez sur les réseaux sociaux, mais est-ce que vous êtes également plus vigilante à ce que vous partagez sur les réseaux ?

R : Souvent, je ne repartage pas pour être sûre que je ne vais pas répandre des choses fausses, et sinon oui, je vais quand même être assez ... je vais y prêter attention parce que je n'ai pas envie que les gens me voient comme quelqu'un qui ne sait pas faire la part des choses et qui repartage des choses, enfin qui tombe dans le panneau.

I : Vous n'avez pas envie de paraître bête ?

R : Oui, voilà, c'est totalement ça.

I : Et considérez-vous les réseaux sociaux comme une source fiable ?

R : Oui et non. J'ai l'impression qu'il y a une dynamique qui commence à vouloir devenir plus fiable et que les gens ont envie de créer une chose qui commence à égaler un peu les journaux où ce sera plus des informations vraies, parce que maintenant les gens recherchent plus de la véracité que des choses fausses, et donc il y a une dynamique pour essayer d'arriver à ça. Mais il y aura toujours des gens qui seront ... Vu que ce n'est pas une source officielle, un endroit réservé, il y a aussi des plus jeunes qui voudront toujours balancer des fausses infos pour attirer l'attention.

I : Et donc, pour les médias traditionnels ou les journaux présents en ligne, par exemple, est-ce que pour vous, ce sont des sources fiables ?

R : Alors, dans ma conscience, oui, car c'est quelque chose d'officiel, mais parfois je me doute qu'il n'y a pas tout de vrai. Enfin ... pas que cela soit faux, mais qu'ils effacent ou cachent des infos.

I : Et alors, selon vos réponses précédentes, quelles sont les sources d'informations auxquelles vous vous fiez et auxquelles vous faites le plus confiance ? Pas forcément en ligne, en général.

R : Euh ...

I : Ça peut ne pas être en ligne, ça peut être en général.

R : En général, c'est des gens de mon entourage que je crois déjà beaucoup plus, parce que je me dis qu'ils font le travail.

I : Quand vous dites qu'ils font le travail, qu'entendez-vous par là ?

R : Qu'ils s'assurent des sources avant d'en parler et que leurs propos sont justifiés, qu'ils ont vraiment une preuve de ce qui est expliqué.

I : Avez-vous tendance à croire plus facilement une information lorsqu'elle est accompagnée d'une image ou d'une vidéo, plutôt qu'un texte ?

R : Un texte tout seul ?

I : Oui, c'est ça. Est-ce que vous allez avoir tendance à croire un texte tout seul, ou s'il y a une image ou une vidéo qui accompagne ce texte, cela vous procure le sentiment que l'information est vraie ?

R : S'il y a une image ou une vidéo, ça dépend dans quel contexte, mais s'il y a une image ou une vidéo qui est réaliste, d'office. Je crois que tout le monde serait plus influencé et plus à même d'y croire. Mais si c'est un vrai texte bien construit et qu'il y a des termes un peu compliqués, ça peut aussi me faire y croire et me faire tomber dans le panneau.

I : Et pourquoi une image ou une vidéo ? Ça pourrait être un peu plus impactant ?

R : Car on le voit de nos propres yeux et on se dit que si c'est filmé ou pris en photo, c'est réel, et on va moins soupçonner tout ce qu'on voit, on n'est pas autant à l'affût. Alors qu'en fait, de ce que je comprends, non. Mais tout le monde n'est pas conscient, et moi non plus, de tout ce qu'on peut faire avec les intelligences artificielles, parce qu'il y a des gens qui savent très bien les utiliser.

I : Et si une même information est répétée plusieurs fois, par exemple par plusieurs comptes ou même sur différents canaux, est-ce que cela va renforcer votre idée que l'information que vous voyez est vraie ?

R : Oui.

I : Pourquoi ?

R : Parce que s'il y a beaucoup de gens qui le disent, c'est que c'est vrai. Ils n'ont pas pu tous tomber dans le panneau. Souvent, si c'est répandu, il y a plus de chances que ce soit vrai que faux.

I : Et si vous n'êtes pas sûre que ce qu'ils partagent est vrai, vous pourriez tout de même y croire ?

R : Tant que j'hésite et que je ne suis pas sûre à 100 % que c'est faux, ça va d'office me faire hésiter. Mais je me dis que les gens, comme il y en a beaucoup, ça fait un effet de nombre et donc que ce qu'ils disent, c'est vrai.

I : Ok, merci. C'est la fin des questions. Y a-t-il quelque chose que vous aimeriez ajouter à ce sujet ? Ça peut être général, une remarque, etc.

R : Est-ce qu'avec chatGPT, on peut faire des deepfakes ?

I : On peut le faire avec chat GPT, sinon il y a plein d'outils plus techniques gratuits ou même payants.

R : Ah oui, ok.

I : Ici s'achève la série de questions. A présent, je vais vous exposer trois images, chacune une à la suite de l'autre. Vous aurez 15 secondes pour observer chacune des images. Elles peuvent être toutes les trois vraies, toutes les trois fausses, une vraie, deux fausses. Il y a vraiment une multitude de combinaisons possibles. Donc vous aurez 15 secondes à chaque fois pour voir la première image, et ensuite me donner votre avis, si pour vous elle est vraie ou fausse, et sur quels éléments vous vous êtes basée. On fera pareil pour l'ensemble des photos. Et à la fin, vous aurez les trois photos l'une à côté de l'autre et vous pourrez modifier ou conserver vos réponses.

R : D'accord.

I : Alors, je vais commencer par la première photo. La voici :



R : (Observe l'image pendant 15 secondes)

I : Selon vous, est-ce que la photo est vraie ou fausse ?

R : Pour moi, elle est vraie.

I : Pourquoi ?

R : Parce qu'il me semble qu'il n'y a pas très longtemps, ils se sont réunis pour parler donc je me dit que la photo aurait pu être prise à ce moment-là. Et donc ça me semble... Enfin, leur posture, leur tête ne m'ont pas l'air bizarre. Et les détails non plus. Après l'intelligence, du coup, je ne sais pas ce que ça fait. Mais ça m'a l'air réel et la situation ne me paraît pas étrange à la réalité. Il y a même les veines et tout. Et les rides donc voilà.

I : Voilà la seconde image.



R : (Observe l'image pendant 15 secondes)

I : Alors, celle-ci est-elle vraie ou fausse selon vous ?

R : Ça me semble bizarre. Ça ne me semble pas étrange. Ils pourraient rigoler. Mais ça ne me semble pas très professionnel. Donc je dirais que c'est faux. Et puis, parce qu'Emmanuel Macron fait un peu peur avec ses dents haha et je ne crois pas que ce soit ça son sourire de base.

I : D'accord. Voici la troisième image.



R : (observe l'image pendant 15 secondes)

R : Elle est fausse. Au départ, rien ne me semblait bizarre, mais en regardant mieux, Emmanuel Macron a une grosse tête. J'ai regardé Donald Trump en premier et je n'ai pas trouvé cela bizarre mais en voyant Emmanuel Macron, la réponse est claire, haha. Ouais, l'intelligence artificielle se voit.

I : Donc selon vous, la troisième image est fausse, pour quelles raisons ?

R : Parce que ça se voit quand même que ce n'est pas une image réelle. Et aussi parce que les têtes sont quand même un peu déformées. Et ça fait plus dessin animé qu'une réelle photo. Je n'ai même pas regardé ce qu'il se passait. Pour être honnête, j'ai vraiment été attirée par le décor et la tête.

I : Et donc maintenant, en voyant les trois l'une à côté de l'autre ... Est-ce que vous confirmez ou vous ajustez vos choix ?

R : La troisième, je suis sûre qu'elle est fake. Et la deuxième, j'hésite parce que le contexte me paraît bizarre, donc elle a l'air aussi un peu fake. Mais moins fake que l'autre. Peut-être que c'est un autre logiciel. Ou alors elle est vraie. Et la première, j'hésite. Est-ce que je peux zoomer ?

I : Oui, oui, à présent vous pouvez regarder plus en détails.

R : (Observe pendant 25 secondes) Ah non mince, la main, elle est bizarre ... Ah ouais, non, elle est fausse.

I : Donc vous gardez la même réponse ou vous la changez ?

R : La première fausse et les deux autres aussi.

I : Sur quels éléments vous êtes-vous appuyée pour faire tous vos choix ?

R : Sur la qualité de l'image globalement. Et sur la probabilité de situation.

I : Vous aviez raison pour la première et la troisième. Mais la deuxième est vraie.

R : Ah ça va, je n'ai pas été trop nulle. Et pour la deuxième, j'ai vraiment hésité, mais le fait qu'ils rigolent, cela m'a induit en erreur.

I : Merci beaucoup pour vos réponses en tout cas.

R : De rien, j'espère que ça ira avec mes réponses.

I : Toutes les réponses sont bonnes à prendre, il n'y a pas de soucis.

Répondant	Âge	Genre
Numéro 4	23 ans	Femme

I : Bonjour et merci d'avoir accepté de participer à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux sociaux numériques. Pour information, les deepfakes sont des contenus photo ou vidéo truqués, créés par des agences artificielles.

Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage de personnes ou encore les éléments d'un décor à partir d'une simple photo ou description. Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes. Bien entendu, vous êtes libre à tout moment de ne pas répondre à certaines questions. Est-ce que vous avez des questions avant de commencer ?

R : Non.

I : Alors, nous pouvons commencer par la première question, étant : Avez-vous déjà entendu parler des deepfakes et, si oui, que savez-vous à leur sujet ?

R : Du coup, j'avais déjà entendu parler du deepfake. Pour moi, c'était surtout la suppression de quelqu'un ou d'un détail sur une photo ou une image totalement inventée par une IA, que ce soit sur une image, une vidéo ou un texte carrément.

I : Et comment en avez-vous entendu parler ?

R : Principalement sur les réseaux sociaux, dans mes cours et des formations supplémentaires et aussi au journal.

I : Et avez-vous déjà été confrontée à des avertissements ou des campagnes de sensibilisation sur les deepfakes ?

R : Pas réellement.

I : D'accord. Et pensez-vous que les deepfakes occupent une place importante sur les réseaux sociaux ?

R : Oui, quand même, parce qu'il y a beaucoup de personnes qui peuvent être sensibles aux deepfakes en pensant que c'est la réalité, alors que pas du tout. On a pu le voir même avec les journaux où un journaliste avait mis un deepfake sur les réseaux sociaux et ça avait été carrément repris au journal, et beaucoup de populations pensaient que c'était la réalité alors que pas du tout.

I : C'était dans le cadre d'un test qu'il avait fait ça ?

R : Non, c'était dans le cadre d'une blague.

I : Et toujours par rapport à la place des deepfakes sur les réseaux sociaux, si vous deviez donner un pourcentage de deepfakes présentés sur les réseaux sociaux ?

R : Je dirais une dizaine de pourcents.

I : 10 % dont vous êtes consciente ou pas ?

R : Oui, principalement en vidéo, surtout sur TikTok, il y a beaucoup de deepfakes.

I : Et pensez-vous qu'en tant qu'utilisateur des réseaux sociaux, vous avez un rôle à jouer avec les deepfakes ?

R : C'est une bonne question ... Principalement, oui, avoir un rôle pour essayer que la jeune génération qui ne connaît pas encore réellement les deepfakes ou la plus vieille, genre les grands-parents qui ne sont pas au courant de leur existence même, là, oui, un rôle à jouer pour éviter qu'ils ne croient n'importe quoi ou ne pensent avoir vu n'importe quoi. Maintenant, je ne sais pas réellement ce que je pourrais faire pour éviter les deepfakes, à part pour moi-même avoir la sensibilité de me dire que non, ce n'est pas possible et d'aller vérifier sur des sources justes. Beaucoup de personnes ne le font pas et je ne peux pas empêcher cette personne de croire ou de ne pas le faire.

I : Et donc, quelle a été votre première réaction ou pensée lorsque vous avez pris connaissance de l'existence des deepfakes ?

R : Euh ...

I : Quelle est la chose qui vous est venue à l'esprit quand vous avez su ce qu'on pouvait faire avec l'IA ?

R : Que c'était bien fait. Principalement avec tout l'avènement des IA, ça permet vraiment de faire des deepfakes où on pourrait même croire que c'est la réalité et que si on ne vérifie pas un minimum ces sources, c'est clairement ça, j'ai failli me faire avoir plusieurs fois.

I : Et donc, c'était plus une pensée positive ou négative ?

R : Positive dans le fait que les technologies nous permettent de faire beaucoup de choses maintenant avec l'évolution constante. Maintenant négative, parce que des personnes mal intentionnées pourraient faire croire aux plus faibles d'entre nous des choses qui ne sont ni logiques ni réelles. Maintenant, si on ne s'en tient pas qu'à une source et qu'on réfléchit un peu plus, normalement ça ne devrait pas poser de problème, même s'il y aura toujours des gens qui croiront aux deepfakes.

I : Et avez-vous déjà eu le sentiment qu'en étant face à une image ou à une vidéo sur les réseaux sociaux que la vidéo ou l'image à laquelle vous étiez confrontée était fausse ? Et si oui dans quel contexte ?

R : Oui, j'ai vu sur mes réseaux sociaux des vidéos avec des animaux que je trouve très mignons, mais c'est beaucoup de l'IA pour l'instant. Des pingouins ou des animaux qui sortent des animaux de compagnie de base, et que du coup ça se voyait sur les contours que c'était de l'IA, mais sinon en tant que tel c'était difficile à voir si c'était un deepfake ou non.

I : Oui, donc si on se confronte, on le voit mais sinon ...

R : Ça passe inaperçu. Si on scrolle juste sur les réseaux sociaux, on peut clairement penser que c'est vrai.

I : Et avez-vous déjà été influencée par une information qui s'est avérée fausse par la suite ?

R : C'est une question difficile.

I : Oui, c'est difficile. Ça peut être une information fausse qui vous a influencée dans une croyance ou même dans un comportement ?

R : Ça, je pensais à un truc. Attends, c'était comment encore. Oh waouh, je pensais à un truc mais je n'arrive pas à retomber dessus. C'était sur une campagne publicitaire avec une crise, une polémique pour une entreprise dont je ne sais plus le nom. Et j'ai mal compris à cause du deepfake, parce qu'il y avait plusieurs versions. Et notamment aussi avec celle du violeur à Gand qu'il y a eu il n'y a pas très longtemps, où certains journalistes avaient apporté des preuves à leurs écrits, pour influencer les gens. Des photos montraient que cela s'était passé

pendant une soirée. Et d'autres informations disaient que c'était à l'hôpital où il avait travaillé. Et du coup je ne savais plus quelle source croire. Et j'avais été ... j'avais un peu mal compris la situation du fait de ces nombreuses sources qui étaient toujours vagues et ne disaient pas la même chose.

I : Et selon vous, qu'est-ce qui rend un deepfake difficile à détecter à première vue ?

R : Si c'est par vidéo ou image, les IA font très bien les choses. Si on n'est pas habitué aux IA et si on ne travaille pas avec, c'est compliqué de voir. Maintenant, je pense que les gens travaillent souvent avec les IA même s'ils ne sont pas dans la communication, la technologie ou quelque chose comme ça. Parce que ça nous aide.

I : Et quand vous dites que cela fait très bien les choses, vous entendez quoi par là ?

R : En soi, si on ne s'y attarde pas, même si on ne s'y attarde qu'une minute sans connaître réellement, on ne peut pas voir les petites erreurs que laissent des IA. Si c'est une vidéo d'un pingouin, ça floute un petit peu, ça grossit les détails par rapport à la réalité. Mais ce sont des mini détails, donc si on n'a jamais vu de bébé pingouin, si on ne fait attention, ou qu'on ne s'est jamais intéressé au sujet des deepfakes, on peut juste penser que c'est la réalité. Parce que quelqu'un qui recueille un pingouin pour faire une bonne action, le sauver etc. on peut penser que c'est vrai.

I : Et donc, comment distinguer un deepfake d'une vraie photo ou vidéo ? Comment feriez-vous pour savoir si c'est vrai ou pas ?

R : Déjà, j'irai chercher sur d'autres sources pour voir si d'autres sources parlent de la même chose. Si c'est de l'actualité que j'ai vue sur les réseaux sociaux, j'irai m'interroger sur des journaux. Aussi, voir ce qu'en disent des experts, ou quelque chose comme ça, d'autres sources scientifiques.

I : Et donc, en connaissance de tout ça, sur une échelle de 1 à 10, à quel point vous sentez-vous capable de reconnaître un deepfake ? Et pourquoi ce nombre ?

R : Je dirai 7, parce que j'en vois beaucoup. Maintenant, j'analyse très peu les photos sur les réseaux sociaux et je pense que j'en laisse beaucoup passer sans plus m'attarder dessus.

I : Et depuis que vous avez connaissance des deepfakes, cela a-t-il modifié votre manière d'utiliser les réseaux sociaux ?

R : Pas réellement. Je n'utilise pas beaucoup les réseaux sociaux, à part pour communiquer avec mes amis et mes proches. J'ai l'air vieille quand je dis ça, haha. Je poste très peu sur les

réseaux sociaux et je regarde souvent très peu le contenu. Souvent, le plus de contenu que je regarde, c'est ceux de mes amis dont je doute qu'ils me mentent.

I : Et donc, à présent, quand vous êtes face à des informations, est-ce que vous doutez plus facilement de tout ? Ou alors juste quand il y a quelque chose qui vous dérange, un sujet qui vous touche plus particulièrement ?

R : Je ne fais pas attention à tous les détails, mais si c'est un sujet qui ne m'intéresse pas réellement, je ne vais juste pas spécialement retenir l'information que j'ai lue. Mais si c'est un sujet qui m'intéresse, je vais plus aller en profondeur et prendre d'autres sources.

I : Et donc, avez-vous pris des mesures spécifiques pour vous protéger contre les fausses informations sous forme de deepfake ?

R : Pas réellement.

I : Et si vous deviez vérifier la source, vous feriez comment ?

R : Honnêtement, à part vérifier si d'autres sources en parlent, je n'ai pas d'idée.

I : Donc vous faites plus ou moins attention à ce que vous regardez, mais est-ce que vous êtes également plus vigilante à ce que vous partagez, ou même à ce dont vous parlez à vos proches par rapport à ce que vous avez vu sur les réseaux sociaux ?

R : Alors partager, je ne repartage rien du coup. Mais par contre, de ce que je dis à mes proches, je fais toujours attention ... Si je dis quelque chose, c'est que j'ai vérifié l'information. Ou si j'ai un doute sur l'information, on en discute, que moi je n'ai pas vu la même chose. Mais seulement si le sujet tombe sur la table ce n'est pas moi qui abordera le sujet, mais on en discutera et puis on ira vérifier d'autres sources.

I : Et donc, considérez-vous les réseaux sociaux comme une source fiable, de manière générale ?

R : Ça dépend de l'utilisation pour moi. Il est clair qu'il y a beaucoup de deepfakes sur les réseaux sociaux, et je pense qu'il faut faire attention, même si à la base les réseaux, ça peut être une source d'infos. Maintenant ... Je ne sais plus la question ...

I : C'était de savoir si vous faisiez confiance aux réseaux sociaux comme une source fiable ?

R : Je pense qu'on peut leur faire confiance si on est vigilant et qu'on fait attention à l'information, parce qu'il y a beaucoup de deepfakes. Donc c'est 50-50, je dirais.

I : Et donc, pour les médias traditionnels ou des journaux qui sont présents en ligne, est-ce qu'il s'agit d'une source fiable selon vous ?

R : Pour les journaux en ligne, je fais confiance, c'est-à-dire que ce sont des journalistes. Souvent, je vérifie quand même, mais beaucoup moins que sur les réseaux sociaux. J'ai plus tendance à vérifier ce que je vois sur les réseaux sociaux que ce que je vois sur un journal, parce que logiquement, si c'est la rtbf.be, ils vont vérifier l'information avant d'en parler.

I : Et donc, selon vos réponses précédentes, quelles sont les sources d'informations auxquelles vous faites le plus confiance pour être sûre de lire une information vraie ? Cela peut être en ligne ou non.

R : Les revues scientifiques, les journaux, les réseaux sociaux, mais je les utilise plus pour des contenus genre des recommandations sur un lieu, la réputation d'un restaurant, etc.

I : Et avez-vous tendance à croire plus facilement une information lorsqu'elle est accompagnée d'une image ou d'une vidéo, plutôt que juste un texte tout seul ?

R : Oui, j'ai plus tendance à croire. Maintenant, je vérifie quand même, mais c'est vrai qu'une photo ou une vidéo permet d'appuyer ses propos et donc que ça paraît plus vrai.

I : Et donc, si une même information est répétée par exemple plusieurs fois, par plusieurs personnes et sur différents canaux, est-ce que cela renforce sa véracité ? Et pourquoi ?

R : Parce que je me dis que la répétition, si plein de gens se répètent, l'information doit être vraie, puisqu'ils sont tous d'accord sur un sujet.

I : Ok, parfait. C'est la fin des questions, je ne sais pas si vous avez envie d'ajouter quelque chose au sujet ?

R : Non, pas vraiment.

I : Donc on va pouvoir à présent, passer à un petit test. Vous allez être confrontés à trois images, une à la suite de l'autre. Vous aurez 15 secondes pour voir chaque image. A la fin des 15 secondes, à chaque fois, vous pourrez me dire si vous pensez que c'est une vraie ou fausse image et pourquoi. Elles peuvent être toutes fausses, toutes vraies ou vraies et fausses. Et ensuite, à la fin des 3 images, vous pourrez voir les trois photos côte à côte et vous pourrez dire si vous confirmez ou ajustez vos réponses. Voici la première image.



R : (observer l'image pendant 15 secondes)

I : Selon vous, est-ce qu'il s'agit d'une vraie ou d'une fausse image ?

R : J'ai l'impression que c'est une fausse.

I : Ok.

R : Mais je ne saurais pas dire pourquoi. La façon dont Emmanuel Macron est mis me semble bizarre. Mais c'est plus une intuition qu'un réel fait. Juste la photo ne permet pas de dire ce qu'ils ont fait et pourquoi ils sont dans cette situation.

I : Ok, je comprends. Voici ensuite, la seconde image.



R : (Observe l'image pendant 15 secondes) Elle est fausse aussi. Fausse, mais plutôt à cause de l'arrière-plan.

I : Que diriez-vous de l'arrière-plan ?

R : En comparaison avec la première image, les chaises sont plus neutres. Elles sont comme si elles ne rebondissaient pas. Comme si ce n'était qu'une planche de bois. Donc, moins de détails, ça veut dire que ça peut être faux. Et aussi l'arrière-plan avec le mur gris. On a l'impression que ce n'est même plus le bâtiment, le truc présidentiel qu'on aurait envie de voir.

I : Entendu, voici la troisième image.



R : (Observe l'image pendant 15 secondes) Celle-là, elle est fausse. Les deux autres, elles ne sont peut-être pas fausses du coup.

I : Pourquoi ?

R : Parce qu'ils ont l'air de ... Mais déjà, sa tête est disproportionnée. La tête d'Emmanuel Macron est disproportionnée et même l'alignement avec son corps n'est pas correct.

I : Et donc maintenant, en voyant les trois, côte à côte, est-ce que vous confirmez tous vos choix ou est-ce qu'au final, vous remettez en doute certaines photos ?

R : Du coup, je remettrai en doute les deux premières, qu'elles sont réellement vraies.

I : Pourquoi changez-vous d'idée ?

R : Parce que, vu la dernière qui est fausse, ça me paraît logique qu'elles soient vraies. Maintenant, les deux premières ... Juste la première, j'ai un problème avec ... Il y a un truc avec Macron qui me paraît très bizarre.

I : Ah, c'est intéressant.

R : J'ai l'impression au niveau de ses jambes qu'elles sont trop nettes. Je dirais que la première est fausse, que la deuxième est vraie et que la troisième est fausse.

I : Je peux à présent vous donner les réponses. La première était fausse, la deuxième vraie, et la troisième bien évidemment fausse.

R : Ah ok, j'ai bien fait de changer alors.

I : Pour faire un récap, sur quels éléments vous appuyez-vous pour faire vos choix ?

R : Pour la deuxième, j'avais surtout regardé les chaises. Et aussi le mur de derrière qui me paraissait bizarre.

I : Mais pas du tout le contexte, alors, celle-ci ?

R : Non, pas spécialement. Les éléments pour moi qui sont importants à regarder, c'est la qualité de l'image, les détails. Oui, et surtout les mains et les visages. Comment le corps est structuré par rapport aux visages.

I : Merci beaucoup pour vos réponses. Je ne sais pas si vous voulez ajouter quelque chose ?

R : Non. Merci, c'était chouette.

Répondant	Âge	Genre
Numéro 5	16 ans	Homme

I : Bonjour et merci d'avoir accepté de participer à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux sociaux. Pour information, les deepfakes sont des contenus photos ou vidéos truqués créés par des intelligences artificielles. Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage d'une personne ou encore les éléments d'un décor à partir d'une simple photo ou description. Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes. Bien entendu, vous êtes libre à tout moment de ne pas répondre à certaines questions ou encore de mettre fin à l'entretien si vous en ressentez le besoin. Avez-vous des questions avant de commencer ?

R : Non, je n'en ai pas.

I : Alors nous pouvons commencer par la première question. Avais-tu déjà entendu parler des deepfakes auparavant ? Si oui, que sais-tu à leur sujet ?

R : Alors, oui. Je suis déjà tombé sur des trucs générés par IA. Ce sont des vidéos ou des photos qui ont été modifiées pour retirer ou ajouter des choses aux photos et aux vidéos.

I : Et donc, comment en as-tu entendu parler ?

R : Des gens sur les réseaux sociaux qui disaient de faire attention à des photos, en disant que c'était généré par des IA et ils donnaient des techniques pour les repérer.

I : Ok, et donc, que sais-tu à leur sujet à part que c'est fait par des IA ?

R : Je ne sais rien de plus, juste que c'est fait par IA, quoi.

I : Ok, pas de soucis, et donc, as-tu déjà été confronté à des avertissements ou des campagnes de sensibilisation sur les deepfakes ?

R : Non.

I : Jamais ? Que ce soit sur les réseaux sociaux ou même en dehors ?

R : A l'école, le prof nous dit souvent de faire attention à ça ...,mais c'est tout quoi.

I : Ok, et penses-tu que les deepfakes occupent une place importante sur les réseaux ?

R : Ça dépend, je sais pas.

I : Tu sais expliquer ?

R : Bah, en fait, ça dépend, parce que ça dépend déjà si toi, tu tombes sur des trucs comme ça et de comment l'algorithme il est fait, quoi.

I : Ok, et donc, par exemple, de manière générale, si tu dois donner un pourcentage de contenu deepfakes qui circule tous les jours sur...

R : Je n'en sais rien du tout. On va dire 10 %, 10-15 %.

I : Ok, donc quand même pas trop quoi.

R : Bah, je ne sais pas, mais de ce que je vois, je dirai 10-15 %. Après, ça dépend si j'arrive à les voir. Il y a des choses où on sait direct que c'est fait par l'IA et d'autres qu'on ne voit même pas.

I : Ok, et penses-tu qu'en tant qu'utilisateur des réseaux sociaux, tu as un rôle à jouer avec les deepfakes ?

R : A part ne pas regarder les trucs générés par l'IA, pour ne pas que ça fasse des vues et que tout le monde le voit ... et dire de faire attention. A part ça ...

I : Ok, à part ça tu ne vois pas ce que tu peux faire alors ?

R : Ouais.

I : Et donc quelle a été ta première réaction ou pensée lorsque tu as pris connaissance de l'existence des deepfakes ?

R : Si on parle genre, des vidéos sur internet, on va se dire ah, ils font de l'argent sans rien faire. Il y en a, ils travaillent des heures, des heures de montage, de tournage, etc. puis t'en as, ils tapent deux phrases sur un logiciel, ils postent le truc et hop ça marche.

I : Donc tu penses plus à ce qu'on peut faire avec les IA pour le travail, par exemple ?

R : Bah ouais, parce que je sais que certains font comme ça.

I : Donc c'est plus une pensée négative, tu penses pas qu'il y a du positif ?

R : Bah, il y a du positif pour le travail, mais sur les réseaux, les plateformes avec les vidéos quoi, ce n'est jamais utilisé pour de bonnes raisons. C'est simplement pour se faire des sous facilement.

I : Ok, et donc si tu devais dire, par exemple, une phrase qui t'es passée par la tête quand t'as découvert l'existence des deepfakes ?

R : C'est stylé, mais sans plus quoi. Je préfère encore que ce soit un truc ... si on parle d'un truc en rapport avec l'art, je préfère clairement que ce soit fait par une vraie personne, qu'un robot quoi.

I : Ok, et donc as-tu déjà eu le sentiment, quand t'étais face à une image ou une vidéo, que la vidéo ou l'image que t'étais en train de regarder était fausse ?

R : Je remarque souvent que c'est faux avant de me poser la question, haha. En fait, des fois, il y a des vidéos, c'est cramé. Il y a des moments où les IA ne savent pas faire les mains. Je ne sais pas si maintenant elles savent, mais avant elles ne savaient pas faire les mains. Il y a aussi des moments où ça bug beaucoup sur les parties du corps, genre l'épaule ou quoi, il y a des moments où il n'y a rien qui va sur le corps. Du coup, là, tu crames facilement que c'est de l'IA. Sinon, il y a aussi la voix faite par IA, c'est insupportable.

I : Ok, donc quand tu parles de la voix, c'est ...

R : Les voix en IA c'est insupportable, c'est ...

I : Qu'est-ce que tu entends par insupportable ?

R : On entend juste la voix d'un robot quoi et ça sonne bizarre.

I : Oui, donc une voix robotique qui fait que tu remarques directement.

R : Non, mais je ne sais pas, même l'accent de la personne il est bizarre.

I : Et as-tu déjà été influencé par une information qui s'est par la suite avérée fausse ?

R : Non.

I : Jamais ? Par exemple, avoir cru à une fausse rumeur sur une personnalité publique, etc. ?

R : Non, mais ça, en général, si cela vient de personnes random, des gens qu'on ne connaît pas qui repostent des infos, ce n'est pas à prendre au premier degré, quoi. Puis il y a des gens qui font de l'actualité, là c'est déjà un peu plus ... fiable.

I : Ok, donc tu fais toujours gaffe à qui poste ?

R : Ouais. Il y en a qui ... genre si c'est Hugo Décrypte ou des chaînes de télé ou quoi. C'est déjà un petit peu plus crédible que si c'était une personne lambda, quoi.

I : Ok, je vois. Et selon toi, qu'est-ce qui rend un deepfake difficile à détecter à première vue ?

R : Bah... Bah, les photos, c'est compliqué, mais genre si c'est une vidéo, il y a des choses que l'IA ne sait pas forcément refaire, elle ne sait pas faire comme si c'était naturel.

I : Ok, dans les ...

R : Toujours sur les parties du corps, par exemple. Genre les mains.

I : Oui, ok.

R : Cela ne fait pas naturel. Surtout s'il y a des mouvements. Et les détails des mains, des doigts ou quoi.

I : Ok, donc si j'ai bien compris, tu trouves cela plus difficile de détecter une photo qu'une vidéo ?

R : Ouais. Une photo, ça dépend surtout de la qualité, mais souvent comme il n'y a pas de mouvements, de voix etc., il y a moins de détails que l'on peut comparer. Et souvent, les photos sont donc mieux faites que les vidéos. Souvent les vidéos, je suis persuadé que c'est faux alors que les photos ça me met le doute.

I : Et donc, si tu dois distinguer une vraie photo d'une fausse, tu fais comment ?

R : Bah déjà il y a ...

I : Donc les expressions, les mains, etc. comme tu m'as dit juste avant, mais est-ce qu'il y a autre chose que tu regardes ?

R : Ouais déjà ça et sinon bah... Je vais voir dans les commentaires.

I : Ok, les commentaires.

R : Parce qu'il y en a qui analysent le truc, et qui expliquent pourquoi c'est faux.

I : Mais dans les commentaires parfois, il y a quand même beaucoup de personnes qui tombent dans le panneau, comment fais-tu la part des choses ?

R : Oui, mais il y en a toujours au moins un qui dit que c'est de l'IA, et donc, dès qu'il y a des doutes, je regarde un peu plus. J'essaie de voir les détails.

I : Ok, super. Sur une échelle de 1 à 10, si tu devais te noter capable de reconnaître un deepfake, tu te mettrais combien ?

R : Entre 1 et 5.

I : Entre 1 et 5, ok, c'est large. Est-ce que t'as un chiffre plus précis ?

R : Moi je dirais 4 et demi.

I : 4 et demi, ok. Et pourquoi ?

R : Parce qu'en fait si le truc est bien fait, ça va être compliqué de le repérer quoi.

I : Ok donc quand tu te dis 1 et 5, c'est ...

R : En fonction de la qualité, quoi.

I : Donc si c'est de mauvaise qualité, tu te dis 5, et si c'est vraiment une bonne qualité...

R : Oui, c'est ça !

I : Et donc, depuis que tu as connaissance des deepfakes, est-ce que cela a modifié ta manière d'utiliser les réseaux sociaux ?

R : Non, parce qu'en général je ne tombe pas vraiment sur des deepfakes.

I : Donc tu n'as rien changé ? Par exemple, tu n'as pas arrêté d'aller chercher des infos sur les réseaux sociaux ?

R : Je n'ai jamais cherché des infos dessus, enfin pas des infos sur l'actualité en tout cas.

I : Et donc, à présent, est-ce que tu doutes de tout ce que tu vois sur les réseaux sociaux ?

R : Bah en gros, c'est genre... Quand ça vient de youtubeurs, par exemple, que je connais, je sais que c'est fiable, donc ça va, je ne doute pas. Après, il faut que ce soit le compte officiel, quoi.

I : Tu veux dire quand il s'agit de personnalités connues ?

R : Connues, non pas toujours, mais tu sais que c'est déjà fiable, tu sais que ... il y a moins de chances en tout cas qu'il y ait des deepfakes.

I : Oui, donc selon toi, les personnes connues, il y a moins de chances qu'ils diffusent des fausses infos comme les deepfakes ?

R : Oui, voilà alors que quand c'est des personnes lambdas, c'est un peu aléatoire je trouve. En fait, je ne suis pas beaucoup de gens différents sur les réseaux, c'est vraiment max 10 personnes dont je suis sûre.

I : Donc, par exemple, si tu vas sur instagram, ...

R : Ouais, non ... Je ne prends jamais rien au sérieux, rien du tout.

I : À part les personnalités connues ou les personnes que tu suis, c'est ça ?

R : Ouais, ouais, voilà. Alors, par exemple, je suis le truc de RTL, bon bah là c'est des sources fiables normalement, c'est quand même des trucs officiels, normalement ...

I : Pourquoi tu dis normalement ?

R : Je ne sais même plus s'il y a des certifications sur Instagram.

I : Si le compte de RTL n'est pas certifié, tu penses que cela n'est pas fiable ?

R : Non, à partir du moment où il y a beaucoup de gens qui suivent le compte, il ne doit pas y avoir de problème, je crois. Ou alors c'est que c'est un compte dédié juste à ça, à l'IA.

I : Ok je comprends. Et donc as-tu pris des mesures spécifiques pour te protéger contre les fausses infos ou les deepfakes ?

R : Bah, je ne prends pas les infos au sérieux, sauf quand je sais que c'est des sources sûres.

I : Et lorsque tu vois une nouvelle vidéo ou une image en ligne, est-ce que tu prends systématiquement le temps de vérifier la source ?

R : La source, c'est genre quoi ?

I : Bah, par exemple, de te renseigner de qui vient l'information ?

R : Bah en général, je passe, je scrolle.

I : Tu regardes simplement le compte, sinon tu passes ?

R : Ouais, voilà. En général, je regarde deux secondes, puis je bouge directement.

I : Et aujourd'hui, donc t'as dit que t'étais plus vigilant avec le contenu que tu regardais, mais est-ce que t'es plus vigilant avec ce que tu partages sur les réseaux sociaux ?

R : Non, ça je m'en fous. Je partage quand même. De toute façon, en général, quand j'envoie des vidéos, ce sont des trucs drôles.

I : Oui, donc ce ne sont jamais des trucs très sérieux ?

R : Non, ça non, je fais gaffe à ça. Je n'envoie que des trucs drôles, donc si c'est créé par l'IA, au pire ce n'est pas grave.

I : Et considères-tu que les réseaux sociaux sont une source fiable ? Et pourquoi ?

R : Alors, c'est fiable. Seulement pour les... les chaînes d'infos. Par exemple, RTL ou la RTBF, les trucs officiels, là, oui, c'est fiable. Par contre, le reste, non.

I : Ok, je vois, tu as anticipé ma prochaine question, haha. Mais selon toutes tes réponses, quelles sont les sources d'informations auxquelles tu fais le plus confiance ?

R : Euh ...

I : Si tu dois aller chercher une information sûre, tu te dirigeras plus vers où ?

R : Moi, je vais voir sur Google.

I : Google ?

R : Oui, je tape le truc, puis ... je vois si les chaînes de télé ou des articles en parlent, je vais voir là-dessus.

I : Est-ce que tu as tendance à croire plus facilement une information si elle est accompagnée d'un visuel ? Donc, par exemple, une image ou une vidéo, plutôt qu'un texte seul.

R : C'est vrai que quand il y a une vidéo qui est crédible, on va moins vite se dire que c'est un truc fait par IA. On a plus facile de faire confiance à un visuel. Par exemple, si quelqu'un raconte une histoire, s'il parle simplement, on peut douter de ce qu'il dit. Alors que s'il montre des photos de ce qu'il s'est passé, c'est plus crédible.

I : Et par exemple, si une information est répétée plusieurs fois, par différents comptes, sur différentes plateformes, etc., est-ce que, pour toi, ça veut dire qu'elle est vraie ?

R : S'il y a des gens lambdas qui en parlent, mais aussi des médias plus connus, alors là oui.

I : Oui, mais par exemple, sans même vérifier les comptes qui ont publié l'info, est-ce que l'effet de masse peut t'induire en erreur ?

R : Euh ... Par exemple, hier, il y a eu le trailer 2 de GTA 6. Au départ, je n'étais pas au courant, mais j'ai reçu plein de notifications et donc je me suis posé des questions. Si c'était vrai ou juste des gens qui s'amusent à dire n'importe quoi.

I : Est-ce qu'il y a quelque chose que tu as envie d'ajouter au sujet des deepfakes ?

R : Non, pas vraiment.

I : Je n'ai plus de questions, on va donc pouvoir passer à un petit test. Je vais te présenter 3 images, une à la fois. Tu auras 15 secondes pour voir chaque image. Et alors, après chaque image, tu me diras si tu penses qu'elle est vraie ou fausse et pourquoi. Elles peuvent être toutes vraies, toutes fausses, ou vraies et fausses. Il y a une multitude d'options. Voici la première image :



R : (Observe l'image pendant 15 secondes)

I : Selon toi, est-ce qu'il s'agit d'une vraie ou d'une fausse image ?

R : Juste la tête de... attends, comment il s'appelle ?

I : Le président français, Macron ?

R : Macron, oui. Juste la tête de Macron qui est un peu bizarre. Même si c'est possible. Ouais, je ne sais pas. En général, quand on les voit, c'est pour la diplomatie. Je ne vois pas le mec en gros fou rire, mais ...

I : Donc, tu dirais que c'est faux ou vrai, selon toi ?

R : Je n'en sais rien du tout.

I : Il faut trancher, haha;

R : Moi, je dis que c'est faux. Ça peut être faux.

I : Ok, pour quelles raisons ?

R : Donc, plus pour le contexte, parce qu'il se tape une barre, et parce que sa tête me paraît bizarre. En général, quand on les voit, surtout pour le moment, avec les problèmes de taxes etc., je n'imagine pas trop qu'ils rigolent comme des amis. C'est surtout pour la diplomatie, pour négocier, ou juste pour faire des discours. Alors, je ne sais pas pourquoi il est... C'est possible. Non, mais je ne vois pas pourquoi il serait en fou rire, quoi. Ça reste quand même très bien fait si c'est de l'IA.

I : La seconde image



R : (Observe l'image pendant 5 secondes à peine) Y a un gros problème au niveau de la tête, là. Y a un problème au niveau des cheveux.

I : Ok. Et donc, voici la dernière photo.



R : (Observe pendant 15 secondes)

Je dirai, que c'est vrai. Euh, attends, non, la main, haha.

I : Tu dirais donc que c'est vrai ou faux ?

R : Faux. Car la main me paraît bizarre.

I : A présent, tu peux regarder les 3 images en même temps, et les analyser plus longtemps pour confirmer ou ajuster tes choix.

R : La dernière image, il y a un gros problème au niveau des bords aussi. Et les mains aussi, on dirait des palmes. On dirait aussi un fond vert, il y a un problème.

I : C'est quoi le problème dont tu parles ?

R : J'ai l'impression, genre, tu sais, qu'il y a un mec qui a mis ça sur Photoshop, et qu'il l'a mis dans un autre décor. Enfin, je ne sais pas, c'est bizarre. On a l'impression que c'est détourné, quoi.

I : Donc, tu confirmes que la troisième, c'est toujours une IA, selon toi ?

R : Oui.

I : Et pour la première et la deuxième ?

R : La première, c'est faux, mais alors très bien fait. Et la deuxième, c'est faux aussi, la tête de Macron c'est pas possible, haha.

I : Donc, pour émettre tes réponses, tu t'es fié à quoi, pour faire un récapitulatif ?

R : La qualité ... Ouais, la qualité. Et aussi la tête des gens. Si on a un peu la tête des gens en tête, on voit vite qu'il y a un problème. Les mains, le corps en général quoi. Et le décor aussi parfois.

I : Ok merci, je te donne les réponses. Du coup, la première que je t'ai montrée, elle est vraie. Le contexte est un petit peu perturbant, mais elle est bien vraie.

R : Ok, oui, j'avais un doute car la qualité était bien.

I : La deuxième, elle est fausse, tu avais bien raison. Et la troisième était fausse aussi.

R : Oui, la troisième, à la base je n'avais pas vu. Je pensais que c'était vrai puis à la dernière seconde la main ... j'ai vu que la main était bizarre.

I : Et bien, voilà, ici s'achève le petit test. Je ne sais pas si tu souhaites ajouter quelque chose ?

R : Je n'ai rien d'autre à dire.

I : Ok, merci beaucoup à toi !

Répondant	Âge	Genre
Numéro 6	53 ans	Femme

I : Bonjour et merci d'avoir accepté de participer à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux sociaux. Pour information, les deepfakes sont des contenus photos ou vidéos truqués créés par les intelligences artificielles. Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage d'une personne ou encore les éléments d'un décor à partir d'une simple photo ou description. Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes.

Bien entendu, vous êtes libre à tout moment de ne pas répondre à certaines questions ou de mettre fin à l'entretien si vous en ressentez le besoin. Avez-vous des questions avant de commencer ?

R : Non.

I : On peut donc passer à la première question. Avez-vous déjà entendu parler des deepfakes et, si oui, que savez-vous à leur sujet ?

R : Oui, j'en ai déjà entendu parler et puis on a eu un bel exemple aux infos dernièrement avec, je ne sais plus le nom de cette femme-là, qui a versé beaucoup d'argent. C'était à ...

I : Brad Pitt ?

R : Brad Pitt, oui.

I : Et donc, la première fois que vous en avez entendu parler, c'était dans quel contexte ?

R : Non, mais en fin de compte, à partir du moment où on a parlé des intelligences artificielles, j'ai bien compris qu'on allait trafiquer les images. Je veux dire, c'est évident et voilà, par déduction.

I : Si vous deviez les définir, vous les définiriez comment ?

R : C'est-à-dire ?

I : Répondre à la question, c'est quoi les deepfakes.

R : Une arnaque. Une belle arnaque.

I : Pour quelles raisons ?

R : Car c'est tellement réaliste parfois qu'on peut tout trafiquer.

I : D'accord, et avez-vous déjà été confrontée à des avertissements ou des campagnes de sensibilisation sur les deepfakes ?

R : Oui, je crois que je suis déjà tombée 2-3 fois en ligne, mais pas beaucoup. Je ne me souviens même plus de quoi ça parlait.

I : Et pensez-vous que les deepfakes occupent une place importante sur les réseaux sociaux ?

R : De plus en plus.

I : Et si vous deviez dire, par exemple, un pourcentage de deepfakes que vous croisez par jour ? Consciemment ou non, peu importe.

R : Je ne sais pas, mais j'en croise tous les jours. Pas beaucoup, mais quand même, j'en croise tous les jours maintenant, alors qu'avant, pas. Donc ça prend de l'ampleur. Ça prend de l'ampleur, bien sûr.

I : Et pensez-vous qu'en tant qu'utilisateur des réseaux sociaux, vous avez un rôle à jouer avec les deepfakes ?

R : Un rôle, c'est-à-dire ?

I : Un rôle dans le processus de propagation, par exemple.

R : Non, non, non, moi je ne partage pas. Non, surtout pas.

I : Ou pour éviter la propagation, par exemple, est-ce que vous avez un rôle à jouer ?

R : Euh ... moi je ne partage rien du tout. Je ne suis pas comme votre génération où on like, on partage, nani, nana, enfin voilà. Et quand je partage, ce n'est pas des deepfakes, c'est certain.

I : Et donc, quelle a été votre première réaction ou pensée lorsque vous avez pris connaissance de l'existence des deepfakes ?

R : Mais ça se voit. Ça se voit. Ça se voit quand même. Je suis sûre que bientôt ça ne se verra plus et on devra utiliser notre bon sens, pour les détecter. Mais pour le moment, ça se voit. On détecte.

I : Donc vous voulez dire que ce n'est pas encore assez performant ?

R : Ce n'est pas encore assez performant, oui.

I : Ok. Et as-tu déjà eu le sentiment, quand tu étais face à une image ou vidéo qui était fausse, sans en être certaine ?

R : Non, là, pour le moment, j'arrive encore à les détecter. Mais je sais qu'à un moment, on n'y arrivera plus. Enfin, il faudra donc faire appel à notre bon sens.

I : As-tu déjà été influencée par une information qui s'est avérée fausse par la suite ?

R : Ça peut influencer une croyance envers la réputation de quelqu'un ou un comportement ... Mais rien qu'aux informations, les infos, les journalistes, je veux dire, ils mettent ce qu'ils veulent dans les articles.

I : Oui.

R : Vous comprenez ? Je veux dire, moi, je ne sais pas vérifier toutes ces infos. Je veux dire, on est un petit peu à la merci de ces journalistes qui ... peut-être, changent les infos. Normalement, non.

I : Et selon vous, qu'est-ce qui rend un deepfake difficile à détecter à première vue, dans votre fil d'actualité, par exemple ?

R : À partir du moment où, toi-même, tu n'es pas suffisamment informé, tu peux tomber dans le piège. Au niveau de l'information, au niveau des médias, au niveau de ce qui se passe dans le monde.

I : Donc, vous ne parlez pas forcément de la qualité des deepfakes ou quoi, mais plus du sujet, du contenu, de l'info.

R : De l'info qu'ils transmettent. Oui, c'est ça.

I : Et donc, selon vous, comment distinguer un deepfake d'une vraie photo ou vidéo ?

R : Je ne sais pas comment expliquer, mais ça se voit quand même. Une personne, par exemple, on voit que les traits sont plus lisses ... comme s'il y avait un filtre. Ça se voit. Et puis, encore une fois, je veux dire, si tu suis l'actualité et que tu es sûre de toi et que tu vois la personne faire un truc complètement à l'opposé, enfin, je veux dire, un truc loufoque, vraiment aberrant ou un truc qui n'a pas de sens.

I : Ok. Donc, le contexte aussi est important ?

R : Oui, le contexte, oui. C'est un tout. C'est plusieurs facteurs qui entrent en compte. Tous ensemble. Mais en fin de compte, dès que ça arrive sur mon écran, il y a plusieurs facteurs qui entrent en compte comme ça en quelques secondes, et c'est évident.

I : Oui, je comprends.

R : Mais je vois ma fille de 12 ans qui tombe dans le panneau. Eh oui, c'est ça, le problème. Les plus petits tombent dans le panneau. Mais bien sûr. Alors, j'essaie de ... lui expliquer, mais c'est pas facile, quoi. Heureusement qu'elle n'a pas encore accès à tout.

I : Oui, en effet, ça peut être un problème pour les plus jeunes générations. Et sur une échelle de 1 à 10, à quel point vous sentez-vous capable de reconnaître un deepfake ? Et pourquoi ce nombre ?

R : Je vais dire 8. Au niveau visuel, c'est 10 sur 10 pour le moment. Je les choppe tous. Enfin je crois que je les choppe tous pour le moment. Mais pour le contenu, comme je disais ... Il faut être informé. Mais alors, un deep fake, au début, c'est visuel de toute façon. Donc, je dois pas non plus trop me pencher sur la question ... Les deepfakes, c'est un peu comme les fake news, sauf que maintenant, il y a un support visuel en plus qui permet d'appuyer les fausses informations, quoi ?

I : Oui, c'est ça.

R : Bon, je vais dire 8, alors.

I : Et depuis que vous avez pris connaissance des deepfakes, est-ce que vous avez modifié votre manière d'utiliser les réseaux sociaux ?

R : Je suis plus vigilante. Je suis au taquet. Ça, c'est certain. Oui, oui, bien sûr.

I : Quand vous dites au taquet, vous faites quoi par exemple ?

R : Je mets mes lunettes, haha, pour être sûre de bien voir.

I : Et donc, maintenant, quand vous êtes sur les réseaux sociaux, est-ce que vous doutez de tout plus facilement ? Par exemple, si tu es face à une photo, tu vas toujours te poser la question, est-ce que ce ne serait pas un deepfake ?

R : Pour l'instant, je ne doute pas. Mais je sais qu'un jour, je douterai.

I : De tout ?

R : Je sais qu'un jour, oui, ils seront tellement performants, ils seront tellement au point qu'il faudra ...

I : Donc, pour le moment, vous doutez uniquement, quand il y a quelque chose qui vous paraît étrange ou qui vous dérange dans l'image, l'information ?

R : L'information. Je me base vraiment là-dessus quand il y a un truc dérangeant au niveau de l'information, je vais regarder plus en profondeur l'image au cas où je n'aurais pas détecté le deepfake.

I : Et donc, à part être plus vigilante, est-ce que vous avez pris des mesures spécifiques pour vous protéger contre les fausses informations ?

R : Non.

I : Non, rien du tout ?

R : Non, personnellement, non.

I : Et est-ce que vous prenez systématiquement le temps de vérifier la source de chaque information que vous voyez, ou alors ...

R : Ça dépend. Ça dépend, mais bien sûr que ça m'arrive, bien sûr, et quand même assez souvent, de vérifier si j'ai un doute.

I : Et si vous devez vérifier la source, comment faites-vous ?

R : Oui, mais tu vois, encore une fois, je me rends compte que je m'égare un petit peu parce que la photo, la vidéo, je ne vérifierai pas parce que je suis sûre de mon jugement, pour l'instant. C'est vraiment le contenu. Là, parfois, j'aurai un doute, et là, j'irai voir sur d'autres sites, j'irai me renseigner ailleurs, oui.

I : Et donc, tu dis que tu es plus vigilante lorsque tu regardes du contenu sur les réseaux sociaux, mais est-ce que maintenant tu es plus vigilante aussi dans ce que tu partages ? Vous m'avez dit que vous ne partagiez rien sur les réseaux, mais, par exemple...

R : Ce sont des blagues, tu vois, je ne partage que des bêtises, et encore, ça reste vraiment ... Ce n'est pas mon truc.

I : Et par exemple, partager sur les réseaux, ce n'est pas trop ce que vous faites, mais par exemple, si vous voyez une information, mais que vous n'êtes pas sûre de celle-ci, est-ce que vous allez quand même en parler à vos proches ?

R : Si j'ai un doute, par exemple, au sein de ma famille, j'en parle avec mon mari.

I : Ok, oui.

R : Parce que mon mari est au taquet aussi au niveau des infos, donc, je veux dire, si je suis tombée sur l'info dans la journée, je sais que lui, pareil. Je passe par quelqu'un qui est informé. Voilà, et je demande son opinion.

I : Ok, et donc, considérez-vous les réseaux sociaux comme une source fiable ?

R : Non, enfin, ce n'est pas fiable à 100 %, on le sait, bien sûr, mais c'est super chouette de les avoir.

I : Et pour les médias traditionnels, aux journaux qui sont maintenant présents en ligne, est-ce que pour vous, ce sont des sources fiables ?

R : Ben non, pas à 100%, justement, non.

I : Et pourquoi ?

R : Ben, parce qu'on se fait manipuler par les journalistes. Enfin, je veux dire, même en politique, je veux dire, ils mettent un petit peu ce qu'ils veulent pour influencer l'opinion publique.

I : Et donc, vous croyez qu'ils pourraient utiliser les deepfakes pour aller dans leur sens ?

R : Je ne sais pas, ça viendra ou alors cela se fait déjà. Je pense que ça ne leur posera pas de problème aujourd'hui.

I : Et selon vos réponses, quelles sont alors les sources d'informations auxquelles vous vous fiez pour être sûre d'avoir une information vraie ?

R : En fin de compte, je fonctionne un petit peu comme si j'allais à la bibliothèque. Je veux dire, à la bibliothèque, tu prends plusieurs bouquins, pour être sûre de ton info. Sur les réseaux, c'est pareil. Je vais chercher un petit peu à gauche, à droite, et je vois si tout va dans le même sens.

I : Et donc, de manière générale, est-ce que vous avez tendance à croire plus facilement une information, si elle est accompagnée d'une image ou d'une vidéo, plutôt que juste une information qui est donnée sous forme de texte ?

R : Non, ça non.

I : Le visuel ne va pas avoir un effet de preuve, par exemple ?

R : Non, pas spécialement.

I : Pour quelles raisons ?

R : Parce qu'encore une fois, je me base beaucoup sur le contenu. Oui, plus sur le contenu que sur l'image. On sait que l'image peut être trafiquée, du coup...

I : Et si, par exemple, une information est répétée plusieurs fois par plusieurs comptes sur les réseaux sociaux ou même différents canaux, est-ce que vous pensez que ça va vous influencer à croire en l'information ?

R : Non. Si beaucoup de gens en parlent et disent tous la même chose ? Non.

I : L'effet de masse ne va pas avoir un impact ?

R : Non, avant oui, mais plus maintenant. Je suis déjà tombée dans ce piège, mais plus maintenant.

I : C'est la fin des questions, merci beaucoup. Est-ce que vous avez quelque chose à rajouter, une remarque, une information supplémentaire ?

R : Je pense qu'il n'y a pas de solution miracle. Non, il n'y en a pas, mais tout ça va prendre de l'ampleur. Il va falloir briefer les jeunes, justement, pour qu'ils ne tombent pas dans le panneau.

I : Vous parlez notamment de votre fille ?

R : Oui, les plus jeunes, ceux qui commencent à utiliser les réseaux sociaux.

I : Super, merci beaucoup. A présent, j'ai trois photos que je vais vous montrer pendant 15 secondes.

R : Je vais tomber dans le panneau. Non, moi, jamais, je ne me suis jamais fait avoir. Ah, mais ce serait la meilleure, haha.

I : Maintenant, je vais vous montrer trois photos. Une première photo durant 15 secondes, ensuite, vous me dites si vous pensez que c'est une vraie ou fausse image. Puis, une seconde image et une troisième. Elles peuvent être toutes les trois vraies, toutes les trois fausses. Il peut y avoir des combinaisons vraies-fausse. Et alors, après, vous les aurez toutes les trois côte à côte, et vous pourrez ajuster vos choix ou les confirmer.

R : Mais c'est des photos sur quoi ?

I : Ce sont des personnages politiques.

Voici la première.



R : (observe l'image pendant 15 secondes)

I : Selon vous, c'est une vraie ou fausse image ?

R : On ne sait pas voir ça sur un grand écran ?

I : Non ici, le but est d'imiter l'apparition de ces photos sur votre fil d'actualité lorsque vous naviguez sur votre téléphone. On utilise beaucoup plus fréquemment son téléphone que son ordinateur portable, par exemple.

R : Ok, donc ce sont deux bouffons, donc ça pourrait être vrai.

I : Vrai ou faux alors ?

R : Vrai, dernier mot. Ça pourrait être une vraie photo.

I : Et pour quelles raisons ?

R : La situation est bizarre, mais eux aussi ils sont bizarres. Donc ça peut le faire.

I : Ensuite, seconde photo.



R : Je peux changer d'avis après ?

I : A la fin des trois photos, vous pourrez les voir côte à côte et confirmer ou non vos choix.

R : D'accord. (Observe l'image pendant 15 secondes) C'est faux. C'est un deepfake.

I : Et pourquoi ?

R : Ben ... Mais non, quoi. C'est ...

I : Qu'est-ce qui vous paraît faux ?

R : Enfin, tu vois, les ... C'est Macron qui pose problème. Il a une tête disproportionnée, on dirait, je crois.

I : Et la troisième photo.



R : C'est une fausse.

I : Et pourquoi ?

R : C'est toujours Macron qui me pose problème. Je ne sais pas pourquoi j'ai ... C'est l'impression que j'ai que Macron a été retouché.

I : OK. Donc, pour vous, c'est faux.

R : Oui. Moi, pour moi, la première photo, elle est vraie. Et les deux dernières sont fausses.

I : Maintenant, je vous laisse voir les trois en même temps.

R : (Observe plus longuement les images) Je suis impatiente de connaître les réponses.

I : Vous gardez les mêmes choix, du coup ?

R : Oui, mais attends, je vais quand même regarder encore un peu. Je peux agrandir ?

I : Oui, maintenant, c'est l'observation plus en détails.

R : C'est bon, je garde mes choix.

I : Et donc, sur quel élément, en général, vous vous basez pour faire vos choix ?

R : On a l'impression qu'elles ont été retouchées. Tu vois ?

I : Vous voyez ça à quoi ?

R : La main, là. On a l'impression que ça a été dessiné, et le contour aussi, là. Ça, ici, on dirait que ça a été coupé et remis dans le cadre. (parle de la troisième image)

I : Donc, pour les réponses, vous avez tout juste. La première, c'est bien une vraie image, même si la situation peut induire en erreur.

R : C'est ce que je disais donc. Si tu connais un petit peu l'info, tu sais. Ce sont deux bouffons. Je ne dis pas qu'ils ne font que de la merde, mais ils font aussi beaucoup de bêtises.

I : Et donc les deux autres, elles sont bien fausses.

R : La deuxième est flagrante, mais la dernière, je sais que ma fille viendrait me voir pour me demander : "Maman, est-ce que c'est vrai ça ?"

I : Oui, je comprends, donc vous pensez que ça dépend vraiment de la qualité de l'image.

R : Ah beh oui, totalement, moi je suis informée et je doute plus de tout, donc je fais attention et je détecte. Mais une image de bonne qualité, pour une enfant qui n'est pas assez informée, oui, ça paraît réel.

I : Super, merci beaucoup pour vos réponses, cela va beaucoup m'aider.

R : Pas de soucis, c'était chouette. j'avais peur de me tromper partout, mais au final, ça va, j'ai bien évalué ma capacité à les détecter, haha.

I : Oui, franchement, vous vous en êtes bien sorties.

Répondant	Âge	Genre
Numéro 7	56 ans	Femme

I : Bonjour et merci d'avoir accepté de participer à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux sociaux. Pour information, les deepfakes sont des contenus photos ou vidéos truqués créés par les intelligences artificielles. Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage d'une personne ou encore les éléments d'un décor à partir d'une simple photo ou description. Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes.

Bien entendu, vous êtes libre à tout moment de ne pas répondre à certaines questions ou de mettre fin à l'entretien si vous en ressentez le besoin. Avez-vous des questions avant de commencer ?

R : Non.

I : Aviez-vous déjà entendu parler de deepfake auparavant ?

R : Non.

I : Vous ne connaissiez pas le nom deepfake, ou même le concept ne vous disait rien ?

R : Le concept et le mot, je ne connaissais pas.

I : Non, du tout ?

R : Non, vraiment pas.

I : Et donc, avez-vous déjà été confrontée par exemple à des avertissements ou des campagnes de sensibilisation sur les deepfakes ?

R : Non, pas du tout. Je ne connaissais pas du tout.

I : Pensez-vous que les deepfakes occupent une place importante sur les réseaux sociaux ?

R : Je ne sais pas du tout, je ne connaissais pas le concept avant. Des fausses images, oui, il y en a, mais combien, je ne sais pas.

I : Si vous deviez donner un pourcentage, à titre indicatif ?

R : Peut-être 10 %.

I : Ok, Et en tant qu'utilisateur des réseaux sociaux, pensez-vous que vous avez un rôle à jouer avec les deepfakes ?

R : Euh oui, quand même, déjà juste apprendre à connaître ce que sont les deepfakes. En savoir plus.

I : Ok oui. Et donc, maintenant, en vous expliquant ce que sont les deepfakes, quelle est votre première réaction ou pensée ?

R : Il faut faire attention à internet. Surtout s'il y a des choses qui sont truquées. Je ferai plus attention.

I : Ok. Et avez-vous déjà rencontré une image ou une vidéo sur les réseaux sociaux que vous soupçonniez d'être totalement fausse ?

R : Oui. Un truc sur Facebook avec un homosexuel qui était marié avec une femme. Je savais qu'il était homosexuel, donc je me suis douté que c'était une fausse photo.

I : Ok. Donc votre connaissance envers la personne a fait que ...

R : Ouais ... a fait que je me doutais que ce n'était pas la bonne image.

I : Et pensez-vous avoir déjà été influencé par une information qui s'est avérée fausse par la suite.

R : Non ?

I : Pas de souvenirs ou jamais ?

R : Pas de souvenirs.

I : Et selon vous, qu'est-ce qui rend un deepfake difficile à détecter à première vue ?

R : Difficile à détecter... Bah oui, c'est vrai que moi, je me suis fait hacker sur mon Facebook. Donc la personne qui avait fait la demande, c'était super bien fait, parce que moi, je ne me suis même pas rendu compte que c'était un truc comme ça. Donc voilà. Les phrases qu'il avait mis j'avais jamais ... Je pensais que c'était la personne proprement dite qui parlait, et pas le hacker ou un truc comme ça.

I : Ah oui, donc par exemple vous avez reçu des messages de quelqu'un qui s'était fait pirater ?

R : Oui.

I : Ok. Donc ça, c'était plus en texte, c'était pas forcément en images ou vidéos.

R : Non, c'était en texte.

I : Et en images ou vidéos, qu'est-ce qui pourrait rendre cela difficile à distinguer le vrai du faux ?

R : C'est de l'intelligence artificielle, je ne sais pas.

I : Non ?

R : Non.

I : Et si vous voyez une photo, par exemple maintenant que vous savez que ça existe, comment essayerez-vous de distinguer le vrai du faux ?

R : J'irai voir des informations sur Google.

I : Google, ok.

R : Sur des trucs Facebook, par exemple, parfois il y a des gens qui mettent que Mylène Farmer est morte, donc d'office, je vais revoir sur Google si l'information est bonne.

I : Ok. Sur une échelle de 1 à 10, à quel point vous sentiriez-vous capable de reconnaître un deepfake si ça apparaissait dans votre fil d'actualité ?

R : 7

I : 7 ? Et pourquoi ce nombre ?

R : Parce que je ne suis pas tout à fait capable de reconnaître. Je ne mettrais pas 9, parce que voilà, comme je te dis, ils sont tellement bien faits des fois. Mais je ne suis pas totalement conne, je veux dire, si je vois l'image, voilà, je m'informe et voilà. Je dirai que je suis entre les deux.

I : Ok, donc vous connaissez quand même le principe de truquer les images.

R : Ouais. Le mot ne me disait rien, et l'explication non plus, mais plus on avance, plus je vois que j'en ai déjà rencontré un ou l'autre.

I : Ok. Et donc, depuis que vous savez ce qu'on peut faire avec l'intelligence artificielle sur les réseaux sociaux, donc modifier les images, truquer ce que dit une personne, etc., est-ce que vous avez modifié votre manière d'utiliser les réseaux sociaux ?

R : Non. Non, je continue.

I : Et doutez-vous à présent de toutes les informations que vous croisez, en vous disant que cela peut être faux, ou alors seulement quand l'information paraît bizarre, si vous avez un doute sur l'image, etc. ?

R : Non, non, je fais attention à tout maintenant.

I : Directement à tout ?

R : Ouais.

I : Ok. Et donc, pourquoi cela ?

R : Parce que j'ai peur d'avoir des soucis. Voilà, on ne sait jamais ce qui peut arriver.

I : Ok. Et donc, avez-vous pris des mesures spécifiques pour vous protéger contre les fausses informations ?

R : Je devrais, non.

I : Ok. Donc vous n'avez pas d'habitudes sur les réseaux sociaux pour éviter de croire en de fausses informations, de fausses photos ou vidéos ?

R : Non, mais je devrais.

I : Et donc, lorsqu'une nouvelle vidéo, une image en ligne circule, est-ce que vous prenez systématiquement le temps de vérifier la source, ou encore une fois, que quand cela vous paraît bizarre ?

R : Oui, quand ça me paraît bizarre. Pas systématiquement.

I : Et donc, si vous devez vérifier la source à partir de, par exemple, Facebook, comment faites-vous ?

R : Bah, avec une vidéo, c'est plus compliqué, ça. Ça, une vidéo, s'il me semble vraiment que c'est un truc bizarre, je ne l'ouvre pas ... du tout. Et je zappe.

I : Ok. Et donc, vous dites que vous êtes beaucoup plus vigilante, quand même, dans ce que vous regardez, mais est-ce également le cas avec ce que vous partagez sur les réseaux sociaux ?

R : Moi, je ne partage pas. Jamais, non.

I : Donc, si vous ne partagez pas, on peut peut-être reformuler la question. Avant d'en parler à vos proches, votre famille, si vous n'êtes pas sûre de l'information, allez-vous tout de même en parler ? Ou allez-vous d'abord vous assurer qu'il s'agit de vraies informations ?

R : Non, pas spécialement. J'en parle quand même et après, on voit. On en parle ensemble.

I : Et considérez-vous les réseaux sociaux comme une source fiable ?

R : Non.

I : Et pourquoi ?

R : Parce que, depuis qu'il y a ça, c'est vraiment ... En plus, maintenant, comme on peut partager avec tout le monde, plus seulement avec nos amis, je vais dire, parce qu'on voit toutes les vidéos de tout le monde sur le réseau, je trouve que les personnes qui sont mal intentionnées, directement, elles peuvent aller sur Facebook, sur Twitter ou plein de choses.

I : Donc, vous n'y faites plus confiance comme une source fiable maintenant qu'il y a les deepfakes, ou c'était déjà le cas avant ?

R : Avant, je ne sais pas, au début, quand j'ai commencé, je trouvais que c'était fiable. Mais maintenant, plus.

I : Et donc, par exemple, les médias traditionnels ou les journaux qui sont maintenant présents en ligne, est-ce que vous leur faites confiance, ou comme c'est en ligne ...

R : Ah oui, non, ça, en ligne, oui. Les journaux, je préfère quand même le papier, mais les journaux, je prends...

I : Et pourquoi, du coup, eux, c'est OK, en ligne ?

R : Parce que je n'ai jamais vu vraiment un truc qui aurait pu être faux dans un journal, à part le 1er avril. Bah, c'est normal. Ils sont quand même censés ... Ils sont censés vérifier leur source, et tout ça.

I : OK. Et donc, selon vos réponses précédentes, quelles sont alors les sources d'informations que vous utilisez au quotidien pour être sûre d'avoir une information vraie ?

R : Les journaux, bien entendu. Ou alors, 7 sur 7.

I : C'est des infos en ligne ?

R : Sur le téléphone, oui.

I : Et par exemple, avez-vous tendance à croire plus facilement une information si elle est accompagnée d'un visuel, par exemple une image ou une vidéo, plutôt que simplement quelqu'un qui va écrire un article sans ...

R : Ah oui, sans photo.

I : Oui.

R : Non, ça n'a aucun impact, non.

I : Ok. Et pourquoi ?

R : Parce que voilà, je lis l'article, et voilà, s'il n'y a pas de photo, il n'y a pas de photo. C'est vrai que ça pourrait être bien aussi, mais bon, voilà.

I : Oui. Quand vous dites que cela pourrait être bien aussi, vous entendez quoi par là ?

R : Oui, le visuel est toujours un peu mieux, parce que voilà, tu vois un peu ... Je pense que le cerveau réagit plus à une photo que de lire un article. Voilà, il y a plus d'impact, je veux dire, visuel si tu as une photo de quelque chose que si tu lis.

I : OK. Et si, par exemple, une information est répétée plusieurs fois par plusieurs personnes sur Facebook, par exemple, est-ce que cela va renforcer votre confiance en la véracité de l'information ?

R : Non.

I : L'effet de masse ne va pas avoir un impact ?

R : Non, parce que je me dis, voilà... C'est peut-être aussi une personne qui a... répété ce qui a été vu par d'autres personnes.

I : Ok, je comprends ce que vous voulez dire. C'est la fin des questions, je ne sais pas si vous souhaitez ajouter quelque chose avant de faire un petit test.

R : Non.

I : Donc, je vais vous montrer plusieurs photos, ce seront trois images. Vous aurez 15 secondes pour observer l'image comme si vous naviguiez sur votre téléphone. Après ces 15 secondes, vous aurez l'occasion de me donner votre avis. Donc, de m'expliquer si cela vous paraît être une vraie ou fausse image. Ensuite, à la fin des trois images, vous aurez l'occasion d'ajuster vos choix ou de les confirmer. Les images peuvent être toutes vraies, toutes fausses ou une combinaison de vraies et de fausses images.

R : Ok, ça marche.

I : Voici la première image.



R : (Observe l'image pendant moins de 15 secondes) Non, ça, c'est vrai, ça. Je pense.

I : Pourquoi pensez-vous que c'est une vraie photo ?

R : Parce que je sais qu'ils ont eu une réunion, tous les deux, il n'y a pas si longtemps que ça. Donc, je pense plutôt que c'est vrai.

I : OK. Pour la deuxième, c'est celle-ci.



R : (Observe l'image pendant moins de 15 secondes) Non, ça, je ne pense pas. Parce que rigoler comme ça, ça m'étonnerait.

I : OK. Donc, c'est plus pour le contexte, celle-ci.

R : Oui, voilà.

I : OK. Et ensuite, la troisième.



R : (Observe l'image pendant 15 secondes) Là, je dirais plutôt qu'on a changé l'image. Parce qu'ils ont l'air d'avoir une tête un peu bizarre.

I : OK. Donc, à présent, vous pouvez naviguer entre les 3 photos et faire ce dont vous avez envie.

R : (Observe attentivement les images)

I : Est-ce que vous confirmeriez vos choix ou souhaitez-vous les changer ?

R : Oui, je confirme.

I : Pouvez-vous me dire, pour faire un récap, sur quels éléments vous êtes-vous basée pour faire vos choix ?

R : La première, l'image, elle est vraiment vraie. Le contexte, les gens et tout ça, je pense que c'est vrai. La deuxième, ça m'étonnerait qu'ils rigolent comme ça à une réunion entre deux présidents. Et la troisième, c'est l'image qui a l'air truquée.

I : OK. Je peux à présent vous donner les réponses. Donc, la première, c'était une fausse. C'est moi qui l'ai faite. Si l'on zoome, on peut voir des défauts au niveau de la main de Macron. Les personnes sont très détournées, il n'y a aucun cheveu qui dépasse.

R : Ah d'accord.

I : La deuxième, c'est une vraie photo. Sortie du contexte, cela peut paraître bizarre, mais c'est bien une vraie photo. Sur cette photo, on voit les petits détails sur les cheveux, les visages, etc.

R : Ok, oui.

I : Et la troisième est bien évidemment fausse.

R : Ok, donc je n'ai trouvé qu'une seule réponse. Mais c'est bien, au moins, je vois que je ne suis pas du tout douée pour les détecter.

I : Merci, je n'ai plus de questions. Voulez-vous ajouter quelque chose ?

R : Non, je vous ai tout dit, haha.

I : Merci beaucoup.

Répondant	Âge	Genre
Numéro 8	52 ans	Homme

I : Bonjour et merci d'avoir accepté de participer à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux sociaux. Pour information, les deepfakes sont des contenus photos ou vidéos truqués créés par les intelligences artificielles. Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage d'une personne ou encore les éléments d'un décor à partir d'une simple photo ou description. Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes.

Bien entendu, vous êtes libre à tout moment de ne pas répondre à certaines questions ou de mettre fin à l'entretien si vous en ressentez le besoin. Avez-vous des questions avant de commencer ?

R : Non.

I : Aviez-vous entendu parler des deepfakes auparavant ?

R : Avant l'intelligence artificielle ?

I : Non, avant que je vous interroge maintenant.

R : Oui, vite fait comme ça quoi.

I : Mais est-ce que vous connaissiez le terme deepfake ?

R : Deepfake, oui. Oui, quand même. Donald Trump, il en parlait beaucoup.

I : Et que savez-vous à leur sujet ?

R : Rien, je n'y fais pas trop attention.

I : Vous savez comment ça fonctionne ?

R : Pas du tout.

I : Vous savez donc juste ce que je viens de vous expliquer ?

R : Ouais, ouais. Je ne sais pas comment ça fonctionne, je sais juste que c'est truqué grâce à l'intelligence artificielle.

I : Et donc, vous avez entendu parler comment ?

R : A la télé.

I : A la télé ? Au JT par exemple ?

R : Oui, on a vu quelques trucs, quelques exemples dont parlait Donald Trump, par exemple.

I : Quels exemples ?

R : Oh, je ne me souviens plus, mais je sais que c'est là où j'ai vraiment compris que cela existait.

I : Et avez-vous déjà été confronté à des avertissements ou des campagnes de sensibilisation sur les deepfakes ?

R : Non, du tout. Du tout. Fin juste à la télé, quand ils en parlent je me dis bien qu'il faut faire attention, mais pas un truc en particulier, non.

I : Et pensez-vous qu'ils occupent une place importante sur les réseaux sociaux ?

R : Pas encore.

I : Pas encore ?

R : Non.

I : Donc vous pensez que par la suite ça pourrait, mais pour l'instant il y en a peu ?

R : Je pense que oui, mais après, je ne suis pas adepte non plus des réseaux sociaux. Peut-être que celui qui est sur les réseaux sociaux H24, il est plus sujet à être confronté à ça, mais moi... Non.

I : Et vous deviez dire un pourcentage de deepfakes que vous pourriez croiser ?

R : Je ne sais pas moi, 2-3 % ? Et encore ? Je ne vais pas être confronté à ça. Moi, c'est beaucoup des bricolages, des trucs comme ça, il n'y a aucun intérêt là-dedans à faire des fausses vidéos ou des fausses photos.

I : Vous pensez que vous ne regardez pas un contenu qui donnerait envie aux gens de créer de fausses photos ou vidéos ?

R : Oui, voilà, cela n'a aucun intérêt.

I : Et pensez-vous qu'en tant qu'utilisateur des réseaux sociaux, vous avez un rôle à jouer avec les deepfakes ?

R : Euh ...

I : Par exemple un rôle à jouer dans leur propagation ou pour stopper leur propagation justement ?

R : Non. Je n'ai pas le pouvoir de faire ça.

I : Quelle a été votre première réaction ou pensée lorsque vous avez pris connaissance de l'existence des deepfakes ?

R : Que je ne suis pas concerné, comme beaucoup de choses.

I : Donc vous pensez que vous n'êtes pas du tout concerné par les deepfakes ?

R : Non du tout.

I : Ok. Et avez-vous déjà eu le sentiment qu'une image ou une vidéo sur les réseaux sociaux pouvaient être truquées ? Donc, par exemple, vous étiez face à une image ou une photo pour laquelle vous aviez un doute ?

R : Non.

I : Jamais ?

R : Je pense que c'est tellement bien fait que... je suis passé à côté, quoi.

I : Ok, je vois. Et pensez-vous avoir déjà été influencé par une information qui s'est avérée fausse par la suite ? Donc vous vous êtes rendu compte par la suite que vous aviez cru à une fausse information, quoi ?

R : Non, parce que je ne fais pas attention, je zappe juste les infos sur les réseaux sociaux.

I : Vous scrollez simplement sans trop réfléchir ?

R : Oui, voilà.

I : Et selon vous, qu'est-ce qui rend un deepfake difficile à détecter sur les réseaux sociaux ?

R : Je n'ai aucune idée. La qualité, comme je dis c'est... je ne sais même pas si je le verrai donc... Je pense que c'est trop bien fait.

I : Ok. Et donc, par exemple, si vous êtes confrontée à une vidéo, comment feriez-vous pour essayer de voir s'il s'agit d'une vraie ou une fausse vidéo ? Même question pour les photos.

R : Je ne connais pas assez l'informatique et tout ça pour... Même en zoomant sur la photo, par exemple... Je ne connais pas assez.

I : Non aucune idée ? Tu n'as jamais essayé de distinguer ...

R : Non, jamais, si c'est faux, je m'en fous.

I : Si c'est une information sur le bricolage, en effet, mais si c'est une information qui concerne un politicien, par exemple ?

R : Je ne sais pas. Je l'entendrai bien à un moment par des collègues.

I : Oui, ok.

R : Mais non, je ne m'intéresse pas à ça. Non. Je n'essaie même pas.

I : Et donc, sur une échelle de 1 à 10, à quel point vous sentez-vous capable de reconnaître un deepfake sur les réseaux sociaux ?

R : Entre 0 et 1.

I : Et pourquoi ce nombre ?

R : Parce que je ne suis jamais confronté à ça. Donc, si vraiment... Allez, c'était dans mon boulot. Voilà, je serais amené à faire attention. Mais là, je ne suis pas concerné du tout. Je ne m'y intéresse pas, je ne regarde pas, je ne fais pas attention. Je prends juste l'information. Et quand même, comme je te dis, s'il y a un truc qui me semble un peu louche, j'hésite ... et je l'entendrai bien d'autres personnes.

I : Ok, oui, je comprends.

R : Tu vois. Si on me dit que Donald Trump est homosexuel, alors je m'en fous, je ne vais pas chercher plus loin dans l'affaire, je l'entendrai bien plus tard.

I : Ok, donc vous allez attendre que cela se dise ailleurs ?

R : Voilà. Ou alors je demanderai à des collègues s'ils ont entendu. Si c'est un truc vraiment... Oui, comme je te dis, si on dit que Donald Trump est homosexuel, tout le monde va en parler. Et je pourrai le confirmer à ce moment-là. Voilà.

I : Et depuis que vous avez pris connaissance des deepfake, est-ce que cela a modifié votre manière d'utiliser les réseaux sociaux ?

R : Pas du tout ... Pas du tout.

I : Et par exemple, doutez-vous à présent de toutes les informations que vous rencontrez sur les réseaux sociaux en vous disant que cela peut être faux ?

R : Non.

I : Non, vous continuez à faire comme d'habitude ? Vous ne doutez que lorsque quelque chose vous paraît bizarre ?

R : Même pas, parce que, comme je te dis, moi, je ne suis pas sur des sites non plus spécifiques ou alors je ne fais pas grand-chose sur les réseaux sociaux en fait. Je regarde

Facebook, c'est tout. Pour le reste, si j'ai besoin d'info sur un produit pour mon bricolage, je vais taper sur Google, c'est tout.

I : Vous n'allez donc pas sur les réseaux sociaux pour vous informer ? Ce n'est que du divertissement.

R : Ouais, ouais. Le matin, je suis en train de déjeuner, je fais défiler. Je vois ce qui se passe, oui.

I : Donc, avez-vous pris des mesures spécifiques pour vous protéger contre les fausses informations ?

R : Pas du tout.

I : Et lorsque vous voyez une vidéo ou une image en ligne, est-ce que vous prenez systématiquement le temps de vérifier la source, si cette information vous intéresse par exemple ?

R : Non.

I : Et si vous deviez vérifier la source, vous feriez comment ?

R : Je ne sais pas. Moi, je le dis, je suis tellement nul en informatique.

I : Vous pensez que vous ne seriez pas capable de le faire ?

R : Même le paiement bancaire, c'est pas moi qui le fait. Je ne sais pas comment on fait un paiement bancaire.

I : Ok, oui je comprends un peu plus votre utilisation.

R : Donc, je suis vraiment nul, nul, nul.

I : Et donc, vous n'êtes pas beaucoup plus vigilant avec ce que vous regardez, comme vous m'avez dit, mais avec ce que vous partagez ?

R : Je ne partage rien.

I : Non, vous ne partagez rien du tout ?

R : Non, je ne sais pas comment on fait.

I : Et donc, par exemple, vous ne partagez pas, mais si vous devez en parler à vos proches, par exemple, mais que vous n'êtes pas sûr de l'information, vous en parlez quand même ?

R : Ah ça non, non, je vais me renseigner. Je ne veux pas dire de conneries. Oui, je vais me renseigner.

I : Comment allez-vous procéder ?

R : Je vais aller voir sur Google. Je n'ai aucun intérêt à donner une mauvaise info à quelqu'un.

I : Mais j'ai envie de savoir, par exemple ... Parfois vous pouvez voir une information et vous vous dites que c'est vrai, vous ne vous posez pas forcément de questions. Dans cette situation, est-ce que vous en parlez quand même ?

R : Oui, mais je ne dirai pas l'info comme si j'en étais sûre. Je dirais, vous avez entendu que ... Je ne vais pas lancer le bazar. Je vais me renseigner. Si c'est un sujet intéressant.

I : Oui, ok. Et considérez-vous les réseaux sociaux comme une source fiable ?

R : Ça dépend de ce qu'on appelle les réseaux sociaux. Si on parle des groupes, des messages ou quoi.

I : Je parle de ce qu'on peut voir sur les différents comptes, sur votre fil d'actualité, etc.

R : 50-50. Parce que les gens mettent n'importe quoi là-dessus. Il y a des gens qui mettent des choses bien et intéressantes et des conneries. Donc, c'est à prendre avec des pincettes, on va dire.

I : OK. Et par contre, les médias traditionnels et les journaux qui sont maintenant présents en ligne, pour vous, c'est une source fiable ?

R : Oui, c'est RTL, par exemple. Parce que c'est des trucs un peu connus. Ce sont des trucs officiels.

I : Et donc, selon vos réponses précédentes, quelles sont les sources d'informations auxquelles vous faites le plus confiance si vous voulez avoir une information vraie ?

R : Internet et la télé.

I : Internet, vous parlez de Google ou des réseaux sociaux ?

R : Google. Pas les réseaux sociaux, Google.

I : Donc, si vous voyez une information qui dit que Donald Trump est homosexuel, par exemple, pour être sûr de l'information, vous allez voir sur Google ?

R : Non, même pas. J'attends d'avoir d'autres infos, d'entendre plus d'infos. Ça ne m'interpelle pas, ce n'est pas ma priorité.

I : OK.

R : Si on me dit que, par exemple, il y a un problème sur toutes les KIA, comme c'est ma voiture, je serai concerné, donc je vais aller vérifier sur Google.

I : Oui donc, s'il y a un sujet qui te touche ?

R : Oui.

I : Et avez-vous tendance à croire plus facilement une information si elle est accompagnée d'une image ou d'une vidéo ? Donc, par exemple, sur Facebook, vous voyez quelqu'un qui dit

« Nouvelle loi cette année », juste en texte, mais s'il y a une photo ou une vidéo qui en parle, vous y croirez plus facilement ?

R : Non, ça ne va pas avoir d'impact.

I : Cela ne va pas avoir d'effet de preuve pour vous ?

R : Non.

I : Et si une même information est répétée plusieurs fois sur différents canaux, cela renforce-t-il votre confiance en sa véracité ?

R : Bah, je ne fais aucun réseau social à part Facebook, donc je ne suis pas sur Instagram par exemple, j'ai pas trop d'autres réseaux, donc je ne sais pas comparer.

I : Et donc même sur Facebook, si, par exemple, vous voyez tout le monde parler de la même chose, cela va-t-il renforcer votre confiance en sa véracité ?

R : Ouais, ouais. L'effet de groupe va faire que je vais me dire qu'il n'y a pas autant de gens qui peuvent dire n'importe quoi. Mais je n'irai pas taper sur un autre réseau, pour voir ce qu'il se dit. J'ai Facebook parce qu'on me l'a mit. Je ne m'amuserai pas à passer ma journée à comparer tout ce qu'il se dit sur Facebook, Instagram, etc. Je m'en fous un peu.

I : Ok, oui.

R : C'est plus quand je suis au boulot, en pause ou quoi, là je vais passer du temps sur Facebook mais autrement je m'en fous un peu des réseaux sociaux. Je les utilise juste pour passer le temps.

I : Oui, plus pour du divertissement alors ?

R : Oui, c'est ça !

I : Ok parfait, c'est la fin des questions, avez-vous quelque chose à ajouter au sujet des deepfakes ?

R : Non.

I : Donc à présent, je vais vous montrer trois photos. Vous aurez 15 secondes pour voir chaque image et me dire si vous pensez qu'elle est vraie ou fausse. Après cela, vous aurez la possibilité d'observer plus en détails, les trois images, et de décider si vous gardez vos réponses ou si vous souhaitez les changer ?

R : D'accord, j'ai compris.

I : Les trois photos que vous allez voir peuvent être toutes fausses, toutes vraies ou une combinaison vraies-fausse.

R : Ok.

I : Voici la première photo.



R : (Observe l'image pendant 15 secondes)

I : Vous pensez qu'il s'agit d'une vraie ou fausse image ?

R : Je dirais, au vu de la tête de Macron, ... Je dirais que c'est faux.

I : OK. Voici la seconde image.



R : (Observe l'image pendant 15 secondes) C'est plus vrai, ça. Sauf qu'ils ne se permettraient pas ça à un rendez-vous.

I : Et donc pour vous, c'est vrai ou faux ?

R : Physiquement, c'est plus vrai, mais est-ce que, réellement, ça se passe comme ça dans les rendez-vous importants ? Je ne pense pas non plus.

I : Donc, selon vous, est-ce vrai ou faux ?

R : Ouais. Il a... Faux par la ... situation.

I : Le contexte vous semble bizarre ?

R : Ouais, par son comportement. Par Macron qui rigole, voilà.

I : Voici la troisième image.



R : (Observe l'image pendant moins de 15 secondes)

Encore, la tête de Macron.

I : Et donc pour vous, est-ce une vraie ou fausse image ?

R : Faux. C'est aussi Macron qui me paraît bizarre, mais je ne saurai pas expliquer pourquoi.

I : Ok, je comprends. A présent, vous pouvez regarder les trois images, plus en détails.

R : Bon, je dirais peut-être que la troisième est vraie, mais non, c'est trop faux.

I : La première vraie alors ?

R : Non. Trois faux.

I : Alors la première, elle est fausse, la deuxième est vraie, et la troisième fausse.

R : En fait, je ne sais pas comment on fait pour voir une fausse photo.

I : Oui, je comprends, si on n'en voit pas beaucoup ou si on est juste pas conscient d'en voir, ça peut vraiment être difficile de les reconnaître.

R : Mais oui, moi, c'est la première fois que j'en vois, en fait. Donc je regardais plus rapidement les personnages que le décor et je n'ai même pas pensé à regarder les détails des mains, du visage ou quoi.

I : Maintenant au moins vous savez que cela existe et que vous pouvez en retrouver sur les réseaux sociaux.

R : Oui, c'est vrai.

I : Et bien merci beaucoup. Avez-vous quelque chose à ajouter ?

R : Non, c'est tout !

Répondant	Âge	Genre
Numéro 9	57 ans	Homme

I : Bonjour et merci d'avoir accepté de participer à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux sociaux. Pour information, les deepfakes sont des contenus photos ou vidéos truqués créés par les intelligences artificielles. Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage d'une personne ou encore les éléments d'un décor à partir d'une simple photo ou description. Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes.

Bien entendu, vous êtes libre à tout moment de ne pas répondre à certaines questions ou de mettre fin à l'entretien si vous en ressentez le besoin. Avez-vous des questions avant de commencer ?

R : Non.

I : Parfait, nous pouvons alors commencer. Aviez-vous déjà entendu parler des deepfakes auparavant ?

R : Oui.

I : Et que savez-vous à leur sujet ?

R : Mais personnellement, je ne suis pas un pro du truc. Enfin, j'en ai fait mon opinion, quoi. Par rapport à ce que je vois et par rapport à ce que j'ai entendu ailleurs, dans les médias, je sais que l'on peut truquer des images ou des vidéos. C'est difficile de dire ce qui est vraiment vrai ou faux maintenant avec l'intelligence artificielle. Mais ma conscience et ma connaissance, ma petite connaissance du monde, font que je peux dire à 70 % que c'est vrai ou pas.

I : Et si vous deviez définir les deepfakes, vous diriez que c'est quoi ?

R : C'est honteux parce que ça jette le discrédit sur tout.

I : Le quoi ? Le discrédit ?

R : Oui, tu ne sais plus dire ce qui est vrai. Tu ne sais plus distinguer le vrai du faux. En fait, ça fait partie de la ... Comment dire ... On essaie de faire croire aux gens des choses pour jeter le discrédit sur une personne célèbre ou une personne politique. Et le contraire est pareil aussi. On essaie de valoriser quelqu'un au moment des élections, par exemple. On voit des trucs, on essaie d'en faire son avis, mais c'est très difficile.

I : Et comment en avez-vous entendu parler ?

R : Sur Facebook et tout ça. Oui, sur Facebook et alors dans les médias. On en a parlé.

I : Avez-vous déjà été confronté à des avertissements ou des campagnes de sensibilisation sur les deepfakes.

R : Oui, à la télé surtout, au JT ou quoi, ils donnaient des exemples. Par exemple avec Brad Pitt.

I : Et pensez-vous que cela occupe une place importante sur les réseaux sociaux à l'heure actuelle ?

R : Oui, énormément.

I : Si vous deviez donner un pourcentage, ce serait lequel ?

R : Il y en a 50 %. C'est 50-50. Il peut y avoir un peu de tout sur les réseaux sociaux, c'est ça le problème.

I : Et pensez-vous qu'en tant qu'utilisateur des réseaux sociaux, vous avez un rôle à jouer avec les deepfakes ?

R : Un rôle à jouer ?

I : Quand je dis un rôle à jouer, c'est par exemple pour les empêcher de se répandre ?

R : Oui, je crois que j'ai un rôle à jouer. C'est de un, ne pas les partager. Ou alors, faire un commentaire disant que c'est fake. Vérifier mes sources.

I : Et quelle a été votre première réaction, votre première pensée lorsque vous avez pris connaissance de l'existence des deepfakes ?

R : Ma première pensée ? C'est qu'on peut faire croire n'importe quoi aux gens. Si tu n'es pas un peu vigilant, tu crois tout ce qu'on va te raconter.

I : Et avez-vous déjà eu le sentiment, en étant sur Facebook, d'être face à une vidéo ou une photo qui était fausse, sans en avoir la certitude ?

R : Oui, ça, ça m'est arrivé souvent.

I : Et du coup, dans quel contexte, par exemple ?

R : Des commentaires d'hommes politiques. Ou alors des lanceurs d'alerte qui font des montages en montrant qu'un tel ministre a dit ça, et que le lendemain, il dit le contraire. Et tu n'es pas certain si c'est bien lui qui l'a dit, parce qu'ils font des montages. Ces gens-là font des montages. Donc, ils prennent des paroles d'un discours qu'un ministre a fait, et ils les remettent dans un autre contexte. Par exemple, un ministre dit « blanc » un jour, et le lendemain, tu l'entends dire « noir ». Mais ce n'est pas la même interview. Ils ont fait des montages. Ils ont changé ses paroles. Encore une fois, c'est comme je dis, quand tu as une certaine connaissance du monde politique, que tu suis un peu le monde politique, tu entends de tout, et tu dois te faire ton idée par rapport à tout ce qu'on entend. Il ne faut pas prendre en compte seulement un élément.

I : Et avez-vous déjà été influencé par une information qui s'est avérée fausse par la suite ? Une information photo ou une vidéo par exemple, un politicien.

R : Oui, ça m'est arrivé. J'ai déjà entendu qu'on disait « celui-là, il a fraudé » ou autre.

I : Et vous y avez cru ?

R : Et j'ai dit « regarde, c'est dégueulasse, ils sont au pouvoir, ils ont fraudé ». Et après, quelques mois après, dans les semaines qui suivent, tu entends que c'est une fake news ou qu'il n'y a rien qui est fondé. C'est encore une personne qui a lancé ça, en faisant un montage. Donc, en réalité, à l'heure actuelle, on n'est plus sûr des informations. Avant, on n'avait pas assez d'informations. Maintenant, on en a trop. Et tu dois faire un fameux tri là-dedans parce qu'il y en a qui racontent n'importe quoi. Tout et n'importe quoi.

I : Et pour vous, qu'est-ce qui rend un deepfake difficile à détecter à première vue ?

R : Quand c'est bien fait. La qualité, quoi. Quand tu as un genre de reportage avec des gens qui parlent en réel et des gens qui en font des montages, comme j'ai dit tantôt. Quand on simule la voix, etc. Et alors, c'est quand même bien fait. Il y a des trucs qui sont très bien faits. Pour une personne qui n'est pas avertie, tu vois ça, pour toi, c'est vrai, automatiquement.

I : Et donc, si vous devez distinguer le vrai du faux, comment faites-vous ?

R : Je vais regarder les traits du visage. Je vais regarder aussi s'il n'y a pas de coupures dans les paroles. Dans la vidéo, oui. Ah oui, parce que c'est là que ... Quand quelqu'un parle et qu'après, le mot qu'il dit après un certain moment ne correspond pas, tu vois que c'est des montages. Ça, encore une fois, c'est mon appréciation. Et alors, j'attends ... En général, on va dire ... Tu seras mieux informé à la TV que sur Facebook, ça c'est sûr. C'est un petit peu plus vrai à la TV, on va dire.

I : Un petit peu plus ? Pourquoi un petit peu plus et pas totalement plus vrai ?

R : Parce que ça dépend du journaliste. Il y a des journalistes qui sont anti-Trump, il y a des journalistes qui sont anti-Macron, et ces gens-là ne vont pas dire du bien de Macron s'ils ne l'aiment pas. Ils choisissent ce qu'ils veulent dire. Pour moi, une personne, un vrai journaliste, il est neutre. Il raconte des faits. Il n'invente pas ce qu'il a dans les journaux. Et il dit les choses quand c'est prouvé ... Le problème, c'est que la justice est tellement lente que c'est prouvé quatre ans après. Et quatre ans après, ça passe sous la trappe. Ça, c'est mon point de vue.

I : Et sur une échelle de 1 à 10, à quel point vous sentez-vous capable de reconnaître un deepfake ?

R : 50 %.

I : Donc 5 sur 10. Et pourquoi ce nombre ?

R : Encore une fois, comme je dis, je ne me laisse pas directement influencer par le truc. J'essaie de voir ... J'essaie de voir des commentaires. Ou alors j'ai vu l'information avant. Et ce que le mec a mis ... J'essaie de regarder plus en détail. Je crois qu'à 50 %, je peux dire si c'est un deepfake ou pas.

I : Et depuis que vous avez pris connaissance des deepfakes, est-ce que cela a modifié votre manière d'utiliser les réseaux sociaux ?

R : Ça ne m'a pas ... Ça n'a pas changé ma manière de regarder les réseaux sociaux. Par contre, ça a quand même changé ma manière de réfléchir. Au moment où tu sais qu'il y a...

Tu te questionnes par rapport à ce que tu fais. Au moment où tu sais que 50 % sont faux, tu ne te fais pas la même idée. Je suis plus vigilant, quoi. Regarde, allez, un exemple. On dit, ouais, Marc Dutroux va sortir, c'est honteux. Il y en a qui mettent ça. Parce que c'est vrai, il pourrait, il aurait le droit de sortir, mais il ne sortira jamais. Ça, c'est juste pour lancer des trucs, tu vois. C'est pour lancer des polémiques, etc.

I : Et à présent, est-ce que vous doutez de toutes les informations que vous rencontrez sur les réseaux sociaux en vous disant que ça peut être faux ?

R : Quasiment.

I : Et pourquoi, du coup ?

R : Parce que tu ne sais jamais plus dire si c'est vrai ou si c'est faux. Même en pensant que, oui, je suis sûr, car j'ai vu l'info ailleurs. Mais l'autre information que j'ai vue, est-ce qu'elle est vraie ? Puisqu'il y a tellement d'informations.

I : Vous voulez dire que même en vérifiant plusieurs sources, vous ne savez plus dire si l'information est vraie ?

R : Pour le moment, tu vois, en France, ils sont en train de dire qu'ils ont un déficit financier faramineux et qu'ils vont prendre dans l'épargne des Français. Donc on va prendre 1000€ à tous les Français, comme ça. Et on en parle partout. Il n'y a pas un jour où je mets Facebook ou Google, et qu'il n'est pas marqué que le gouvernement va prendre dans vos... dans vos économies. Et on n'arrête pas de parler de ça. Moi, dans ma pensée, c'est de un, interdit. Je ne vois pas pourquoi on pourrait venir te prendre 1 000 € comme ça. Sauf si c'est une taxe. Et de deux, on n'arrête pas d'en parler sur différents réseaux sociaux. Est-ce vrai ou pas ? Je ne sais pas. Mais selon mes connaissances à moi, je me dis que ça ne peut pas être vrai. Pour moi, ça ne peut pas être vrai.

I : Avez-vous pris des mesures spécifiques pour vous protéger contre les fausses informations ?

R : Je suis simplement plus vigilant. Mais non, rien d'autre.

I : Et lorsque vous voyez une nouvelle vidéo ou une image en ligne, est-ce que vous-prenez systématiquement le temps d'en vérifier la source ?

R : Non.

I : Et si vous deviez le faire, comment procéderiez-vous ?

R : Je regarderai plus de témoignages. Je regarderai plus d'émissions là-dessus. Enfin, j'essaierai de voir un petit peu ... En fait, la seule manière où je peux trouver si c'est plus ou

moins vrai, c'est d'abord mon opinion, par rapport à ma connaissance, et par rapport à d'autres commentaires. Et alors aussi, la source. Par exemple, quand tu as Mediapart qui parle de choses, en général, c'est vrai. Je regarde beaucoup Cash Investigation. Ce sont des journalistes indépendants qui font des reportages sur des firmes qui polluent, etc. Et en général, quand ça se passe à la télé, il y a de grands soupçons que ça soit vrai.

I : Ok. Donc, vous êtes plus vigilant dans ce que vous regardez sur les réseaux sociaux, mais est-ce que vous êtes plus vigilant aussi dans ce que vous partagez ?

R : Oui, je vais regarder. Je ne vais pas partager une connerie. Je ne vais pas partager le fait qu'un mec dise « Ouais, celui-là, il a fait ci, il a fait ça ». Non, je ne vais pas faire ça.

I : Ok.

R : Les trucs que je partage, ce sont des trucs rigolos ou alors des trucs plutôt émouvants. Des trucs qui, je pense, ne peuvent pas avoir un impact directement sur les gens. Ou alors des trucs prouvés. J'ai déjà publié des trucs de Mediapart, mais ce sont des choses prouvées. Ils ont fait des recherches. Ça, oui. Mais je ne vais pas mettre « Trump, c'est un voleur ». Sans aucune preuve. Ça, ce n'est pas mon genre.

I : Et considérez-vous les réseaux sociaux, à l'heure actuelle, comme une source fiable ?

R : Non.

I : Pourquoi ?

R : Pour toutes ces raisons-là. C'est que tout le monde met tout et n'importe quoi. Tu ne sais pas dire ce qui est vrai. La plupart du temps, tu vois des trucs ... Ce sont tous des montages. Il y a des gens pour se faire voir, pour avoir des vues et tout, ils font des... des faux trucs, inventent des fausses rumeurs, des fausses attaques. Ils mettent en scène. Ils mettent en scène. Il y a beaucoup de gens qui mettent leur vie en scène ou je ne sais pas quoi. Je ne comprends pas trop l'intérêt, mais ça marche pour faire parler d'eux.

I : Oui, je comprends. Et par contre, les médias traditionnels qui sont maintenant présents en ligne, les considérez-vous comme une source fiable ?

R : Oui.

I : Pourquoi ?

R : Parce que le journaliste est neutre. Oui. Même si certains, les BFMTV, je me méfie parce que ... Tu ne sais pas. Leur patron, c'est des multinationales. Ils ont peut-être d'autres intérêts derrière.

I : OK. Vous faites donc plus confiance à des journalistes indépendants qui font des investigations qu'à autre chose.

R : Oui. Tu as des chaînes ... la plupart, 50 % des journaux sont tenus par des milliardaires. Ils vont peut-être moins parler de social, eux. Ils vont parler de fric.

I : Vous voulez dire que vous n'aurez pas toutes les informations que vous voulez ?

R : Non, en fonction de leurs opinions et de leurs intérêts, quoi.

I : Et selon vos réponses, à quelles sources d'information accordez-vous le plus de confiance pour avoir une information vraie ?

R : Honnêtement, la plupart du temps, comme je t'ai dit, tu le sais quelques années après, parce que le gars a été condamné avec des preuves. Avant ça, ce n'était que du blabla.

C'est-à-dire qu'un dit comme ça, l'autre dit comme ça, etc. Soit ça passe au journal, ou alors, comme je dis, j'ai eu beaucoup de confirmations par Cash Investigation, les documentaires, et Mediapart. Ce sont des trucs indépendants.

I : Et est-ce que vous avez plus tendance à croire plus facilement une information si elle est accompagnée d'une image ou d'une vidéo plutôt qu'une information au format texte uniquement ?

R : Oui. Parce que les images font beaucoup.

I : Que voulez-vous dire par "font beaucoup" ?

R : Parce que déjà, il y a beaucoup de gens qui ne lisent pas quand il n'y a pas d'image. C'est vrai. Alors que si tu mets une image choc, ça marche. Je ne sais pas, on va dire que Macron prend de la drogue. Si on a pas de preuve on ne saurait pas s'imaginer si c'est vrai ou faux. Alors que si on le voit en vidéo ou en photo prendre de la drogue, on va se dire plus facilement, ok, oui. En tout cas, cela aura plus d'effet. Et surtout quand ça parle de personnes connues.

I : Et si une information est répétée plusieurs fois sur différents canaux, par différentes personnes, est-ce que cela renforce votre confiance en sa véracité ?

R : Oui, automatiquement.

I : Pourquoi ? Pour l'effet de masse ?

R : L'effet de masse, oui. Et qu'il y ait plusieurs ... On va dire ... des journalistes indépendants, différentes chaînes qui disent la même chose. En général, pour moi, un journaliste, quand il dit quelque chose, il a des preuves. C'est prouvé, tu ne lances pas un truc comme ça. Après, ça dépend des personnes qui publient, ou alors il y a un fameux soupçon.

I : Merci pour vos réponses. On arrive à la fin des questions, avez-vous des remarques ou quelque chose à ajouter ?

R : Non.

I : On peut donc à présent passer à un petit test. Je vais vous montrer 3 photos. Les trois peuvent être toutes vraies, toutes fausses, ou vraies et fausses. Vous aurez 15 secondes pour les regarder, une à une. Vous me direz ensuite si vous pensez qu'elles sont vraies ou fausses, et pourquoi. Ensuite, vous aurez l'occasion de les revoir plus longuement et plus en détails afin de confirmer vos choix ou de les ajuster. Voici la première image.



R : (Observe l'image pendant 15 secondes)

I : Pensez-vous qu'elle est vraie ou fausse ? Et pourquoi ?

R : Pour moi, elle est fausse. Parce que Macron n'a pas cette coiffure-là, à ce moment-là. Quand il est allé voir Donald Trump, il n'avait pas cette coiffure-là, en tout cas. Ça, c'est ma première impression. Il n'avait pas la tête qu'il a maintenant. Mais je ne vois pas bien, je n'ai pas mes lunettes.

I : Oui, vous pouvez aller les chercher pour que ça soit mieux pour vous.

R : Voilà. Mais ici, c'est uniquement une photo, il n'y a pas d'information, de contexte. Maintenant, tout le monde sait que Macron et Trump se sont vus dernièrement. Est-ce que c'était cette photo, cette année ? Ça, on ne sait pas.

I : Ok, je comprends. Vous vous fiez plus au contenu qu'à la photo.

R : Oui, je regarde les deux en fait. Sans le contexte, parfois, on peut croire que c'est une vraie photo. Les deux jouent, quoi.

I : Voici la seconde image.



R : (Observe l'image pendant moins de 15 secondes) Celle-là, elle est vraie. Parce que ces images-là, on les a vues en direct à la télé. Il est allé voir le président. Macron est allé voir Trump. Et j'ai vu cette image-là.

I : Donc vous avez vu la même image à la télé ?

R : Oui, et si à la télé, on ne met pas les vraies images, alors là, ça ne tourne plus rond.

I : Voici alors la dernière image.



R : (Observe l'image pendant moins de 15 secondes) Celle-là, c'est un fake total. Macron, c'est pas lui. On ne dirait même plus lui.

I : Et donc, à présent, en regardant les trois images, souhaitez-vous changer ou garder vos réponses.

R : Celle-là, elle est vraie, je suis certain. Et celle-là, elle n'est pas vraie. Regarde, il a la même tête que dans la deuxième. Il est rasé.

I : Donc vous confirmez vos choix ?

R : Je confirme mes choix.

I : Et donc, pour faire un petit récap, sur quels éléments vous êtes-vous basé pour faire tous vos choix ?

R : Surtout l'apparence actuelle des présidents et si j'avais déjà vu l'image.

I : Vous n'avez pas pris d'autres éléments en compte ?

R : Non, parce qu'il n'y a pas de contenu, tu ne sais pas dire à quelle date. Maintenant, celle-là pourrait être vraie, parce qu'au tout début, Macron, il avait cette tête-là. Il était coupé très court. Maintenant, tu vois, il a une autre coiffure. Oui. On le voit quasiment tous les jours sur Facebook, Macron. C'est peut-être le vrai Macron, une vraie photo mais une ancienne, et un montage.

I : Et bien, vous avez tout juste. La première et la troisième sont bien des deepfakes que j'ai créés à l'aide d'une intelligence artificielle.

R : Je suis un pro de l'informatique en fait, aha. Un vrai pro. Je suis pas si mal.

I : Non, bien joué. Et merci. Merci beaucoup.

Répondant	Âge	Genre
Numéro 10	36 ans	Femme

I : Bonjour et merci d'avoir accepté de participer à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux sociaux. Pour information, les deepfakes sont des contenus photos ou vidéos truqués créés par les intelligences artificielles. Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage d'une personne ou encore les éléments d'un décor à partir d'une simple photo ou description. Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes.

Bien entendu, vous êtes libre à tout moment de ne pas répondre à certaines questions ou de mettre fin à l'entretien si vous en ressentez le besoin. Avez-vous des questions avant de commencer ?

R : Non, pas de questions.

I : Donc la première question est : Aviez-vous déjà entendu parler des deepfakes auparavant ?

R : Oui.

I : Que savez-vous à leur sujet ?

R : Rien de plus que ce que tu viens de dire. Parce que, du coup, j'en ai entendu parler dans Ici tout commence.

I : Donc, la première fois que vous en avez entendu parler c'est dans la série Ici tout commence ?

R : Non, sur les réseaux sociaux quoi, mais ils en parlent aussi dans Ici tout commence. Il y a un truc où un papa a disparu et un frère, pour ne pas faire stresser sa sœur, il va faire une fausse vidéo en animant une photo de son papa et faire genre qu'il parle.

I : Donc animer les personnes, etc. c'est l'exemple que vous en avez ?

R : Oui, c'est ça.

I : Et avez-vous déjà été confronté à des avertissements ou des campagnes de sensibilisation sur les deepfakes ?

R : Oui.

I : Comment c'était ? Dans quel cadre ?

R : Ce sont des gens qui font des vidéos sur les réseaux sociaux pour prévenir, dire de faire attention. Ils mettent en garde et disent que c'est peut-être faux, que ce ne sont pas les vraies personnes ou pas la vidéo originale.

I : Donc ce sont des avertissements qui sont faits par des influenceurs ou des personnes lambdas ?

R : Oui, les deux, mais souvent les gens lambdas qui donnent leur avis sur les vidéos qui circulent.

I : Et pensez-vous que les deepfakes occupent une place importante sur les réseaux sociaux ?

R : Oui.

I : Si vous deviez me donner un pourcentage ?

R : Que moi je croise ?

I : Oui, par exemple, consciemment ou pas.

R : Je ne sais pas, je vais peut-être dire 35 % alors, pour moi.

I : Dont vous êtes consciente ou non ?

R : Oui, dont je suis consciente.

I : Et pensez-vous qu'en tant qu'utilisateur des réseaux sociaux, vous avez un rôle à jouer dans la propagation des deepfakes ?

R : Non, pas spécialement.

I : Vous pensez que vous ne pouvez rien y faire ?

R : À part signaler la publication ou quoi, c'est tout. Je ne suis pas influenceuse à faire des vidéos là-dessus, ou quoi que ce soit.

I : Je comprends. Et quelle a été votre première réaction ou votre première pensée lorsque vous avez pris connaissance de l'existence des deepfakes ?

R : Aucune idée en vrai.

I : La première phrase qui vous est venue à l'esprit, cela peut être positif ou négatif, peu importe.

R : Je ne saurais pas dire. Dans un sens, c'est bien. C'est marrant dans le sens où tu prends une photo et que tu peux l'animer. Je trouve ça marrant. Et en même temps, c'est un peu comme avec une Brad Pitt, ça arnaque les gens.

I : Donc vous vous dites que c'est marrant, mais en même temps, que cela peut être dangereux pour certaines personnes.

R : Oui, voilà, tout dépend de comment les deepfakes sont utilisés.

I : Et avez-vous déjà eu le sentiment, en étant face à une image ou une vidéo, qu'il s'agissait d'un deepfake ?

R : Non ... Deepfakes , non, je vois plus des vidéos faites par l'IA.

I : Comment ça, des vidéos faites par l'IA ?

R : Genre, ce sont des vidéos, je ne sais pas, de lions, mais tu vois que ce n'est pas un vrai lion, quoi.

I : Ah oui d'accord. Et bien, typiquement, ce genre de vidéos, ce sont des deepfakes. Ce sont des vidéos créées de toutes pièces, ou modifiées à l'aide de l'IA.

R : Ah bah voilà, alors oui haha. Au début, je regarde la vidéo, puis je me pose la question si c'est vrai ou non. Et souvent, ça ne l'est pas car le lion ressemble plus à une peluche qu'à un vrai lion.

I : Et avez-vous déjà été influencée par une information qui, par la suite, s'est avérée fausse ?

R : Ouais, non, je ne pense pas.

I : Vous n'avez jamais cru quelque chose sur une personnalité publique suite à une vidéo ou à une photo, qui par la suite s'est avérée fausse ?

R : Non, je ne pense pas, ou alors je ne me souviens plus.

I : Et selon vous, qu'est-ce qui rend un deepfake difficile à détecter à première vue ?

R : Son aspect réaliste.

I : Comment décririez-vous ce réalisme ?

R : Mais dans le sens où on ne voit pas tout de suite que c'est faux, on ne se dit pas, oh là, il a un sixième doigt, c'est bizarre. Tu vois, tu peux le voir, mais c'est super discret, quoi.

I : Ok, donc ce qui le rend difficile à détecter est uniquement son réalisme, ou vous souhaitez dire quelque chose d'autre ?

R : Tu peux poser la question ?

I : Donc, selon vous, qu'est-ce qui rend un deepfake difficile à détecter à première vue ?

R : Difficile à détecter ? Oh non, je dirais ça.

I : Ok. Et donc, comment faites-vous pour distinguer un deepfake d'une vraie photo ou d'une vraie vidéo ?

R : Je regarde les doigts. Les doigts, ou alors, tu vois, par exemple, dans le cadre du lion, les contours aussi, parfois.

I : Qu'est-ce que vous voulez dire par les contours ?

R : Tu vois, un peu comme quand on modifie parfois les photos, quoi. Parfois, il peut manquer un petit bout de la personne qu'on a modifiée, ou alors le contour est trop net.

I : Oui, je vois ce que vous voulez dire.

R : Ou alors, maintenant, parfois, les vidéos, tu vois des ... Comment on appelle ça ? Les trous. Quand on fait des montages, tu vois, ... les transitions ! Les transitions sont parfois méga mal faites, on voit qu'il y a un bug à un moment dans la vidéo.

I : Ok, oui. Et sur une échelle de 1 à 10, à quel point vous sentez-vous capable de reconnaître un deepfake ?

R : Bon, je vais dire 5.

I : Et pourquoi 5 ?

R : Je ne pense pas être experte non plus.

I : Pas experte, mais pourquoi 5, du coup ?

R : Bah, entre les deux, ça dépend des moments aussi. Mon niveau d'attention, mon niveau de fatigue. Ouais, parce que parfois, je ne cherche pas à comprendre, quoi.

I : Oui, donc votre cerveau peut assimiler des informations sans même que vous y prêtiez attention.

R : Quand je suis fatiguée, complètement.

I : Et depuis que vous avez connaissance des deepfakes, est-ce que cela a modifié votre manière d'utiliser les réseaux sociaux ?

R : Non. Et je me rends compte que c'est un peu contradictoire. Je dis que je fais plus attention, alors qu'au final je fais tout comme avant.

I : Et à présent, est-ce que vous doutez de toutes les informations que vous rencontrez sur les réseaux sociaux ?

R : Ça, je le faisais déjà avant, en soi. On m'a toujours appris à avoir un esprit critique par rapport à ça. Donc, à chaque fois, même au ... Oh mon Dieu, la journée à l'hôpital, ça ne fait pas du bien. Les articles des journalistes, là, tu vois ?

I : Oui.

R : De base, ils sont censés être neutres. Et bien, aujourd'hui, faut essayer de citer un journaliste neutre. Donc, faut toujours un peu être l'avocat du diable. Vérifier tout ce qu'on voit.

I : Donc, vous doutez de tout sur les réseaux sociaux, pas uniquement quand il y a quelque chose qui vous paraît bizarre ?

R : Oui, vraiment de tout.

I : Et avez-vous pris des mesures spécifiques pour vous protéger contre les fausses informations ?

R : Je regarde les sources, quoi.

I : Et comment faites-vous pour vérifier vos sources ?

R : Je vais regarder qui poste. Je vais aller voir son contenu, si c'est un truc qui sort totalement de son domaine, entre guillemets, je vais dire, ou pas. Si un mec, il est expert ou pas dans tel domaine. Enfin, dans les petits trucs qu'on peut écrire au-dessus, là, dans la bio.

I : Et prenez-vous systématiquement le temps de vérifier vos sources lorsque vous voyez une vidéo ou une photo en ligne ?

R : Non, je regarde plus pour tout ce qui est fait d'actualité qu'autre chose. Le reste, je ne vérifie pas spécialement. Parce que, au final, j'y prête pas plus attention que ça, quoi. Je ne vérifie que quand l'info m'intéresse.

I : Ok, oui, je comprends. Et donc, aujourd'hui, vous êtes plus vigilante avec ce que vous regardez sur les réseaux sociaux, mais est-ce également le cas avec ce que vous partagez sur les réseaux sociaux ?

R : Oui. Mais je ne me dis pas ... Enfin, c'est pas le fait que ça soit faux ou vrai qui va me dire que je partage ou pas, tu vois.

I : C'est quoi, du coup ?

R : C'est plus le fait que les personnes vont peut-être pas forcément être d'accord, du coup je ne partage pas. Je garde pour moi. Ou alors, je partage juste avec les personnes qui vont être d'accord avec moi, tu vois. Je n'ai pas envie non plus d'entrer dans des débats.

I : Oui, ok, intéressant. Et donc, vous ne partagez rien à cause de ça ?

R : Ouais, c'est rare en tout cas.

I : Et considérez-vous les réseaux sociaux comme une source fiable et pourquoi ?

R : Oh ... Ça, c'est un sujet de dissertation. Euh ... Ouais, mais j'aurais envie de dire les deux.

I : C'est votre avis, cela peut être fiable et non fiable à la fois. Pourquoi cela ?

R : Du coup, je dirais fiable, dans le sens où tu peux avoir des comptes certifiés, etc. Et donc, vérifier les sources aussi. Il y en a qui mettent directement la source parfois, donc ça, c'est fiable. Et en même temps, pas fiable, parce que tu peux balancer tout et n'importe quoi, quoi.

Il n'y a pas spécialement de vérifications qui sont faites, donc ...

I : Quand vous dites qu'il n'y a pas de vérifications qui sont faites, vous voulez dire par les réseaux sociaux ?

R : Oui, les réseaux sociaux ne font rien pour éviter que les fausses informations ne circulent..

I : Ok, oui. Et donc, par exemple, les médias traditionnels qui sont maintenant présents en ligne, est-ce que vous les considérez comme une source fiable ?

R : Je vais dire oui et non. Parce que, du coup, si je prends RTL, chaque journaliste va avoir une tendance à expliquer un fait d'une certaine manière, qui peut t'orienter dans une autre direction.

I : Vous voulez dire que RTL peut traiter l'information d'une certaine manière, et TF1, par exemple, d'une autre manière, ce qui va modifier le sens de l'information ?

R : Oui, c'est ça. Oui, en fonction de leur intérêt ou de leur affinité avec le sujet, etc., quoi.

I : Oui, ok, oui. Et selon vos réponses, quelles sont alors les sources d'informations vers lesquelles vous vous tournez pour obtenir une information vraie ? Quelles sont les sources d'information auxquelles vous vous fiez pour obtenir une vraie information ?

R : Vers quelles sources d'informations je me dirige ?

I : Oui, c'est ça. En ligne ou non, c'est comme vous le sentez.

R : Moi, par exemple, sur Instagram, je suis le compte Dormir au courant. Il met toutes les petites infos d'actualité et, en général, il met la source à chaque fois.

I : Et donc, c'est une personne lambda ou c'est un journaliste ?

R : Je ne sais même pas. Ou alors c'est juste une personne lambda qui relaie ce qu'elle voit dans les journaux, etc., peut-être, je ne sais pas. Je ne saurais même pas dire, parce que du coup, il n'y a pas de prénom, tu vois. C'est sous le nom "Dormir au courant". On peut contacter la personne derrière le compte par mail, pour avoir plus d'infos sur un sujet, donc je suppose que c'est quelqu'un de qualifié, ou de très informé.

I : À part ce compte, est-ce que vous consultez d'autres choses, d'autres sources ?

R : Je regarde directement les comptes politiques.

I : Les comptes des personnes politiques ou des ...

R : Des partis. Pour avoir de l'information d'un peu partout, peut-être avec différents points de vue, etc.

I : Ok, oui. Et avez-vous plus tendance à croire une information si elle est accompagnée d'une image ou d'une vidéo plutôt que d'un simple texte ?

R : Oui, c'est sûr.

I : Et pourquoi ?

R : Parce que, du coup, ça fait un peu effet de preuve plutôt que juste des faits dits. Il faut que cela soit en rapport avec ce que la personne a dit quand même, mais oui.

I : Et si une même information est répétée plusieurs fois sur différents canaux, par différentes personnes, est-ce que cela va renforcer votre confiance en sa véracité ?

R : Je pense que oui. Inconsciemment, je pense que oui.

I : Pourquoi ?

R : Pour le nombre de personnes qui en parlent. Même si je ne m'y intéresse pas forcément, à force d'entendre la même chose, je pense que je finis déjà par capter l'info et du coup, me dire que oui, c'est vrai.

I : Ok, merci beaucoup. Je suis arrivée à la fin des questions. Avez-vous quelque chose à ajouter sur le sujet, une remarque ou autre ?

R : Non, pas spécialement.

I : Pas de soucis. On peut à présent passer à un petit test. Je vais vous montrer trois photos. Vous aurez 15 secondes pour observer, une à une, les photos. À la fin des 15 secondes, vous devrez me dire si vous pensez qu'il s'agit d'un deepfake ou d'une vraie photo, et pourquoi. Les trois photos peuvent être toutes vraies, toutes fausses, ou une combinaison de vraies et de fausses. Une fois les trois photos vues séparément, vous aurez l'occasion de les voir côte à côte pendant le temps nécessaire, et de me dire si vous confirmez ou ajustez vos réponses.

Voici la première image :



R : (Observe l'image pendant moins de 15 secondes) Elle est fausse. Parce que la tête d'Emmanuel Macron, elle est super bizarre. En fait, ils se serrent la main, mais ils ne se regardent pas. Genre, Emmanuel Macron, t'as l'impression qu'il regarde dans le vide. Et Donald Trump, t'as l'impression qu'il regarde quelque chose, d'autre, quoi. Donc, ce n'est pas logique.

I : Ok, merci. Voici la seconde image :



R : (Observe l'image pendant 15 secondes) Vraie ? Les mains sont quand même bizarres. Donc je dirai fausse. La qualité est bien, mais les mains, comment elles sont mises, ça m'a l'air bizarre, elles sont tordues comme ça.

I : Ok, et la troisième image :



R : (Observe l'image pendant 15 secondes) Celle-ci est de bonne qualité, donc je dirai que ... Ah non c'est faux. La main de Macron est bizarre.

I : Ok, à présent vous pouvez regarder les trois en même temps, autant de temps que vous le souhaitez, pour me confirmer ou ajuster vos réponses.

R : Je vais confirmer mes réponses. Je trouve que la deuxième est quand même plus réaliste, mais voilà. Je ne sais pas.

I : Ok, donc vous avez un doute pour la deuxième pour les mains. Ce n'est pas le contexte de la photo qui vous met en difficulté ?

R : Non c'est pour les mains. La situation ... je sais qu'ils sont capables de faire ça. Je regarde souvent les mains.

I : Et pourquoi regardez-vous souvent les mains ?

R : Parce que je trouve que c'est le plus dur à reproduire, dans les différentes positions, etc. Tu vois, par exemple, la troisième, c'est pas bien découpé entre les doigts. En plus, les doigts, c'est quelque chose de droit mais avec des détails, des bosses pour les phalanges, etc.

I : Ok, oui, super intéressant. Donc vous regardez souvent les mains et les doigts car vous pensez que c'est à cet endroit que l'IA peut faire des erreurs ?

R : Oui, c'est ça. C'est là que l'IA peut faire plus facilement des erreurs.

I : Ok, merci. Donc, pour les réponses : la première est fausse, la deuxième est vraie, et la troisième est fausse.

R : Oui, la deuxième, j'hésitais. Cela ne m'étonnait pas d'eux qu'ils rigolent, mais j'ai vraiment hésité sur les mains. Mais bon, du coup, c'est vrai. Et la troisième en la revoyant, même le nez de Macron on dirait qu'il est un peu découpé. Il a été mal découpé.

I : Et bien, c'est quand même un bon score. Je ne sais pas si vous souhaitez ajouter quelque chose ?

R : Non à part que c'est vraiment un super bon indicateur, les mains.

I : Merci beaucoup pour votre aide en tout cas.

Répondant	Âge	Genre
Numéro 11	20 ans	Homme

I : Bonjour et merci d'avoir accepté de participer à cet entretien. Je suis étudiante à l'UCLouvain et je mène une étude sur les réactions et comportements que peuvent susciter les deepfakes auprès des utilisateurs des réseaux sociaux. Pour information, les deepfakes sont des contenus photos ou vidéos truqués créés par les intelligences artificielles. Ces contenus paraissent cependant très réalistes. Les intelligences artificielles permettent par exemple d'animer les traits du visage d'une personne ou encore les éléments d'un décor à partir d'une simple photo ou description. Cet entretien durera environ 30 minutes et toutes vos réponses seront anonymes.

Bien entendu, vous êtes libre à tout moment de ne pas répondre à certaines questions ou de mettre fin à l'entretien si vous en ressentez le besoin. Avez-vous des questions avant de commencer ?

R : Non.

I : Avais-tu déjà entendu parler des deepfakes auparavant ?

R : Non, je n'en avais jamais entendu parler.

I : Tu connaissais le concept ou même pas ?

R : Oui, je connaissais le concept.

I : Et donc, comment as-tu entendu parler du concept ?

R : A mon stage, par exemple, on utilise ChatGPT, on paye pour qu'ils nous créent des images pour faire les plans de communication. Je ne sais pas si c'est vraiment la même chose que les deepfakes, parce que ...

I : Oui, c'est la même chose. Je suppose que tu écrivais un prompt, tu écrivais une description et tu voulais que ChatGPT génère ce que tu lui décrivais.

R : Ouais, ouais, ouais.

I : Donc c'était clairement des deepfakes. C'est juste que tu les utilisais à bon escient, mais du coup ...

R : Oui, ce n'était pas moi qui les faisais, mais j'avais entendu dire qu'ils faisaient ça.

I : Ok. Et avais-tu déjà été confronté à des avertissements ou des campagnes de sensibilisation sur les deepfakes ?

R : Ah, pas du tout. Non.

I : Et penses-tu que les deepfakes occupent une place importante sur les réseaux sociaux, à l'heure actuelle ?

R : Euh ... Important, je ne sais pas, mais je sais que... C'est vrai que j'en vois pas mal passer sur, par exemple, TikTok ou Instagram. Ce sont des personnes qui n'existent pas. Ça se voit. Mais je n'en vois pas tous les jours, mais c'est vrai qu'ils sont présents, quoi.

I : Ok oui. Et si tu devais donner un pourcentage, par exemple, de contenus que tu croises, on va dire par jour, de deepfakes.

R : Je dirais que j'en vois déjà d'office deux par jour, en tout cas que je détecte. Et je dirais le pourcentage, quand même, 50 %. Je ne sais pas si c'est important, mais... ça a l'air quand même de plus en plus présent.

I : Et penses-tu qu'en tant qu'utilisateur des réseaux sociaux, tu as un rôle à jouer avec les deepfakes ?

R : Un rôle, c'est-à-dire ?

I : Par exemple, un rôle dans leur propagation ou un rôle pour éviter leur propagation ?

R : Peut-être que j'ai un rôle, justement, pour éviter leur propagation, mais j'avoue que... Moi, je suis déjà assez ... pas préventif, mais j'évite tout ce qui peut être un peu faux, pour éviter tout ce qui est arnaques et tout. Donc j'évite tout ce qui est faux, pour éviter toute fraude ou pour ... J'avoue que je m'en fous, et je préfère faire passer l'info que, tu sais, signaler le truc pour dire que c'est faux.

I : Donc tu signales ?

R : Non, justement, genre ... Je ne fais pas confiance en ce genre de choses, et donc j'évite tout contact avec l'image ou quoi que ce soit, pour ... ne pas avoir de problèmes.

I : Et donc tu penses que t'aurais peut-être un rôle à jouer, mais comme t'as peur ...

R : Ouais, voilà. C'est vrai que c'est très bien de signaler pour dire, voilà, c'est faux ou ... Si c'est une publication pour te faire vendre quelque chose, tu vois. Je ne sais pas, j'imagine qu'il y a des trucs derrière, en mode les gens vont leur envoyer des photos, j'en sais rien. Tu sais, les trucs bizarres sur Internet, là. Donc peut-être qu'il se passe ça, mais moi je tombe pas dans le panneau, perso, et ... Pour le moment, parce que ça se voit que c'est encore assez faux. Donc voilà, j'espère que j'ai répondu à ta question.

I : Oui, pour être sûre que j'ai bien compris ... Tu veux dire que, par peur et comme tu sais que pour le moment tu arrives encore à les détecter, tu ne penses pas jouer un rôle dans leur propagation ?

R : Oui, voilà.

I : Et quelle a été ta première réaction ou pensée lorsque t'as pris connaissance des deepfakes ?

R : Rien, parce que ... l'intelligence artificielle existait, même avant les deepfakes. Et donc, juste, je sais qu'il y a eu ça en plus. Et donc, quand j'ai été face à mon premier deepfake, je savais ce que c'était.

I : Donc tu ne t'es pas dit quelque chose de positif ou de négatif au sujet des deepfakes ?

R : Non, je me dis tellement que la technologie, elle évolue vite que ... C'est normal que ça arrive, avec tout ce qu'on entend et tout ce qui est en train d'évoluer. Bientôt, je ne sais pas, ce seront sûrement des vidéos ... Même ça existe déjà, des vidéos deepfakes, on va dire. Donc non, ça m'a pas choquée parce que ça devait arriver. Mais c'est vrai que ça peut être dangereux si c'est utilisé avec une mauvaise intention derrière.

I : Et as-tu déjà eu le sentiment qu'une image ou une vidéo sur les réseaux sociaux pouvaient être truquée ?

R : Est-ce que j'ai eu un doute sur une image qui était deepfake ou pas ? Non, parce que souvent je trouve que ça se voit. Pour le moment, ça se voit.

I : Et as-tu déjà été influencée par une information qui, par la suite, s'est avérée fausse ?

R : En vrai, sûrement, parce que j'avoue que tout ce que je lis, je le crois un peu, on va dire. Mais est-ce que c'était faux ? Au final, ça, je n'en sais rien. Parce que j'avoue que des fois, je ne vais pas m'informer plus longuement sur le sujet. Mais c'est vrai que ça peut m'arriver de lire des choses qui peuvent être, au final, fausses. Parce que je lis beaucoup pour essayer de m'informer. Mais peut-être que des fois, c'est totalement faux.

I : Et selon toi, qu'est-ce qui rend un deepfake difficile à détecter à première vue ?

R : Peut-être parce qu'ils se basent sur des choses qui existent déjà. Je ne sais pas s'ils ont aussi connaissance des algorithmes, de tout ce que les gens aiment regarder, ou font sur les réseaux sociaux. Et du coup, peut-être qu'ils essaient de créer un peu les mêmes images que les personnes voient au quotidien. Et donc, les personnes ne voient pas de différence. Parce que c'est un peu tout ce qu'elles voient dans leur fil d'actualité tous les jours.

I : Tu veux dire par rapport aux intérêts des gens et tout ?

R : Oui, comme Instagram qui nous propose toujours des trucs qu'on regarde souvent. Je ne sais pas si l'intelligence artificielle a accès à ça, mais j'imagine peut-être que oui. Et qu'ils mettent à chaque fois des contenus du même genre. Pour moi, ce serait le plus difficile, c'est que ce seraient des trucs qu'on a l'habitude de voir et donc qu'on ne s'en méfie pas.

I : Et selon toi, comment distinguer un deepfake d'une vraie vidéo ou photo ?

R : Déjà la qualité. Je trouve que c'est toujours assez lisse comme image. Quand tu regardes les éléments, tout est trop lisse, trop parfait. Il n'y a pas de défauts. Je trouve que, quand ce sont des personnes, ça se voit plus, parce qu'il y a zéro défaut. Et tu vois que ce n'est pas humain, mais au niveau des objets ou des images ... je ne sais pas. Mais j'ai quand même l'impression que ça se voit. L'aspect lisse, je ne sais pas pourquoi, je ne sais pas comment expliquer.

I : Sur une échelle de 1 à 10, à quel point te sens-tu capable de reconnaître un deepfake ?

R : Je dirais 7. Mais après, je viens de penser que si tu me fais un test, j'ai l'air bête. Je dirais 7.

I : Et pourquoi du coup ?

R : Parce que je ne me sens pas si sûr de moi sur l'image que je vois. Comme je te dis, je les zappe quand même souvent. Donc, je ne m'attarde pas non plus trop dessus. Donc, je ne regarde pas trop quels sont les éléments qui font que ce n'est pas une IA. Et 7, parce que je trouve que c'est une bonne moyenne et que je sais que je sais les détecter, mais je ne suis pas sûre de moi non plus. Entre 7 et 8.

I : Ok. Et depuis que tu as connaissance des deepfakes, est-ce que tu as modifié ta manière d'utiliser les réseaux sociaux ?

R : Non. Juste avant, je passais les pubs, et maintenant, je passe ça aussi.

I : Et doutes-tu à présent de toutes les informations que tu rencontres sur les réseaux sociaux ou seulement quand il y a quelque chose qui te paraît étrange ?

R : Je tiens juste à rajouter que, comme je passe souvent ça, j'en vois moins du coup. Donc, ça ne change pas non plus mes habitudes des réseaux sociaux, parce que je n'en vois pas tout le temps, tous les jours comme je t'ai dit. Souvent, quand même, une fois par semaine.

I : Tu veux dire que comme tu montres que tu n'es pas intéressé, les algorithmes font qu'ils ne t'en montrent plus trop ?

R : Oui, je trouve que je n'en ai plus beaucoup.

I : Et donc, à présent, doutes-tu de toutes les informations que tu rencontres sur les réseaux sociaux ou alors que quand il y a quelque chose qui te paraît étrange ?

R : Alors moi, je suis du genre à tout lire. Et donc, je vais me renseigner, je vais lire, ça va rentrer dans ma tête et si vraiment ça m'intéresse, j'avoue que je vais aller voir sur Google plus de détails. Donc, j'avoue que je doute des informations quand ... Non, enfin, je ne doute pas ... Attends, comment expliquer ça ?

I : Et maintenant que tu sais qu'il y a des deepfakes, tu ne doutes pas forcément des informations parce que tu es capable de les détecter ?

R : Oui, c'est ça. Par exemple, ça peut être quelqu'un qui dit : “Donald Trump est devenu homosexuel “ et tu vois Donald Trump qui tient la main à un mec, par exemple. Ouais, alors là, ce genre de choses, j'y crois évidemment jamais.

I : Et si c'est quelque chose de plus subtil ?

R : Non, je ne doute pas plus, parce que des fois, ça se voit quand c'est complètement faux. Et juste moi, si je doute et que ça m'intéresse vraiment de savoir ça, je vais voir sur Internet.

I : Donc, tu fais confiance à Google pour aller vérifier tes sources ?

R : Ouais.

I : Et as-tu pris des mesures spécifiques pour te protéger contre les fausses informations ?

R : Je vais voir sur Google pour vérifier, mais sinon, est-ce qu'il y a d'autres mesures de sécurité pour ... Il y en a, mais je ne sais pas si ... Non, perso, à part faire ça, je ne prends aucune mesure.

I : Et donc, quand tu dis que tu vas sur Google, c'est pour vérifier tes sources ou c'est pour avoir un complément d'infos ?

R : Vérifier.

I : Oui, donc, tes mesures à toi, c'est de vérifier tes sources ?

R : Ouais, genre, par exemple, ici, je vois plein de TikTok sur l'affaire P. Diddy. Et du coup, le problème, c'est qu'il y a tellement de trucs qui sont dits et moi, ça m'intéresse, mais du coup, j'ai été voir sur Google.

I : Donc, au début, ça ne t'intéressait pas, mais il y a tellement de gens qui en ont parlé que tu t'es dit : "Je vais vérifier."

R : Ouais, je voyais plein de TikTok. Et puis après, le problème, c'est que ce n'était pas toujours la même chose. Évidemment, pour faire du buzz, mais donc, du coup, j'ai été voir sur Google.

I : Et lorsque tu vois une nouvelle vidéo, une nouvelle image en ligne, est-ce que tu prends systématiquement le temps de vérifier la source ?

R : Ouais, je crois que c'est un réflexe chez moi. Genre, je défile et je tombe sur le nom, sur le nom que la personne s'est donné sur les réseaux. C'est un réflexe, parce que je regarde aussi le contenu qu'il y a en dessous, la description de la vidéo. Donc, ouais. Après, je regarde pas pour me dire est-ce que c'est fake ou pas.

I : Ok, donc, si tu regardes le compte, comment fais-tu pour savoir si c'est une bonne information ?

R : Quand je vais sur le compte, je défile, je tombe sur l'image et je regarde qui a publié l'image ou la vidéo que je viens de voir.

I : Comment tu sais si, du coup, tu peux lui faire confiance ou pas ?

R : Ben, en regardant ses images, en regardant ce qu'il publie. Je vais défiler un peu tout ce qu'il a déjà créé, tout ce qu'il a déjà mis. Je vais voir aussi le nombre d'abonnés. Je pourrais savoir s'il y a beaucoup de gens qui font confiance à ça. Et puis après, en fonction de s'il y a 50 abonnés ou 50 000, ou si ses photos n'ont pas l'air louches et que la description, c'est ok, je fais confiance.

I : Et donc maintenant, tu es quand même un peu plus vigilante avec ce que tu regardes sur les réseaux sociaux. Mais est-ce que, dans ce que tu partages aussi, tu es vigilante ?

R : Oui. Je trouve que tout le monde ne connaît pas assez bien les réseaux, ne connaît peut-être pas encore tout ce qui peut arriver pour escroquer, oui. Quand je partage quelque chose, je m'assure que c'est bien une bonne chose. Je mets même souvent un petit mot, tu sais, quand j'envoie un lien, je mets toujours un petit mot, en mode t'inquiète, c'est bien moi qui t'envoie ce message. Par exemple, j'ai dû faire un questionnaire pour un travail, et

j'essayais vraiment dans le lien d'avoir un truc normal pour éviter que les gens pensent que c'est un truc faux.

I : Et par exemple, pour revenir un peu plus aux deepfakes, est-ce que, par exemple, si une vidéo attire ton attention et que tu souhaites partager cette vidéo avec tes abonnés ou tes proches ...

R : Alors, je ne suis pas très story, je ne poste pas grand chose. Et le partager avec mon entourage, ouais, je pense que je serais capable, parce que je ne fais pas trop attention à ce que j'envoie à ma famille ou amis.

I : Et considères-tu les réseaux sociaux comme une source fiable à l'heure actuelle ?

R : Non.

I : Et pourquoi, du coup ?

R : Au niveau des données personnelles. Ils volent nos données ou qu'ils prennent des données et qu'ils nous écoutent, et qu'ils vendent après tout ça.

I : Je voulais plutôt tourner cette question dans le sens ... Est-ce que pour avoir de l'information, penses-tu que les réseaux sociaux sont une source fiable ?

R : Non. Les réseaux sociaux, le problème, c'est qu'ils aiment trop raconter n'importe quoi pour avoir des vues, pour avoir des likes, pour avoir des sous. Du coup, sur Google, il y a quand même toujours tous ces articles qui sont normalement à peu près fiables. Après, il y a des comptes que je suis pour l'actualité. Mais c'est un faible nombre de gens fiables.

I : Et donc, les comptes que tu suis sur les réseaux sociaux, ce sont des journalistes ou ...

R : C'est un journaliste ou un truc qui est très connu. Je serais capable, peut-être, de m'abonner à la RTL.

I : Donc, les médias traditionnels qui sont maintenant présents en ligne, ça, ce serait une source fiable pour toi ?

R : Oui, parce que, comme je te dis, les autres, j'ai l'impression qu'ils racontent toujours tout ce qu'ils veulent pour avoir des vues et des likes. Et que, du coup, c'est un peu compliqué de savoir si c'est vraiment fiable ou pas.

I : Donc, RTL, pour toi, il n'a pas besoin de faire des vues et des likes, donc c'est fiable ?

R : Oui. Et c'est une chose que je connais, tu vois. Quand je ne connais pas ou que je n'entends pas parler, je ne préfère pas... Par exemple, Gossip Room, on m'en a parlé, et on m'a dit, il y a plein d'actualités là-dessus, même si c'est en France. Et du coup, je suis allé m'abonner, tu vois. Parce que je me dis que ça doit être un truc fiable.

I : Oui, donc, par exemple, tu vois un nouveau journal qui commence en ligne dont tu n'as pas entendu parler, tu n'y ferais pas confiance ?

R : Non, pas d'office.

I : Oui, donc tu te fies vraiment à ton entourage, à ce qu'ils suivent, eux, et un peu à l'effet de groupe, s'il y a beaucoup de gens qui suivent le compte, etc.?

R : Oui, si j'en entends vraiment parler, je me dis, ok, ça peut être intéressant pour moi de suivre un peu l'actualité avec ce compte.

I : Ok. Et selon tes réponses, du coup, quelles sont les sources d'informations auxquelles tu fais le plus confiance ?

R : Ça va être Google, et quand j'ai la flemme, ça va être ChatGPT, mais à ce moment-là, je demande un lien pour pas qu'ils me racontent des bêtises.

I : Et as-tu tendance à croire plus facilement une information lorsqu'elle est accompagnée d'une image ou d'une vidéo plutôt que d'un simple texte ?

R : Ouais, non, images et vidéos, je trouve que ça rend plus fiable. Pour moi, ça aura plus d'appui. D'abord je verrai la photo, puis je lirai le texte et je me rendrai peut-être compte que c'est faux après. Mais je verrai plus la publication si je vois une image ou une vidéo.

I : Ouais, donc ça va plus attirer ton attention ?

R : Ouais, ouais, certainement.

L : OK, et par contre, en fonction de la qualité, tu te diras si c'est vrai ou faux, comment ça se passe ?

R : Ouais, parce que vraiment, c'est ... D'abord, je vais voir la photo, et puis après, je vais lire ce qui se dit. Et je pourrais y croire si la photo est bien faite, mais à ce moment-là, je vais voir dans les commentaires aussi. Je vais beaucoup dans les commentaires pour voir ce qu'il se passe. Dans les commentaires, les gens disent souvent si c'est vrai ou faux. Donc, ouais, parfois j'ai quand même des doutes. Je crois que tu m'avais posé la question si j'avais des doutes. Donc, c'est vrai que je peux avoir des doutes, mais j'essaie de m'assurer de ne plus en avoir, que mon doute soit effacé, tu vois, que j'obtienne une réponse claire.

I : Et si, par exemple, une même information est répétée plusieurs fois par différentes personnes, est-ce que cela va renforcer ta confiance en sa véracité ?

R : Ah, ouais. Si ce sont des personnes qui sont assez hautes, qui sont assez célèbres. Alors, ouais, je croirais. Si c'est plusieurs personnes assez connues qui le disent, ouais.

I : Connue, tu veux dire, par exemple, Madonna qui met un truc, c'est forcément fiable.

R : Oui, parce que je me dis qu'ils peuvent pas non plus poster n'importe quoi. Ils ont quand même une communauté ... Tu sais, s'ils disent n'importe quoi. Oui, il y a beaucoup de gens qui vont croire ce qu'ils vont dire, quoi. Ils peuvent influencer un grand nombre de personnes. Donc, d'un côté, c'est vrai que ça peut être dangereux mais, d'un autre côté, je me dis qu'ils ne veulent pas ruiner leur réputation non plus.

I : Et si ce sont des personnes lambdas ? Enfin, est-ce que tu vas quand même te poser la question, en te disant : "Mais il y a quand même beaucoup de gens qui en parlent donc c'est peut-être vrai ?"

R : Ouais, mais c'est quelqu'un qui peut être mal informé. Mais après, souvent, cette information, elle tombe à un moment donné dans les mains des journalistes. Oui, donc si c'est vrai, ils vont la reprendre à un moment donné.

I : Donc tant que l'information n'est pas reprise par des journalistes, ty n'y crois pas ?

R : Oui, mais par contre je pourrais croire mon entourage, pas de souci.

I : C'est la fin des questions, je ne sais pas si tu souhaites ajouter quelque chose ?

R : Non, pas spécialement, j'espère que j'ai bien répondu.

I : Oui, merci. A présent, je vais te montrer trois photos. Tu auras donc 15 secondes pour voir chaque image, et me dire ensuite si tu penses qu'elle est vraie ou fausse. Pareil pour la deuxième, pareil pour la troisième image. A savoir que les trois photos peuvent être toutes fausses, toutes vraies ou une combinaison vraies-fausse. Et donc, une fois que tu auras fait ça pour chacune des images individuellement, tu pourras les regarder toutes les trois en même temps et me dire si tu confirmes ou ajustes tes réponses.

Voici la première image :



R : (Observe l'image pendant 15 secondes) Alors moi je dirais qu'elle est fausse. Il y a un truc qui ne va pas. La morphologie de la tête de Macron n'est pas la même. Et le problème, c'est que, quand je disais que tout est parfait, ici, c'est le cas. Et c'est très net, etc. Donc là, pour moi, c'est un deepfake.

I : Voici la deuxième image :



R : (Observe l'image pendant 15 secondes) Elle est vraie. Il n'y a rien qui me choque.

I : Tu te bases souvent sur s'il y a un élément qui te tape à l'œil ?

R : Oui, et là, il n'y a rien de bizarre à première vue.

I : Voici la troisième image :



R : (Observe l'image pendant 15 secondes) En fait, c'est un peu compliqué. Je pense aussi au fait que Macron et Trump, se marrer comme ça, c'est bizarre. Il y a aussi le fait que je me demande si c'est possible qu'ils rigolent comme ça. Faux pour la situation, mais au niveau de la qualité et tout. Les trois photos ont une bonne qualité, quoi.

I : A présent, tu peux les regarder toutes en même temps et me dire si tu gardes ou changes tes réponses.

R : La première, je n'ai pas de doute, ce n'est pas lui haha. Mais c'est vrai que ça, je vois souvent sur les réseaux sociaux, des gens qui essaient de faire des deepfakes qui ressemblent à des célébrités, mais ce n'est pas encore ça. L'intelligence artificielle n'est pas encore au top pour ça. Quoique, Trump est très bien fait. Même ils regardent tous les deux dans le vide alors qu'ils se serrent la main. Pour la deuxième photo, la qualité me fait dire que c'est une vraie image. Macron a une bonne tête, etc. Donc je dirai Après, quand je zoome sur la deuxième photo, la main de Macron, il y a un truc chelou. On dirait qu'il a été copié d'une autre image. Il a été copié-collé à la photo. C'est possible ça ? C'est aussi des deepfakes ?

I : Oui, oui, totalement. Tu peux créer des deepfakes à la base d'un texte, d'une description, mais aussi sur base d'une photo ou vidéo. Et l'IA peut reprendre exactement les personnes et les ajouter à un autre décor.

R : Ah bah alors, c'est faux. Mais c'est vrai que maintenant je ne vois que ça, la main. De loin ou de près, je le vois maintenant. Et la dernière image, je dirai du coup que c'est vrai par rapport aux autres images où j'ai trouvé un truc bizarre. Là, il n'y a rien de bizarre sur la

photo ou la qualité, c'est juste le contexte qui m'a fait douter. Mais maintenant, en vrai, je me dis que c'est possible.

I : Donc tu as tout bon à présent. C'était bien faux, faux et vrai.

R : J'ai bien fait de changer, haha. C'est vrai qu'en regardant vite fait, on peut tomber dans le panneau.

I : Oui, c'est là toute la difficulté. Et bien, merci !

Résumé

Dans notre société actuelle, les innovations technologiques transforment profondément notre manière de communiquer, d'accéder à l'information et de percevoir la réalité. Parmi elles, les deepfakes. Ces hyper-trucages, d'un grand réalisme, générés à l'aide de l'intelligence artificielle, fragilisent la frontière entre la réalité et la fiction, et mettent à l'épreuve la capacité de discernement des utilisateurs des réseaux socio-numériques. Cette recherche explore un phénomène aussi prometteur qu'inquiétant. S'il a été démontré que les deepfakes sont utilisés à des fins malveillantes, les répercussions sur le rapport à l'information sont encore assez floues. En adoptant une méthodologie qualitative articulée autour de deux approches complémentaires, cette recherche tente d'analyser les réactions suscitées par les deepfakes chez les utilisateurs des réseaux socio-numériques et leurs répercussions sur la réception, la transmission, l'interprétation et la crédibilité des messages. Des entretiens semi-directifs permettent d'explorer les attitudes cognitives, affectives et conatives des utilisateurs, tandis qu'un test d'awareness les confronte à des visuels truqués et authentiques, afin d'évaluer leur réelle capacité de discernement.

Mots-clés : Deepfakes, Réactions, Répercussions, Désinformation.