Fake News and Deepfakes Handout

The Problem of Fake News is Old: The phenomenon of fake news is old, as partially fabricated stories were common in the "yellow journalism" of the late nineteenth century, and even historical figures such as Benjamin Franklin and John Adams deliberately produced fake news stories. For example, Franklin created a fake issue of a real American newspaper that contained a story about native Americans being employed by the English crown to scalp innocent people, which he successfully had reported in several real newspapers. Even the medium, the copying of a real Boston newspaper, echoes the way in which contemporary purveyors of fake news mimic the layout of more reputable media outlets and often use a URL that suggests legitimacy (e.g. 'news-cnn.com', intended to borrow the legitimacy of the real Cable News Network (CNN)).

How Digital Technologies Exacerbate the Problem of Fake News: Levy notes that while fake news has always existed in some form, new technology has made it possible for fake news to be disseminated on a scale and with a speed that was never before possible:

- Fake news is not just a problem because it spreads false information, but also because it is often motivated by financial gain. Many individuals produce and disseminate fake news for advertising dollars or are paid directly to do so. This proliferation of fake news is then amplified by "dupes, bots, and a number of media outlets indifferent to truth and eager to swing debates."

- Levy highlights the significant impact that fake news can have on political events, noting that it has been credited with swinging both the Brexit referendum and the 2016 US presidential election.

- There are ethical issues to consider at multiple levels surrounding fake news, including the individual user, networks of users, social media companies, and regulators. For example, Levy poses questions about whether the consumption or sharing of fake news is a manifestation of vice, and whether such behavior is blameworthy. Levy also asks whether sharing fake news is permissible when it is explicitly labeled as such.

Main Argument of Article: Levy focuses on the effects of fake news on our **epistemic environment,** which refers to the social and institutional factors that influence our belief formation. Levy argues that forming accurate beliefs is essential for our flourishing, and fake news corrupts our epistemic environment in a way that has not been widely appreciated. **Instead of believing false narratives, we become skeptical of all narratives, which leads to a flattening of the epistemic landscape, characterized by a loss of cues to credibility and higher-order evidence**. Levy proposes possible responses to this flattening of the epistemic landscape and evaluates their ethical permissibility.

I. What is Fake News?

Fake News is Not Synonymous with False News: News is, roughly, reporting of purported facts of broad interest about the world. Fake news would seem to be news that is not true (Gelfert 2018). But lack of truth does not seem sufficient for (apparent) attempts at reporting to qualify as fake news (Jaster and Lanius 2018). Donald Trump loves to characterize CNN and The New York Times as 'fake news', but even when these media organizations get things wrong (as, of course, they do from time to time), they do not count as fake news. Equally, The Onion and other satirical newspapers do not count as fake news.

Fake News as 'Bullshit' View: What distinguishes fake news from genuine news is at least in part its aim. Fake news purveyors either seek to deceive—aiming to convince audiences that the world is, in some respect, different from how it actually is—or they are indifferent to the truth because they are motivated almost entirely by financial considerations. Frankfurt (2009) famously defined *'bullshit'* as assertion made without regard to truth. Several philosophers have suggested that fake news is closer to bullshit than to lies: its purveyors are often indifferent to truth, rather than seeking to deceive. For Mukerji (2018), fake news is dissimulated bullshit: it is bullshit that is passed off as news. Of course, people's intentions are hidden, and we might disagree about the intentions of a particular person or organization. However, attention to the processes whereby stories are generated helps to settle such disputes.

Discussion Question: In your experience online, do you think most purveyors of fake news are liars or bullshitters? Generally speaking, is it ethically worse to be a liar or a bullshitter?

Should We Get Rid of the Term 'Fake News'? Levy discusses the nature of fake news and argues that while reputable news organizations can sometimes be motivated by financial considerations and deceive their audience, they are still more reliable than typical purveyors of fake news. For instance, Donald Trump's favourite example of fake news, The New York Times is given a reliability score of 47 by the respected Ad Fontes Media organization, while Fox News cable TV is given 23, Breitbart 20.5 and Infowars. Whatever their motivation, reliable news organizations use a process that leads to significantly more reliable reports than those who aim to deceive or are indifferent to the truth. Levy notes that the accusation of fake news can be weaponized to dismiss any report that the accuser does not like but argues that dropping the term altogether would not solve the problem of unreliable claims. Instead, Levy suggests that the focus should be on improving media literacy and critical thinking skills to help people better discern between reliable and unreliable sources of information.

II. The Importance of an Unpolluted Epistemic Environment

Why Fake News is a Threat to Democracy: In democracies, voters are given the power to decide important policy issues such as taxation, peace, war, and human rights. However, for democratic control to be exercised responsibly and effectively, voters must be adequately informed. This requires a favorable epistemic environment, which means that voters must have a basic understanding of the direction in which different parties would take the country, even if they do not know the details of the policies they would implement. Levy acknowledges the democratic deficit, which refers to the ways in which actual democracies and their institutions fall short of the principle that important decisions should be subject to democratic deliberation. However, Levy emphasizes that well-functioning democracies leave many decisions subject to

some degree of democratic control, and that an unpolluted epistemic environment is crucial for this control to be exercised effectively.

The Concept of Epistemic Pollution: Levy discusses the concept of "epistemic pollution," which refers to features of the environment that make it difficult to form accurate beliefs. Levy notes that most countries have laws in place to minimize such pollution.

Real-Life (Instead of Online) Examples of Epistemic Pollution: The example is given of a soft drinks manufacturer deliberately mimicking a well-known brand to make it difficult for consumers to distinguish the two, or a seller of unproven "miracle cures" for terminal illnesses obtaining a fake accreditation and publishing in a predatory journal to appear credible. These tactics are aimed at lowering people's guard and making it easier to deceive them. Levy suggests that reducing epistemic pollution is essential to promoting a decent life.

Legislative Countermeasures to Epistemic Pollution: include trademark laws, which reduce the capacity of the unscrupulous to mimic familiar brands, and laws against deceptive advertising.

Institutional Responses to Epistemic Pollution: Beyond the law, institutional responses to epistemic pollution include the accreditation of universities by organizations set up for that purpose and formal and informal mechanisms for distinguishing reputable academic journals from disreputable ones.

Why Contemporary Fake News Versus Yellow Journalism: Yellow journalism declined as journalism professionalized and was able to self-regulate. But contemporary fake news is produced anonymously, circulates outside traditional channels, and is able to mimic real news.

III. Fake News and False Belief

Evidence that Fake News Produces False Beliefs: fake news can have a negative impact on democratic decision-making by leading people to hold false beliefs. While fake news may not always be motivated by an intention to deceive, it often consists of false claims. This can result in some individuals being taken in by fake news and acting on false beliefs, which can have negative consequences. For example, the person who fired shots at Comet Pizza believed false claims about nefarious activities taking place there. Similarly, conspiracy theorists who believed that the Sandy Hook mass shooting was a false flag operation subjected parents who lost children to online trolling and threats in person. Studies have shown that a significant number of people report believing fake news, such as the "Birther" conspiracy theory about Barack Obama's birthplace.

Evidence that Surveys Overestimate the True Extent of Belief in Fake News: Surveys that measure belief in fake news may overestimate its true extent. Survey responses are often generated on the spot, and people may use partisan heuristics or motivated reasoning to answer questions. *As a consequence, surveys may play a role in producing the beliefs that they report.* Respondents answer the question they are asked not only by bringing to mind prior beliefs, but also by using (often partisan) heuristics or motivated reasoning (*'Do I believe Hilary Clinton*

gave uranium to Russia in return for donations to the Clinton Foundation? That sounds like the kind of thing. she'd do, so I will answer yes').

The Concept of Expressive Responding: Survey participants may also engage in *expressive responding,* where they give a response not to report their beliefs but to express support for a policy, party, or person. For example, people may report that they believe Hilary Clinton gave uranium to Russia, not because they believe it but because they want to express their dislike of her. Levy cites a study that demonstrated the role of expressive responding in survey responses. The study asked participants to identify which photo showed a larger crowd at Trump's inauguration and found that 15% of Trump voters picked the photo of his inauguration as depicting a larger crowd, even though it was not true. In summary, he suggests that survey responses may not accurately reflect people's true beliefs about fake news, and that factors like partisan heuristics and expressive responding can lead to overestimates of belief in fake news. Levy further suggests that expressive responding is likely more common than previously thought and that people may exaggerate their beliefs for a variety of reasons, including trolling or expressing their dislike for a particular person or group.

Discussion Question: How often do you think fake news causes digital consumers to have false beliefs? Do you think the phenomenon of expressive responding is widespread?

IV. Flattening the Epistemic Landscape

The Epistemic Landscape: There are many sources of information in our epistemic environment. We accord greater credibility to some than to others. Think of more credible sources as rises in the epistemic landscape. Those sources to which we assign a great deal of credibility are peaks. Those we regard as just as likely to be wrong as right represent the ground level in the landscape. There may be depressions, or even valleys: they are sources that are more likely to be wrong than right (and which we can therefore utilize as sources of information by raising our credence in the negation of what they assert). *These sources of information may be individuals or institutions* (your mother; the American Medical Association; Fox News) non-human animals (you may think the flight of crows foretells the future, as the Romans apparently did, or perhaps you learn about the onset of spring from the song of the cuckoo), and so on. Whatever the source, the greater the credibility appropriately given to it, the higher the peak in the epistemic landscape it represents.

The Importance of Having Peaks in the Epistemic Landscape: Levy emphasizes the importance of peaks in the epistemic landscape for the proper functioning of our epistemic agency. He emphasizes that humans cannot evaluate each and every claim for themselves, as some claims are about events that are far from us in time or space, or require expertise that we may lack. Therefore, throughout human history, knowledge has always been distributed across multiple agents, and humans have always relied on one another for knowledge. *This dependence on testimony means that most of what we know is because we have been told.* However, we can only rely on testimony if we have cues on the basis of which to filter it. Multiple competing sources of testimony exist in the epistemic environment, and we rely on cues to credibility to distinguish reliable from unreliable reports. Therefore, peaks and troughs in the epistemic

landscape are essential for much of our knowledge. Levy suggests that fake news can have pernicious effects on the epistemic landscape, as it can undermine the cues to credibility that we rely on to distinguish reliable from unreliable reports.

Disorientation as the Goal of Fake News: Levy discusses the effects of fake news on the 2016 US presidential election and argues that Russian fake news was only effective because it was amplified by a right-wing media ecosystem that already engaged in similar tactics. This ecosystem was relatively self-contained, with consumption patterns on the right involving sampling from unreliable sources and little sampling from reliable sources. The ecosystem pursued the aim of disorientation rather than inducing false belief, and prior to the Russian initiative, it produced and echoed conspiracy theories and similarly unbelievable claims. These claims were not designed to be believed but rather to disorientate. Disorientation is produced not by bizarre claims on their own, but by the fact that they are apparently taken seriously by those with epistemic standing. There is a large body of experimental evidence showing that much of what we accept we do not really understand, leaving us epistemically vulnerable. Disorientation seems to be a primary goal of propaganda aimed at Russians themselves by state-linked agencies. The aim is not to convince viewers of some state-sanctioned narrative but rather to leave them confused, paranoid, and passive, living in a Kremlin-controlled virtual reality that can no longer be mediated or debated by any appeal to "truth."

Post-Truth and the Flattening of the Epistemic Landscape: Levy discusses the relationship between fake news and the flattening of the epistemic landscape, which is contributing to the era of post-truth. When fake news disorients people, it makes all sources of information suspicious, which means that people have lost trust in the mainstream media and no longer have a means of choosing between different narratives. *In this environment, expressive responding can take over, and people tend to choose the narrative that aligns with their beliefs and casts themselves and their allies in the best light while portraying their opponents in the worst.* Therefore, people tend to repeat and assert belief in the most congenial story, even if it may not be entirely true, leading to the era of post-truth.

The Disarmament Response to Fake News: Rini has previously proposed tightening up online norms of assertion to eliminate ambiguities and hold people accountable for spreading fake news. If we were to eliminate the ambiguities surrounding tweets, for instance (is retweeting endorsement or not?) we would be able to hold one another to account for spreading fake news and people would exercise more care to check veracity before posting or retweeting. However, Rini is now pessimistic about this response because people use social media for diversion and may not have an appetite for epistemic care.

- **Rini Quote on Disarmament Solution:** "In an ideal world, we would disarm the threat. We would focus on educating social media users to make thoughtful testimonial decisions, or on pressuring the platforms to provide tools for better testimonial hygiene. But I am increasingly doubtful that this approach will work. Individual users seem to have little motivation toward epistemic scrupulousness" (50).

The Detonation Response to Fake News: involves rejecting social media claims altogether and becoming universally skeptical of them. *Levy argues that Rini's approach would flatten the epistemic landscape by treating all social media claims as equally unreliable.* This approach

may be less threatening to Rini because she believes that the dissemination of fake news leads to a decrease in trust among citizens. Levy notes that there is not yet much evidence to support this claim but acknowledges that it is plausible. If this is the case, then Rini's solution may be successful in addressing the problem of decreasing trust. By treating all social media claims as unreliable, people would not come to distrust each other based on their social media content.

- **Rini Quote on the Detonation Solution:** "the best option may be to blow it all up. Rather than try to inculcate responsible use of social media testimony, we might do best to embrace testimonial skepticism. If we can convince people to become universally skeptical of social media content – to disbelieve everything conveyed through the medium, not just what their partisan affiliation makes disagreeable – then we can at least blunt the effects of induced turmoil, and slow the Ilyinic souring of democratic discourse. On this strategy, social media needs to be brought down to the epistemic level of children's cartoons; a source of diversion, but not a place any adult would think to go for information (50)."

Discussion Question: What is your opinion on the disarmament versus detonation response to fake news? Do you find one more plausible than the other? Do you have any ideas about how we go about mitigating the effects that fake news is having in flattening the epistemic landscape?

Levy's Concerns with Rini's Detonation Strategy: Levy argues that while it is important to arrest the decrease in trust between people to function as democratic citizens, this alone does not address the more narrowly epistemic problems that require reliable information about the world, society, and options for democratic deliberation. Levy argues that social media is used routinely to make assertions and report news events, and that it would be unrealistic to expect people to stop making assertions or using social media as a news source. Levy also questions the feasibility of asking people to be selectively skeptical and to adopt a skeptical attitude towards potentially fake news, as it is difficult to demarcate such news from straight news or routine social media assertions. Levy suggests that a response to fake news that does not require sacrificing epistemic value for the sake of trust should be sought.

V. Fake News and Higher-Order Evidence

The Concept of Higher-Order Evidence: higher-order evidence refers to evidence about our evidence. source and context can provide higher-order evidence that can affect how credible we perceive a claim to be. For example, a claim made by someone with a medical degree may be perceived as more credible than the same claim made by someone who is an accountant. Similarly, the context in which a claim is made, such as the platform on which it is presented, can also affect its perceived credibility. Levy notes that different types of higher-order evidence can be equally significant and that we routinely modulate the credibility we give a claim by reference to its source or context. Levy also notes that higher-order evidence is important in assessing the credibility of claims, especially in cases of disagreement. He argues that the prevalence of fake news in our current epistemic environment can be understood in terms of the corrosive effects of such misinformation on higher-order evidence.

How The Concept of Higher-Order Evidence Helps Us Respond to the Problem of Fake News: Understanding the epistemic landscape in terms of higher-order evidence provides a better sense of the permissible options for responding to fake news. Levy explains that it is

generally seen as unacceptable to be paternalistic with competent agents within the sphere of their competence. Controlling the epistemic environment might be seen as epistemic paternalism or authoritarian. For instance, suppressing arguments against the consensus view on evolution might improve people's epistemic position, but it could be seen as infantilizing people. Levy argues that epistemic paternalism is sometimes acceptable, but embracing it is a cost to any theory in terms of plausibility and the extent to which it can be expected to win over people. However, if peaks in an epistemic landscape are appropriately thought of as providing higher-order evidence, certain ways of controlling the epistemic environment might not count as epistemically paternalistic. Levy has previously argued that there are grounds for refusing certain speakers a platform because providing a platform generates higher-order evidence in favor of the claims they make. Similarly, there may be good reasons to control the epistemic environment to the extent that doing so avoids generating misleading higher-order evidence and instead generates reliable higher-order evidence.

Restoring the Topography: The problem of fake news requires a multidisciplinary approach from fields such as political science, media studies, sociology, computer science, law, and philosophy. One suggestion is to establish a fact-checking website with a small number of accredited facts, checked by a panel of experts with diverse political allegiances. The aim is to establish a stock of claims accepted as common knowledge across the political spectrum, which can serve as an accepted background for political debates. Over time, the stock can expand and help restore a variegated epistemic landscape in which people can reasonably trust competing sources of information. The proposal is limited in ambition but may be an important first step in addressing the problem of fake news.

Discussion Question: What do you make of Levy's modest proposal for helping to restore the topography of the epistemic landscape? Do you think it would be successful? Why or why not?

VI. Deepfakes

I. Introduction to Deepfakes

A. Definition: Deepfakes are synthetic media that use artificial intelligence (AI) to generate or manipulate images, videos, or audio. More specifically, deepfakes are artificially created video or audio recordings made using machine learning technology. These recordings involve superimposing a person's facial or vocal characteristics onto another individual, resulting in seemingly authentic videos of well-known people doing or saying things they never did.

B. How they work: Deep learning techniques, such as generative adversarial networks (GANs), are used to create hyper-realistic content.

C. Evolution: From simple face-swapping to full-scale video manipulation and voice cloning.

II. Deepfakes: Potential and Perils

A. Positive use cases

1. Art and entertainment: Filmmaking, animation, and virtual reality experiences

2. Historical and educational purposes: Reviving historical figures or events

3. Business applications: Advertising and personalized customer interactions

B. Negative use cases

1. Misinformation and propaganda: Spreading false information to manipulate public opinion

2. Cyberbullying and harassment: Non-consensual and harmful uses, such as revenge porn

3. Impersonation and identity theft: Exploiting someone's identity for malicious purposes

III. Ethical Concerns Surrounding Deepfakes

A. Consent and privacy: Deepfakes can be created without the knowledge or consent of the individuals involved.

B. Authenticity and trust: Deepfakes undermine the credibility of media and can lead to public distrust.

C. Legal implications: The lack of clear legal framework surrounding deepfakes complicates matters of intellectual property and defamation.

IV. Strategies for Detecting and Combating Deepfakes

A. Detection technology: Developing AI-based tools to identify and flag deepfake content

B. Digital watermarking: Implementing watermarking techniques to verify the authenticity of content

C. Media literacy: Educating the public on how to identify deepfakes and be critical consumers of information

D. Legal and policy approaches: Updating laws and regulations to address the challenges posed by deepfakes

V. Recommendations for Responsible Use of Deepfakes

A. Informed consent: Obtaining explicit consent from individuals involved in creating deepfake content

B. Transparency: Clearly labeling deepfake content to avoid misleading viewers

C. Ethical guidelines: Adhering to industry standards and ethical guidelines when creating and distributing deepfake content

VI. Regina Rini on Deepfakes

Deepfakes and Social Epistemology: While current deepfakes may not be entirely convincing, their potential to improve poses a threat to democratic processes of information-sharing and debate. Rini aims to address this gap by raising concerns about deepfakes and the potential erosion of knowledge in democratic societies before it becomes a reality.

Audio and Video Recordings as Epistemic Backstops: Rini discusses the role of audio and video recordings in regulating testimonial practices, which she refers to as an "epistemic backstop." Testimonial practices involve the transmission of knowledge through say-so, and recordings provide a way to verify or disprove testimonies. Recordings have two primary functions: actively correcting errors in past testimony and passively regulating ongoing testimonial practices. Acute correction involves situations where a recording reveals the truth, while passive regulation refers to the knowledge that being recorded might make people more sincere and competent in their testimonies. Rini argues that the role of recordings in testimonial practices has been historically overlooked by analytic philosophers. However, the ubiquity of recording technology in contemporary society means that it is an essential part of our social epistemic practices. The epistemic role of recordings applies primarily to testimonies about public events, but this category is already large and politically significant.

Deepfakes and Public Deception: Deepfakes gained public attention in December 2017 when an article by Samantha Cole highlighted the technology's use in pornography. The term "deepfakes" comes from a Reddit user who provided a free software tool called FakeApp, allowing people to create their own deepfake videos. Reddit eventually banned the deepfake community for violating its "involuntary pornography" rules. Although deepfakes pose ethical concerns, Rini focuses on their potential for epistemic mischief, especially in politics. Journalistic entities have published political deepfakes, such as Donald Trump's face superimposed on Angela Merkel's body or Barack Obama appearing to insult Trump. As deepfake technology improves, it becomes more difficult to distinguish real from fake, posing a risk to the credibility of recordings.

The Deepfake Arms Race: While deepfake detection technology might help counter the issue, it could also lead to an arms race between fakers and detectors. Furthermore, even if reliable detection technology exists, corrections may not travel as widely or be as readily believed as the initial fakes, resulting in epistemic chaos.

Backstop Crises: The primary concern is not that people will believe deepfakes, but that they will come to distrust all recordings, undermining the corrective and regulative functions of recordings. Rini refers to "backstop crises" as moments when the public realizes there is no longer a "smoking gun" tape, leading to a growing sense of displaced epistemic reality. As people become increasingly skeptical of video and audio recordings, the passive regulatory role of recordings may be lost, eroding the motivation to be responsible testifiers. The gravest danger of deepfakes, then, is the potential elimination of the epistemic credentials of all recordings,

which could lead to a decrease in trust for testimonial practices and the eventual unraveling of norms that secure anticipated cooperation.

The Epistemology of Recordings- Perceptual Versus Testimonial Knowledge: Rini discusses the epistemic roles of photographs and paintings, comparing them to audio and video recordings. She references Kendall Walton's transparency thesis, which argues that photographs allow direct, literal perception of their objects. She also cites Robert Hopkins' idea of a necessarily veridical seeing-in relation and Dan Cavedon-Taylor's claim that photographs generate perceptual knowledge while paintings only support testimonial knowledge. She posits that audio and video recordings, like photographs, provide perceptual evidence that has stronger presumptive authority than testimonial evidence, making them suitable for providing an epistemic backstop. The concern is that deepfakes may erode the epistemic status of recordings, demoting them to mere testimonial evidence and undermining their role in regulating testimonial practice. *She also addresses the question of why digital manipulation of photographs hasn't led to an epistemic catastrophe, questioning whether society will adapt similarly to deepfakes.*

The Distinctive Threat of Deepfakes: Rini discusses the history of photo manipulation and the concern surrounding deepfakes. She notes that photo manipulation dates back to the early 20th century, as demonstrated by Arthur Conan Doyle's publication of photographs of fairies taken by two young girls, which were later revealed to be staged. She highlights four reasons deepfakes pose a distinctive threat:

1. **Psychological difference:** People are more psychologically invested in audio-video recordings, as they feel more like reality than photographs.

2. **Efficient response to challenges**: Deepfakes can be produced more efficiently than traditional photo manipulation, allowing for a quicker response to doubts about authenticity.

3. **Mass production**: Machine learning enables deepfakes to be produced on a large scale, allowing malefactors to "spam" the environment with false information.

4. **Role as an epistemic backstop:** Audio and video recordings function as backstops for still photographs, and if their reliability is compromised, it exposes decay in our epistemic relations to photos and testimony.

Discussion Question: Given that we have had photoshopped images for some time and this has not lead to a huge epistemic crisis, do you think Rini overexaggerates the epistemic threat of deepfakes?

Frankenporn and Virtual Domination: Deepfakes offer a disturbing form of power over others and have been used to create involuntary pornographic content featuring celebrities and non-consenting individuals. This has reignited debates around the objectification of women in pornography and the ethics of such content. Deepfakes exacerbate the objectification issue as they involve no consent, literally reducing women to interchangeable and commodified body parts. Furthermore, deepfake porn consumers often target specific individuals, suggesting a desire for sexual domination rather than mere physical gratification. The creation and sharing of

deepfake pornography represent an extreme expression of sexual objectification, turning real people into digital toys and perpetuating harmful attitudes toward women. ‘Frankenporn’ involves digitally fusing the parts of different women to create pliable characters incapable of giving consent to their depiction.

Deepfakes and Illocutionary Harm: Illocutionary harm occurs when a person is illegitimately compelled to perform an undesired speech act. Deepfakes can cause this harm even if they are debunked or not believed. Rini presents an example of a deepfake featuring a public figure, forcing the person to address a fabricated situation, which harms their capacity as a speaker. Illocutionary harm can occur even when the faked speech is consistent with the speaker's beliefs but they did not wish to express it publicly. Rini provides examples of existing technology, such as maliciously edited videos, causing similar harm. While it acknowledges that the harms discussed can occur without deepfakes, the accessibility and ease of creating deepfakes make them a distinct ethical problem. The widespread availability of deepfake technology raises concerns about the impact on ordinary lives and day-to-day relationships.

Panoptic Gaslighting and Existential Trauma: Rini discusses the potential harms of deepfakes on personal memory and identity. Deepfakes involve the creation of fabricated recordings that can deceive both the target and the audience. They can even make the target doubt their own memories, which can lead to several serious consequences.

Psychological studies show that fake photographic and video evidence can manipulate autobiographical memories and cause people to falsely remember events or actions. *This can result in material harms, such as being tricked into giving up something or fulfilling unremembered obligations.*

Gaslighting, a form of emotional manipulation, can occur due to deepfakes. It involves inducing a sense of doubt in someone's reactions, perceptions, memories, or beliefs. Deepfakes can lead to *"panoptic gaslighting,"* where a person's grip on reality is deliberately ruined by creating subtly altered videos, causing them to doubt their memory altogether. Existentially, this can be harmful, as losing faith in one's memories can lead to dislocation, disintegration, and a loss of self-respect. Personal identity is often linked to the ability to recall earlier experiences, and doubting those memories can result in a form of identity fracture.