Gabriella A. Macias

Oxford Scholars

12 April 2026

Keeping People Safe: Battling AI Misuse Through Responsibility and Education

It is no surprise that since the rise of technology and social media in the 2000's, we have seen an increase in media use and consequently in the unethical use of online platforms and programs. The development of easy access AI chatbots like ChatGPT which only came about a little over 3 years ago has not only allowed it to infiltrate almost every aspect of society, such as schools, workplaces, and media, but has also increased the risk of misuse and harm to individuals. Problems with AI arise when we stop using it for learning or to understand a work better and start using it to do our own work for us, or even for criminal purposes. Pope Leo writes to us Catholics on the ethical use of AI, remarking that we need "not to halt digital innovation, but to guide it," toward the betterment of our souls. Artificial intelligence should not be restricted but guided by responsibility and education, ensuring that its capabilities are not exploited for criminal purposes, ultimately empowering people to strengthen their capacity for knowledge and moral judgment.

We often hear about the immoral uses of AI in the workplace and school for things such as cheating, plagiarizing, and more, but more serious implications of AI being weaponized include deeply harmful and personal attacks against persons. Deepfakes, a form of identity fraud that uses other people's appearances and wrongly deposits them on different bodies or in different scenarios, are harmful and invasive to the person whose identity is being stolen. According to Reuters, in one year, around "1.2 million children had their own images turned into explicit deepfakes," which signals that this is a global and criminal concern needing to be

addressed. In one such instance, the Guardian reported that a company lost nearly $500,000 after criminals used a fake AI video to impersonate the company's directors. Experts at the Guardian warn that deepfake fraud is now "widespread, inexpensive, and highly accessible." Impersonation scams and fraud attempts have increased by 2,137% over the last three years, according to SQ Magazine, and they specifically target vulnerable groups, indicating that AI is actively used in financial crimes. Deepfakes are one branch of cybercrime, and unfortunately, there are multiple ways AI can be used criminally, especially in automated hacking, election tampering, or creating intelligent malware that can evade detection by antivirus software. It is clear that AI's remarkable intelligence and easy access have increased the scale and speed at which these crimes occur, and will continue to occur if responsibility is not managed.

People often don't anticipate the risks of AI until they are actively being harmed by it, and a lack of education about it makes people more vulnerable to deception and misuse. AI can be deceptive in two ways: the first way has to do with deepfakes and AI-impersonation, the second way is when you think you are completing an assignment or task to the best of your ability but really you are just copying AI's thought process. This limits your capacity to think independently and thus limits your intelligence. When it comes to the former, according to studies conducted by Cornell researchers, 66% of participants could not identify artificially generated audio, and 43% could not identify artificial videos, ultimately highlighting a major risk for personal or financial harm. Cornell researchers call AI deception a "systematic inducement of false beliefs in the pursuit of an outcome other than the truth." Digital literacy is essential for informing users and keeping them from falling victim to AI scams or using the service irresponsibly. For this to be possible, we must adapt as a society and integrate specialized AI education into schools and workplaces. This solution, along with education in critical thinking,

can help eliminate the unethical use of AI even in non-criminal cases. The ultimate goal of AI should be to expand human creativity and free our conscious minds, not to deceive ourselves or others when used unethically.

Our responsibility in AI management also calls into question the source of these issues, and a lack of accountability can create a dangerous environment where the government or state cannot check this harmful power. If we as a society cannot dictate who is responsible for the crimes, such as the user, developer, or platform, then we cannot begin to instate laws and regulations for appropriate use. Furthermore, the rise of agentic AI, systems capable of acting independently from human input, raises even more ethical concerns because this shifts responsibility. Agentic AI can interact, complete tasks, and conduct research autonomously by looping its thinking and continuing to process information without guidance, ultimately taking action rather than simply answering questions. To keep people safe from autonomous systems, we need to continue pushing for digital literacy across generations. Cornell researchers outlined three possible solutions to AI deception: regulatory frameworks that conduct risk assessments of systems and hold users legally accountable for criminal misuse, bot-or-not laws implemented by policymakers that prohibit undisclosed AI communication, and relevant research funding to prevent future harm. In addition to these solutions, watermarking AI content has proven to help users identify fabricated media, and the government and technology companies could take this a step further by implementing detection and verification systems that label AI media. Furthermore, stricter identity protection laws can deter people from using AI for fraud, and, along with other laws and regulations, can significantly reduce the likelihood of deception.

To work with AI in the future, we must acknowledge that AI can help at times, but must be guided by the proper education and use of such tools. People will not comply with AI

regulations or use the service properly if they don't even understand AI's capabilities. However, the knowledge of such capabilities must also be capped by accountability measures to prevent misuse. If we combine our responsibility to support unique human creativity and our responsibility to educate people on AI, we can strengthen societal resilience and ensure people are not harmed by it. AI education should go above and beyond to teach students how to recognize AI-generated media, evaluate the credibility of information, and understand the limitations of their own use of AI. In the workplace, employees should be guided in the responsible use of AI, especially when handling sensitive data or making important decisions. If we can support the ethical use of AI alongside critical thought, as the UCLA Comparative Lit AI class taught by Professor Stahuljak does by generating an AI textbook and learning thoughtfully, we can teach future generations the importance of thinking for themselves, only by expanding their minds.

AI will only continue to grow and expand its capabilities, as seen in the rise of automated systems, which means we will have to work even harder to keep people safe. As future and current generations learn more about AI, it will become imperative that we educate people on responsible use. Our growing responsibility when it comes to AI means understanding ways to protect individuals and ourselves from an industry that operates in ways we don't quite understand. Through proper education and regulation, we can empower people to use AI thoughtfully instead of wrongfully. This means integrating AI programs into schools and workplaces and increasing our commitment to individuality and uniqueness. AI should not be a replacement for human effort but a supporter of it, and should never be a source of harm to others. The more we understand our responsibility for ethical growth and development through

the lens of AI, the better we can realize that the future need not depend on technology but on human values and experiences.

Works Cited

"AI Deception: A Survey of Examples, Risks, and Potential Solutions." *arXiv*, 2023, arxiv.org/abs/2308.14752.

"AI Voice Cloning Fraud Statistics." *SQ Magazine*, sqmagazine.co.uk/ai-voice-cloning-fraud-statistics/.

"Audio and Video Deepfake Detection Study." *arXiv*, 2025, arxiv.org/abs/2502.00961.

"Con Quotes: Artificial Intelligence Debate." *Encyclopaedia Britannica*, britannica.com/procon/artificial-intelligence-AI-debate/Con-Quotes.

"Deepfake Fraud Taking Place on an Industrial Scale, Study Finds." *The Guardian*, 6 Feb. 2026, theguardian.com/technology/2026/feb/06/deepfake-taking-place-on-an-industrial-scale-study-finds.

"UNICEF Calls for Criminalization of AI Content Depicting Child Sex Abuse." *Reuters*, 5 Feb. 2026, reuters.com/legal/government/unicef-calls-criminalization-ai-content-depicting-child-sex-abuse-2026-02-05/.