

## Artificial Intelligence (AI), Adverse Media, and Defamation: Navigating the Digital Age of Reputation Risk

---

YOU ARE FALSELY ACCUSED AND YOU CANNOT EVEN OPEN A BANK ACCOUNT?

What happens when artificial intelligence—capable of sifting through billions of media sources in seconds—gets it wrong and harms someone’s reputation? Imagine being falsely accused of embezzlement or misconduct in an AI-generated news summary, only to find your name splashed across the internet before you even know about it. *(For more information about adverse media, please visit our YouTube channel at: [https://www.youtube.com/watch?v=sJdfkVYhW\\_U&ab\\_channel=PerspectivesEYE](https://www.youtube.com/watch?v=sJdfkVYhW_U&ab_channel=PerspectivesEYE))*

YOU WILL NEVER KNOW THAT YOU HAVE ADVERSE MEDIA, UNTIL YOU CHECK YOURSELF IN AN AI SEARCH

In today’s digital world, adverse media—negative news about individuals or organizations—can spread at lightning speed, amplified by AI-driven tools that collect, analyze, and sometimes misinterpret vast amounts of data. With the rise of generative AI, the risk of defamation—false statements that harm reputation—has become more complex and far-reaching than ever before.



AI is transforming how organizations screen for adverse media, but it also introduces new challenges and legal questions about defamation and responsibility for false information.

ARTIFICIAL INTELLIGENCE SCREENING IS CHANGING THE GAME

AI is revolutionizing adverse media screening by automating the collection and analysis of news articles, **social media posts**, regulatory databases, and more. Advanced natural language processing (NLP) and entity recognition allow AI systems to extract relevant information from unstructured text, identify key entities, and filter out duplicates or low-quality content.



YES, SOCIAL MEDIA POSTS,  
YOU READ RIGHT!!!

BENEFITS OF AI IN ADVERSE MEDIA SCREENING

- AI can process millions of sources across multiple languages, expanding coverage and efficiency.
  - Accurately identifies individuals, organizations, and locations mentioned in adverse media.
  - Detects trends and risks, helping organizations stay ahead of emerging threats.
-

## THE OTHER SIDE, THE DARK SIDE

Despite its power, AI is not infallible. Generative AI models can produce confident but false statements, that can be damaging if taken at face value and disseminated. For example, in a recent case, A generative AI model falsely accused a radio host of embezzlement, leading to a high-profile defamation lawsuit.



### **AI ON TRIAL: LEGAL MODELS HALLUCINATE IN 1 OUT OF 6 (OR MORE) BENCHMARKING QUERIES**

<https://hai.stanford.edu/news/ai-trial-legal-models-hallucinate-1-out-6-or-more-benchmarking-queries>

Defamation occurs when a false statement harms someone's reputation. Traditionally, defamation law has focused on human publishers and media organizations. But AI, as a non-human actor, complicates the legal landscape.

## WHAT TO DO

- Contact the AI provider or Platforms
- Request Correction or Removal: Reach out to the company behind the AI tool or the platform where the content appears. Ask them to correct the information or remove it if possible
- Review Terms of Service: Check the terms you agreed to when using the service, as they may contain clauses about dispute resolution, liability limits, or arbitration requirements
- Check yourself in all the AI platforms you can. Write your name and ask if there is any background information about that name
- Seek Legal Advice

The law is still catching up to the realities of AI-generated content, and future cases may set new precedents.

---