International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

The Unseen Risks of Humanizing Technology

B Nidhi Rathore

Student
Mumbai University

Abstract

AI is no longer just a tool, it's becoming a part of our daily lives. It shapes how we communicate, work, and even form relationships. As AI gets smarter, human interactions with it are becoming more personal, more complex, and sometimes, disturbingly inappropriate. From AI-driven harassment and sexualization to deep-fake exploitation and emotional manipulation, we're entering a dangerous territory where AI can be misused in ways we never anticipated.

This blog breaks down these risks, using real-world cases like Replika's chatbot controversy and California's deepfake legislation, showing why we can't afford to ignore the dark side of AI interactions. In absence of clear laws, and strict content moderation, AI could become a new tool for exploitation.

The question isn't whether we should regulate AI, it's how fast we can act before it's too late.

A New Kind of Grey Area

Artificial intelligence seems utopian. But its in fact, the reality in today's revolutionary digital. With this innovation comes a new, unsettling reality - AI's intersection with sexual harassment.

As technology evolves, so do the risks of its misuse.. Should existing regulations be adapted, or is it time for entirely new laws to govern inappropriate human-AI interactions?

Recent years have witnessed **disturbing cases of AI-driven abuse**, such as the creation of non-consensual explicit content using AI-generated deepfake technology [1].There have been legislations, signaling a growing awareness of the need for legal intervention. Unfortunately, such measures remain **a patchwork solution** in an increasingly complex digital world.

With time, AI systems are becoming more deeply embedded in personal and professional environments, and with that, the potential for unethical and harmful interactions grows. Who should be held accountable when AI is weaponized for harassment? Can AI itself be complicit in wrongdoing, or does the responsibility lie solely with its human operators? Most importantly, are the **current laws and legislations** robust enough to address these emerging threats?

[1] Susan Rose-Ackerman, *Transnational Business Deals and the Professional Obligations of Lawyers*, 54 Colum. Hum. Rts. L. Rev. 673(2022)

It is essential to break down the **urgent need to define boundaries in human-AI interactions, so** that technological advancement does not come at the cost of human dignity and safety

More Than Just Machines

Human-AI interaction is the dynamic exchange between people and AI-powered systems. These interactions can range from basic command-driven inputs (i.e. using a chatbot for customer service), to more advanced engagements, like AI-generated content creation. The ultimate goal of human-AI interaction is to design systems that are intuitive, efficient, and beneficial to users.

Human-AI interactions are influenced by certain factors like **natural language processing (NLP), machine learning, user intent, and AI adaptability**[2]. As AI capabilities improve, interactions are shifting from simple one-way commands to complex, human-like dialogues where AI can predict and respond intelligently.

Dimensions of Human-AI Interaction

To create meaningful and effective AI interactions, various particulars must be considered:

1. **Agency and Initiative:**
   AI systems vary in their level of autonomy. Some operate purely reactively (e.g., voice assistants responding to commands), and others take proactive initiatives (e.g., AI-powered recommendation systems suggesting actions based on user behavior). Striking a balance between human control and AI autonomy prevents many ethical dilemmas, and unintended biases.

2. **Transparency:**
   AI systems make decisions that impact individuals and organizations. It is essential e that users understand how and why AI reaches specific conclusions. Explainability of AI brings about trust and accountability, especially in sensitive areas such as finance, healthcare, and criminal justice.

3. **Ethics and Bias:**
   AI is only as unbiased as the data it is trained on. Ethical considerations play a vital role in shaping human-AI interactions. Biased AI decisions can reinforce societal inequalities, and lead to ethical oversight.

4. **Reliability:**
   Human trust in AI systems depends on their consistency and predictability. If an AI system produces unreliable outputs or exhibits erratic behavior, users are less likely to adopt and rely on it. AI functions correctly under diverse scenarios, including stress-testing in real-world conditions.

[2] Cabitza F, Campagner A, Sconfienza LM. *Studying human-AI collaboration protocols: the case of the Kasparov's law in radiological double reading*. Health Inf Sci Syst. 2021 Feb 5;9(1):8. doi: 10.1007/s13755-021-00138-8. PMID: 33585029; PMCID: PMC7864624.

5. **Adaptability:**
   Many AI systems have learning capabilities. Machine learning models evolve based on user interactions, and continuously improve their accuracy and responsiveness. But unregulated adaptability may lead to ethical risks, more so if an AI system learns inappropriate behaviors from users.

The Growing risks

Even through significant advancements, several challenges remain in regulating human-AI interactions:

- **Misinterpretations:**
  AI may struggle with understanding nuances, emotions, and cultural contexts in human language, which might lead to incorrect or inappropriate responses (e.g. AI chatbots sometimes misinterpret sarcasm or humor, and in ways that feel unnatural or robotic.)

- **Overreliance and Bias:**
- Users may become overly dependent on AI. They might assume its recommendations or decisions are always correct.[3] This is especially concerning in high-risk areas like medical diagnoses, legal decision-making, and financial transactions.

- **Privacy Concerns:**
  AI systems often collect and analyze vast amounts of personal data. Without proper safeguards, these interactions may lead to data privacy violations,or unauthorized surveillance. Ethical AI development must prioritize transparency and consent in user interactions.

- **Emotional and Psychological Effects:**
  AI is increasingly designed to simulate human emotions. This also at times leads to emotional attachment to AI systems. Users may form psychological bonds with AI, which raises ethical questions about the responsibilities of AI creators.

The ELIZA Effect

A well-documented phenomenon in human-AI interaction is the **ELIZA effect**, where users mistakenly attribute human-like intelligence and emotions to AI systems. The effect is named after the early AI chatbot *ELIZA*, developed in the 1960s. This effect demonstrates how people project human characteristics onto AI, while assuming it possesses deeper understanding or reasoning abilities than it actually does.

For example:

- A person may engage with a chatbot as if it has real empathy, even though its responses are pre-programmed.

[3] Ineffective Human-AI Interactions and solutions, The Oxford Review Briefings,(30-01-2025, 7:20pm),https://oxford-review.com/ineffective-human-ai-interactions-and-solutions/

- Users may believe AI-generated advice is as reliable as human expert guidance, without realizing the AI lacks real-world contextual awareness.

This overestimation of AI capabilities can lead to misplaced trust, and unrealistic expectations regarding AI's role in society.

Blurred Boundaries and Limits

Day by day, AI is becoming a bigger part of our daily lives. Our interactions with it are becoming more personal and complex. Though AI offers many benefits, it also brings new ethical, psychological, and legal challenges when these interactions cross inappropriate boundaries. These issues raise serious concerns such as the sexualization of AI, AI-driven harassment and the creation of non-consensual explicit content,. Below, we explore the different types of inappropriate human-AI interactions and real-world case studies that showcase the urgent need for clear regulations.

The Types of Disturbing Interactions

1. **Sexualization & Objectification**

AI chatbots and virtual assistants are often anthropomorphized. At times, this leads to users developing personal or even sexualized relationships with them. Its true that some AI systems are designed for companionship, but problems arise when:

- Users engage in explicit, degrading, or abusive conversations with AI.
- AI entities are designed to cater to hyper-sexualized or gender-stereotyped roles.
- [4]AI companionship leads to **dependency**, where users develop unhealthy emotional bonds that impact their real-life relationships.
- Ethical concerns arise about **whether AI can consent** and what psychological effects these interactions may have on users.

2. **AI-Facilitated Harassment and Abuse**

AI-powered tools can be exploited to **harass, manipulate, or spread harmful content**. These abuses include:

- AI chatbots being used to send **non-consensual explicit messages** to real users.
- Automated AI bots creating **offensive, racist, or misogynistic content**.
- AI-generated text-based **cyberstalking** and **harassment** on social media.

3. **AI-Generated Non-Consensual Content**

The emergence of **deepfake** technology (AI-generated images, videos, and audio) has led to **serious ethical and legal concerns**:

- The creation of **AI-generated pornographic content** using real people's faces without consent.

[4] Rui Zhang, Christopher Flathmann, Geoff Musick, Beau Schelble, Nathan J. McNeese, Bart Knijnenburg, and Wen Duan. 2024. *I Know This Looks Bad, But I Can Explain: Understanding When AI Should Explain Actions In Human-AI Teams.* ACM Trans. Interact. Intell. Syst. 14, 1, ART. 6 (March 2024), 23 pages.

- [5]The manipulation of public figures' images for **political misinformation** or **blackmail**.
- The rise of **"revenge porn" deepfakes**, which disproportionately target women.

4. Emotional Manipulation and Psychological Harm

U**nregulated AI systems can cause harm** when:

- Users develop **emotional dependency** on AI companions, which causes **social withdrawal** or mental health issues.
- AI chatbots inadvertently **reinforce suicidal thoughts** or **manipulate vulnerable users**.
- AI-generated interactions **blur reality**. This causes users to make decisions based on **fabricated emotions**.

The integration of artificial intelligence (AI) into various aspects of society has led to advancements, but it has also introduced legal and ethical challenges concerning sexual harassment and interactions.

[6] As AI systems become more sophisticated and ubiquitous, it is important to address these challenges for the protection of individuals and the ethical use of technology.

From Loopholes to Laws

The emergence of AI-generated inappropriate content has prompted legislative bodies to consider new laws or amend existing ones to address these issues. In 2024, the U.S. Senate unanimously passed the DEFIANCE Act[7], which allowed victims of non-consensual intimate images created by AI to sue creators for damages up to $150,000, and $250,000, if they were related to sexual assault, stalking, or harassment.

Similarly, California enacted legislation to protect minors from AI-generated sexual content. [8]This clarified that child sexual abuse material is illegal even if AI-generated. The legislation enabled prosecutors to pursue individuals who possess or distribute such material, closing previous legal loopholes.

[5] Charlotte McDonald-Gibson, 'I was just terrified': Parents battle Big Tech over AI nudes, THE TIMES, (29 Jan 2025, 8:00PM),https://www.thetimes.com/world/us-world/article/i-was-just-terrified-parents-battle-big-tech-over-ai-nudes-qmdvhc2wk?utm_source=chatgpt.com&region=global .

[6] Yoonsuck Choe, Lama Nachman, Philipp Michel, Bendert Zevenbergen, In Kwon Choi,Human - *AI Collaboration Framework and Case Studies* ,Collaborations Between People and AI Systems (CPAIS) (29-01-2025,6:29PM),https://partnershiponai.org/wp-content/uploads/2021/08/CPAIS-Framework-and-Case-Studies-9-23.pdf.

[7] Lauren Fein, The Senate passed a bill cracking down on sexually explicit deepfakes,, THE VERGE, July 24, 2024 at 11:17 PM GMT+5:30, https://www.theverge.com/2024/7/24/24205275/senate-passes-defiance-act-non-consensual-intimate-ai-deepfakes?

[8] Tran Nyugen, California governor signs bills to protect children from AI deepfake nudes, APNEWS, March 21, 2024 at 7:00PM, https://apnews.com/article/ai-deepfakes-children-abuse-7dcf5c566e2a297567f1e148ac2074a4

Even with these efforts, legislation is often not at par with technology. Many jurisdictions lack specific laws addressing AI-generated non-consensual explicit content. This leaves victims with limited legal recourse, and solidifies the need for comprehensive legislation that addresses these unique challenges.

Where Do We Stand?

India's primary legislation addressing workplace sexual harassment is the Sexual Harassment of Women at Workplace (Prevention, Prohibition and Redressal) Act, 2013 (PoSH Act). The Act mandates the establishment of Internal Complaints Committees (ICCs) in organizations and explains the procedures for addressing complaints. But the PoSH Act primarily focuses on human-to-human interactions and does not explicitly cover scenarios involving AI systems.

The Information Technology Act, 2000, along with its amendments, addresses certain facets of digital misconduct, including the transmission of obscene content. Yet, it lacks specific provisions related to interactions between humans and AI that may be deemed inappropriate or harassing.

Ethical Considerations

The ethical concerns surrounding AI and sexual harassment are profound. The creation and dissemination of such explicit content without consent violates individual autonomy and dignity. It propels gender-based violence. Studies have shown that women are disproportionately targeted, with 96% of deepfake videos online being pornographic and 99% of the victims being women. [9]

Discrimination:

AI systems used in recruitment and performance evaluations, can inadvertently perpetuate biases present in their training data [10](e.g if an AI tool is trained on data that reflects gender biases, it may favor male candidates over female ones)

Accountability:

It is difficult to determine who is responsible when an AI system behaves inappropriately. Is it the developer, the operator, or the organization deploying the AI? Clear guidelines are necessary to delineate accountability in such scenarios.

Prevalence of AI-Generated Explicit Content

[9] Deepfake Pornography, WIKIPEDIA- THE FREE ENCYCLOPEDIA (25 Jan 2025,7:45PM),https://en.wikipedia.org/wiki/Deepfake_pornography?utm_

[10] Avik Biswas, Ivana Chaterjee, Advika Madhok , *Ethical implications of artificial intelligence at the Indian workplace* , INTERNATIONAL BAR ASSOCIATION, 1, 1-3, Dec 2023, ethical-implications-of-artificial-intelligence-at-the-Indian-workplace

It has become increasingly easy to create realistic yet fabricated explicit images and videos without the consent of the individuals depicted. A notable instance occurred in January 2024, when AI-generated sexually explicit images of American musician Taylor Swift were widely disseminated on social media platforms, including X (formerly Twitter) and 4chan. One post was viewed over 47 million times before its removal. This is an indication of a rapid spread and significant impact of such content.

This incident is an example of a broader trend where AI is used to produce such explicit material, often targeting women and minors. The Internet Watch Foundation reported a significant increase in AI-generated child sexual abuse imagery, noting the spread of realistic deepfake videos. These developments depict a disturbing evolution in the misuse of AI.

Time for a New Rulebook

At present, there's an urgent need to revisit and possibly amend existing laws to involve scenarios including AI. Some measures may be :

- **Expanding the Scope of the PoSH Act**: Incorporating provisions that address harassment facilitated by AI systems.

  Data Protection Legislation: Implementing data protection laws that regulate the collection, storage, and use of personal data by AI systems, and safeguard individual privacy.

- **Ethical Guidelines**: Forming ethical guidelines for AI development and deployment. AI systems should be designed and used in ways that protect human dignity.

Balancing Innovation & Safety

The regulation of inappropriate human-AI interactions needs a balanced approach that addresses legal and ethical considerations. Legislation should be enacted to criminalize the creation and distribution of non-consensual AI-generated explicit content. Concurrently, ethical frameworks should guide the development[11] and deployment of AI technologies, showing respect for individual rights, fairness, and accountability.

International cooperation is also vital, as AI-generated content can easily cross borders. This will complicate enforcement efforts. Harmonizing standards across jurisdictions can improve the effectiveness of regulations and provide protection for individuals worldwide.

The Case for AI Regulations

Implementing regulations to govern these interactions offers several significant advantages:

1. Protection of Individual Rights

[11] Ahmad, S.F., Han, H., Alam, M.M. *et al.* Impact of artificial intelligence on human loss in decision making, laziness and safety in education. *Humanit Soc Sci Commun* 10, 311 (2023)

Regulations can safeguard individuals from content that infringes upon their rights and privacy. As an example, the creation and distribution of AI-generated explicit images without consent have become a pressing concern. In response, a few countries have enacted laws to protect their citizens from AI-generated sexual content.

2. Clear Accountability

Regulations define the responsibilities of AI developers, users, and other stakeholders. This helps in providing accountability for inappropriate interactions, identifying liable parties and enforcing penalties.

3. Prevention of Discrimination

Regulations can mandate fairness and non-discrimination measures in AI development and deployment.

4. Promotion of Ethical AI Development

Development of AI systems that adhere to ethical standards, will be encouraged. This will increase public trust in AI technologies. By setting guidelines for ethical AI behavior, regulations make sure that interactions align with societal norms.

5. Safe Reporting Mechanisms

Safe and effective channels will exist for reporting inappropriate human-AI interactions. [12]Nowadays AI-based moderation tools are being explored to address emergent harassment in virtual environments, granting opportunities to detect abusive behaviors proactively.

6. Adaptation to Advancements

As AI evolves, regulations can be updated to address new forms of misuse. The increasing prevalence of AI-generated inappropriate actions, has prompted legislative actions worldwide to criminalize the creation and distribution of such content[13]

Therefore, regulation is essential to protect individual rights, prevent discrimination, promote ethical development, and adapt to technological advancements. These regulations mitigate potential harms and pave the way for a more trustworthy society.

The Tough Road to Regulation

The rapid evolution of artificial intelligence (AI) has outpaced existing legislature. Thus, there are several key obstacles in effective regulation:

[12] Kelsea Schulenberg, Lingyuan Li, Guo Freeman, Samaneh Zamanifard, and Nathan J. McNeese, *Towards Leveraging AI-based Moderation to Address Emergent Harassment in Social Virtual Reality*, CHI Conference on Human Factors in Computing Systems (CHI '23), Article 514, Association for Computing Machinery, New York, NY, USA, 2, 1–17, 2023
[13] Jessy Lin, New directions for human-in-the-loop machine learning, Rethinking Human-AI Interaction, (30 Jan 2025,3:30PM), https://jessylin.com/2020/06/08/rethinking-human-ai-interaction/

1. Definitional Ambiguities

Traditional legal definitions of harassment and abuse are predicated on human-to-human interactions. The emergence of AI complicates these definitions, as AI entities lack consciousness and intent. Determining what constitutes inappropriate behavior towards or by an AI system requires a reevaluation of existing terminologies.

2. Jurisdictional Variances

AI technologies operate across global networks, but jurisdictions are confined by national or regional boundaries. The disparity here leads to inconsistent regulations, where an action deemed unlawful in one country may be permissible in another.

3. Technological Evolution Outpacing Legislation

The swift progression of AI capabilities often outstrips the legislative process. By the time a law is enacted, it may already be outdated due to new technological developments. This lag creates a regulatory gap, and leaves certain AI behaviors unaddressed.

4. Deterrence of Innovation

Overregulation may stifle innovation, and deter investment. Conversely, underregulation can lead to misuse and harm. Creating a balance between innovation and ethical use of AI is a delicate endeavor.

5. Enforcement Challenges

Even when regulations exist, enforcing them poses difficulties. Identifying perpetrators of AI misuse, mostly in anonymous or decentralized networks, is complex. Assigning liability whether to developers, users, or the AI systems themselves remains a contentious issue.

6. Ethical Considerations

AI systems can perpetuate or even increase existing biases, leading to discriminatory outcomes. AI interactions adhere to ethical standards involving addressing these biases, which are often deeply embedded in training data. Moreover, the creation of such content, like deepfakes, raises significant concerns.

7. Public Awareness and Education

There is a general lack of public understanding regarding the capabilities and limitations of AI. Educating the public about responsible AI interactions is crucial for effective regulation and prevention of inappropriate conduct.

Addressing these challenges requires a collaborative approach. It has to include policymakers, technologists, ethicists, and the public.

Understanding the Present Dilemma

The integration of artificial intelligence (AI) into society has led to advancements, but has also introduced challenges regarding sexual harassment and the misuse of AI-generated content. Currently, there is a troubling increase in the creation and distribution of non-consensual explicit material through AI, which raises questions about the adequacy of existing frameworks and the need for new regulations

Impact on Victims and Society

The proliferation of such content has great psychological and social impacts on victims. Individuals depicted in non-consensual explicit material often experience significant mental health issues, including anxiety, depression, and post-traumatic stress disorder.

[14] The violation of privacy and the loss of control over one's image can lead to long-term emotional distress and reputational damage.

To top it off, ease of creating and distributing such content perpetuates a culture of harassment and abuse. This harms individuals and reinforces harmful power imbalances, So it can be concurred that though recent legislative measures represent positive steps toward addressing the misuse of AI in this regard, more efforts are necessary to keep pace with technological advancements.

What Needs to Change?

To effectively address the challenges posed, a unidirectional approach wont suffice. The following recommendations aim to guide policymakers, organizations, and stakeholders in developing strategies to reduce risks associated with AI-related sexual harassment.

1. Establish Clear Legislation

- Legislation should explicitly prohibit the creation, distribution, and possession of AI-generated sexual or obscene material without consent.
- Clarify the responsibilities of developers, distributors, and users of AI technologies in cases of misuse. This includes accountability for those who create or disseminate harmful AI-generated content.

2. Development Practices

- Developers should integrate safety features into AI systems to prevent misuse. This may be done by using content moderation tools and restricting the generation of such material.
- Conduct regular audits of AI systems to identify and curb potential risks. Transparency in AI development processes can help in building trust and making sure that ethical standards are maintained.

[14] Han Li, Renwen Zhang, *Finding love in algorithms: deciphering the emotional contexts of close encounters with AI chatbots*, Vol 29, Issue 5 , JOURNAL OF COMPUTER MEDIATED COMMUNICATION , 1, 2-10, Dec 2023

3. Promote Public Awareness

- Launch initiatives to inform the public about the ethical use of AI and the legal consequences of creating or sharing such AI-generated content.
- Bring about programs to educate individuals, especially young people, about the risks associated with AI and how to protect themselves.

4. Collaboration Among Stakeholders

- Encourage AI companies to collaborate in developing best practices and sharing information on managing AI risks. Several companies have voluntarily pledged to promote the safe, secure, and transparent development of technology.
- Facilitate partnerships between government agencies and private companies to develop strategies for preventing AI misuse.

5. Strengthen Support Systems for Victims

- Demarcate clear and accessible channels for individuals to report instances of AI-related harassment or abuse.
- Provide resources and support for victims seeking legal recourse, or even counseling services and guidance.

6. Encourage International Cooperation

- Work towards reaching international standards for AI development and use with a unified approach to preventing misuse.
- Propel the sharing of information and best practices between countries to effectively address the global nature of AI-related challenges.

Final Verdict- Should Laws Regulate Inappropriate Human-AI Interactions?

The advent of artificial intelligence (AI) as a transformative technology has sparked numerous ethical debates, none more significant than those surrounding inappropriate human-AI interactions. AI systems are becoming increasingly realistic in everyday life, and the potential for misuse, including instances of sexual harassment and exploitation, has risen to the forefront of societal concerns. AI cannot "feel" harm in the way humans do, but the ramifications of unregulated interactions have ethical, psychological, and social repercussions. After examining the arguments for and against regulation, we arrive at a definitive conclusion:
laws must regulate inappropriate human-AI interactions to maintain ethical boundaries, protect human dignity, and decrease the likelihood of societal harm.

A possible solution would incorporate:

1. **Clear Definitions and Boundaries**
   Legislators must collaborate with technologists, ethicists, and sociologists to define inappropriate interactions. These definitions should account for societal values while allowing room for

technological innovation.

2. **Focus on AI's Societal Impact**
   Regulations should prioritize addressing the consequences of AI misuse on individuals and society, such as deepfake pornography or the reinforcement of harmful stereotypes, rather than dictating all aspects of human-AI interactions.

3. **Encouraging Ethics**
   Policies should incentivize developers to add safeguards into their systems, such as preventing explicit misuse and guaranteeing transparency in AI functionalities.

4. **Education**
   Public awareness campaigns can help individuals understand the impacts of wrongful human-AI interactions and promote responsible behavior without relying solely on legal measures.

Regulation of inappropriate human-AI interactions is also a moral imperative. AI itself cannot be harmed, but the ripple effects of unregulated misuse impact real people and society. Laws addressing these interactions would provide clarity, and prevent harm without stifling innovation. However, these regulations must be carefully crafted for still preserving the benefits of AI.

Hence, the rise of AI calls for proactive governance to solve upcoming challenges. By structuring thoughtful, well-defined regulations, society can harness the utility of AI while safeguarding its foundations. This approach makes sure that technology serves humanity responsibly. It lays the groundwork for a future where AI enhances, rather than compromises, values.