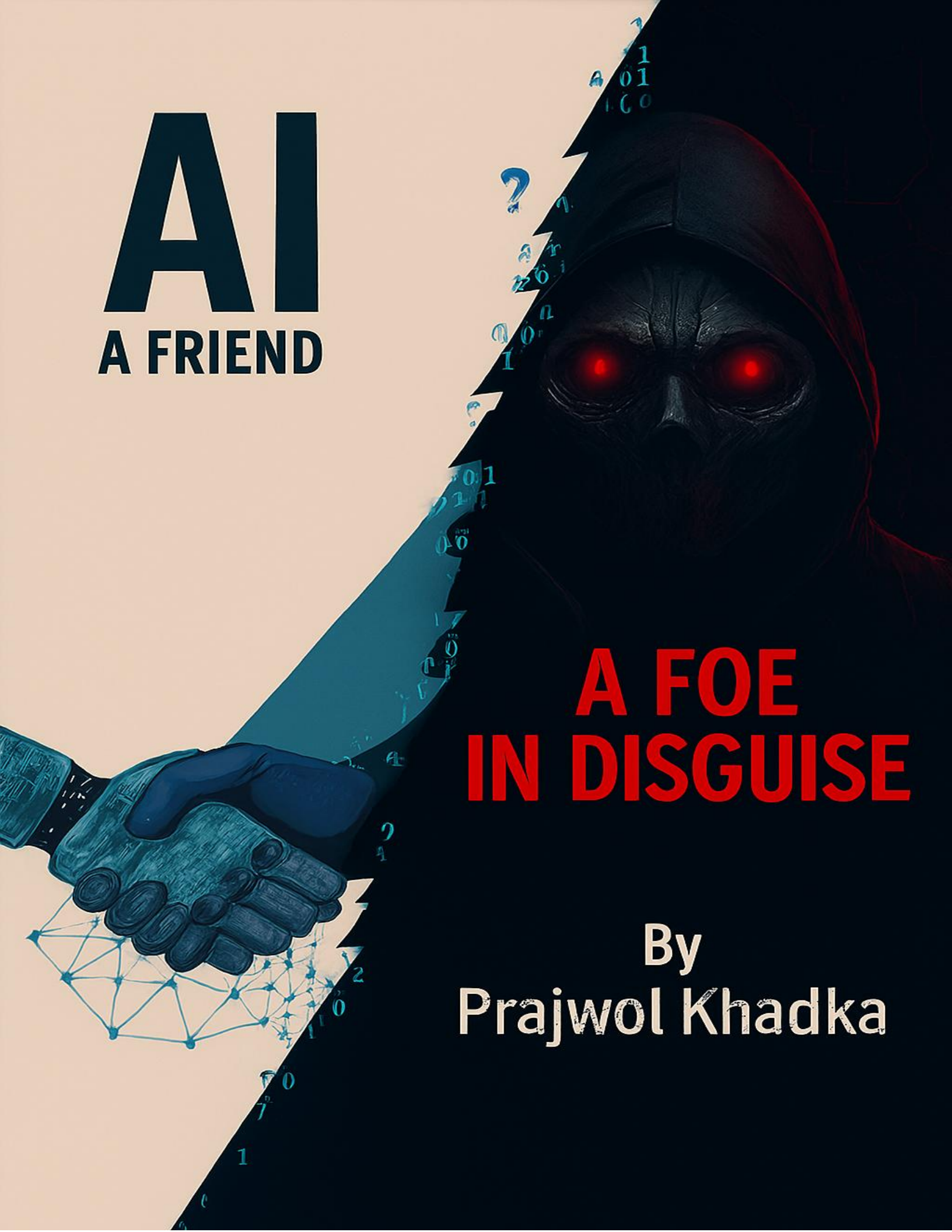
AI

A FRIEND



A FOE IN DISGUISE

By
Prajwol Khadka

Hello!

My name is **Prajwol Khadka** and I sincerely acknowledge my journey of learning and exploring Artificial Intelligence and software development. Through my academic pursuits at Softwarica College of IT & E-commerce, personal projects such as an AI-powered adaptive quiz system and VR Chemistry Lab, and my dedication to understanding ethical AI use, I have developed both technical skills and a passion for responsible technology.

This article and the accompanying workshop reflect my commitment to sharing knowledge and promoting awareness about AI, its benefits, and its ethical considerations among students.

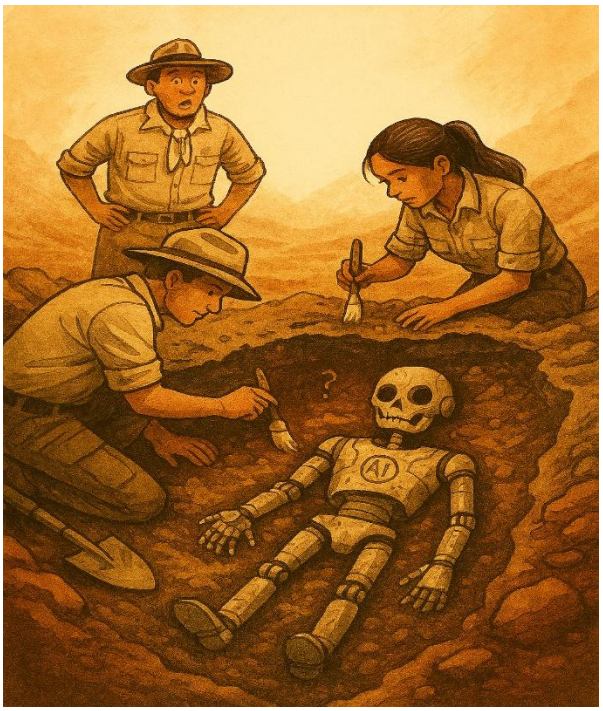
What's in this report?

This report is an explanatory summarization of my views, news articles, blog posts, public opinion articles, the ongoing trends, and factual source articles. This report is not to go against anyone's opinion or idea but a presentation of my views and research. Further, the report is an attempt to answer the following questions:

- What is AI (History, Current Trend, and Future Estimation)?
- How is AI helpful (Education, Productivity, and Society)?
- What are the risks of AI (Bias, Misinformation, Dependency, and Misuse)?
- Are we safe (Privacy Concerns)?
- What are the Dos and Don'ts (Ethics, and Fair use of AI)?
- What are the laws implemented by The Government of Nepal (Legalities, Offenses, and Consequences)?

What is AI?

Artificial Intelligence (AI), if we are to define Artificial Intelligence as the intelligence given to computers by humans (artificially) is correct on a surface level but might not give the idea we actually need. AI if have to be defined should be defined as a technology that enables computers and machines to simulate human learning, comprehension, problem solving, decision making, creativity and autonomy.



The history of Artificial Intelligence (AI) is much older than we realize, AI dates back to the early 20th century. It began in the early 1900s with early machine like Leonardo Torres y Quevedo's chess-playing machine which could play simple chess game automatically. Moving forward to 1921, the word "Robot" was introduced which also inspired the idea of machines that could perform tasks, later the first digital computer was introduced (Atanasoff-Berry Computer) in 1939 that used binary data and electronic circuits. During the 1940s and 1950s, scientists like Alan Turing and John McCarthy developed foundational ideas for AI further, in the 1960s and early 1970s, programs such as ELIZA, DENDRAL, and robots like Shakey demonstrated problem solving, reasoning, and natural language understanding but quickly fell cold with rising criticism and funding cuts. Luckily, in the 2000s AI advanced rapidly thanks to powerful computers, large volume of data, and better algorithms. Early AI machines (robots) like Honda's ASIMO robot could walk, run, and interact with humans. Later in the year 2016 robot Sophia caught widespread attention for its human-like appearance, ability to mimic some human expressions, and conversational abilities, it set an example of how AI can simulate human interaction. Simultaneously, virtual assistants like Siri (iPhone), Alexa (Amazon), Google Assistant became a part of everyday life by using AI to understand and respond to human language. Further, in the year 2018 OpenAI released GPT (Generative Pre-trained Transformer) models which could generate human like text, answer questions, and even create stories, write codes, write poems and lyrics. As the time went by OpenAI released more versions of GPT (GPT-3, GPT-4) which demonstrated advanced reasoning, creativity, and language understanding, making AI widely accessible for learning, productivity, and entertainment. Similarly, other companies like Meta, Google, Microsoft, etc. introduced their own AI models, each with its own unique value. Today, AI is steadily growing in almost every field ranging from healthcare and education to entertainment, finance, and robotics – showing both the potential and ethical challenges of machine intelligence in our daily lives (Mucci, 2025).

How is AI helpful?

AI just like anyother technologies, was invented to enhance human productivity. It acts as a smart assistant that helps us complete task faster, more accurately, and efficiently. AI can process large amounts of data in very less of a time, predict and provide insights, and automate repetitive tasks. Virtual Assistants, AI tools for education, healthcare, entertainment, and creativity, AI used in autonomous driving – supports humans by reducing effort, saving time, and improving decision-making.

AI can generate codes, write stories and essays, generate images and videos, and AI can also generate music – helping humans in many sectors like education, healthcare, productivity, and society. Students are no strangers to the use of AI in completing their assignments or to seek help for their queires, AI generates answer to anything without judging the user unlike a human might increasing comfort in the use of AI. Also, AI is the ultimate shortcut to solve any problem as they reduce the effort drastically to solve a problem. Further, AI helps writers, teachers, speakers, and even politicians generate or polish their literature with better grammer, correcting flaws, and providing ideas for better delivery.

Additionally, AI tools in law and enforcement helps with facial recognition and predictive analytics help in criminal activity investigation to help case solving be more efficient while ethical use of the tool is important in order to protect one's privacy. Overall, AI is an integral part of our society – touching nearly every aspect of life. While challenges and learning experience may arise, AI is expected to have a generally positive impact – helping society work smarter, safer, and more efficiently.

What are the risks of AI?

As we discussed in the previous pages, AI is just like any other technologies, they are there to help us work smartly, and more efficiently. So, the big question is:

WILL AI TAKE OVER HUMANITY?!

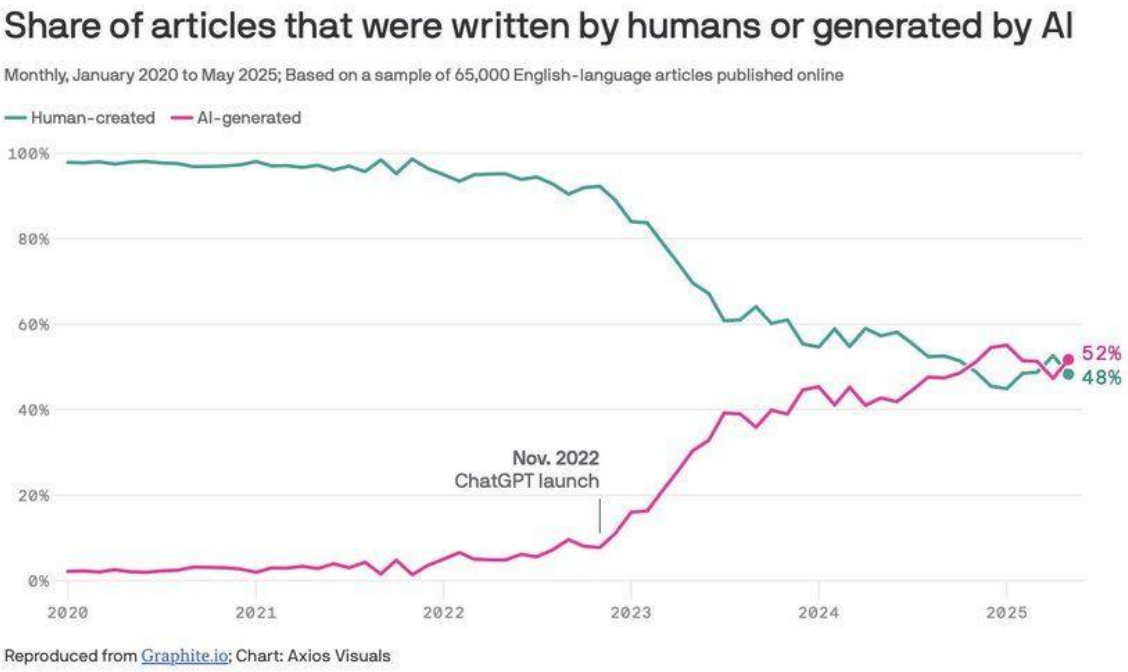
No! AI will not take over humanity. The world is yet very far away from attaining the superhero fantasy movie level AI threat – still very far from developing and deploying AI systems that possess emotions, consciousness, or independent thinking ability or motives like humans do. In very simple terms, Marvel's Ultron is not happening any time soon. Today's AI works strictly within the limits of human programming and data and it cannot act beyond its designed purpose.

Once again, AI mimics Human like intelligence and thinking but it cannot replicate the way we actually think and work. So, we can be assured that AI will not have us hostage because we bullied it at 3am asking to fix a bug in our code. That said, AI does come with real risks that must be addressed. AI can hallucinate, it can generate false or misleading information presented as fact, as it tries to fill in the gaps or misinterpreting patterns in its training data meaning it can make incorrect predictions, false positives (identify something as being a threat while it is not), false negatives (fail to identify something as being threat when it is).(*What Are AI Hallucinations?*, n.d.)

AI are designed, developed and trained by humans, and humans are biased naturally. We have our opinions and views, bias is disproportionate weight in favor of or against idea or thing usually such that it is inaccurate, closed-minded, prejudicial, or unfair. For example, assuming men are better at math or engineering than women, expecting women to handle all domestics tasks, judging someone with their skin color, only seeking information that supports your beliefs, etc.

When it comes to AI, bias in AI is not merely a technical issue but a societal change. AI systems are increasingly integrated into decision-making processes in healthcare, hiring, law enforcement, and other critical areas. Therefore, it is very important to address biasness as it is not only about improving technology but also fostering ethical responsibility and social justice (*Bias in AI* n.d.). Some examples of biasness in AI are, AI recruiting tools favoring male candidates over female candidates as it was trained on historical hiring data from companies where men were predominantly hired, facial recognition system struggling with darker skin tone as it was trained with fairer skin tones, etc.

The trend of humans depending on AI is steadily growing and it is very concerning. Humans are said to have superior intelligence due to our ability to think rationally, we have creativity that an AI cannot replicate. Students, writers, journalists, and content creators are seen to be depending too much on AI – which will contribute negatively to their own thinking and problem-solving ability.



This image illustrates a concerning trend, in 2020, only about 5% of online content was AI-generated. By May 2025, that number reached 48%, and projections show it could exceed 90% next year. The reason is simple as AI generated articles are much cheaper than the human generated articles with AI costing about 1 cent and human written one costs $10 to $100. This might seem fine as it is just saving money but the real challenge is not the price. When AI systems learn from AI-generated material, quality deteriorates. It's like making a photocopy of a photocopy. The details fade, originality disappears, and everything starts to sound the same.

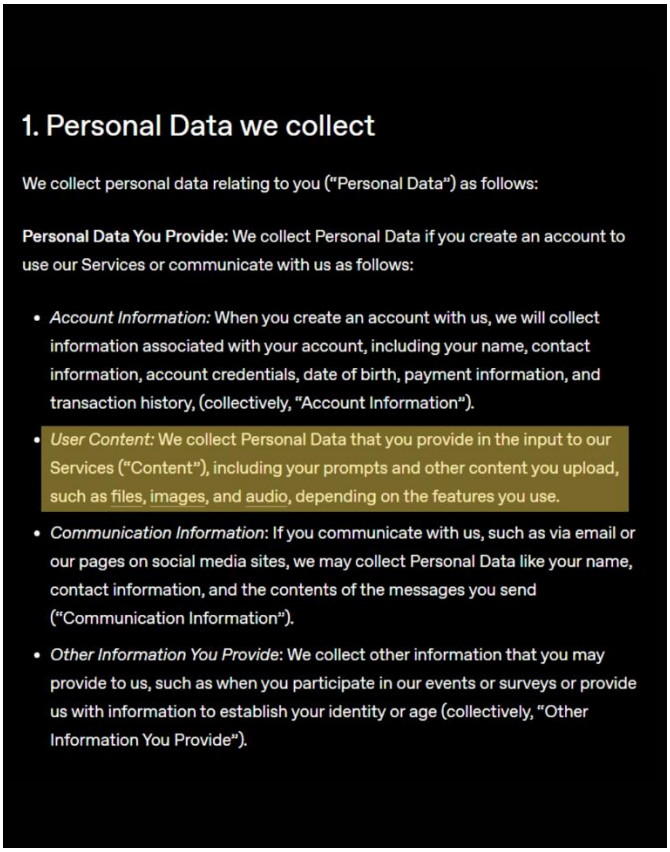
Popular Indian actress Rashmika Mandanna was once a victim of AI misuse, in 2023 a 24-year-old male from India made an AI generated deepfake video of the actress was being surfaced all over social media platforms (Tribune News Service, 2024). The video was made and shared online with the intention to defame the actress; the accused was arrested shortly and now faces charges as per the decision of Indian Court. This is just one example of misuse of AI, the technology can be and is being misused in various ways from cheating in examinations, generating explicit images, carelessness in terms of factual accuracy and many more. AI should be used to make our life easier and not more difficult than it already is.

Are we safe?

"*If you are getting something for free, you are the product"*

Tech giants like Meta, Google, Amazon, Microsoft, Tiktok, etc. are some popular services we use. Meta charges Rs. 0 from us to upload photos or videos on Facebook or Instagram, make high quality video calls on Messenger, Instagram, and Whatsapp. Google charges Rs. 0 for using its services until we decide to upgrade, Amazon charges nothing except for purchases, and Tiktok charges nothing. All these services have absolutly no reason to be allowing us to use their services for free but they do allow us to use it as they generate revenue from the user as they use their services. For example, Meta allows advertisers to pay to access your behavior and attention to target Ads based on personal data. Which means, we are the product and not the customer.

Unlike these big tech companies AI has a different approach of generating money and playing around with out data. OpenAI for example mainly generates its revenue from model subscription and API usage and not data selling but it does make use of user data, it the data from your conversation to improve the way it solves problem for you.



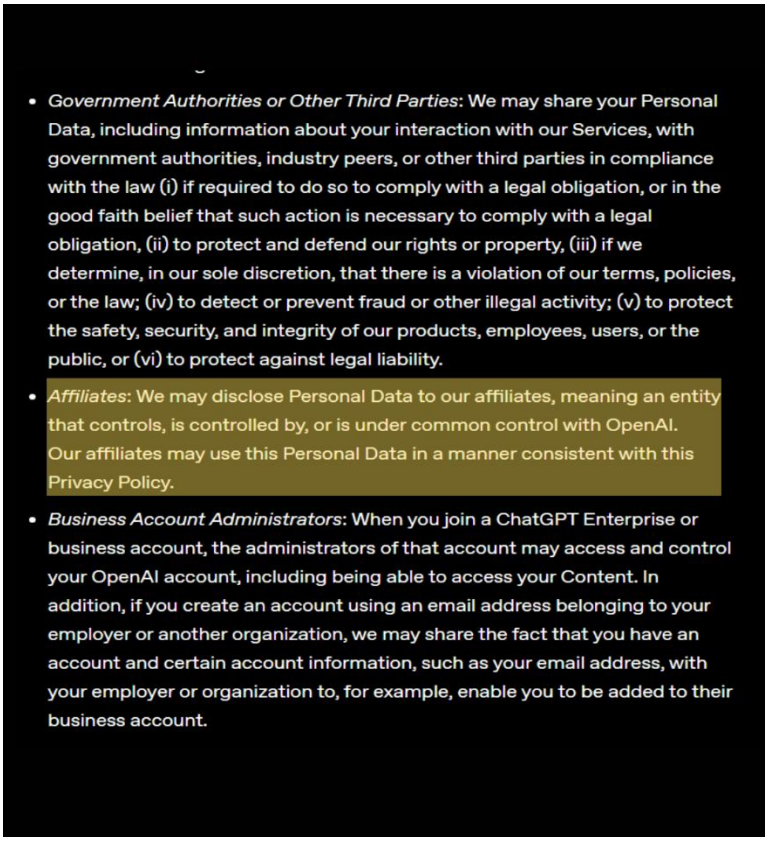
1. Personal Data we collect

We collect personal data relating to you ("Personal Data") as follows:

Personal Data You Provide: We collect Personal Data if you create an account to use our Services or communicate with us as follows:

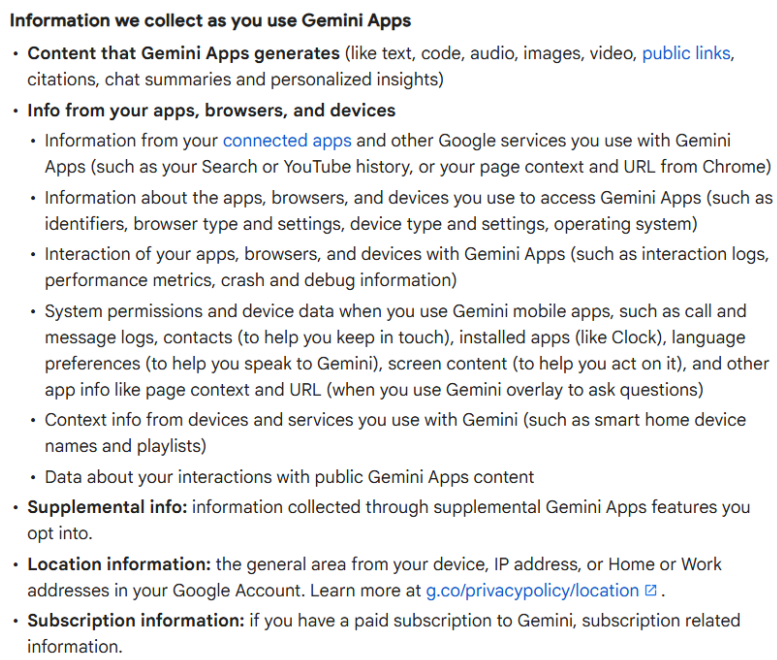
- *Account Information:* When you create an account with us, we will collect information associated with your account, including your name, contact information, account credentials, date of birth, payment information, and transaction history, (collectively, "Account Information").
- *User Content:* We collect Personal Data that you provide in the input to our Services ("Content"), including your prompts and other content you upload, such as files, images, and audio, depending on the features you use.
- *Communication Information*: If you communicate with us, such as via email or our pages on social media sites, we may collect Personal Data like your name, contact information, and the contents of the messages you send ("Communication Information").
- *Other Information You Provide*: We collect other information that you may provide to us, such as when you participate in our events or surveys or provide us with information to establish your identity or age (collectively, "Other Information You Provide").

This is a part of OpenAI's privacy policy (the same thing we never read while signing up to any services). This part says that the user's data are collected via account information, prompts, images, files, audio. These data are not sold outside unlike meta does but these are not any better as well, OpenAI allowes OpenAI's trained staff to review some conversations to improve model accuracy or saftey. The company claims that these information are anonyomous but still sharing personal data like address, phone, password, ID number, sensetive photos or files should be avoided as far as possible.



- *Government Authorities or Other Third Parties*: We may share your Personal Data, including information about your interaction with our Services, with government authorities, industry peers, or other third parties in compliance with the law (i) if required to do so to comply with a legal obligation, or in the good faith belief that such action is necessary to comply with a legal obligation, (ii) to protect and defend our rights or property, (iii) if we determine, in our sole discretion, that there is a violation of our terms, policies, or the law; (iv) to detect or prevent fraud or other illegal activity; (v) to protect the safety, security, and integrity of our products, employees, users, or the public, or (vi) to protect against legal liability.
- *Affiliates*: We may disclose Personal Data to our affiliates, meaning an entity that controls, is controlled by, or is under common control with OpenAI. Our affiliates may use this Personal Data in a manner consistent with this Privacy Policy.
- *Business Account Administrators*: When you join a ChatGPT Enterprise or business account, the administrators of that account may access and control your OpenAI account, including being able to access your Content. In addition, if you create an account using an email address belonging to your employer or another organization, we may share the fact that you have an account and certain account information, such as your email address, with your employer or organization to, for example, enable you to be added to their business account.

Affiliates – companies that are legally related to OpenAI, the company may share your personal data with these affiliated companies but only in the same ways OpenAI itself can. But the actual concern here is would you like your data to be used like this even though it is anonymous, it is to provide services, improve models, or for security purposes.



Information we collect as you use Gemini Apps

- **Content that Gemini Apps generates** (like text, code, audio, images, video, public links, citations, chat summaries and personalized insights)
- **Info from your apps, browsers, and devices**
  - Information from your connected apps and other Google services you use with Gemini Apps (such as your Search or YouTube history, or your page context and URL from Chrome)
  - Information about the apps, browsers, and devices you use to access Gemini Apps (such as identifiers, browser type and settings, device type and settings, operating system)
  - Interaction of your apps, browsers, and devices with Gemini Apps (such as interaction logs, performance metrics, crash and debug information)
  - System permissions and device data when you use Gemini mobile apps, such as call and message logs, contacts (to help you keep in touch), installed apps (like Clock), language preferences (to help you speak to Gemini), screen content (to help you act on it), and other app info like page context and URL (when you use Gemini overlay to ask questions)
  - Context info from devices and services you use with Gemini (such as smart home device names and playlists)
  - Data about your interactions with public Gemini Apps content
- **Supplemental info:** information collected through supplemental Gemini Apps features you opt into.
- **Location information:** the general area from your device, IP address, or Home or Work addresses in your Google Account. Learn more at g.co/privacypolicy/location .
- **Subscription information:** if you have a paid subscription to Gemini, subscription related information.

Gemini as an another example sets a perfect example to not trust AI as your personal councellor, therapist, or editor. Just like OpenAI, Google's Gemini collects information from user like text, code, audio, images, video, public links, citations, chat sumamries, and personalized insights. It also observes how users interact with the apps, devices, and browsers, collecting system information such as device type, browser, operating system, settings, performance logs, and crash data. Additionally, Gemini gathers context from connected Google services, location data, and subscription information. This data is used to provide, maintain, and improve Gemini's services, personalize and customize user experiences, communicate with users, measure performance, and protect users and the system.

So the big question here is are these aesthetic “Saree” edits of yours actually safe and are these images misuse proof?

Reference

Mucci, T. (2025, October 10). History of artificial intelligence. *The history of artificial intelligence*. https://www.ibm.com/think/topics/history-of-artificial-intelligence

What are AI hallucinations? (n.d.). Google Cloud. https://cloud.google.com/discover/what-are-ai-hallucinations

Bias in AI. (n.d.). Chapman University. https://www.chapman.edu/ai/bias-in-ai.aspx

Tribune News Service. (2024, January 21). Delhi Police arrest Rashmika deepfake video creator in AP - The Tribune. *The Tribune*. https://www.tribuneindia.com/news/delhi/delhi-police-arrest-rashmika-deepfake-video-creator-in-ap-583115/