Keep your personal data secure from AI

Rajagopal Tampi

Introduction

That AI is seeping into all areas of our life is well known. I had written a paper on the 'Why and how to control AI' . This is a sequel to that article with a different focus. If you decide to follow these words, you are choosing a path balanced between adoption of productivity enhancement that AI offers and your own security. As with all things there is a cost associated with securing yourself. In the long term it should turn out to be a wise decision because regulation on AI and data security is woefully lagging. When the law does catch up, those who have been careful will have paths open to them which will not be available to others.

One cannot be blamed for being fearful and thinking that AI will harm us using our personal data. While that is a possibility, it is not as urgent as humans themselves who will exploit AI by using personal human data for extortion, obfuscation, creating divisions, and other harmful purposes against us. Humans will use AI to enhance their abilities to achieve intended purposes which will include nefarious ones. How will training of AI on our most personal and intimate data harm us? Which is the intimate personal data that we should not share with AI?

Human Uniqueness Data (HUD):

We know that every human being is unique. As far as AI is concerned, our personal data is what differentiates us and makes us unique humans, Human Uniqueness Data (HUD) is what we must protect from AI. HUD can be defined as comprising of Personal Profile Data (PPD)[1], Personal Real-Time Data (PRTD)[2], Experience data, Karma Quadrant© Data and our preferences (Fig1). There are great dangers in allowing AI to train on uniqueness data representing individual human beings. These dangers include:
1. Human behaviour in any circumstance can then be predicted by the AI.
2. Psychological profiles can be used to plan tactical and strategic opposition to individual human's plans, choices and actions.
3. Weakness of individual humans can be predicted by AI and exploited.
4. Strengths can be predicted by AI and eroded by counter actions recommended by the AI.
5. Normal human societal interactions can be jeopardised by using AI generated knowledge of society members to intentionally vitiate, create divisions and communal disharmony.
6. Such uniqueness data will be sold by unscrupulous vendors to terrorists, kidnappers, hackers, political strategists, intelligence agencies and more just as financial details and account passwords are peddled on the dark net.

Why HUD should not be shared

The training of AI on HUD will result in the AI being able to predict an individual human's behaviour in any circumstance. This is not a desired goal for the reasons to be explained below. The comparison of the human solution for a problem against an *original* AI solution for the same problem may be done by individual humans to decide which is the superior solution to adopt.

The ethically correct design of AI should generate its original (own) best in class solution to any problem without training on HUD. The industry should be constrained to find innovations to make this possible.

AI training on HUD data of an individual serves only one end *i.e. to counter/thwart/ outmanoeuvre the individual for some purpose at some chosen time by someone who stands to benefit from it.* The need for use of HUD to train AI in the name of technology advancement, productivity increase and innovation is being advanced by industry. This is a patently ***duplicitous explanation***. The fact is gullible public are giving their personal data for their own commercial exploitation by industry collusion between vendors and other disparate buyers. The process has been started covertly by technology companies. A natural outcome of sharing HUD is the empowerment of unethical employers and other such people on whom the common human is helplessly dependant daily. HUD of personalities/individuals can be sold on the darknet to thieves, anti-social elements and enemies of individuals or nation states for the exploitation of humans.

Sensible people should opt not to offer the industry a chance to use their own HUD for training AI. Be assured that every individual in a democracy has the right to protect their own HUD. This is the right time for making that decision as a global human collective. Here are some important ways in which we can prevent the sharing of our own HUD with any technology company to build and train AI products and services. Regulation needs to be legislated to prevent collection of HUD by companies.

Understanding Human Uniqueness Data (HUD)

Fig 1 below shows the Human Mind model from a Data perspective. Each human mind is unique and therefore the HUD is also unique resulting in the uniqueness of character of each human being. The ‘Karma Quadrant©[3]’ represents the innermost mind related characteristics of the human mind. It is represented by the four quadrants of thoughts, values, judgements and attitude (nature). For details of these aspects please refer to my paper ibid. The other elements are static and real-time personal data (PPD© and PRTD© respectively) and our preferences. These items together constitute the data perspective of the human mind model.

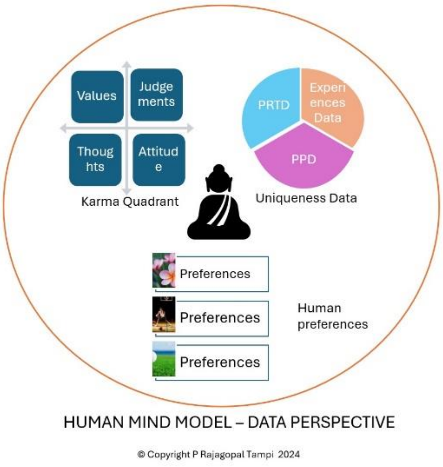

Fig 1: Human Mind Model – A Data Perspective

Steps to safeguard ourselves

Journal applications

These days some companies offer Journal apps for writing electronic journals which will be stored on the cloud. In journals we capture our daily thoughts, activities, decision making, conflict resolution, logic, reasoning, imagining and conceptualisation experience data. They are our most confidential data which reveals the functioning of our minds and our innermost natures. This constitutes the Karmic Quadrant©[3] and uniqueness data explained above. If the data in the journal is captured, it is likely to find its way as data for training AI in higher order capabilities relating to the human mind which are currently not possible in AI. One of the harmful applications of this data can be to counter/thwart the person whose data it is. Therefore, it is advisable not to use journal applications to store such confidential data.

LLMs and AI Image Generators

Do not offer your HUD while seeking any assistance from LLMs through prompts or by way of explaining specific situational challenges to the LLM. Queries which impinge on personal and situational aspects should be addressed in a generic manner in the prompts that are created.

An indicative list of information which should not be shared with LLMs is given below. This may not be an exhaustive list, but it gives the reader a fair idea of the type of data which should not be shared with LLMs.

Any matter regarding your priority management. Any matter regarding your value systems and choices made. Reasons for making choices in any circumstance. Courses of action in the case of work matters, courses of action in the case of personal matters. Knowledge boundaries, capability boundaries, emotional strengths and weaknesses, personal qualities like honesty, patience, affability, selfishness, kindness, love, nature, breakdown of emotional control and reasoning, logical processes, incidents during childhood, early life and growing up experiences, medical histories, social nature, capability for reasoning, the cultural, preferences and nature endowed traits of each human being which makes us unique individuals. Psychiatric, psychometric, psychotherapy, psychology, medical data including our relationships with people. Learned experiences, attitudes, thoughts capture– preferences, analysis, plans, actions and openness to feedback/learning receptivity. These constitute the Karmic Quadrant©[3], Uniqueness and Preferences data.

Use of personal Agents

It is a great idea to use AI agents for our convenience and productivity improvement. The precautions to take are to avoid the sharing of Karmic Quadrant©[3], uniqueness and preferences data as explained in the case of LLMs above.

Social media applications

It is advisable not to post photographs of yourself and your family on these applications along with explanation of the occasions. One disconcerting application with the emergence of AI, is the linking up the information to create a profile of your social circle and the influence wielded by each one in your network. This information has value for your competitors, detractors and fraudsters. Do not express political views on social media as Government agencies are tracking such comments. Social media statements and posts can be used as evidence in courts subject to authenticity verification and certain other factors. Due to the massive speed and computing power growth of AI, tracking and actions can be taken *in real time in important cases*.

Web-site Marketing feedback using cookies

If there is a choice to reject all cookies, then that option should be chosen. If not, it is advisable to enable only 'essential' or 'legitimate interest' cookies only. Do not allow tracking of sites visited and sharing of your data with other vendors. The paragraph below illustrates one possible dangerous situation created by sharing of one's location information while accepting all cookies option.

Let us say a person while responding to cookies notice from a malicious website, leaves it to the discretion of the AI system by accepting "ALL COOKIES" option. That is, she does not exercise her choice and leaves it to the AI system to exercise discretion for her. Such a situation .. could become the object of bias or personal harm or threats to the person's life. Such a situation is perceivable when the malicious website sells the persons GPS location information as a stream to terrorists or criminals who may want to kill or rob the person[3].

Federated ID logins

Federated ID logins for websites and other applications should best be avoided. Most of the smaller app creators choose to offer federated id logins using the big players email ids like Google, Meta and Apple logins. This is offered as a convenient way to quickly create an account. The cost of this convenience is the risk of loss of your data to the very same big players who are not the vendor of the service or the application you are using. Providing access to your work or personal data to vendors other than the direct product vendor compromises your own security and is best avoided.

Passwords management

Avoid using browser stored passwords. The best option is to use a password management solution with 2 factor authentication. For bank and other important passwords create passwords manually.

Location Information

For our own security, it is necessary to be aware of the importance of our GPS/location information. This is generated by mobile phones, wi-fi stations and tracking drones including IoT devices and indoor positioning systems. Browsers can generate our location using wi-fi points, html – 5 geo-location, IP address and other methods. Lax web-browser settings are an important loophole used by service providers for collecting information about individuals and their preferences. It will be time well spent to choose the desired settings in your web browser.

Conclusion

The value of our personal data has shot up in the age of AI since it exposes our nature and character which can be used as training data for AI analytical models. These human aspects will form the next set of goals for AI innovations in its search for Artificial General Intelligence. The eight steps outlined above are a balanced and safe way to use AI ensuring the harnessing of both productivity gains while being safe. Big Tech companies should innovate other techniques to train AI in the higher orders of intelligence rather than training AI on UHD. Regulations should be brought out to prevent technology companies from training AI on UHD.

References/Notes:

1. Personal Profile Data (PPD): Personal data comprises personal profiling data including psychological and psycho-metric profiles, value systems, experiences, beliefs, personal biometric data, emotional states, iris scans and medical data etc.

2. Personal Real-Time Data (PRTD): Data of how daily life situations are handled or managed by individuals and their ways of thinking, imagining, conceptualization, reasoning logic, emotions and other such mind related data.

3. 'Karma Quadrant' © Copyright P Rajagopal Tampi 2024.

4. Applied Human-Centric AI, *Clarity in AI Analysis and Design,* Rajagopal Tampi

About the Author

https://aipathfinder.org

Disclaimer:

This paper is forward looking and represents the author's own views. The paper does not try to cast aspersions on any company, the IT industry or national governments.