A practical solution for the conundrum of AI and personal data

Introduction

This paper assumes that the reader has sufficient knowledge of data related problems which the world is facing due to the emergence of AI/ML. An acceptance of the need for data control particularly for personal data for AI/ML applications is also assumed. The paper provides a definitive solution for personal data protection from AI for consideration by national regulators and technology companies. The words “tech titans” is used in this article to refer to technology giants such as Google, Apple, Microsoft, Meta, Amazon, Open AI et cetera, who control the AI markets and are responsible for setting directions for AI technological progress.

The problem statement

There are multiple levels of control of data management for AI/ML applications. This categorization of data is based on data ownership. These levels of control are predicated upon data storage locations, data processing and data interchange interfaces. The classification of data consists of (from the highest level of control to the lowest) personal data, corporate data and government data, private data and public data. Corporate and Government data are at about the same level of control. Private data may be any data that is declared explicitly as ‘private’ and is not part of any of the higher levels of data.

There are other forms of classifications of data from the Government Regulation point of view which are not relevant to this discussion. These classifications may be found in Chapter 13, Fig 1 dealing with AI and Data Management Governance Tasks.[1]

One alarming case of how personal data is being captured is through Active Listening[2] used by social media and tech titans. They use AI to eavesdrop on our telephonic conversations to create revenue streams through marketing opportunities for themselves at the cost of capturing real-time user intentions and creating personal preferences and profiles of users without their knowledge and consent. Combine this with AI assistants being developed in every app, product, and website and the lives and psychological profiles of each human being is an open book which is being monitored instant by instant for commercial profit. AI/ML related data captured for all these purposes unbeknown to gullible users resides on the cloud servers belonging to the

tech titans. People and their personal data only represent revenue generating opportunities for these companies. This open scorn, one-sided and unethical practice being inflicted on people by tech titans must stop.

The elephant in the AI room is where the data is stored and processed. Unfortunately, this elephant has been manoeuvred stealthily into the preferred location to solely serve up revenues to technology companies at the cost of people's free will, freedom of choice, privacy and liberty despite the latter being our most precious rights and life possessions.

With the progress of the IT industry since the mid 1960's we have witnessed the arrival and departure of many technologies starting with mainframes to minicomputers to personal computers to client-server computing, distributed processing to server-based processing /ASP.net to the mobile and internet to private data centres to virtual servers and virtual operating systems to the cloud. Data which was managed privately on owned or leased data centres of businesses was migrated to the tech titan owned clouds and data centres. For past decade and more, data migration formed unending and growing waves of revenues for technology companies. The trend for moving data to the cloud was embraced by customers because of the benefits of managed services offered alongside the cloud migration such as database management, automated monitoring tools, virtual servers, managed backups and security offered by the technology companies and Data centre service providers.

It becomes very difficult to be able to track, control and audit data manipulation (of all data including personal data) and its usage once the data is uploaded on the cloud. Even if data tracking audit reports are mandated by law, these can be manipulated by the service provider unless the watermarking technology is completely foolproof. Tech titan cloud service providers are themselves responsible for developing and implementing water marking technologies leaving a loophole that is practically impossible to fill.

The solution for AI and personal data

A logical approach to avoid data manipulation of personal data is to mandate the storage and processing of such data on the individual's own device itself and ensuring that the data does not ever leave the device. The backups and restore can be done by the individual himself on his own offline media. This is the simplest and the most effective solution for safeguarding personal data, personal profiles, psychology and intent revealing thoughts and preferences of humans from being exploited for

commercial purposes. It is indeed the fool proof and universally accepted method where the security of personal valuables of an individual owner are stored in the owner's locker with the key held by the owner.

Tech titans developing the AI products need access to the data only for enabling the computing of AI functions of the product. They do not need possession or retention of the personal data at all. There are technologies existing since the last many decades to achieve this end. And this is where tech titans should adapt their AI systems design and architecture to support the larger good of the right to free-will, freedom to choose, privacy, security and safety of human beings.

Architecture for the AI/ML processing of personal data

The personal data storage and AI processing will happen only on the user's device. The user's device will now perform the AI/ML computational functions hitherto performed by servers located in the cloud. The cloud program will be modified to a thin client which will only check whether the accessing device and API is licensed, personal device configuration supports the AI vendor's product or service environment requirements, the product is permitted to run and provide the necessary IP clearance for execution. Thus, all AI functions using personal data will run on the user's personal device itself.

AI functions which do not process personal data will continue to run on the cloud-based servers with the client running on the mobile or other personal device unchanged. Interfaces with personal devices should strictly take place through auditable and trackable APIs only.

Personal AI Data (PAID) standards and architecture

- Develop new AI architectures and components to run personal AI apps only on the user's handheld devices and laptops and store the data only on the user device with offline backups done by the user.
- Develop new powerful and miniaturised processor chips and AI/ML models for running AI efficiently on personal devices.
- Develop new methods and processes for bug fixing data to be analysed on the user's own device and use standardised API's to transfer reports to the cloud. Alternatively, simulate the bugs with synthetic data on the cloud servers for developing a fix.
- Personal data should not be taken online to the cloud for processing or for any other reason.

- PAID to be developed with human security, free will, free choice and prevention of access to personal data as the primary considerations.
- All AI training for personal applications will be done on the vendors servers on the cloud using synthetic data only. In the rarest of the rare cases depending on necessity, AI training may be done on anonymised personal data from a voluntary test group with case-by-case approval of national governments.
- PAID architecture will support Single Point of Accountability and Responsibility principle (SPAR). In the sense that if multiple vendors are involved in providing any AI functionality on a personal device, the method for integration between the functionality supplied by each vendor should be through APIs which are designed to support the SPAR principle so that accountability for failures can be established with full clarity and certainty based on audit trails.
- User activation of personal database and personal AI use cases: PAID architecture stresses on user (himself) activating every use case dealing with personal data on any of his personal devices. This works based on all personal AI functionality being switched off in the default factory settings of a new device at the time of purchase. The releases and updates and maintenance of personal AI products section in this paper explains this feature in greater detail.

Developing Data Expertise

1. The set of personal data which needs to be banned by law for collection or storage for AI training should be delineated.
2. Develop miniaturised, specialised Personal Data (PD) tracking chips for use on personal devices.
3. Develop advanced Data tools for data auditing, tracking and watermarking.
4. Develop Data that vanishes after configurable time setting and prevents copies from being made in case of personal data.
5. The set of personal use-cases which need to be categorised under PAID be identified and delineated.

Releases, updates and maintenance of personal AI products

AI product design should incorporate AI features in mobile phones, tables, laptops and other personal and home devices which are by default switched off at initial purchase. These functions should be activated only by the owner by actively choosing the AI feature and using their own credentials for activation. The activation will be logged. This will ensure users have a choice of activating AI features of their own free will and assume the risks associated with those features. Users have the choice to exercise their free will. This means that users may decide not to activate any or all personal database, AI use cases that their device offers to them. This approach is the opposite of features which come preloaded and fully activated today on phones and laptops and tablets at

the time of initial purchase. Suitable strategies for updates and maintenance support should be innovated by tech titans in compliance with the PAID architecture principles.

Conclusion

AI self-regulation by technology companies has not been an effective solution. The cases of fines imposed on Meta[3] ($1.3 bn), Amazon[3] ($781mn), Google[4] ($270mn) are but three cases in point where self-regulation have failed. It may be challenging and more costly for to tech titans develop PAID architectures, yet it is precisely that which represents the cardinal need and cost of upholding personal free-will, free choice, privacy, security and safety of human beings.

National Governments should study the PAID architectural proposal for AI processing personal data suggested in this paper and initiate discussions with tech titans. In the interest of preserving the free will, choices, rights, safety, and security of its citizens, Governments should convert these recommendations into mandatory AI/ML laws for compliance by tech titans and the industry.

Please sign the petition to build Government support for introducing appropriate AI Data Regulations at https://www.change.org/personal_data_protection

References:

1. "Applied Human-Centric AI" by Rajagopal Tampi.
2. Is your phone listening to your conversations? Yes, and there is finally proof: Story in 5 points (msn.com)
3. https://dataprivacymanager.net/5-biggest-gdpr-fines-so-far-2020/
4. https://www.nytimes.com/2024/03/20/business/france-google-fine.html

Rajagopal Tampi

The author is an M. Tech (Computer Sc) from IIT Bombay. He is an entrepreneur and pathfinder for AI/ML. He is the author of "Applied Human-Centric AI", a book in which he has explained how to overcome the major problems posed by AI using the 'AIDP Model' of analysis and design. His opinions and writings on AI can be found at HTTPS://aipathfinder.org