

# Personal Data Center

Haolin Chen

November 2020

"Data scientists are involved with gathering data, massaging it into a tractable form, making it tell its story, and presenting that story to others" – Mike Loukides, VP, O'Reilly Media. Why not let data present that story to the individuals?

I am a data scientist intern at DiDi, which is the world's leading multi-modal transportation platform, provides app-based transportation services, and also is the biggest competitor to Uber in the international market. My daily work is analysing data from the driver side to help business activities by providing vital data insight. During daily work, I know how business companies are good at using data to precisely improve profit and normative driver's behaviour. DiDi and almost all the business companies collect and analyse users' data to optimize platform operation. They know you from different views, model your user portrait. They even know you better than yourself. At here, I need to thank General Data Protection Regulation (GDPR), which regular the data processing flow, making those companies stop invading user privacy for no reason. However, the law is the lowest standard for ethics. Meanwhile, I was thinking, is there any company established to protect users' data? Let data be controlled by the user and served for the user. Nowadays, the business structure is that data collected, analysed, and output result by different companies, which will lead to seriously personally sensitive information leak and the users cannot enjoy the convenience from data directly. What I want to do is let data serve the user, let users be their own data's controller and processor, while restricting companies' rights to hold sensitive personal data and avoiding advertising boom and scam mail. It is time for everyone to have a Personal Data Center.

I will address Personal Data Center in three parts:

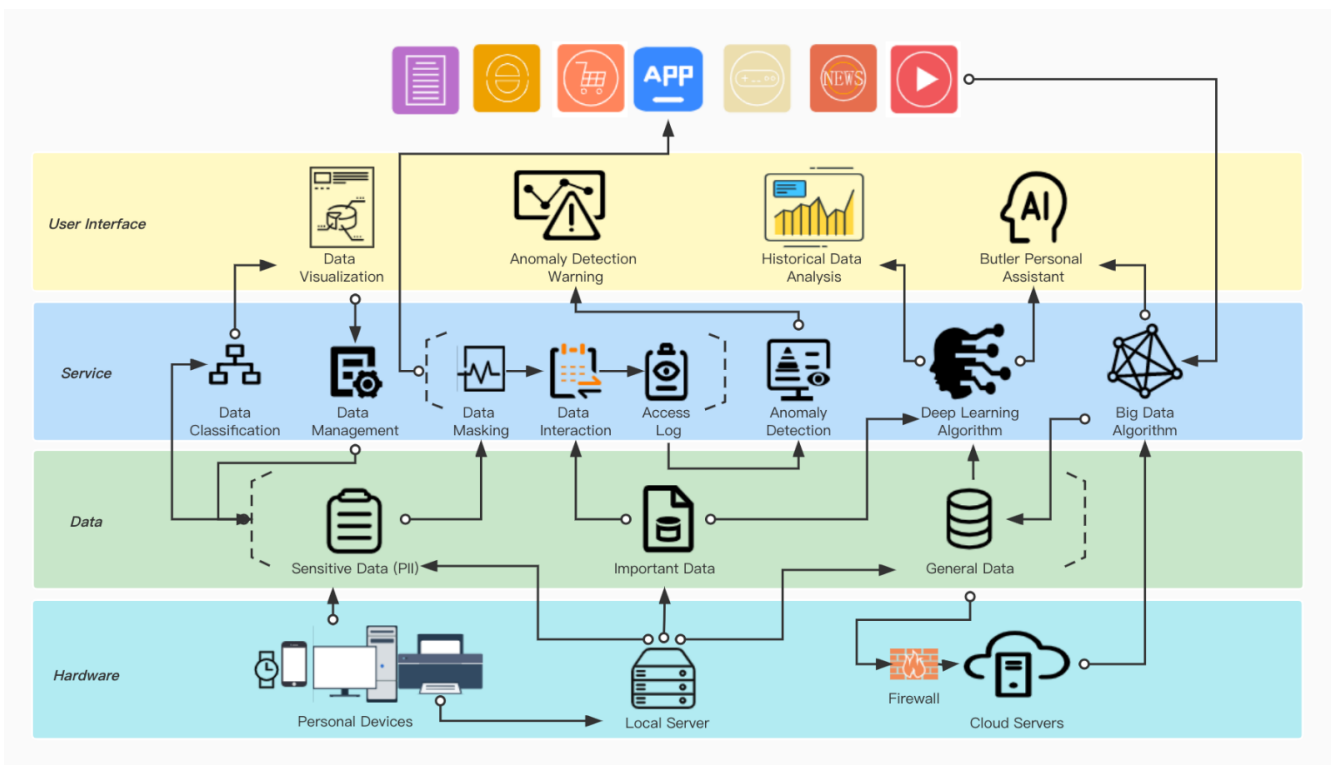


Fig. 1. Business structure and data flow

- (a) Business structure: The main goal is to protect personal data property, and based on safety, let data come out the result to serve the user. Personal Data Center is the first barrier before users connect with the internet. It can automatically store and classify users' sensitive data, important data and general data. When applications apply for data interaction, Personal Data Center will obey Data Minimization Principle and combine it with this application's characteristics to provide the necessary data after masking. When applications apply for reading pictures or videos, Personal Data Center will give those pictures or videos watermarks that can be invisible for tracking. When applications apply for location, Personal Data Center will choose to provide a precise location or

vague location. And meanwhile, Personal Data Center will record all the data activities.

- (b) Basic services: As we can see from Figure 1, in the beginning, Personal Data Center will classify users' data into three levels: sensitive data, important data and general data. It will provide a data management function to add, delete, change and check the data log. After that, Personal Data Center provide sensitive data masking and tracking services to make sure data security. The data masking technology includes vague locations, virtual email and virtual phone numbers that can redirect by platform. The data tracking includes data colouring, watermarking, and data log etc. Based on the data log, Personal Data Center also can provide an anomaly detection function. The sensitive personal data will be stored in the local devices or local server to avoid leaking or being analysed for business reasons.
- (c) Further services: After users' agreement, Personal Data Center will collect data from personal devices. Firstly, It will visualize data and come out dynamic description. Secondly, analyzing historic data for anomaly detection and users' behaviour diagnosis can warn in advance. Thirdly, predicting users' behaviour to plan schedule and guide users' behaviour correctly. In the end, based on big data algorithm and neural network algorithm, provide a "Butler Personal Assistant" to help users make a better life, which will provide advice for users' life habits, physical health, and career development.

For now, this sounds like a utopian illusion because we cut the Big Tech's monopoly on data and give users more rights to deal with their data. How to make those companies compromise? If Personal Data Center only charges a service fee, that cannot help this company survive and grow. How to feed this company? Last but not least, this company needs to build very strong trust with users. How to let users trust this company?

First of all, if we only limit this ideal to an application or platform, which will feel like in the dead end. However, putting Personal Data Center into an operating system, like Apple's IOS, creates a whole different environment. That sounds like a vast concept, but it also will be cutting-edge innovation when people, before connecting with the internet, your operation system protect your data at the first step.

About trust, how to guarantee data safety? First, sensitive personal data is stored by encrypted algorithms and computed at local devices. We can develop an encrypted transmission protocol to synchronise data between different devices. Our company also provides a Personal Data Center Server (PDCS) to store their family's data, including pictures, videos, movies, etc. Our system only maintains and add functions or algorithms. The result only can be seen by the user. We can make part data transport code open sources to make it more trustable. Currently, protecting users' privacy and mining users' value seems like a paradox, but I think when technology develops, we can improve them both.

There are more details I want to discuss in this letter, due to page limitation I must stop at here. I desire to join EIT to find more people who have the same dream as me. When a creative collision happens, it is the moment some great thing was born. You can feel my passion for technology innovation in this letter. I was born to be an entrepreneur. My father told me that we could observe the world closer and use knowledge to make it better. Therefore, when I watch how the world works, I always thinking what we can do more? The difference between entrepreneurs with others is this one step more. Participating in Innovation and entrepreneurship competitions, high-tech events, and speech contests. I always prepare for being an entrepreneur. There are two essential qualities that an entrepreneur needs. The first one is the angle to see the world. The second one is leadership. The reason why I want to leave my hometown to choose to study in Europe is that I need diversity inclusive culture that also contributes to diversity of thought and perspectives that make learning more interesting and dynamic. Living in a strange environment can train my communication ability. When we interact with others, we can build bridges to trust, respect, and understanding across cultures. In conclusion, EIT is perfect for me. It will be my chance to build a strong digital Europe.