Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/14/20230777

App design based on the protection of personal privacy information in China

Xiangru Qin

School of Art and Design, Goldsmiths of University of London, London, UK

Xqin003@campus.goldsmiths.ac.uk

Abstract. In today's world, digital technology is rapidly evolving and the way people live their lives is changing dramatically. For example, tools such as computers, the internet and mobile phones are bringing us new experiences like never before, while digital technology is also bringing unprecedented advances and innovations to business and science. New opportunities are growing and traditional hierarchies are beginning to fall apart - and with them comes a breakdown in trust. Google, Instagram, Wechat, Xiaohongshu - these disruptive platforms,networks and technologies have changed the current status of their respective industries, making our lives easier while there are bad actors in the unseen grey areas stealing and exploiting our privacy and data. The emergence of COVID-19 in recent years has forced us to use and rely on digital devices with high frequency, especially in China, which has accelerated our transition to a digital world, with both advantages and disadvantages. This article will explore how to better protect the privacy and data of digital nomads within reasonable and lawful limits, and provide an idea as a solution. The article presents the idea of a data protection app called Data Butler, which is an application dedicated to protecting the privacy of its users. The aim is to enable users to enjoy the benefits of big data while protecting their private data. Data Butler is designed to enable users to enjoy the benefits of big data while their private data is protected.

Keywords: user privacy data, personalised customer, interface app, user profiling, data protection.

1. Introduction

1.1. The context of user privacy data security in the era of big data

People have changed the way they work in post epidemic China, switching from offline face-to-face offices to online communication meetings, and in the midst of the epidemic quarantine, much of our identity has been digitised online: the interaction between people is almost entirely through digital means. Now, more than ever before in the history of traditional media, people are aware of their digital presence: specifically in the digitisation of services, online commerce, digital learning, virtual health, payments, which are now linked to our digital identity. At the same time, what is behind people's digital identities is the big data economy that drives socio-economic development. While enjoying the convenience brought by big data, individuals form a deep interaction and dependence on the Internet of Things, and people inevitably leave their user preferences, user behavioural information, user bio-information, user privacy information, etc. within the Internet. At the same time, most companies in China are using big

DOI: 10.54254/2755-2721/14/20230777

data algorithms to create a unique user profile of each individual user in order to provide a better 'experience' for the user. It is a vital part of many digital marketing applications. Online user profiling is one of the profitable prospects and businesses should take advantage of it to realise the full potential of "big data." according to McKinsey & Company [1]. User analytics extracts users' goals, opinions and then uses these personal and private data to make business decisions, helping companies to better analyse users' personal preferences for accurate business marketing promotions and profits [1].

As of March 2023, the followings are the latest statistics reproduced, According to the China Internet Network Information center (CNNIC) published the 51st 《Statistical Report on the Development Status of the Internet in China》. The number of Internet users in China increased from 35.49 million in December 2021 to 1.067 billion, with an Internet penetration rate of 75.6%. According to the 50th 《Statistical Report on the Development of the Internet in China》, the proportion of Internet users who have encountered personal information leakage was 21.8%, implying that one out of every five persons in China who use the Internet would encounter such issues. In an era of frequent data privacy breaches, most consumers are anxious about their online privacy and many feel they do not have effective enough measures to protect their privacy [2].

According to the Law of the People's Republic of China on the Protection of Personal Information [3], the "Personal information should be gathered to the degree necessary to fulfil the processing goal, and personal information shall not be collected unnecessarily". However, the definition of "minimum scope" has not yet been clarified, which has led to the excessive collection and misuse of citizens' personal information by some companies under the banner of "improving the consumer experience" [4].

Meanwhile, in a grey area invisible to the law and people, users' private data may have been trafficked many times. In the light of the above analysis, there are certain companies in China that are taking advantage of gaps in the law that are not yet perfect to excessively collect users' private data to maximise their economic interests.

1.2. How do companies or grey areas gain access to users' private data?

As the design focus of this article is on the app design process and output, it will not focus on backend operations. First of all this article will introduce two concepts of data background operation：

SDK: The full name of SDK is Software Development Kit which is a set of kits provided (usually) by the manufacturer of the hardware platform, operating system, or programming language, helping programmers to create or amend the app in an easier way. API (application program interface), in fact SDK can be understood as a series of software development kits provided by a system to facilitate other systems to connect to their API. The purpose of an API is to enable two systems/platforms that need to communicate and share data to send and receive information to and from each other in a way that is understandable and recognised by both parties.

Through JavaScript tracking code, just like Baidu statistics, it belongs to this category, there is a section of JAVASCRIPT at the end of the page, When the user accesses the web page, it is activated and then sends certain browser information to the server. This type of data analysis is used to aid website management as well as APP optimisation.

Through API, just like some weather interfaces, there are many platforms in this field in China. Aggregation is one of them, having a lot of interfaces. This kind of, usually real-time, newer data is available on a pay-as-you-go basis.

Through the crawler, like the Baidu spider, or similar to our Octopus collector, as long as the Internet is open data can be collected, this type of product has several versions, for different groups, each has its own characteristics. They can automatically help you identify elements on the web, automatically help you speed up, etc.

The Buried Point, similar to JavaScript, generally refers to the APP, such as magic strategy, GrowingIO and so on, this principle is nested in an SDK inside the APP. This article pays more attention to the collection of user information by mobile APP, so here we focused on the technology applied in mobile APP, namely SDK.

DOI: 10.54254/2755-2721/14/20230777

1.3. The significance of the fair use of citizens' data privacy

The era of big data is developing rapidly and its value is recognized by the state and provides convenience to citizens' lives. Although it plays a huge role in many industries, but there are two sides to everything. The rapid development of big data has been accompanied by the frequent leakage of personal data, which has sounded the alarm for the development of the industry and also brought various impacts to the public. For example, spam messages waste people's time, telecommunication frauds take away the money of many people who are not aware of self-protection, and personal bio-data privacy information can even threaten people's lives. Personal private data is a virtual asset of individual persons and is legally protected. Today, China's data privacy rules and regulations are not yet integrated, allowing certain unscrupulous elements and corporations to use personal private data to do something illegal and detrimental to society and citizens. The development of technology should be people-oriented and technology-based, rather than using citizens as a tool for development and extraction, which is inherently detrimental to the long-term development of the big data economy and can cause a degree of public fear and backlash. Chen argues that the personal privacy of users on social media is a mutually rational exchange [5]. But reasonable privacy disclosures are also an important part driving big data. The personalisation of Big Data increases user satisfaction by reducing the time spent searching for accurate data and information [6, 7]. Consumers may release private information about themselves in exchange for personalised services or useful information. Only by protecting citizens' data privacy in a reasonable and legal way can some of their concerns be removed to a certain extent, completely eliminate to dispel worries is not possible. As long as there are data issues, there will inevitably be controversial. Only by providing good service quality measures and feedback to citizen users within a limited scope can the citizens enjoy the convenience brought by big data as much as possible, while promoting the development of big data economy.

Citizens have no control over their personal data to the extent that it is lawful and reasonable to do so. As a tool for data transmission between devices and service providers, SDK brings personalized and convenient services to users, but also inevitably puts the security of users' personal information in danger. In order to protect user data and enjoy the convenience brought by big data analysis, we hope to generate new technologies based on the existing information anonymity technology and collection technology, which can standardize SDK, prevent SDK from disclosing personal information, prevent information from tracking individuals, return the management right of personal data to individuals, simplify the management of personal data, visualize the flow of personal data, etc.

The rest of the paper is as follows. First, the literature will cite several cases of data privacy breaches and in a second step summarise the views of experts in the relevant research areas. The research methodology and results section then details the design methodology and analysis of the literature, and presents the final design results with images. Finally, we discuss the findings, implications, recommendations and limitations associated with this design study, as well as ideas for further research.

2. Related perspective in privacy data

2.1. Examples of user privacy data breaches

In September 2018, Facebook announced that 30 million users' information had been compromised as a result of a hacking attack due to a security system breach. Hackers gained access to the private data of 14 million people. Of which there were names, contact information, search history, login locations, and other sensitive information. On December 14, it was once more discovered that a Facebook software flaw may have resulted in the exposure of 68 million users' private images. Particularly, between September 13 and September 25, a flaw in its picture API gave access to users' private photographs to around 1,500 applications. Normally, applications that require user authorization can only view shared photographs, however this flaw makes it possible for apps to read photos that the user has made private. The end result of the handling was an apology from Facebook's CEO for the data breach and several appearances at hearings. Various events since the beginning of 2018 have affected Facebook's share price, which has fallen 29.70% as of December 25, 2018. With this breach in December 2018, the Irish

DOI: 10.54254/2755-2721/14/20230777

Data Protection Commission, Europe's privacy authority, initiated an inquiry that may result in a $1.6 billion punishment for Facebook [8].

In China in 2018, a user with the ID "f666666" began selling 1 billion pieces of courier data from the Yuan tong Express Company on the dark web in June 2018, claiming that the data was from late 2014 and the data information included the sender's name, phone number, and address. (recipient). After 1 billion data were checked and processed, It was discovered that the data duplication rate was less than 20%, and the data was sold by the user in a package for 1 bitcoin. Finally, several netizens checked the authenticity of some data and discovered that the acquired "single number," name, phone number, address, and other information were correct [9]. Examples of corporate data breaches like the one above happen every year [10].

2.2. Theoretical research by the world's first trust and technology researcher

Some researchers have concentrated on privacy theory. Rachel Botsman-The world's first trust researcher specialising in the relationship between trust and technology and a leading expert on trust in the modern world. As a Trust Fellow at Oxford University's Saïd Business School, this study aims to challenge and change people's perceptions of trust and related topics such as power, influence, truth and faith [11]. At the point of this work, Facebook can change its algorithms, but fundamentally, people don't trust their promise for privacy around data and they don't trust their business model. Focusing on product integrity rather than transparency, transparency reduces the need for trust because it is reducing the number of unknowns, which requires fewer leaps of trust. Secrecy is not the enemy of trust; the enemy of trust is deception, and transparency is almost impossible to guard against. Think of transparency as a tool rather than a core need and concern. This work believed that people should pay more attention to the integrity of the operational process and the integrity of the various related policies in the design process. Designers should be ahead of the users to provide a better user experience for the users.

3. App design process for securing user privacy data

3.1. The concept of safeguarding the security of users' private data

This stage will combine background research and user pain points to come up with initial ideas for the design, focusing on the completion of the overall service initiative. After brainstorming a design idea that mentioned data migration and node access, this led the design team to think about VPN, as VPN can migrate the user's IP and address, could VPN technology be used to protect the user's privacy?

Firstly, most Chinese people are currently unaware of VPN technology, which is a relatively new technology for China's large population base. This article introduces the definition and benefits of VPNs as well as discusses which age group is the primary user for the initial product rollout. If this idea holds true, the product could protect the privacy of user data as well as the personal information of certain criminals. However, most of the VPN's on the market are currently illegal in China and the use or manufacture of VPN's in China is criminally liable. With these questions in mind, the designer take the next step in research and analysis.

VPN (Virtual Private Network) refers to the ability to establish a secure network connection when utilising public networks. VPNs encrypt users' internet traffic and conceal users' identity online. This makes it more difficult for third parties to follow users' internet activities and steal information. The encryption is done in real time [12].

VPNs can be securely encrypted. A VPN server essentially acts as a proxy for users on the Internet. Because the demographic location data comes from a server in another country, there is no way to determine users' actual location. Furthermore, most VPN providers do not keep logs of their users' activities. Some providers, on the other hand, log user behaviour but do not share it with third parties. As a result, any potential logs of user conduct are permanently hidden. Regional content can also be accessed. It is possible to access websites that are banned in users' region.

DOI: 10.54254/2755-2721/14/20230777

But VPN is currently perceived as an illegal node in China, is there any way to make use of this node technology within reasonable limits? And, what age group of users would accept this technology?

3.2. Design method

3.2.1. Questionnaire research data. This part of the questionnaire took the form of a cross-tabulation of two questions. This image visually shows the impact of age on new technology products and it is clear from the graph that the highest combined level of acceptance is found among 26-35 year olds (see Figure 1 (a)). This image shows the effect of gender on whether or not a good product is recommended to others, as can be seen from the graph, the proportions of male and female recommendations are roughly the same (see Figure 1 (b)). This graph shows the effect of age on recommending good products to other users, as shown in the graph each age group will recommend good products to others (see Figure 2 (a)). This graph shows the effect of educational attainment on the acceptance of new technologies, as shown in the graph the highest overall acceptance is for those with a university degree or higher (see Figure 2(b)). This graph shows the level of acceptance of new technologies by gender, which indicates that men and women have similar levels of acceptance as shown in the graph (see Figure 3(a)). This graph shows the effect of age on whether or not users like to learn about and understand new technology, as shown in the graph (see Figure 3(b)).

Figure 4 is questions that used open-ended responses. As shown in the graph, a total of 213 people gave open-ended responses, this graph only lists the 15 most frequently occurring words.(see Figure.4)

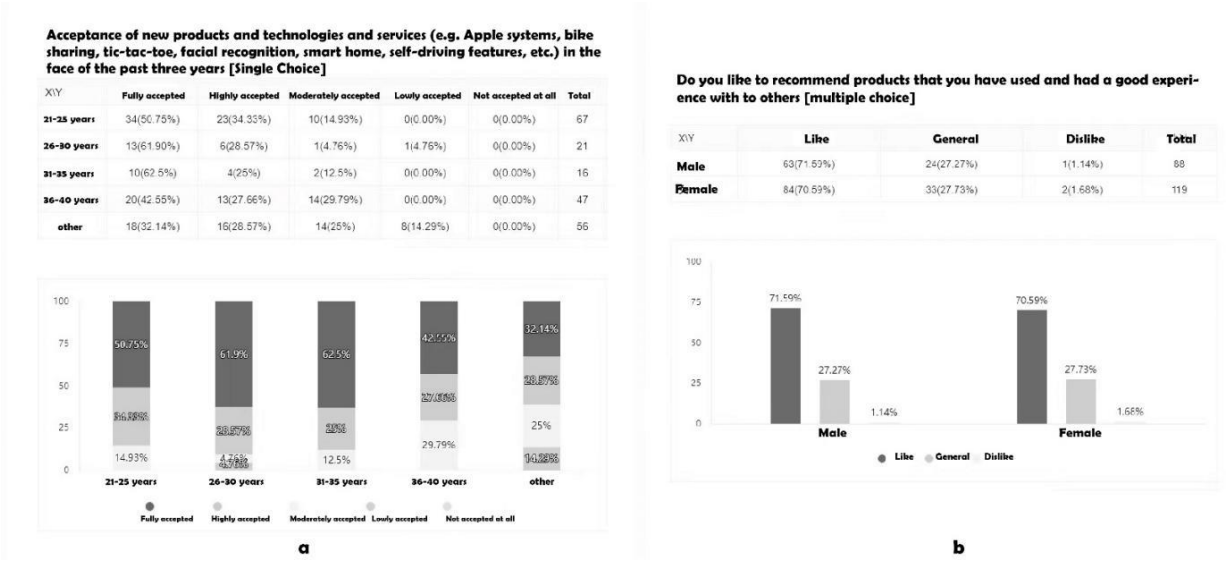

Figure 1. (a) Age acceptance of new technology ; (b) The influence of gender on recommending good products.

DOI: 10.54254/2755-2721/14/20230777

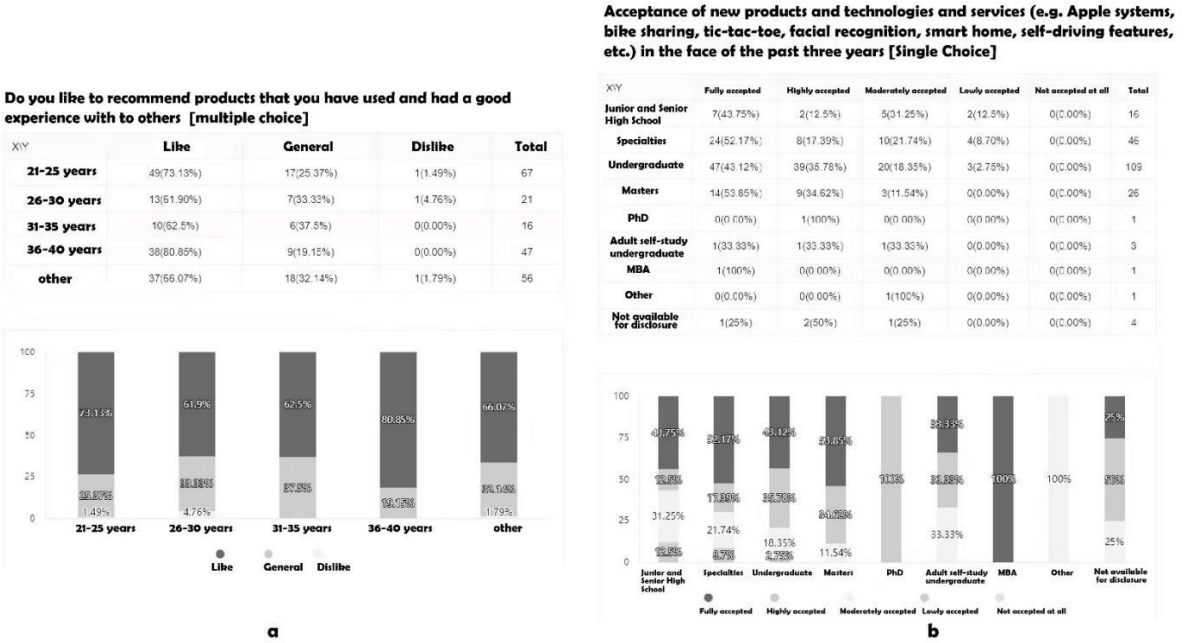

Figure 2. (a) The effect of age on recommending good products; (b) Acceptance of new technologies by academic qualifications.

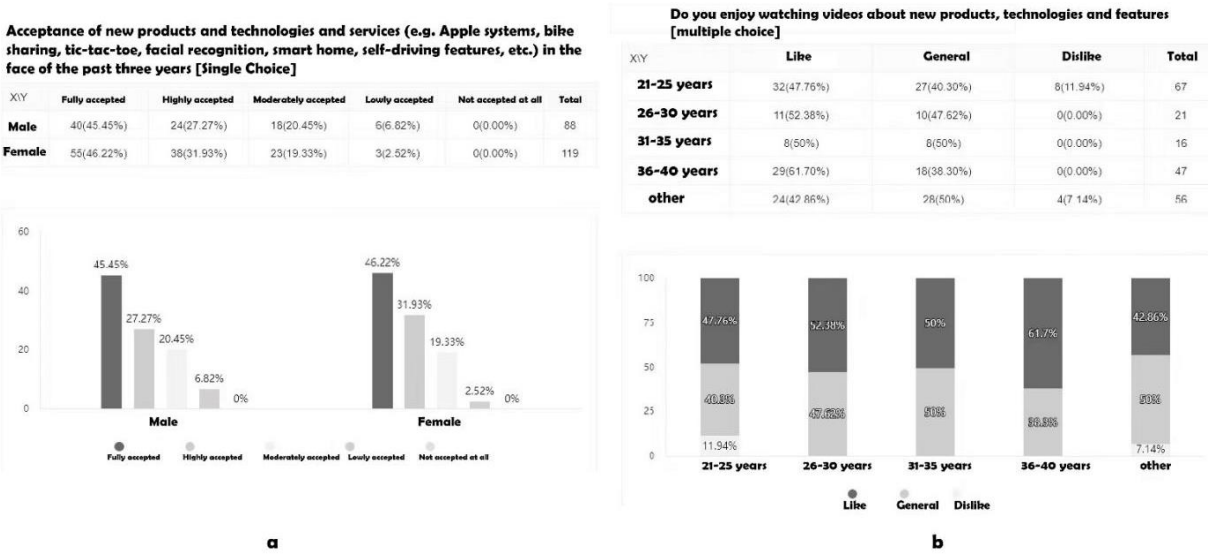

Figure 3. (a) Gender acceptance of new technologies; (b) The impact of age on knowledge and understanding of new technologies.

DOI: 10.54254/2755-2721/14/20230777

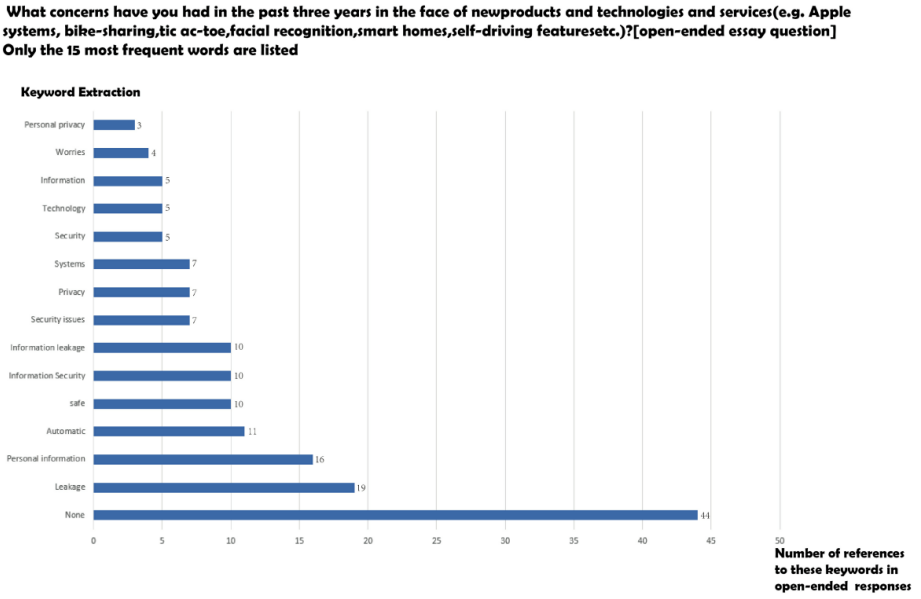

Figure 4. Users' concerns about new technologies in the last three years.

3.2.2. Interviews.

Figure 5. Interviews with officials from the government data department.

Due to the specificity of VPN in China, questionnaire research and secondary research could not meet the needs of the project, so the author decided to adopt a face-to-face interview design method, and finally interviewed officials from local data security departments at the municipal level in China to answer questions and solve confusion with government officials in a face-to-face format, and consulted the following questions according to the concept, and intercepted some of the core content here. The following questions were asked based on the concept, and some of the core content is included here. This also solves the confusion I have been having about the use of VPN technology in China (see Figure 5).

At the same time, some researchers proposed that more attention should be paid to the role of the personal information protection supervisory authority, focusing on major violations of the law. In the face of powerful processors of personal information, individual subjects are not able to effectively

DOI: 10.54254/2755-2721/14/20230777

counteract them. Therefore, even if the law gives individuals the right to give their final consent, in practice it may be reduced to a "right on paper". The effective protection of personal information depends not only on the effective exercise of individual rights, but also on strict regulation at the national level [13].

3.2.3. User prototype test feedback. Using the 'user usability testing' design method, new products with high uncertainty can be tested with prototypes drawn on paper or wireframe prototypes, avoiding wasted manpower and material resources in revisions. For products with a high degree of certainty, high-fidelity prototypes with interaction design can be used for testing, which covers both interaction design and interface design and is more comprehensive. This is a way to reduce the waste of human, material and financial resources, but also to obtain a more comprehensive test results.

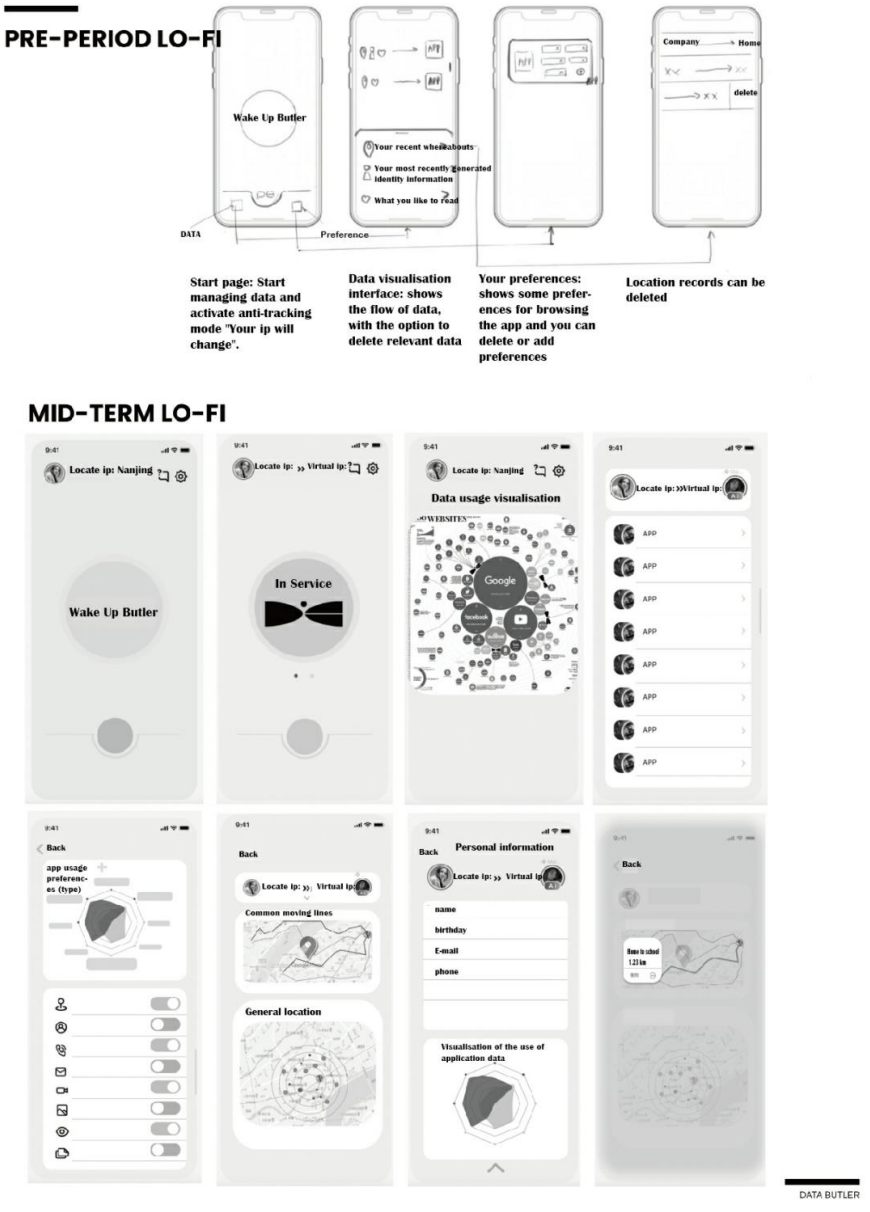

Figure 6. App design sketching process.

This step will create simple app models (low fidelity models) and invite people with certain needs to test the layout and flow of the feedback, while conducting the test, this work also used the "Think Aloud"

DOI: 10.54254/2755-2721/14/20230777

method, so that the test user in the use of the app This allowed the author to understand the cognitive activity going on inside the user at each stage of the process. For example, when using the wireframe model of the main interface during testing, the user cannot trigger the function of personal data privacy management, the user does not know how to jump from the main interface to my personal privacy settings interface, and the test user will say I don't know how to jump to the personal privacy settings interface, and will have doubts why this work can't find this interface, and where this interface should be placed appropriately (see Figure 6).

This allows the author to obtain a message that there is a problem with the interface's operation logic and navigation logic, and to improve this in the next version of the model, before inviting users to come in for usability testing, and to test until there are no problems with the operation and layout allocation of each interface before the final product appears. It is not until it has received a high degree of positive feedback from the users who participated in the test and the overall smoothness of the operation process that it can proceed to the final product design presentation.

3.2.4. User profile construction.

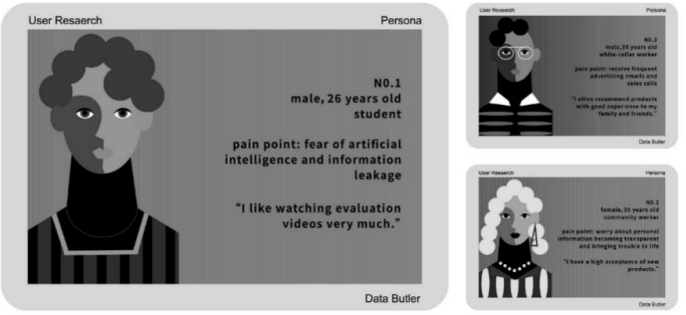

Figure 7. User profiling.

Based on the results of the research, and in conjunction with friends and students who had been researched, this step used a user portrait design technique to help pinpoint the needs of the users, and a total of three virtual user portraits were constructed, with the three users having different pain points and needs, in order to make a more targeted design (see Figure 7).

4. Result

4.1. Findings from the questionnaire

1) In terms of age, people aged 26-35 are highly receptive to new products; In terms of educational background, graduate students have the highest acceptance of new products, followed by junior college students. Gender makes little difference; In terms of educational background, the acceptance of new products by master's degree students is the highest, followed by junior college, junior high school and bachelor's degree students.

2) In terms of age, people aged 36-40 like to watch evaluation videos the most, followed by 26-35; In terms of gender, men are more likely to watch evaluation videos than women, but the difference is not large. In terms of educational background, junior college students prefer to watch evaluation videos, followed by undergraduate students, middle and high school students.

3) In terms of age, people aged 36-40 are more likely to recommend products with good experience to others, followed by those aged 21-25. Gender makes little difference; In terms of educational background, junior colleges are more willing to recommend, followed by undergraduates.

Generally speaking, people aged 26-35 have the highest acceptance and certain radiation ability, and 36-40 is the best radiation range, so it can be used as the secondary promotion target.

4) As shown in the graph, a total of 213 people gave open-ended responses, this graph only lists the 15 most frequently occurring terms, "None" appears 44 times, after a subjective summary of keywords there are words related to "data and security" mentioned The terms "data and security" were mentioned

DOI: 10.54254/2755-2721/14/20230777

101 times, namely leakage, safe, personal information, information security, information leakage, security issues, privacy, security, worries, personal privacy. personal privacy (see Figure 4).

4.2. Conclusions of the face-to-face interviews

Some VPN technology is legal in China and is used by the Chinese government. Data Butler designs the service process and nodes so that the service backend (B-side) is under the supervision of government agencies to protect the product from being used by unscrupulous elements.

4.3. Data butler app interactive product output

4.3.1. Operation flow chart. Headshot area: this area is mainly responsible for switching virtual IP, editing location and personal preference information. Big dots area: opens up data protection and visualises the management of personal data flows, making it easy for users to see how much personal information is accessed by applications. Bottom flag area: manages data permissions for different applications, permissions for private data transfers, adds interference with tagging preferences, influences the calculation of private user profiles by applications (see Figure 8).

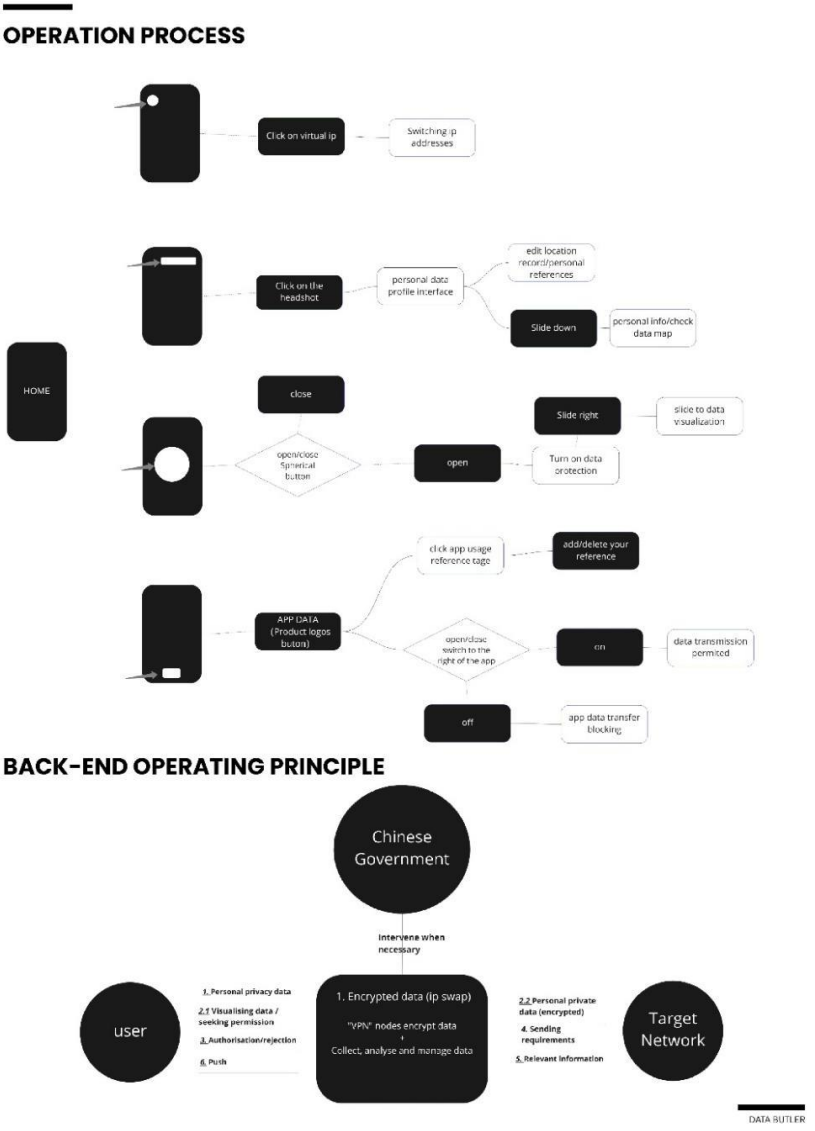

Figure 8. App workflow principles.

DOI: 10.54254/2755-2721/14/20230777

4.3.2. *App interface design.* The user interface is designed around "seriousness", "trustworthiness" and "professionalism". The overall design is in black, white and grey, with a simple interface that highlights the professionalism and calmness of the Data Butler app's service. The use of design elements is more dynamic in comparison and is intended to appeal to a younger target group. The use of design elements is more dynamic in comparison and is intended to appeal to a younger target group (see Figure 9).

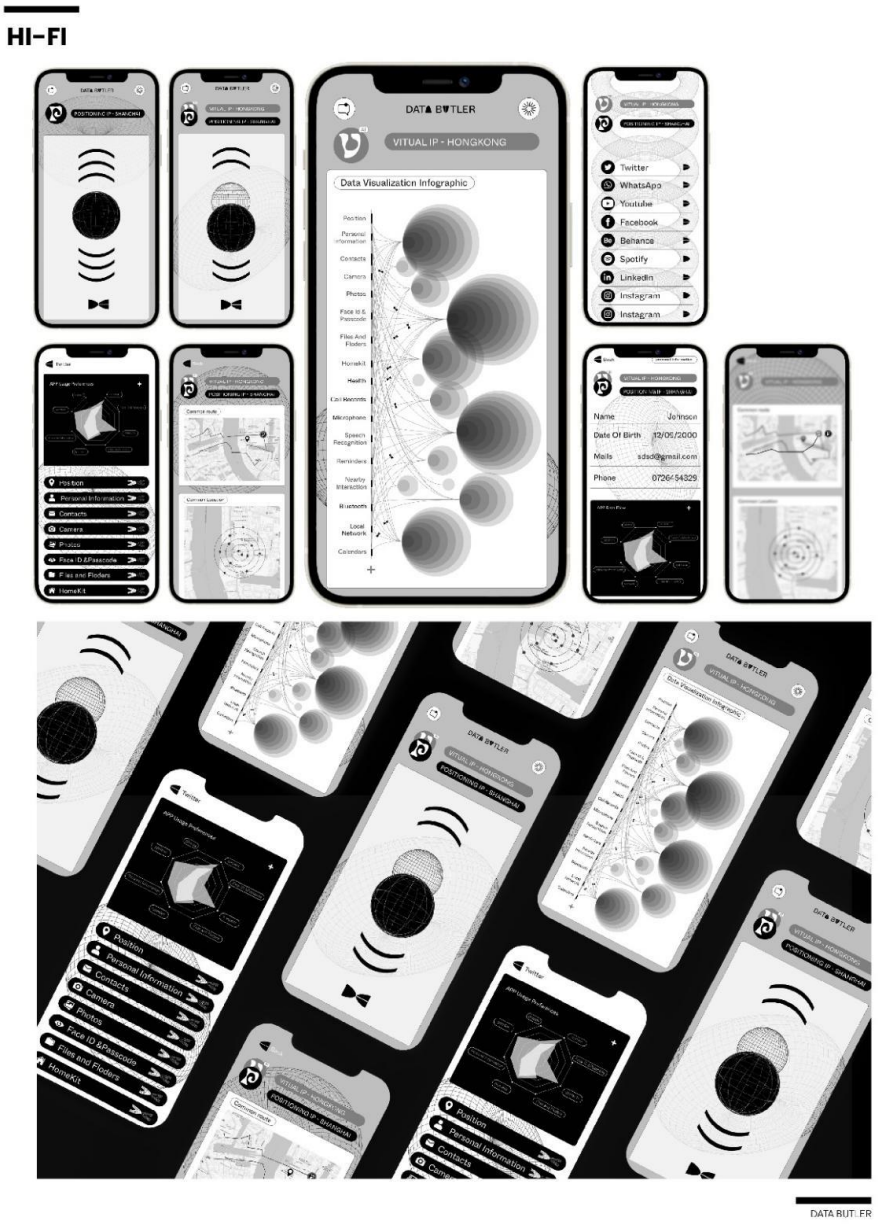

Figure 9. App interface design.

4.3.3. *Other design presentations.* The app logo "DB" is the initials of Data Butler and resembles the bow tie of a butler: the user interface is designed in black white and grey to emphasise the professionalism and calmness of the Data Butler App's butler service; The use of design elements is more energetic and in tended to attract a younger target group. The product is a new product in the field, so it needs a certain amount of publicity and promotion, and for this reason a promotional poster and peripheral products have been designed (see Figure 10).

DOI: 10.54254/2755-2721/14/20230777

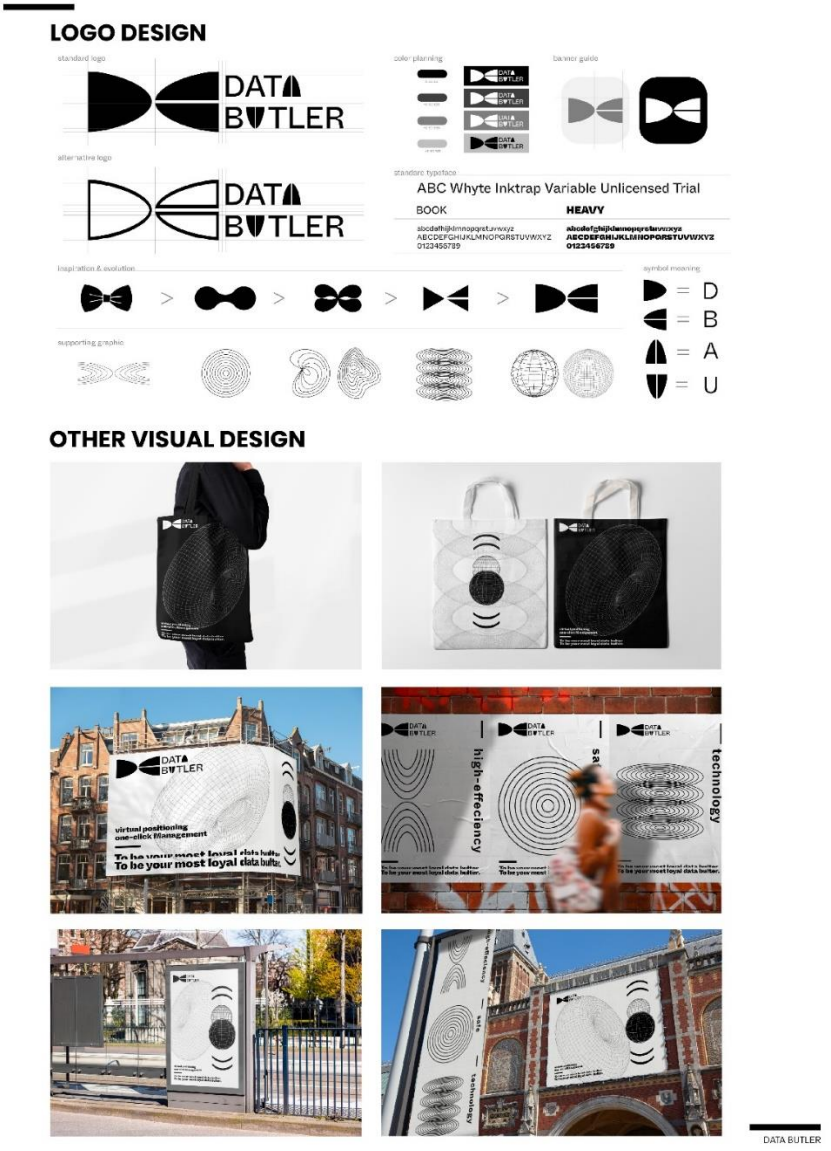

Figure 10. Other related designs.

4.4. Design description

Data Butler is an app dedicated to protecting users' privacy. Its logic is designed to replace the user's real lP with a virtual lP, to visualise how various platforms read personal network data, and to provide a way for users to view and manage their personal network data.

4.5. Function description

Using VPN-like technology, it encrypts user, making user information and data inaccessible to third parties for sale to advertisers. Third parties do not have access to user information and data that they can sell to advertisers. However, in order to enable users to enjoy the convenience of obtaining information brought by big data, our services and products will collect and analyse user data and give feedback to users in the form of visualization. Under the control of users, we will transmit and authorize necessary data (to advertisers but try to avoid the disclosure of personal information). In most cases, the user agrees to tell the advertiser what information is needed, and the advertiser sends it to a VPN node and then sends it to the user). To realize legalization, we are willing to cooperate with relevant government

DOI: 10.54254/2755-2721/14/20230777

departments under the necessary condition. we may say that we do not disclose personal data, but if the netizen has violated the laws related to network, we will cooperate with relevant government departments to track illegal Internet users.

5. Conclusion

Judging by the current online interaction practices of most young social users, the majority of young social users are not aware of how to effectively manage and control their privacy data due to their education level, social environment and personal awareness. But at the same time, Chinese internet users are becoming more aware of privacy and data protection. Through comprehensive research, there is no such product on the market in China, Data Butler completes this gap in the market. Under the regulation of the Chinese government, users are encrypted using VPN-like technology so that user information and data cannot be accessed by third parties nor sold to advertisers. However, in order for users to enjoy the easy access to information that big data brings, data butler collects and analyses user data and feed it back to the user in a visual format. Under the control and authorisation of the user, data butler will transmit and authorise the necessary data.The big data economy is an important part of the development of all industries, and data protection apps are currently a gap in the market, while the innovation and implementation of such products is a complex process. It is expected that Data butler will provide a new solution to the growing social problem of data protection.

References

[1] Trusov, M., Ma, L., & Jamal, Z. (2016). Crumbs of the Cookie: User Profiling in Customer-Base Analysis and Behavioral Targeting. Marketing Science (Providence, R.I.), 35(3), 405-426.

[2] Anderson, B., Vance, A., Kirwan, C., Jenkins, J., & Eargle, D. (2016). From Warning to Wallpaper: Why the Brain Habituates to Security Warnings and What Can Be Done About It. Journal of Management Information Systems, 33(3), 713-743.

[3] Gao, Y. (2022). The development of big data in the context of the Law of the People's Republic of China on the Protection of Personal Information[J]. Library Theory and Practice, (4): 4-11.

[4] Shao M. (2021). Research on the legal protection of personal information in data applications[J]. Legal Expo, (05): 65-66.

[5] Chen, H. (2018). Revisiting the Privacy Paradox on Social Media With an Extended Privacy Calculus Model: The Effect of Privacy Concerns, Privacy Self-Efficacy, and Social Capital on Privacy Management. The American Behavioral Scientist (Beverly Hills), 62(10), 1392-1412.

[6] Honerkamp, V. (2020). Predictors of avoidance towards personalization of restaurant smartphone advertising [Summary].

[7] Park, J. (2014). The effects of personalization on user continuance in social networking sites. Information Processing & Management, 50(3), 462-475.

[8] ur Rehman, I. (2019). Facebook-Cambridge Analytica data harvesting: What you need to know. Library Philosophy and Practice, 1-11.

[9] Huang, L. (2021). On the Protection of Privacy in the Internet Era[J]. Business Intelligence, (8): 254

[10] Kuhn, M. (2018). 147 Million Social Security Numbers for Sale: Developing Data Protection Legislation After Mass Cybersecurity Breaches. Iowa Law Review, 104(1), 417-445.

[11] Botsman, R. Who can you trust?: how technology brought us together–and why it could drive us apart. Penguin UK, 2017.

[12] Ding, Z., Yang, R., Cui, P., Zhou, M., & Jiang, C. (2022). Variable Petri Nets for Mobility. IEEE Transactions on Systems, Man, and Cybernetics. Systems, 52(8), 4784-4797.

[13] Liu, Y. (2023). The dilemma of consent notification rules for personal information processing and suggestions for improvement: interpretation and reflection based on the Law of the People's Republic of China on the Protection of Personal Information[J]. Media, (1): 73-76.