

Personal Information Management based on the Era of Big Data

Jia KANG¹, Zheng-Hong WANG^{1,*}

¹Beijing Institute of Technology, China

*Corresponding author

Keywords: Information security, Personal information management, Big data.

Abstract. This paper is mainly for discussing the personal information management in new period. Through the method of literature review, it sums up the correlative research findings. According to the deficiencies of existing systems, we proposed a new general model, and put forward some constructive opinions to future research.

Introduction

Under the background of big data, global information is growing at an exponential rate. The manifestation and carrier of information also have profound changes. The big challenge we're faced with currently is to find the target information in the massive information as well as privacy protection. The concern about the information gives birth to a new field: personal information management.

The Era of Big Data

The Concept of Big Data

There are still no uniform definitions of big data. Wikipedia believes that big data usually includes data sets with sizes beyond the ability of commonly used software tools to capture, curate, manage and process within a tolerable elapsed time. However, Gartner defines it as high-volume, high-velocity and high-variety information assets which demands cost-effective, innovative forms of information processing that enables enhanced insights, decision-making, and automation [1].

It can be seen that big data is the phenomenon that technology develops to a certain stage. Many organizations, companies and governments invest in related industries of big data. It becomes a hot topic in the Internet information technology field. However it should be mentioned that the key point is not the amount of systems developed. Finding out the connotative and potential values of data matters most. As examples, analyzing the data produced by lights, conditioners and fans can bring down energy consumption. Even the analyses of consumers' behavior data lead to more popular products. The data value is more important than the data itself.

Personal Information

Personal information means the information produced by the individuals, including personally identifiable information, document information, account information, space information, etc. William Jones holds the view that personal information mainly includes three types: users save and limited to their own use, related to the individuals but controlled by other

organizations or agencies (like citizen identity information controlled by the government departments), produced in the process of external communication but not controlled by individuals (information left after one browsing the website) [2].

Personal information is showing new features. Firstly, the volume of information is growing fast. People produce information during their work or life. Most of the information is worthless, resulting in information redundancy directly. Secondly, the relevance of information is not strong enough. The sources of personal information are diversified because of the widespread uses of different electronic devices. Mutually independent information hampers the unified management. Thirdly, the security of personal information is being threatened. Hacking, information leakage events occur much more frequently than before, which reflects the existing measures are not perfect. All of these traits definitely raise the difficulty in personal information management.

Research on Personal Information Management

Overview of Personal Information Management

The concept of personal information management was first proposed by Bush in 1945 [3]. As a result of the development of relevant theories and techniques, the personal information management field has scored important achievements. PIM Workshop 2009 agreed that personal information management was the study and practice one in a series of activities, getting or creating, storing, organizing, saving, retrieving, using and distributing all kinds of information that people need, in order to achieve many real-world goals and fulfill their roles and obligations in life [4]. As can be seen from the definition, it's the operation of useful information. And it's closely associated with people's needs and roles.

The recent literature shows that, the domestic works on personal information management mainly concentrate on relevant tools, patents and information security. The studies abroad go far, which also consist of interdisciplinary research and information integration. On the tools' side, more and more systems turn to B/S (Browser/Server) structure thanks to the constraints of C/S (Client/Server). To solve the problem of compatibility and version inconsistency of mobile terminals, researchers have tried Semantic Desktop. Speaking of patents, Tencent ranks first among local companies. In the aspect of information security, the development of foreign countries is relatively mature. Remarkably, personal information management is associated with behavior science and psychology at first, and tightly related to organizations as well as information later. Integration management is at early stage, which is proposed for fragmentation. The major mean is to add function modules of integrating personal information. Although personal information management has obtained some results, there is still a long way to go.

Personal Information Management System

In the context of big data, people use mobile devices to manage information in daily work and life. The mobile environment is so different from the traditional desktop environment that users need a more direct and natural way to interact. In addition, today's users don't have much professional knowledge in computer, which requires personal information management to be easily operated. So users can master the skills of collecting, processing and retrieving personal information in a short time on their own. On this basis, personal information management systems divert from traditional WIMP (Window, Icon, Menu, Pointer) interface

to Post-WIMP interface. Ming-xuan Li designed a personal information management system for mobile environment [5].

The whole system is divided into four layers: data layer, method & technique layer, information management layer and interface layer. Among these, the interface layer converts users' handwritings into commands that a machine can recognize. And the information management layer is to save, extract, retrieve personal information of users and establish relationships. While method & technique layer matches the handwritings or signs with corresponding user information. From here, method & technique layer provides essential support for information management. Data layer is the database of users' personal information, containing documents, e-mails and tags. There is no doubt that it is an excellent system, but one small drawback exists: that is, it works only with users' handwritings or signs. What's worse, handwritings can be imitated, then information security loses.

Based on relevant studies, the author proposes a general model of personal information management. The whole model includes two parts: operation layers and databases. Databases mainly refer to users' commands and information. Meanwhile, operation layers show the complete process of information processing. Users' commands database establishes corresponding relationships between requests and machine languages, so the system can identify users' real purposes. The next stage is to generate tasks. If they are enclosed in the task layer, then execute, false otherwise. All the commands are related to personal information stored in information space, a considerably safe and organized place. Details are shown in Fig. 1. Before actions, users need to go through multiple authentication, like iris recognition. If not, it will be locked and send error report. It is simple but practical.

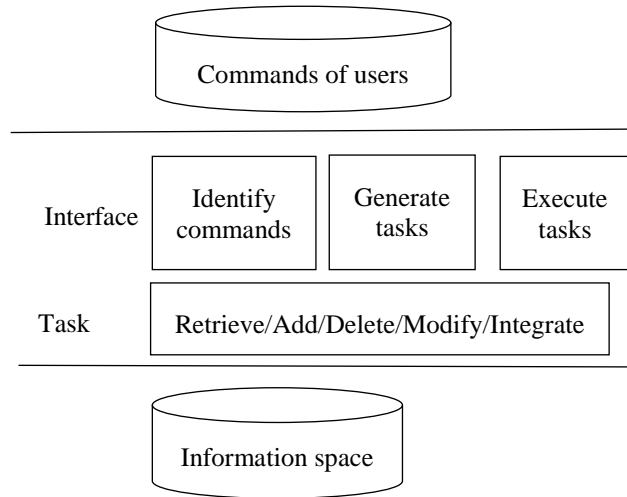


Figure 1. General model of personal information management.

Development Trends

More Attention to Information Security

In the information society, personal information has become a valuable commodity. Gender, age, income and position could be used improperly. Recommendations of goods or services, even blackmails are of frequent occurrence. At the same time, cloud storage and transmission

also face security risks, information leakage and theft are pretty common. Almost nine hundred thousand users' resume information leaked while fake wi-fi stole users' privacy [6]. It's hard to imagine, but they did happen. These events alarm us. It's a challenging job to keep the information safe.

Everyone has a great deal of information. It's difficult to ensure information security. Therefore, some researchers suggest classifying the information into different levels of protection based on its severity. The key information to identify individuals, such as ID number, phone number, address, and confidential information should be listed as the highest level. Nevertheless, it hasn't reached an agreement on importance rankings of personal information up to now. Further researches may put emphasis on the comprehensive classification of information.

In view of the current situation of information leakage, it's really pressing to develop corresponding information security technologies. Anonymity technologies, encryption technologies and access control technologies are regarded as essential ways to protect personal information security. Information security technologies should run through the whole process of generating, transmitting, saving, modifying and deleting information, for the purpose of protecting information beforehand instead of making up later. Only proven technology can promote the establishment of perfect integration models and protection systems: then personal information security will be guaranteed.

More Integration of Personal Information

In the era of big data, personal information is distributed on various terminals. The chaotic information has brought a lot of inconvenience to our work and life, thus the integration of personal information is gradually put on the agenda.

The integration of personal information represents classification and collation of information. It aims to make the operation effective and reduce the time costs. First, achieve the goal of information synchronization of different devices. The information is stored on the third-party platform. Once the passwords match, the personal information can be transferred from one device to another. Again, we can file the information based on its importance. It's also feasible to classify according to the topics, such as project, e-mail, time and form of information. In order to facilitate the retrieval, it is constructive to divide the personal information into two parts: work and life, then place it in different information spaces with regard to its importance.

Information integration is the foundation of information management. Nowadays, the pursuit of efficiency is so popular that it seems critical to integrate information. How to exchange the information with other systems is one of the development trends.

More Intelligent Personal Information Management Platforms

With the large amount and diverse carriers of personal information, individuals can't manage information just by memory. Therefore, some certain tools are needed. Platforms in mobile environment are not limited to computers; instead, they extend to mobile terminals. The reason for this is that mobile devices have the advantages of small size, light weight, simple operation, etc. The 21st century is not only the era of big data but also the era of mobile computing.

In view of information security and information processing functions, platforms will adopt more advanced and sophisticated user authentication methods. Digital tattoos and brain waves recognition are likely to be applied. The combination of many approaches is uncertain. Users

can handle information only by voices or brainwaves. The entire platforms tend to be task-oriented and take users' preferences into consideration, achieving the goal of intelligence and meeting clients' demands.

Moreover, personal information management platforms are inclined to specialization. Each function exists as an independent system, to strengthen the depth of operability for each task. Whether AI and VR will be deeply incorporated in or not, it's still a question. But to be sure, there will be more smarter and varied platforms in the near future, it is a matter of time before the transition of information from individuals to society is realized.

Conclusions

Personal information is an important asset in 21st century. We need to make use of certain tools when protecting and utilizing it. This paper briefly introduces Post-WIMP model for personal information management, and proposes general model. Moreover, information security, information integration are main topics of further studies.

Acknowledgement

Thanks to Zheng-hong Wang, the director of Technology Department of College of Continuing Education, for his guidance in the process of composing this paper.

References

- [1] M.L. Tao, Z.C. Wang, Study on personal information protection mechanism in the era of big data, *J. Information Research*, 1 (2016) 3-14.
- [2] Y. Zhang, Study on development tendency of personal information management, *J. Information Research*, 7 (2013) 90.
- [3] V. Bush, As we may think, *J. The Atlantic Monthly*, 7 (1945) 101-108.
- [4] X.F. Tang, S.J. Chen, Development research of personal information management tools, *J. Modern Information*, 33 (2013) 163.
- [5] M.X. Chen, Y.Y. Jiang, F. Tian, G.Z. Dai, Ruby: a personal information management system for mobile environment, *J. Computer-Aided Design & Computer Graphics*, 22 (2010) 1477.
- [6] S. Hu, Personal information management and protection based on the era of big data, *J. Software Guide*, 14 (2015) 9