Takasaki, Haruo; Nakamura, Toru; Kiyomoto, Shinsaku

Conference Paper

Proposal for a privacy policy manager as architecture for a new privacy-enhancing platform

14th Asia-Pacific Regional Conference of the International Telecommunications Society (ITS): "Mapping ICT into Transformation for the Next Information Society", Kyoto, Japan, 24th-27th June, 2017

Provided in Cooperation with:

International Telecommunications Society (ITS)

Suggested Citation: Takasaki, Haruo; Nakamura, Toru; Kiyomoto, Shinsaku (2017) : Proposal for a privacy policy manager as architecture for a new privacy-enhancing platform, 14th Asia-Pacific Regional Conference of the International Telecommunications Society (ITS): "Mapping ICT into Transformation for the Next Information Society", Kyoto, Japan, 24th-27th June, 2017, International Telecommunications Society (ITS), Calgary

This Version is available at:
https://hdl.handle.net/10419/168544

Proposal for a privacy policy manager as architecture for a new privacy-enhancing platform

Haruo Takasaki, Toru Nakamura, Shinsaku Kiyomoto
KDDI Research, Inc., 3-10-10, Iidabashi, Chiyoda-ku, Tokyo, 102-8460, Japan

Keywords: Privacy, Personal Data, Data Protection, PDS, Privacy by Design

Abstract

Refinement of the legal framework that balances the protection and usage of personal data in online and offline services has been discussed globally. On the other hand, because of information asymmetry between users and service providers and bounded rationality of users, the existing notice-and-consent mechanism (Solove's privacy self-management model) has proven inadequate in its effectiveness. This research contributes to solving the abovementioned issues using a complementary technical system (named Privacy Policy Manager). The Privacy Policy Manager creates users' privacy policies based on their preferences and supports them in controlling data disclosure and data usage in accordance with these policies.

1. Introduction

In recent years, utilization of big data, especially personal data, is expected. Utilization of personal data for the creation of new industries or new services is expected, and thus such personal data are considered a source of growth. This expectation has become a global trend since 2011. The World Economic Forum report (Schwab et al. 2011) pointed out that personal data are the 'new oil' and 'new currency' of the world market.

The up to data of legal framework for data protection to fit into Big Data era has been discussed globally. Therefore, 2012 has been called the 'Privacy Year' (Miyashita 2012). The European Commission published new drafts of data protection regulations in January 2012, and they were adopted as the General Data Protection Regulations in April 2016. In the United States, the Obama administration unveiled the Consumer Privacy Bill of Rights in February 2012 as part of a comprehensive blueprint to improve consumers' online privacy protection. Moreover, in OECD countries, the 1980 OECD Guidelines were amended in July 2013.

On the other hand, in Japan, under the Abe administration, the 'Declaration to be the World's Most Advanced IT Nation' was published in June 2013 (Cabinet Decision 2013). It clearly stated that it aimed to put an end to the 'lost two decades' following

the collapse of the bubble economy in 1991, promote private sector access to public data, and create a political environment for promoting the use of personal data that was expected to provide considerable value in spurring the creation of new businesses and services that use big data. Under this declaration, discussion regarding a new legal framework for balancing the utilization and protection of personal data has taken place intensively in the Japanese Cabinet's IT Strategic Headquarters. The 'Policy Outline of the Institutional Revision for Utilization of Personal Data' (IT Strategic Headquarters 2014) was published, and a draft proposal was compiled for the amendments of the Act on the Protection of Personal Information. Then, through diet deliberations, the Amended Act on the Protection of Personal Information ('Amended Act') was adopted and was made public in September 2015 (in full effect from May 2017).

The Amended Act clarifies the definition of personal information to resolve the ambiguity in the legal interpretation of the scope of personal information. It sets forth the creation of the Personal Information Protection Commission ('PIPC') as an independent authority. It introduces 'anonymously processed information' that can be utilized by data holders without the consent of data subjects. In addition, it imposes a limitation on data transfer to third parties in foreign countries. As preparation for the full enforcement of the Amended Act from May 30, 2017, a variety of legal documents have been created, including a Cabinet Order, commissioner's rules, guidelines, and FAQs.

One appreciable characteristic of the Japanese people is that they generally have a law-abiding spirit. Thus, even if the guidelines and FAQs have no legal binding force, business sectors want the PIPC to prepare the burdensome guidelines and FAQs that present the best practices of business sectors.

2. Constraints of the Legal Protection of Privacy

The abovementioned legal environmental changes are increasing business sectors' expectations regarding the utilization of personal data in their businesses. At the same time, it was pointed out that because of reputation risks and users' serious concerns regarding data leakage, business sectors are hesitant to utilize personal data (IT Strategic Headquarters 2016). Privacy breaches have been a major concern for users of personalized services, not only online web services but also offline real services. Online-to-Offline (O2O) is a new direction for commercial services; however, privacy concerns have become serious due to the expansion of service collaborations. Users have been very concerned when diverted to services with which they were unaware of having any relationship. In fact, some research results (Korolova 2010) have suggested that Internet ads personalized using privacy data leak users' private

information. On the other hand, it has been suggested that the creation of privacy awareness can assist users in dealing with context-aware services without harming their privacy unintentionally (Deuker 2009).

Another issue is the burden of checking on and maintaining privacy policies (Knijnenburg, Kobsa, & Jin 2013). Users must check the privacy policies of a service that is presented by a service provider before using that service. Each service provider prepares a privacy policy for each service, so users must often check on many privacy policies. Furthermore, it is troublesome that users cannot determine or customize the privacy policies for themselves. If a user does not agree with the privacy policy of a service, the user cannot use the service. Many early studies doubt the effectiveness of privacy policies. They state that privacy policies contribute to neither the improvement of reliability for the service provider nor to their service promotion (Berendt, Günther, & Spiekerman 2005; Metzger 2006). Moreover, many users do not read the privacy policy at all (Kohavi 2001, Tuunnainen, Pitkänen, & Hovi 2009).

Solove (2013) suggested that the privacy self-management model cannot achieve the goals demanded of it, and it has been pushed beyond its limits, while privacy law has been relying too heavily upon the privacy self-management model. In his paper, issues involved in giving consent to a privacy policy are identified as related to (1) developing a coherent approach to consent, one that accounts for social science's discoveries about how humans make decisions about personal data and (2) developing more substantive privacy rules. In the Big Data era, consumers show strong distrust towards service providers regarding transparency and data control as consumers feel they cannot virtually control their personal information (Kobsa 2015; Tsai et al. 2011).

An experimental result by Acquisti and Grossklags (2005) shows that there is lack of knowledge about technological and legal forms of privacy protection when consenting to a privacy policy. Their observations suggest that several difficulties obstruct even concerned and motivated individuals' attempts to protect their own private information. Gates et al. (2014) also suggested that users are not familiar with technical and legal terms related to privacy. Moreover, it was suggested that users' knowledge about privacy threats and technologies that help protect their privacy is quite inadequate (Acquisti & Grossklags 2005).The Platform for Privacy Preferences Project (P3P; Cranor 2003; W3C 2002) enables websites to express their privacy practices in a standard format that can be retrieved automatically and interpreted easily by user agents. The project provides user agent modules that allow users to be informed of site practices and to automate decision making based on these practices when appropriate. However, in practice, it is not used by online and offline services due to complex policy definitions, even though some browsers have a module for privacy matching. Furthermore, the module can only be implemented on web browsers.

In this study, we consider an architecture for personalized services and present solutions to privacy problems related to personalized services. The architecture separates data storage from access control based on a privacy policy, and it supports privacy policy management by users. We design a core module named Privacy Policy Manager (‘PPM’) that provides two functionalities: identity management and privacy policy management.

3. Towards Privacy-Preserving Personalization

In this section, we introduce the background of our study and clarify issues that arise in designing the architecture.

3.1 Personal Data Service

A personal data vault has been presented as support for a user-transparent architecture that can control information flow (Estrin et al. 2010). It is a secure container to which only the individual has complete access. It decouples the capture and archiving of personal data streams from the function for sharing that information. The personal data vault would then facilitate the selective sharing of subsets of the information with various services. There are some platforms that manage a personal data vault. An individual can execute the following functions in the personal data vault: controlled push and informed pull. Each platform is managed by a company, so individuals must trust the service provider of the platform. To solve this problem, the concept of the personal data service (‘PDS’) has been presented[1], and some research projects have profited tools for realizing individual-based management of private information.

The PDS is a platform that allows users to control their own information by themselves. It is used for sharing personal data with friends and organizations that are trusted. The PDS holds an individual’s sensitive data such as address, credit card number, and employment details and provides the functionality of user access control. The concept of the PDS is an individual-centric model, meaning that centralized access control by each individual should be provided on the individual’s own terminal. Both the access control mechanism and storage of sensitive data are implemented in

[1] In Japan, the Cabinet Office’s IT Strategic Headquarters are studying the creation of new data flow platforms including PDS, information banks, and data trading markets. PDS is a concept similar to personal data stores. Information banks are almost similar to the concept of National Information Market presented by Laudon (1996). Information banks hold personal data of users and manage data such as money deposits. Data trading markets determine the price of data, match the demand and needs, and guarantee the credentials of data transactions (IT Headquarters 2017).

a program (such as a web browser) on the terminal.

By using PDS, users can securely manage their own information and control data flows of the information. Higgins[2] is a browser extension including modules for PDS, and it supports PDS for browser extension including modules for PDS, and it supports PDS for browser interactions and web client interactions. The VRM (Vendor Relationship Management) project is a research project that aims to provide a platform and tools for realizing a PDS (Searls 2013). The project defines five principles for customers who use privacy-preserving services:

- *Customers must enter relationships with vendors as independent actors.*
- *Customers must be the points of integration for their own data.*
- *Customers must have control of the data they generate and gather. This means they must be able to share data selectively and voluntarily.*
- *Customers must be able to assert their own terms of engagement.*
- *Customers must be free to express their demands and intentions outside of any one company's control.*

On the other hand, there is a problem as an individual must manage all functionalities for protecting and controlling his/her private information. Thus, a more user-friendly architecture is required. We will formalize the issues for personalized services based on the above principles in the next subsection.

3.2 Issues for Personalized Services

There are some issues in the existing PDS services when they are used for handling an individual's private information. The PDS solves some problems outlined in this subsection, but some issues remain for constructing a user-friendly architecture. We should clarify these issues before designing an architecture for personalized services. The four issues are summarized as follows:

1. Current service providers issue their own privacy policies for each service. Users must examine and accept a huge amount of information in multiple policies before even beginning to use the services. In the era of Internet of Things, the existence of these services may be invisible and ambiguous for users. Therefore, it could be almost impossible for users to individually determine their own agreements for a multitude of services.

[2] The Eclipse Foundation (2008). Eclipse Higgins. https://projects.eclipse.org/projects/technology.higgins

2. Generally, privacy policies of service providers are difficult for users to understand. This discourages users from giving their consent (no service use), or users may be obliged to click the consent button despite being unsatisfied with the policy.
3. Users do not understand how their data is gathered and utilized by service providers. As transparency of service providers towards users is decreasing, distrustfulness of users towards service providers is increasing.
4. Users have no way delete their data gathered by service providers. Generally, the right to be forgotten is not guaranteed.

4. Proposal for the Privacy Policy Manager

We have been working on the system design and development of the PPM as a personal data platform to solve the issues mentioned in section 2.

The PPM functions as a web portal, provides access control policies, and includes access control policy and preference databases (DBs) as well as data providing logs (see Figure 1).

・Users set their own privacy preferences through the web portal function. These preference data should be stored in the preference DB. When a user wants to begin a service of Service Provider A, he/she will give consent towards the service contract and privacy policy through the web portal function. At this time, users can choose the personalized policy based on previously set preferences.

・Based on the content consented by users, the PPM creates the access control policy and stores the data in the access control policy DB. The control of the first and secondary uses of personal data are managed by the access control policy. Upon receiving the users' consent towards the service contract and privacy policy of their service, the service provider (Service Provider A) can then make first use of the personal data of users by using the access control function provided by the PPM for receiving the personal data.

・The PPM provides the access control policy to users who inquire about it and determines the possibility of providing personal data to service providers only if it can be provided. The data are then provided through the access control function to service providers (Service Provider A).

When data are transferred to Service Provider A, a transfer log is sent to the PPM, which stores it. This data transfer log is provided to a user through the web portal function. When the data are transferred during secondary data use from Service Provider A to another service provider (Service Provider B), the same access control function provided by the PPM is used.

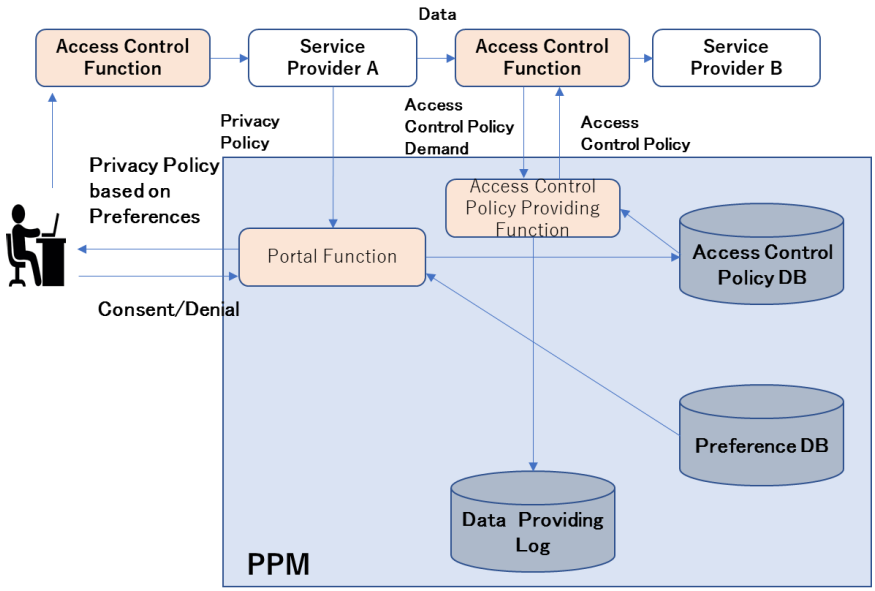

Figure 1: System architecture for the Privacy Policy Manager ('PPM')

We will now describe the overall functions of the PPM for solving the issues mentioned in section 2 as follows:

【Solution for Issue 1】: Integrated Management of Consent Information

The PPM manages user information concerning data transfer consent to each service provider. In accordance with this information, PPM adequately manages data access control and allows the data flow control to be based on user consent. Moreover, the PPM deals with information in an integrated manner and makes it possible to use this enhanced technology for reducing the user's consent management burden.

【Solution for Issue 2】: Comprehensible Display of Personalized Privacy Policy

When a user wants to use a service of Service Provider A, the PPM extracts the source data of the service contract and privacy policy from the service information DB and shows it to a user for getting consent regarding personal data disclosure. At this time, the PPM shows to the user a personalized privacy policy of each service provider based on the user's preferences. The PPM highlights the items of the service contract and privacy policy about which the user is concerned.

【Solution for Issues 3,4】: Management of Personal Data Transfer Log

When the PPM permits personal data disclosure upon request of a service provider, it maintains a personal data disclosure log and allows the user to browse this log. Moreover, in accordance with this personal data disclosure log, the data deleting function, which specifies the data, will be provided to users.

5. Development and Implementation of the PPM

We have been engaged in developing a prototype of the PPM as part of the government-funded New Energy and Industrial Technology Development Organization (NEDO) project from 2012 to 2013. We implemented the prototype on the Home Energy Management Systems' (HEMS) information platform funded by the Ministry of Economy, Trade and Industry from 2014 to 2015.

5.1 NEDO Funding Project

We developed a prototype of the PPM as part of the NEDO Funding Project in 2012 and 2013. During development, we designed the web consent display shown in Figure 4, in which check boxes for choosing whether data should be disclosed are shown at the top of the consent screen, and a button for consenting to the privacy policy is shown at the bottom of the screen. Based on his/her preferences, the user can choose the (1) data items with check boxes, (2) granularity of the privacy policy content, and (3) important items in the privacy policy content.

(1): When a user provides consent to a user agreement or privacy policy of a service provider, the user can control the disclosure of personal data items by checking which personal data are to be disclosed among those demanded by a service provider. We assume the voluntary data and essentially disclosing data. Only if a user checks all the essential data, will the user agreement be deemed lawfully valid. As for voluntary data, a user can agree with the user agreement without checking the voluntary items. In this case, the corresponding data will not be disclosed to the service provider. When a user gives consent to a user agreement and privacy policy, the PPM refers to the preferences set in advance by a user and checks the designated items by default. When a service provider demands disclosure of personal data without the consent of a user, the PPM shows the user a display screen in which the data items are not checked; whether the user chooses to check the boxes depends on the user's judgment. In Figure 2, we show the consent screen for the privacy policy.

We can expect that a user can substantially reduce the consent process burden by checking the expected items in advance and setting preferences.

(2),(3): By improving the Standard Information Sharing Label (SISL) of the Kantara Initiative [8], we have developed display functions for customizing the privacy policy in accordance with the user's preferences. In Figure 3, we show the screen for the implemented privacy policy.

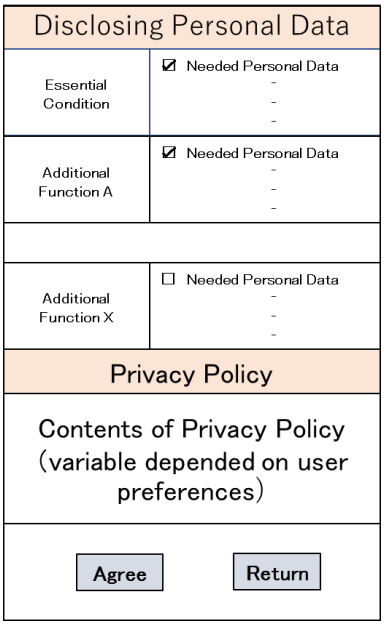

Figure 2: Consent Screen for Privacy Policy

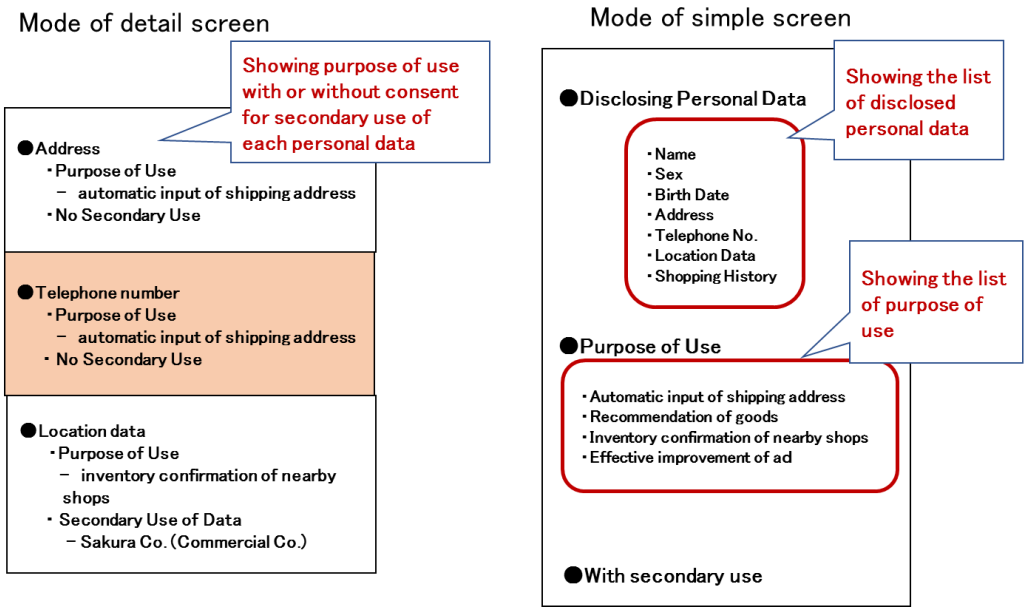

Figure 3: Formation of the Privacy Policy

For evaluating user acceptance of the PPM concept and its implementation as is, we conducted a questionnaire investigation using our developed prototype system. The method of investigation was as follows.

Ⅰ. Prior questionnaire: We asked information on the attributes of subjects, IT literacy, and feelings towards utilization of personal data and privacy policy.

Ⅱ. Explanation of the PPM: We provided a general explanation and introduced the functions of the PPM, including viewing of the PPM demonstration, no real device functioning, viewing of a video on our device handling.

Ⅲ. Post questionnaire: We asked about relief users experienced after using the PPM system and their willingness to use the PPM system.

We conducted the above investigations in November 2013 with 227 subjects and obtained results from 198 subjects using the prior and post questionnaires.

Here, we describe a part of the results of our questionnaires. Regarding the question about users' relief from anxiety about personal data utilization by service providers, we received affirmative responses from about 44% of the subjects, which is twice the percentage of negative responses (see Figure 4). This means that the PPM can relieve users' anxiety about unintended personal data utilization by service providers.

We asked subjects who responded that the PPM relieved their anxiety about the reasons for feeling such relief (see Table 1). 'Very easily confirms the service provider's privacy policy' and 'Easily confirms used data and helps in easy deletion of data' are the most common answers given by over 50% of the subjects. As for commercial viability, the affirmative response rate was 51%. This means that the existence of use needs for commercial realization of PPM. A part of the questionnaire results are shown in Figure 4 and Table 1.

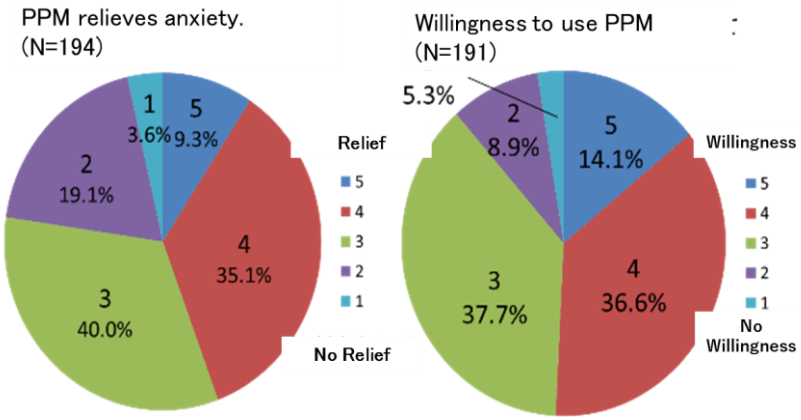

Figure 4: Questionnaire Results (1)

Very easily confirms the service provider's privacy policy	50%
Easily confirms used data and helps in easy deletion of data	48.7%
PPM confirms the discrepancy of own privacy policy and the service provider's policy	42%
Clearly understand own privacy policy	38.7%
Lose the anxiety about unintended use of data	16%

Table 1: Result of Questionnaire (2)

5.2 HEMS Project

Our motivation for designing privacy-preserving personalized services is driven by an online survey (i energy communication booklet 2016) conducted with HEMS users living in and around the Tokyo area; central Japan; and the Fukuoka, Shizuoka, and Fukushima prefectures. Among the 14,000 HEMS users, 6,648 responded to the questionnaire sent to them during August and September 2015.

The questionnaire was focused on the following privacy-related points: (1) which features a user checks while providing private data, (2) whether a particular set of data is sensitive for the user's privacy, (3) the user's feelings about usage of HEMS raw data by third parties, (4) the user's feelings about usage of his/her lifecycle data estimated from HEMS raw data by third parties, and (5) the user's opinion on the trade-off between service quality and privacy.

The survey results are shown in Figures 5-8. In summary, users are not comfortable sharing their personal identifiable information; they prefer checking the features of the data-sharing platform before deciding to share their private information and prefer privacy protection over service quality. Based on the survey results, we decided to design a platform for users through which they can manage the share, access, and usage policies of their private data collected through HEMS devices.

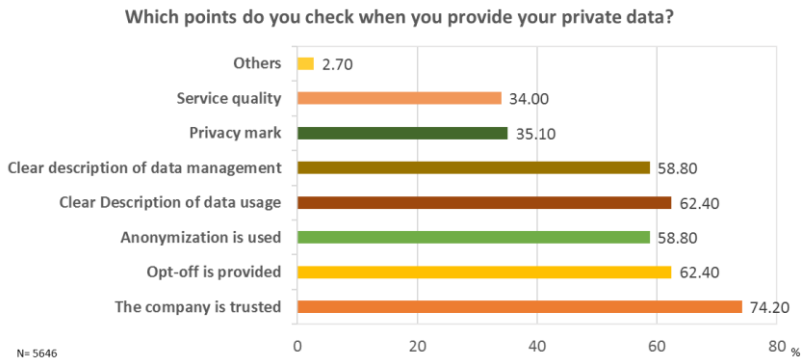

Figure 5: User check points

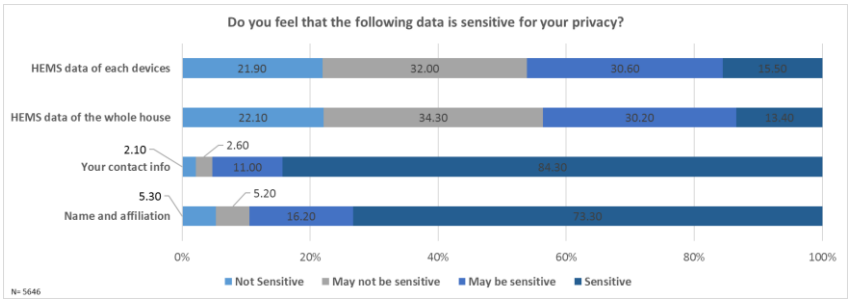

Figure 6: Feelings about sensitive data

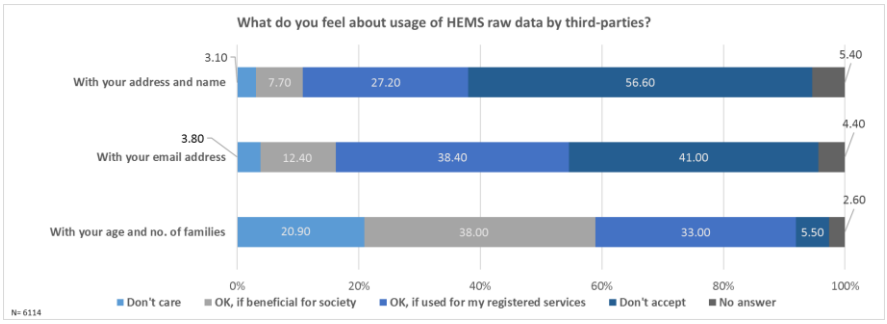

Figure 7: Feelings on HEMS raw data usage

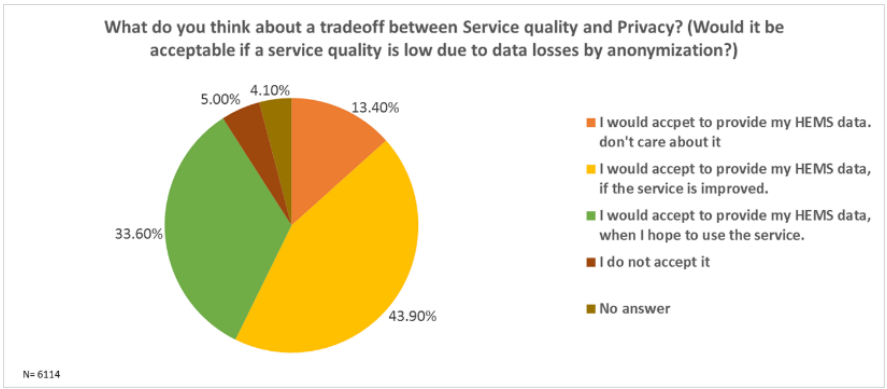

Figure 8: Feelings about service quality and privacy trade-off

6. Discussion and Conclusion

We have presented an architecture for privacy-preserving services and designed a core module called PPM. The PPM supports privacy management by users and acts as a proxy that checks the flows of private information and records them. Our concept involves the delegation of access control and policy management, which are inconveniently complex for users, to a trusted third party, even though our architecture is based on the concept of PDS. Kobsa (2007) suggested that there exists no silver bullet for radically enhancing the privacy-friendliness of personalized systems, neither technical nor legal nor social/organizational.

As stated in section 1, the Amended Act would be fully in effect from May 30, 2017. The Amended Act should be expected to be more proactive towards personal data protection. Meanwhile, issues such as introduction of the individual legal relief clause, privacy impact assessment procedure, and restriction on profiling should be studied till the next amendment three years after the full enforcement.

Therefore, taking a closer look at improving the legal framework, we should continue to verify the complementary effectiveness of the PPM for user-friendly privacy protection. We should expand our PPM implementation fields where the level of protection and data sensitivity felt by users may be changing in accordance with the context. Through these implementations, we can refine the PPM systems. Moreover, we can contribute to academic conferences or standardization organizations for harmonization and development with the other similar systems. Lastly, we should consider the business model of the PPM. How can the PPM be installed in the business flow, and how can it be monetized?

References

Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. IEEE Security & Privacy, 3(1), 26-33.

Andrieu, J. (2012). The Standard Information Sharing Label, Kantara Initiative Draft Report.

Berendt, B., Günther, O., & Spiekermann, S. (2005). Privacy in e-commerce: stated preferences vs. actual behavior. Communications of the ACM, 48(4), 101-106.

Cabinet Decision (2013). Declaration to be the world's most advanced IT nation. http://www.itdashboard.go.jp/en/Achievement/index#100

Cranor, L. F. (2003). P3P: making privacy policies more useful. IEEE Security & Privacy, 99(6), 50-55.

Deuker, A. (2009, September). Addressing the privacy paradox by expanded privacy awareness-the example of context-aware services. In IFIP PrimeLife International Summer School on Privacy and Identity Management for Life (pp. 275-283). Springer, Berlin, Heidelberg.

Estrin, D., Chandy, K. M., Young, R. M., Smarr, L., Odlyzko, A., Clark, D., ... & Hölzle, U. (2010). Participatory sensing: applications and architecture [internet predictions]. IEEE Internet Computing, 14(1), 12-42.

i energy communication booklet (2016). i-consortium Web Page. http://www.ienecons.jp/pdf/ienecons Letterzine Vol4.pdf

IT Strategic Headquarters (2014). Broad outline of regal framework reform for personal data usage. http://www.kantei.go.jp/jp/singi/it2/info/h260625_siryou2.pdf

IT Strategic Headquarters (2016). Investigation commission on data flow environmental consideration: data usage working group in AI & IoT era. http://www.kantei.go.jp/jp/singi/it2/senmon_bunka/data_ryutsuseibi/detakatsuyo_wg_dai1/gijisidai.html

Knijnenburg, B. P., Kobsa, A., & Jin, H. (2013, April). Preference-based location sharing: are more privacy options really better?. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2667-2676). ACM.

Kobsa, A. (2015). Privacy support for users beyond transparency and control. http://cra.org/ccc/wp-content/uploads/sites/2/2015/05/Alfred-Kobsa.pdf

Kobsa, A. (2007). Privacy-enhanced personalization. Communications of the ACM, 50(8), 24-33.

Kohavi, R. (2001, August). Mining e-commerce data: the good, the bad, and the ugly. In Proceedings of the seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 8-13). ACM.

Korolova, A. (2010, December). Privacy violations using microtargeted ads: a case

study. In Data Mining Workshops (ICDMW), 2010 IEEE International Conference (pp. 474–482). IEEE.

Metzger, M. J. (2006). Effects of site, vendor, and consumer characteristics on web site trust and disclosure. Communication Research, 33(3), 155–179.

Miyashita, H. (2012). Privacy Year 2012: international trends of privacy and data protection in Big Data era and Japanese challenges. Nextcom, 12, 32–41.

Schwab, K., Marcus, A., Oyola, J. O., Hoffman, W., & Luzi, M. (2011). Personal data: the emergence of a new asset class. In An Initiative of the World Economic Forum. http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf

Searls, D. (2013). Project VRM–vendor relationship management. Project of the Berkman Center for Internet Society, Harvard University, Boston, MA.

Solove, D. (2013). Privacy self–management and the consent dilemma. Harvard Law Review, 126, 1880–1895.

Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: an experimental study. Information Systems Research, 22(2), 254–268.

Tuunainen, V. K., Pitkänen, O., & Hovi, M. (2009). Users' awareness of privacy on online social networking sites–case Facebook. In Proceedings of the 22nd Bled eConference 2009 (pp. 1–17). Bled, Slovenia.

W3C (2002). The platform for privacy preferences 1.0 (P3P 1.0) specification. In Platform for Privacy Preferences (P3P) Project. http://www.w3.org/TR/P3P/