Managing Personal Information: A Telco Perspective

Ewelina Szczekocka*, Justyna Gromada*, Agata Filipowska†, Piotr Jankowiak†, Piotr Kałużny†, Arnaud Brun‡, Jean Michel Portugal‡ and Jacopo Staiano§

*Orange Polska S.A., Warsaw, Poland
‡Orange Labs Research, Paris, France
†Department of Information Systems, Poznan University of Economics, Poznan, Poland
§LIP6 - UPMC, Sorbonne Universities, Paris, France

***Abstract*—While paving the way for novel and exciting application scenarios, the foreseen large-scale deployment of *connected objects* poses a number of ethical concerns which, if not timely addressed, can potentially dampen the full potential of this drive. In particular, issues related to privacy and control of the data collected by sensors can undermine users' *trust*, hence hindering several application scenarios due to the perceived sensitivity of the data required. In this work, we describe a first prototype which addresses such issues by granting users full ownership and control of the data their sensors produce. The proposed solution - the Personal Information Management platform - aims at *breaking the data silos*, allowing the spontaneous emergence of ad-hoc and decentralized communities and applications, and *gaining user trust* by providing transparency on personal data and its usage with a user-centric approach. Furthermore, we provide a sample use case application based on the platform, enabling social activities of a community supported by data from objects belonging to Internet of Things and personal information inferred from this data. Finally, we discuss current limitations and elaborate on future developments of the described technology.**

I. Introduction

The current wave of commercial offerings regarding the *Internet of Things* and *connected objects* opens exciting possibilities for designing applications tackling a variety of problems in everyday life. For instance, it will be possible to propose, implement, and test novel approaches to issues of high societal value (e.g. energy consumption optimization, food waste reduction). Nonetheless, it can be speculated that the efficacy of such solutions will be highly influenced by human factors: these connected sensors will allow for extensive collection of sensitive and personal information, as the raw data gathered will encode details about people's behavior, habits, preferences, interactions, social activities and so forth. Looking back at the last years, it is clear that the emergence of the "Web 2.0" and the wide adoption of smart-phones have, on the one hand, allowed for significant technological breakthroughs (e.g. Big Data solutions) while, on the other, the dominant business models have relied on offering "free" services in exchange for the commercial exploitation of personal data: for instance, it is very common to implement user profiling in order to serve targeted advertisements.

Although these models are still widely in use, the striking unbalance in terms of effective data control between the data producers (the users) and the data retainers (the service providers) has lately generated several concerns, bringing policy-makers to start dealing with the issues of privacy and data control, researchers to devise possible solutions, and an increasing amount of users to grow awareness on the matter and find ways to protect their privacy. The focus of this work is on personal data of customers and on building a platform which offers full control over usage of personal data in customers' social activities supported by different services. Moreover, this platform enriches social communication services by injecting into them knowledge derived from personal data in a *trust-by-design* fashion. We believe that this approach can create significant value from personal data for both customers and operators. Previous related work focused on modeling customers' relations based on data from their communications – a first step in building a value from their personal data; it soon became clear that this is not enough to provide a complete offer to a customer. Service providers, and particularly Telco operators, nowadays find themselves in the unfavorable position of losing their customers' *trust*. In this work we propose a prototype solution – the Personal Information Management (PIM) platform – able to grant its users (Telco customers) full control on their data, and to allow them to take conscious and informed decisions to e.g. make their data available to other users or services.

Preliminary user studies have shown significant lack of user awareness of privacy risks, e.g. during their internet activities. Hence, the proposed approach is expected to also contribute towards an increase in customers' awareness on the effects of their data-sharing choices, on how their data is used and by whom.

The main contributions of this paper can be summarized as follows:

- we describe the Personal Information Management platform and depict factors like methods, algorithms, tools allowing for building a smart system predicting and prompting to a user different usages based on his social context;
- we report on a first deployed application based on the

platform along with the feedback obtained in the associated user studies;
- we elaborate on current limitations and future developments of the proposed platform.

II. Related Work

In this section, we provide details on research activities related to this work. In order to contextualize the proposed platform, and to highlight the importance of allowing behavioral data collection in a *trust-by-design* paradigm, we first briefly report on works harnessing the potential of big data processing and we highlight how to adopt and use results of this processing from a single user perspective; then, we describe proposed solutions for granting users control on the data they produce.

A. Behavioral Data Research

In order to obtain useful insights on their customer base and provide advanced analytic capabilities [1], Telcos have in the last years heavily relied on user profiling techniques, which proved particularly valuable to design personalized services and modeling customer groups [2]. The profile of a user can be seen as a *machine processable description* [3] of his/her behavior, encoding several facets such as calling/messaging patterns, i.e. the user's network activity and derived information ([4], [5], [6]), and mobility information, i.e. places the user visits, derived by the location of the Radio Base Stations (RBS) his/her mobile phone connects to ([7], [8]). Furthermore, additional information can be derived by the user's position in the graph of Telco customers [9]; The most employed source of data for these studies is represented by Call Detail Records (CDRs)[1], which hold a variety of information[2] allowing also to estimate the user's current location.

Previous works have shown how the mobility of people is characterized by very similar patterns [10], and display distinct motives over depending on the temporal resolution adopted [11]; these findings confirmed the intuition that people tend to visit only a few places on a regular basis, hence further research has focused on defining and identifying such important locations[3] ([12], [13]). Based on these anchor points, several approaches have tried to uncover user daily travels/trips ([8], [14], [15]) by identifying the places where users stopped, those he/she passed by rapidly (probably not having any activity in there), and so on. Moreover, researchers have exploited area labelling and Points of Interest (POI) to estimate the activity of a user in a given place [16]. Accounting for temporal information, it is then possible to derive even more precise information about user's habits and preferences: in [17], the authors investigated the relations between the important locations in users' lives (home, work, other) and their social interactions, hinting at the impact that mobility profiling can have on deriving information on the users social sphere.

Besides focusing on profiling individuals or customer segments, recent efforts have used aggregated mobility information, derived by CDRs or Global Positioning System (GPS) traces, to tackle problems of societal relevance, such as floodings and emergency response ([18], [19]), mapping the propagation of diseases such as malaria [20] and H1N1 [21], and to predict crime hotspots [22]. These works well exemplify the enormous advantages of big data processing for society, and highlight the need of devising data-sharing solutions able to meet the privacy demands of the data producers.

B. Privacy and Data Control

Research conducted in living-lab settings has recently shown the desire of internet and smartphone users of being granted more control on their own data ([23], [24]). In particular, users have been found to associate higher relative monetary value to their location information, hence considering it the most sensitive behavioral trace. Several privacy-preserving platforms have been developed in the last few years: in particular, the OpenPDS [25] provides a "SafeAnswers" mechanism, allowing data access to collectors only at the user discretion (and at a level of aggregation specified by the user); moreover, systems inspired by crypto-currencies have been developed in a decentralized manner, such as Ubiquitous Commons[4] and Enigma [26].

III. The PIM Platform

Under current paradigms, massive amounts of personal data are automatically collected by service providers: users leave their digital footprints in several systems whenever they e.g. perform a web search, install a mobile app, make a call or message someone, post on social media and so on. Surely, all this data can be used to the users' benefit, by e.g. providing an improved, more responsive experience. Nonetheless, a striking unbalance between users and service providers exists: in fact, the former have no power to act on the uploaded data (e.g. to delete it) nor can control the way their data is exploited by the providers and third parties. With the currently increasing public awareness on privacy matters, this issue seems very likely to undermine customers' *trust*: even if a Telco operator would treat its customers data with the highest caution and responsibility, without transparency it would not be perceived differently from other, less cautious, providers, hence not improving its reputation despite its efforts.

In this work, we present a possible solution to this problem, from a Telco perspective. Based on user-centric data services, the *Personal Information Management* (PIM) platform is a technological facility devised to grant users control over their data: the PIM allows them to visualize the data they produced, to allow for selective actions on it (e.g. partial deletions), to control who has access to it and to do what, and to foster the bottom-up emergence of novel applications by promoting data reuse.

[1] Also called "billing logs".

[2] Who, using which service, has communicated with whom, during which time and where the event happened.

[3] Also referred to as "anchor points".

[4] https://m i u. pshuntong.com/ur

Figure 1 provides a graphical representation of the amount and variety of data collected by Telco operators: data coming from service usage (event logs, mobility, contracts, etc.). It should be noted that collection and storage of such data is delegated to several subsystems (*silos*), and can legally exclusively be used for the purposes they were collected. For instance, this means that billing data can only be used for billing purposes in agreement with a contract between a telco operator and a customer, e.g. when a customer agreed for his data to be used for marketing purposes, his profile is harnessed to propose additional offers or service improvements. Usually, these operations rely on statistical approaches and simple customer segmentation.

The scope of PIM, hence, becomes two-fold: on the one hand, it is instrumental for granting customers full control over the data they generate; on the other, by providing them with the choice of opening their data for specific purposes, it can be seen as an enabler for novel applications of societal and/or business value. The goal of the proposed system can thus be summarized as making the Telco operator data space *secure*, *trustable*, *open*, and *social*.

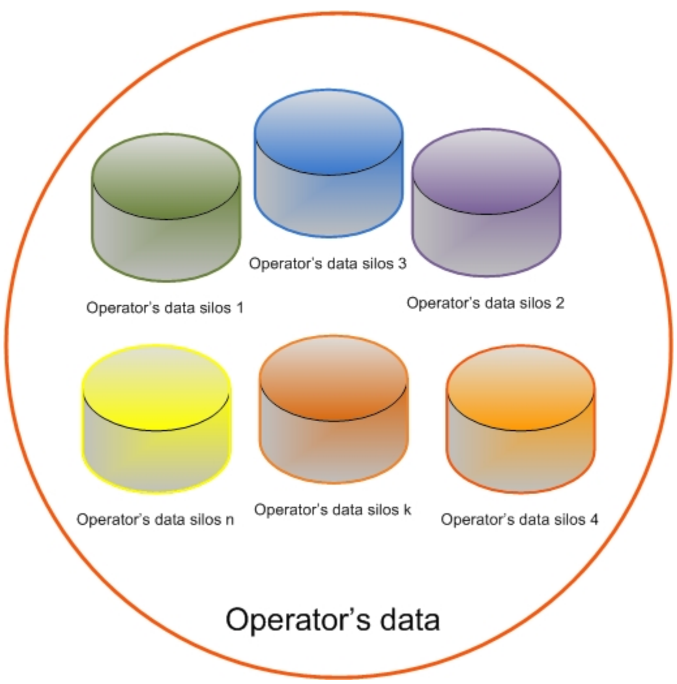

Figure 1. Data collected and protected by Telco operators.

Furthermore, we envision PIM as the platform which will allow Telcos for transition from the model shown in Figure 1 to the one represented in Figure 2. In the latter, all data produced belongs and is managed by the specific customer who produces them: thus, each customer would be enabled to state the following: "I am the owner of data produced by me - my small data belongs to me", "I authorize people, services and applications to view or process all or selected portions of my data", "I let my data being reused when I see my personal or common (social) benefit". A crucial challenge

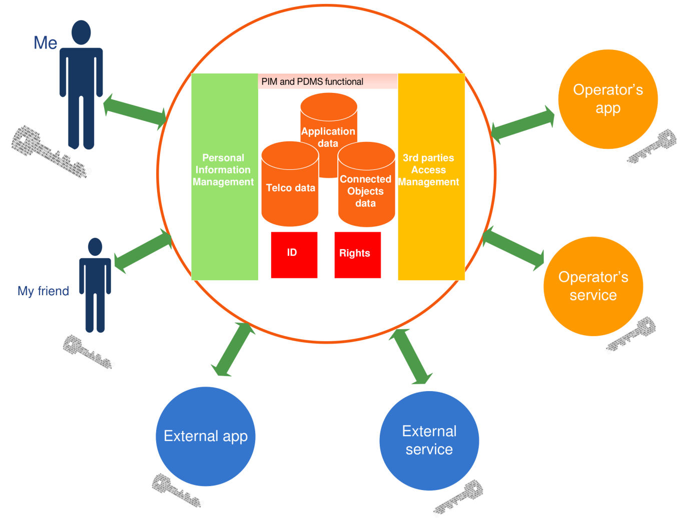

Figure 2. The PIM user-centric data control model.

is therefore represented by the need of finding technological solutions able to reduce the burden of data access settings configuration: the system is required to be smart enough to streamline the personal data management experience. For instance, by recommending services relevant to the specific user, or providing adaptive and personalized interfaces.

A. *PIM Functionalities*

Our goal is to provide each customer with a secure solution that takes care of privacy of his data, extracts knowledge from this data and lets him/her control usage of both his data and the derived knowledge. An important success factor is linked to the customers awareness of the benefits that can derive from their data and appropriate knowledge modeling. Ideally, both the customers and the Telco operators will in the end benefit from this approach.

The most adopted online social networks currently offer centralized infrastructures in the Web [27]. These solutions are said vertical, i.e. they serve users' data only in scope of the same social network. The main disadvantage is the lack of clear rules of personal data and information management: in fact, a user is not aware of how his data is stored and processed, and of how it is used and managed.

Taking into account all of the above, this work is dedicated to elaboration of Personal Information Management, devised as a component of a social communication services platform and as an enabler for social communication services. Previously, we have dealt with the ways of enabling social Telco applications based on customer behavior and relations between customers [28]; in this paper, we focus on personal data aspects.

The PIM lets customers manage their personal data and grants them control over them, and includes a sub-system able to derive knowledge from data. In order to properly collect, store, maintain and transform data, two functional layers were designed, as represented in Figure 3:

- a logical layer, the Personal Information Management, responsible for data processing, knowledge extraction, and management of privacy policies;
- a physical layer called Personal Data Management, which performs data maintenance and operations (e.g. collection, storage, retrieval, physical security) on raw Telco data.

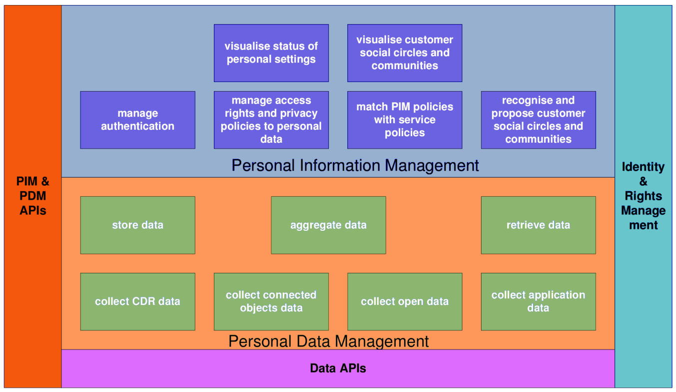

Figure 3. The two-layers architecture of PIM.

An important feature of the sub-system is concerned with ensuring safety of users' data. Personal Information Management functionalities being an integral part of social communication services platform (Social Tool for Telecommunication), are exposed to the external world through a set of APIs. The current implementation provides a set of 22 APIs used to e.g. register a service, retrieve service information, subscription.

In brief, the functional requirements of PIM are:

- to allow user to manage personal data with circles of contacts or communities which govern with own rules of information gathering, sharing etc.;
- to offer smart solutions with a simple interface and let customer manage his/her personal data (e.g. configure privacy settings, check them, confirm system prompts of settings based on user profiling, access visualizations);
- to offer control of different privacy levels based on user data and knowledge derived via customer dynamic profile (i.e. adaptive to the user's evolving interests and needs), social cartography and context (e.g. user relations, usages, location, availability, mood);
- to achieve interoperability among different telco operators' systems.

Usually, a significant part of online social activities consists of sharing of data, information, content which to a great extent has a personal mark. Sharing is in most cases happens through an application allowing the user to expose data to the external world (e.g. friends, classmates, co-workers, students, different groups and communities of people). The subject of sharing can be any kind of customer data (collected or created by him or bound with him by default – e.g. name, address). Specific novel kind of data is related to sensor readings which can include highly personal data (as exemplified by the Quantified Self movement), as well as enviromental and situational measures, which can also be exploited for profiling a particular person's behavior.

Currently, the first implementation of the Personal Information Management system offers relatively simple functionalities concerned with personal data sharing, via different applications to different people belonging to the customer's social circles and communities. Ongoing work will soon allow the introduction of more intelligence to the management of personal information, along with advanced functionalities. Present functionalities of PIM allow:

- to manage authentication of a customer allowing for entering well identified customer into a system;
- to control customer social network visibility;
- to match PIM policies with specific service policies;
- to let a customer set up access rights and privacy policies to his assets (like data, photos, devices) via several services;
- to manage access to customer's data, photos and devices and policies in the context of services and people (e.g. customer contacts' rights within particular services);
- to recognize and propose different social circles and communities for sharing;
- to visualize the status of customer's data access settings;

B. Algorithms, methods and tools for Personal Information Management

For the purpose of introducing intelligence into PIM platform several algorithms, methods and tools should be applied. The customer perspective fits the "small data" perspective elaborated in [29]; nonetheless, Big data methodologies have a significant role for e.g. inferring a customer's background and habits.

In order to provide a customer with personalized and contextual services that can be used on platforms like PIM the raw data needs to be transformed and preprocessed. Basic telecom data is available mostly in form of logs with specific Ids connected to the most important characteristics of a network event (e.g. place, user, service). Data safety and validity is achieved thanks to state of the art algorithms; then, multiple methodologies are adopted to transition from raw call detail records to aggregated data, focusing on different aspects of the logs themselves.

Several algorithms are in the process of being integrated into the platform. These include methods for dynamic behavioral profiling, anomaly detection, social action prediction, and community discovery. By working in the customer's context, such information will be used to provide privacy protection features: for instance the customer will be notified in case of anomalous data-sharing through a given service. One of the algorithms extended, integrated in the platform, and currently under validation is EVABCD [30] for dynamic behavioral profiling. Other methods adopted concern detection of communities based on social strenght of relations, which allow to differentiate contacts between users.

An algorithm that determines home and workplace of a customer based on CDR data can be seen as privacy threatening: nonetheless, a naive solution might consist in labeling such information with unique IDs – disconnected from the users personal data; this way, it is possible to derive higher level knowledge on the customer base while limiting the privacy issue. Following this example, the higher level information derived (where people live and work) can be used in:

Big Data perspective - obtaining information about certain user groups (e.g. identifying students based on POI visits) can give valuable information considering school transportation and providing authorities with suggestion of pedestrian crossings or speed limits on most visited areas.

Small Data perspective - information itself can be used to contextually provide users with offers of meeting people with similar habits and profiles that have agreed for it. For example people living or working in similar areas both use social app focused on certain hobby and receive suggestion of contact. This can be further expanded on with social network analysis providing "friend of a friend" features implemented in Telco network without the need for logging in external apps. Provided with this data it is then aggregated and that there are no ways to identify a given customer, preprocessed data like this aims to represent a real person as much as possible to draw conclusions regarding population based on it.

Many studies have concluded that the standard call logs are a good proxy to infer population characteristics; furthermore, aggregating the data provides us with data sets that can be of a high value for the analysis of:

- Population mobility - traffic, carbon dioxide emission;
- City centered analysis - urban planning [31], use of public transport and its optimization, characteristics of city districts and cross district movement;
- Tourism – identifying tourist movement models and prediction, identifying nationality focused "heatmaps" (e.g. where users from country X mostly visit);

From the academic point of view, several research fields focus on creating models of human behavior, spatial habits or social networks. In these studies, data is used differently from the standard "value oriented" approach (i.e. when user gives his personal data in exchange for personalized advertisements or special offers). This creates opportunities for the user to decide for what purpose his data is analyzed. The social aspect of influencing global research for which data are inaccessible by other means could show users the social benefit of their data preprocessing. On the other hand having easily understandable user information can open new ways of sharing user personal data based on the context. That way user can be more in control of his data, wishing to e.g. share information only when he is at work or only during the time he is travelling between places (for instance, if he wishes to participate only in a transportation study). In this example, the user could be in control not only of the type of the data that is collected, why it is gathered but also when the collection takes place. Such user-centered approach, providing increased transparency on data usage, can encourage users that otherwise would not share their personal logs or other data to do so.

IV. A Use Case: Social Gardening

In this section, we present a first use case implemented using the PIM platform, with the goal of illustrating the platform functionalities: *Social Gardening*, a service to support collective gardening in micro-communities (e.g. between friends, neighbors, or people who share a common interest). Social Gardening is a data-based service: it uses data coming from connected objects (sensors) deployed in a shared garden, information on users location, contacts of people etc., in order to streamline the shared garden management for the users that collectively take care of it. Personal data are thus required: the PIM platform enables the development of these kinds of apps, allowing the reuse of these resources in the service and the sharing of data with selected people for specific purposes.

Thanks to an application, usable from a mobile, a tablet or a computer, gardeners share data provided by connected objects to help take care about plants through social activities. The community can also share content relative to the garden.

The PIM represents the distinguishing feature for the Social Gardening application, and manages information and content sharing in a consistent user-centric fashion.

Figure 4 shows its interface.

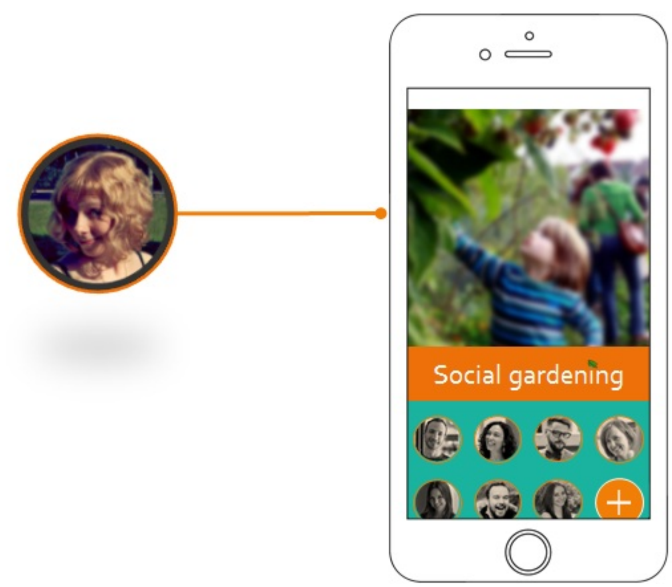

Figure 4. The Social Gardening front-end.

A. Use Case Description

The main functionalities offered by the Social Gardening app can be summarized as follows:

- Real time view of the shared garden: see who is available in the garden, connected objects present, etc.

- Information sharing: know the vegetables' needs (data coming from connected objects equipped with sensors) in real time, check weather forecasts, etc.
- Alert/Messaging: receive alerts and notifications from sensors, find volunteers to take care of the garden, etc.
- Socialize: share content (photos, etc.) and news with co-gardeners, organize social events in the garden, etc.

The gardens managed by the application are real world gardens, where people can grow fruits, vegetables, and flowers in a collective fashion. This use case allows us to bridge the world of mobile phones to the emerging scenarios made possible by the increasing commercial availability of sensors, i.e. the Internet of Things. The application is designed to leverage data coming from the sensors the garden is equipped with (cameras, weather stations, humidity and hydration sensors, etc.). Thanks to these sensors, gardeners can keep track of what is happening in the garden and the micro-community can self-organize to obtain the best results: for instance, if the plants need water those in vicinity can take direct action and communicate in real-time with their peers. Moreover, data reuse from other communities of *social gardeners* can allow for the spread of best practices. Functionalities of the Social Gardening app are summarized in Figure 5.

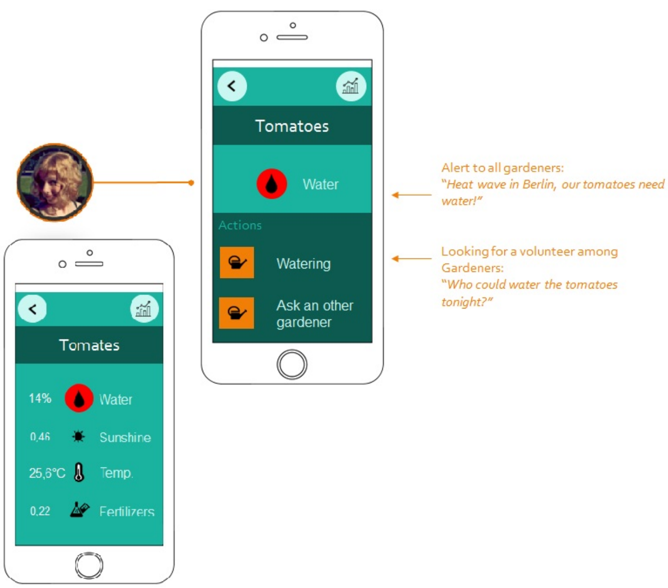

Figure 5. Social Gardening service functionalities.

A use case that can illustrate the need for such service can be the following: Emma, Charles and Léon live in Paris and cultivate together a shared garden in the 20th district. Emma grows tomatoes, Charles some apple trees while Léon takes care of flowers there. When Emma goes for her two-week holidays to London, she can monitor via the Social Gardening application what happens in the garden. For instance, she notices that it is boiling hot in Paris and receives a notification that her tomatoes need to be watered. In this case, she can send a group request for help to his co-gardeners but she can also monitor in real-time who is in the garden (Léon for instance) and ask him directly to take action.

B. User Studies

We ran user studies dedicated to the PIM platform and the Social Gardening service proposal, further evaluations are planned for the next future. Major objectives of the studies were:

- to learn the opinions of potential users on the concepts of new services;
- to determine the factors encouraging people to use the presented services and the possible barriers as to their use;
- to assess the appeal and uniqueness of the concepts;
- to identify potential benefits from the services and the possibilities for their application;
- to gather clues for further work on the services.

The studies were performed in qualitative mode as focus group interviews (FGI), in two sessions. The concept sequence was rotated at each interview; results are summarized below.

A sample of approximately 40 individual customers, divided in two groups, all of Polish nationality, accepted to participate in the user studies; they were selected according to the following criteria: aged 25 to 50; at least with secondary education; regular smartphone/tablet users; interested in technology and open to novel applications; active social media users (often sharing their location, or activities); users of technology for both practical and entertainment reasons; participating in the cultivation of allotment gardens (or house-adjacent gardens). The sample was balanced in terms of gender.

Regarding Social Gardening results can be summarized as:

- the service very well addresses the need for quick and easy access to information about the condition of the plants grown by the user;
- it makes it possible to better organize one's free time (easier decision about the need to go to the garden) or to decide to ask a friend to take care of the plants grown by the user;
- users are worried about the cost of sensors and the risk of them being stolen from the area of an unmonitored garden.

Regarding Personal Information Management can be summarized as:

- for the first group, the service seemed not to evoke major interest, mostly because of limited awareness on privacy issues; they maintained that they did not need a service to manage the sharing of their information; however, after additional questions it turned out that in some situations it was an important aspect for them (e.g. sharing the pictures of their children, HR checking information about them, sharing sensitive pictures via Snapchat;
- the second group of people perceived Personal Information Management quite positively and confirmed interest in this kind of service (offering adjustment of customer

settings with a dedicated interface to Personal Information Management); it must be noted that PIM version presented during this iteration was more mature;
- interviewers noted that while presenting the idea of the service, special attention needs to be paid and emphasis needs to be put on the customers' ability to manage those privacy aspects which are the most sensitive to them.

Interestingly, the results of the user studies about PIM reveal that several respondents seem not too excited about the possibility of managing and having control on their data. Digging deeper into the interviews, though, it becomes evident that the great majority of our respondents were not aware of the reasons why they should care: in fact, as interviewers exposed exemplar situations in which personal data might be misused, they appeared to realize the importance of the problem of privacy and data control. These results seem to point at a lack of information and awareness, rather than a lack of interest.

To summarize, platforms like PIM can only be adopted once at least two conditions are met: 1) users are aware of privacy and data control issues; 2) the platform should provide a streamlined and smart data management solution.

V. Insights to potential business models

Taking into account all the elements presented, in this section we evaluate the business alternative for the introduction of the proposed paradigm into the market: how to monetize the possibility given to people to manage their personal data?

The first possibility, the most classical and obvious that we can call "the Telco option", is to introduce it as a complementary service side-linked with the access subscription. Telcos are already selling access to data networks, hence we can easily imagine that they could introduce additional options to present and secure the personal data. Nonetheless, apart from the feasibility of introducing such a service into the provisioning/billing information system, the Telco will bear the full cost and generate no obvious new value. The willingness to accept an increase by X percent of the price of subscription by the customers is quite uncertain. It can also be assumed that not all the Telcos will deploy the solution without some specific incentive to do it. Such incentives could come in form of pressure from the regulators and policy makers, but this remains to be seen. Moreover, specific applications justifying such a new investment should devised.

Conversely, the second possibility that we can call "the Startup option" consists into growing a specific service that will build a dedicated infrastructure and sell it at a certain price with the promise of simplicity of usage and safety of the data. Once again, the customer is expected to pay some specific cost in exchange for proper and substantial services. In this case, the key success factor from the customer point of view should be further evaluated. From the network side, i.e. Telcos and OTTs, the solution must be seamless and provided through public APIs in order to get a universal impact. We may assume that if any possibility to do it appears, the market will provide many different solutions.

Several other possibilities exist between these two extreme ones. For instance, we could discuss a mixed model, that could be labeled as "the Vertical option". In this model, a specific solution dedicated to one specific activity or interest, such as Health, Books or Banks is developed and provided by a startup, and syndicated by a professional organization, leading to a mixed model:
- a specific offer with a price paid by the final customer;
- a cost shared by the professional using the solution and more or less visible on their invoices to the final customer.

These three examples show that the business model spectrum for PIM is wide and open. We may hence conclude that it provides an incentive to the authors to propose a truly cooperative approach to set up the technical solution, its evaluation, and explore the value created for the different potential stakeholders.

VI. Limitations and Future Developments

The PIM platform presented in this paper aims at providing a technological facility to *break the data silos*, i.e. to allow the emergence of novel services in a horizontal fashion while establishing the customers' rights of ownership and control over the data they produce. In the current and preliminary implementation, data encryption is managed by the Telco operator – hence, the goal of providing customers with full control over their data is only partially met. Several solutions, such as the aforementioned Enigma [26], are in active development and may be considered for integration into the platform to fully achieve this goal.

Future developments of this work will focus on the extension of current functionalities in accordance to the insights derived from the user studies presented. In particular, much attention will be devoted to devise strategies for raising awareness of customers, and to the development and integration of machine learning capabilities needed to ease the burden of managing data access for the end users. Further user studies will adopt a mixed (quantitative and qualitative) methodology to investigate customers' acceptance of the alternative solutions proposed.

VII. Conclusion

In this paper, we described a first implementation of the Personal Information Management platform and discussed its motivations; we reported on a first sample service built on top of it, leveraging connected sensors jointly with mobile phone data, and discussed potential business plans centered on PIM as well as its limitations and further development.

The PIM aims at effectively breaking the data silos currently present in the Telco industry, allowing the deployment of novel services of societal and business value while providing customers with full control over the usage of their data. The presented solution, hence, represents a first step moving from vertical to horizontal solutions *within* a Telco operator, while granting its customers with rights and power on the data they produce. One goal of this paper is also to start a wider discussion *between* Telco operators on the possibilities opened by allowing customers to manage their own data: indeed, a solution like PIM can have a great impact in the Telco ecosystem if this vision is shared among operators. To this end, we plan to put efforts into establishing cooperation with other interested operators, in order to produce a first reference implementation to be adopted by more than one operator. In case of success, Emma, Charles and Léon (in the sample use case above) will be able to use and leverage the same service even in the case they are customers of different Telco companies.

References

[1] P. Russom *et al.*, "Big data analytics," *TDWI Best Practices Report, Fourth Quarter*, 2011.

[2] C. Zhao, Y. Wu, and H. Gao, "Study on Knowledge Acquisition of the Telecom Customers' Consuming Behaviour Based on Data Mining," in *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM'08. 4th International Conference on*. IEEE, 2008, pp. 1–5.

[3] E. Nidelkou, M. Papadogiorgaki, B. Bratu, M. Ribiere, and S. Waddington, "User Profile Modeling and Learning," 2009.

[4] R. A. Becker, R. Cáceres, K. Hanson, J. M. Loh, S. Urbanek, A. Varshavsky, and C. Volinsky, "Clustering anonymized mobile call detail records to find usage groups."

[5] V. Frias-Martinez and J. Virseda, "On the relationship between socio-economic factors and cell phone usage," in *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development*. ACM, 2012, pp. 76–84.

[6] Y.-A. de Montjoye, J. Quoidbach, F. Robic, and A. S. Pentland, "Predicting personality using novel mobile phone-based metrics," in *Social computing, behavioral-cultural modeling and prediction*. Springer, 2013, pp. 48–55.

[7] S. Isaacman, R. Becker, R. Cáceres, M. Martonosi, and I. C. M. Simulation, "Human Mobility Modeling at Metropolitan Scales 2 . Spatial and Temporal Parameters for Mobility Modeling," *Acm*, pp. 239–251, 2012.

[8] P. Widhalm, Y. Yang, M. Ulm, S. Athavale, and M. C. González, "Discovering urban activity patterns in cell phone data," *Transportation*, vol. 42, pp. 597–623, 2015. [Online]. Available: http://link.springer.com/10.1007/s11116-015-9598-x

[9] J. Staiano, B. Lepri, N. Aharony, F. Pianesi, N. Sebe, and A. Pentland, "Friends don't lie: inferring personality traits from social network structure," in *ACM UbiComp*, 2012.

[10] M. C. González, C. a. Hidalgo, and A.-L. Barabási, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008. [Online]. Available: http://www.nature.com/doifinder/10.1038/nature06958

[11] A. Sevtsuk and C. Ratti, "Does Urban Mobility Have a Daily Routine? Learning from the Aggregate Data of Mobile Networks," *Journal of Urban Technology*, vol. 17, no. 1, pp. 41–60, 2010. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10630731003597322

[12] R. Ahas, S. Silm, O. Järv, E. Saluveer, and M. Tiru, "Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones," *Journal of Urban Technology*, vol. 17, no. 1, pp. 3–27, 2010. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10630731003597306{\#}.VaoNNaTtmko

[13] S. Isaacman, R. Becker, M. Martonosi, J. Rowland, and A. Varshavsky, "Identifying Important Places in People ' s Lives from Cellular Network Data 1 Introduction," *Pervasive Computing*, vol. 6696, pp. 133–151, 2011.

[14] F. Liu, D. Janssens, J. Cui, Y. Wang, G. Wets, and M. Cools, "Building a validation measure for activity-based transportation models based on mobile phone data," *Expert Systems with Applications*, vol. 41, no. 14, pp. 6174–6189, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2014.03.054

[15] D. Maldeniya, S. Lokanathan, S. Lanka, A. Kumarage, and S. Lanka, "Origin-Destination Matrix Estimation for Sri Lanka Using 2 . the Four Step Model," no. May, pp. 785–794, 2015.

[16] S. Phithakkitnukoon, T. Horanont, G. D. Lorenzo, R. Shibasaki, and C. Ratti, "Activity-Aware Map : Identifying human daily activity pattern using mobile phone data," *Human Behavior Understanding*, vol. 6219, pp. 14–25, 2010. [Online]. Available: http://www.springerlink.com/index/JJ21508881433584.pdf

[17] M. Picornell, T. Ruiz, M. Lenormand, J. J. Ramasco, T. Dubernet, and E. Frías-Martínez, "Exploring the potential of phone call data to characterize the relationship between social network and travel behavior," *Transportation*, vol. 42, no. 4, pp. 647–668, 2015. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84930821078{\&}partnerID=40{\&}md5=c06377dc8f1b54834cc86cb75222807e

[18] D. Pastor-Escuredo, A. Morales-Guzmán, Y. Torres-Fernández, J.-M. Bauer, A. Wadhwa, C. Castro-Correa, L. Romanoff, J. G. Lee, A. Rutherford, V. Frias-Martinez *et al.*, "Flooding through the lens of mobile phone activity," in *Global Humanitarian Technology Conference (GHTC), 2014 IEEE*. IEEE, 2014, pp. 279–286.

[19] X. Song, Q. Zhang, Y. Sekimoto, and R. Shibasaki, "Prediction of Human Emergency Behavior and Their Mobility Following Large-scale Disaster," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '14. New York, NY, USA: ACM, 2014, pp. 5–14. [Online]. Available: http://doi.acm.org/10.1145/2623330.2623628

[20] A. Wesolowski, N. Eagle, A. J. Tatem, D. L. Smith, A. M. Noor, R. W. Snow, and C. O. Buckee, "Quantifying the Impact of Human Mobility on Malaria," *Science*, vol. 338, no. 6104, pp. 267–270, 2012. [Online]. Available: http://www.sciencemag.org/content/338/6104/267.abstract

[21] E. Frias-Martinez, G. Williamson, and V. Frias-Martinez, "An Agent-Based Model of Epidemic Spread Using Human Mobility and Social Network Information," in *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, Oct 2011, pp. 57–64.

[22] A. Bogomolov, B. Lepri, J. Staiano, E. Letouzé, N. Oliver, F. Pianesi, and A. Pentland, "Moves on the Street: Classifying Crime Hotspots Using Aggregated Anonymized Data on People Dynamics," *Big Data*, vol. 3, no. 3, pp. 148–158, 2015.

[23] J. P. Carrascal, C. Riederer, V. Erramilli, M. Cherubini, and R. de Oliveira, "Your browsing behavior for a big mac: Economics of personal information online," in *Proceedings of the 22nd international conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2013, pp. 189–200.

[24] J. Staiano, N. Oliver, B. Lepri, R. de Oliveira, M. Caraviello, and N. Sebe, "Money Walks: A Human-centric Study on the Economics of Personal Mobile Data," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '14. New York, NY, USA: ACM, 2014, pp. 583–594. [Online]. Available: http://doi.acm.org/10.1145/2632048.2632074

[25] Y.-A. de Montjoye, E. Shmueli, S. S. Wang, and A. S. Pentland, "openPDS: Protecting the Privacy of Metadata through SafeAnswers." *PLoS ONE*, vol. 9, no. 7, 2014.

[26] G. Zyskind, O. Nathan, and A. Pentland, "Decentralizing Privacy: Using Blockchain to Protect Personal Data," in *Security and Privacy Workshops (SPW), 2015 IEEE*. IEEE, 2015, pp. 180–184.

[27] R. Boutaba, "What's next on online social networking?" in *Intelligence in Next Generation Networks (ICIN), 2015 18th International Conference on*. IEEE, 2015.

[28] A. Filipowska, M. Mucha, and B. Perkowski, "Towards social telco applications based on the user behaviour and relations between users," in *Intelligence in Next Generation Networks (ICIN), 2015 18th International Conference on*. IEEE, 2015, pp. 95–102.

[29] D. Estrin, "Small data, where n = me," *Commun. ACM*, vol. 57, no. 4, pp. 32–34, Apr. 2014. [Online]. Available: http://doi.acm.org/10.1145/2580944

[30] J. Iglesias, P. Angelov, A. Ledezma, and A. Sanchis, "Creating Evolving User Behavior Profiles Automatically," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 24, no. 5, pp. 854–867, May 2012.

[31] M. De Nadai, J. Staiano, R. Larcher, N. Sebe, D. Quercia, and B. Lepri, "The death and life of great italian cities: A mobile phone data perspective," in *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2016.