

Trusted, Fair Multi-Segment Business Models, Enabled by a User-Centric, Privacy-Aware Platform, for a Data-Driven Era

Iosif Alvertis, Michael Petychakis, Romanos Tsouropis, Evmorfia Biliri, Fenareti Lampathaki, Dimitrios Askounis
National Technical University of Athens
Athens, Greece
{alvertisjo, mpetyx, rtsouropis, ebiliri, flamp, askous}@epu.ntua.gr

Timotheos Kastrinogiannis, Andreas Daskalopoulos, Theodoros Michalareas
VELTI S.A.
Athens, Greece
{tkastrinogiannis, adaskalopoulos, tmichalareas}@velti.com

Eric Robson, Donal MacCarthy, Christine O'Meara
TSSG, Waterford Institute of Technology
Waterford, Ireland
{erobson, dmccarthy, comeara}@tssg.org

Lukasz Radziwonowicz, Robert Kleinfeld
Fraunhofer FOKUS,
Berlin, Germany
{lukasz.radziwonowicz, robert.kleinfeld}@fokus.fraunhofer.de

Abstract. Today's mobile applications, spanning from file storage and syncing, place checking in and tagging, multimedia sharing, streaming and recommendation, to social and professional networking, rely a lot on cloud-based services, which in turn offer increasingly more functionalities through their publicly available APIs. Users have greeted this abundance of services with even greater demand for innovative applications to address personal and business needs, but are still unaware of the full power that could be leveraged from their data, currently scattered on various platforms. In this complex ecosystem of mobile applications OPENi offers a way to facilitate developers into building applications over the multiple and diverse APIs but also provides users with their own personal cloud storage space, the "Cloudlet", equipped with strong privacy controlling mechanisms, accessed through its open-source Graph API platform. New business models can thus be deployed that leverage the full potential of the available data and metadata, offering application developers the means to create powerful but also privacy-aware services, respecting the requirements of both parties, i.e. service providers and data-owners.

Keywords: APIs, Cloud-based Mobile Applications, Generic Graph API, Context, Personal Data Privacy, Privacy-by-design.

1. Introduction

A new data economy has started to emerge that makes aggregated data valuable, and the conflict is around who owns personal data and how businesses can process or build additional value over this information. Data mining and analytics advancements provide more revolutionary insights than ever to understand and even predict where humans focus their needs, attention and activity at the individual, network and global level. According to the World Economic Forum in 2011 [21], personal data is becoming the new “oil”, a valuable resource of the 21st century that “will touch all aspects of society”. Under this economy companies act as data brokers and get paid for data sets with or without users’ permission; for example, the US data brokerage industry is worth in the region of \$15bn [6]. But this is not the first time doing such activities, traditional brokers exchange customer data, like contact details and demographics for decades, used in phone marketing, market research businesses, or for evaluating prospective customers for their credit risk. The players of note operate globally and include Experian[7] and Bluekai[11].

The main reasons are the rise of web 2.0 applications and then the dominance of smartphone applications and cloud services, which have given boost to multi-segment business models evolved since the web 1.0 era; users’ personal data coming from free-of-charge services are offered through advertising services to business customers who wanted to sell targeted advertisements in the social media age. Nevertheless, there is great difference since web 1.0 services: personal data is fragmented, stored across various services, in walled silos, exposed under proprietary APIs and not available on the broader Internet via web standards. Thus companies like Google and Amazon continue their traditional businesses, but the last decade belongs to companies like Facebook and Twitter which do not contribute on the Web but ask to build mobile, web and desktop applications through their APIs, which keep information fragmented, unrelated and controlled under strict policies that change once a while based on the strategy of the companies.

Nevertheless, new business models addressing the needs and worries of users, as well as new policies introduces, have started changing the market related with personal data. Motivated by this emerging and constantly changing landscape, OPENi [1][2] has designed and delivered a novel consumer-centric, privacy-by-design, open source, cloud-based development platform, serving as a catalyst for new applications era. This paper aims at pointing out a new business model, where end-users own and control their data, developers build applications on a distributed authorization mechanism (i.e. no central authority, thus less business risks) and enterprises can host such services in bundled telecommunication organisations, e.g. cloud hosting provided by carriers.

In section 2 there is a detailed description of the privacy-concert business models emerging lately. Section 3 describes relative work, with other solutions in the data-privacy aware area. Section 4 describes the OPENi solution with its architecture and major components. Finally section 5 has the conclusions of that work and future research.

2. Current landscape

These new business modes and terms for building a platform for new, cloud-based, social-driven and mobile enabled applications, have moved the value away from users who create data and should own them, while companies exploited business ideas of third party developers until they grow their communities. Such approach has created (a) worried, socially overexposed and exploited users who afraid how their data is used, (b) policy makers and data privacy organisations who worry and try to control how such companies use personal data, and (c) frustrated developers who either look to develop their own community with social end-points or work to adapt in any change in enterprises' API terms of use that constantly change.

Relatively to addled consumers, a recently published PEW report [4] revealed that 91% of adults agreed or strongly agreed that consumers had lost control over personal information collected and used by companies. Moreover, 80% of those who use social networking sites say they are concerned about third parties accessing data they share on these services. For that reason customers start turning into new privacy respectful tools, such as the Epic browser [9] or DuckDuckGo [10] search engine. However, individuals are often willing to compromise on privacy in return for benefits; 55% of respondents in the aforementioned PEW research indicated that they would be willing to share some information about themselves with companies in order to use online services for free. Other pain-points experienced by individuals on their online presence include headaches associated with digital identity and asset management and generally online fraud, cybercrime and identity theft.

From the side of active regulators, EU-based data regulation is about to be released and for the first time all member states will be bound by a common data privacy policy. Some aspects include the requirement of consent, required high levels of transparency with explicit communication on all uses/sharing of data, the need to keep data within EU jurisdiction, the right to be forgotten and some hefty fines for non-compliance (up to 2% of annual turnover). At US governmental level, Smart Disclosure[12] is a policy initiative designed to help individuals access personal information in formats they can use and make better decisions based on them.

Table 1. Forces changing typical business models dealing with personal data

Addled Consumer	Active Regulator	Opportunistic Enterprise
<ul style="list-style-type: none"> • Fraud fear • Privacy concerns • Value aware • Digital identity headaches 	<ul style="list-style-type: none"> • Imminent EU regulation • US FTC recommendations 	<ul style="list-style-type: none"> • New services/business models • Enterprise cost to manage data • Leveraging the data asset • External pressures

This shift in individual and public level around data privacy has woken up global players, who try to build on this emerging trend and build more balanced, compliant with the policies business solutions [15]; estimations are that there is a new player

every week in this area [3]. Such services can be categorised according to several types of proposition: (a) Storage & Utility, (b) Transparency & Trust, (c) Marketing & CRM Tools and (d) New World Data Traders. However, the market is still fragmented and there is no clear market leader among the market entrants, as organisations have to solve the multiple challenges including disparate views of the individuals and legacy platform integration challenges.

Thus, legacy business models may be under threat by new, innovative propositions [8]. The 2013 UK MiiData Pilot[5] in this area identified some critical factors for adoption, which are: (a) data must flow, thus incorporate multiple compliant sources, (b) consumer participation is essential, (c) needs-driven approach is better than data-driven focus at early market stage, (d) the proposition should focus on value, not on data, (e) convenience is important, (f) customer control is a benefit in its own right, and (g) trust and safety are of fundamental importance.

3. Relative Work

In this section an analysis of the personal data storage solutions and businesses takes place. Most of the reviewed systems are commercial platforms and some are open source solutions. In general they are middle-ware systems (so called Backend-as-a-Service) or cloud data storages which abstract external APIs and add scalability, management capabilities and other value-adding services to them. The systems act as a proxy and can rewrite request and response formats to create an interoperability layer between the system and the back-end APIs.

Cayova is an emerging social networking site set up to empower its users in the context of data privacy and control. In their privacy policy it states there are no assurances as to the security of those data but that they do their utmost to protect them. All interactions exploit HTTPS to protect users. Transactions are also protected using SSL. On the other hand, FreedomBox chooses a different approach by having a distributed one, and it is completely open source. Gigya is another commercial company that facilitates the gathering, transfer and storage of user data. This commits to not selling or renting personally identifiable information to third parties. In a similar direction, Personal Inc, provides secure storage as a service to its consumers. Similar approaches like Mydex, OwnCloud, Pidder, Qiy or even Privowny try to accomplish the same goals with a variety of similar methods. Most of those are based on standards and open source projects. What is more, in the analysis provided by Abbas [18] and that of Rahimi [19] we can see that they are not that far from the ones we already analysed above. Finally, a really interesting work that is worth mentioning is that of Zhu [20] that makes a considerable effort into the privacy field.

Figure 1 compares personal data storage services similar to OPENi's under a number of criteria. The diagram shows that OPENi is grouped with the companies that allow their users the most control through its authorization mechanism that gives permissions both per object and per instance, it is the most interoperable as it operates under various standards (i.e. Schema.org, JSON LD), and it has the most support for dynamic data as it can be extended with more objects and connections through an API builder. However, the capability for providers to build new applications over a

personal storage, even if the users can control what the developer can see, creates some compromises relatively to the way the developers use users' data; technologically it is not yet possible, and legally blur, to track data across various services and check how they are used.





	least					most
Privacy	CAYOVA OwnCloud	Gigya		Pidder Mydex Personal Privowny	Freedombox	
Control	Gigya OwnCloud		CAYOVA		 Pidder Mydex Personal Privowny Freedombox	
3rd Party Interoperable	Freedombox Pidder Privowny		CAYOVA Gigya Mydex Personal OwnCloud			
Dynamic Data Support	CAYOVA Freedombox Pidder Privowny	OwnCloud	Gigya	Mydex Personal		

Figure 1 - Personal Data Storage Services

4. Personal Cloudlets exposed under a unified API platform

OPENi aims to offer both data owners - thus web and mobile applications users - and developers the means to leverage the power of the vast amount of collected information, in a privacy-aware way. The envisioned "Cloudlet", a personal storage space with fine-grained permission mechanisms, holds a central role in this new mobile applications ecosystem.

As seen from the user perspective, the increasing usage of web services has led to a large amount of personal data, like place visits, images, purchases etc., being created and stored in various web storage spaces, each with its own privacy terms and conditions. OPENi brings all this information under the same roof, with central privacy mechanisms enforced by the data owner. To that end, after careful analysis of many popular APIs and of the data that they exchange, multiple APIs (e.g. Activity API, Media API, Location API, Product and Services API etc.) and relative objects been created to support the large variety of information being stored in the Cloudlets. Contextual information is also stored together with every OPENi object, under a dynamic Context API that extents meta-data attached in every object. The power of this data aggregation is thus revealed to its owner who can better appreciate its

potential, as well as developers who can find rich semantically information to build better applications.

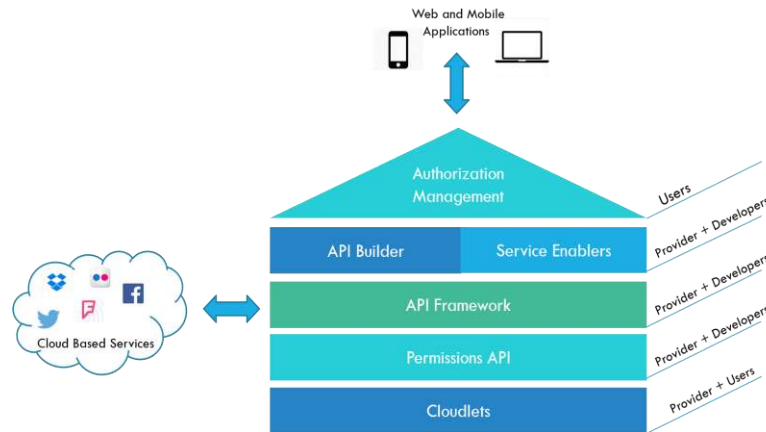


Figure 2. OPENi layering of services and stakeholders

On top of this storage system, a multi-level access control mechanism has been developed, through which users can decide which types of data, or even which data instances, will be accessible by specific OPENi applications. As a privacy-intensive project, OPENi's primary objective is to produce an innovative solution that integrates personal data storage and cloud-based services and gives the end-user maximum control on their data. This requirement is addressed through a number of technical solutions, security and privacy policies: AES and RSA, HTTPS and SSL are deployed to provide transport and storage description; Commercial providers claim different law compliances (mostly US) to their customers, such as HIPAA, PCI, Safe Harbor and ISO27001 compliance as well as different forms of SLAs. While our approach is not to share the focus on US certification, there is a great focus on a compliance with European law. The systems expose a HTTP REST APIs and communicate JSON. OAuth 2.0 is commonly used for access control. However, as OAuth 2.0 is not designed to be a narrow, interoperable standard but rather a framework, the outlined solutions are not interoperable with each other. OAuth 1.0A is supported by some API platforms to access external cloud based services. With the exception of the Intel platform, none of the API solutions target the end-user as a customer. The data belongs to the developer who obtained it. OPENi has a considerably different focus and provides fine-grain access control and detailed audits to end-users to provide them with a user-centric privacy concept.

From the developer perspective, OPENi offers a unified Graph API framework [13] that handles all interactions with the Cloudlets through RESTful services [16], enabling faster and easier application creation. The advantages OPENi has to offer to developers, spanning from its user community, the "Cloudlet" paradigm and the usage of one common unified Graph API to handle all Cloudlet interactions has been discussed in [14]. It is worth mentioning that OPENi also offers its "Builder", a platform addressed to developers, through which they can extend OPENi Objects and APIs, keep versions of their APIs and connect the Objects to other CBS Objects creating mashups, making OPENi sustainable in this rapidly evolving ecosystem [17].

In order to leverage OPENi Cloudlets beyond their storage capabilities, some native OPENi applications, namely the “Service Enablers”, have been developed to show the power of this information aggregation in a privacy-aware environment. It is expected that both users and application developers will benefit from the provided Service Enablers, which will function as trusted agents, enabling smarter applications to be built on top of OPENi, following its transparent privacy rules. Four of the Service Enablers (SE) are briefly discussed:

1. The Advertising SE aims at enabling advertisers, under partnerships with service providers, to use OPENi opted-in users’ anonymised, personal data in order to enable proficient mobile marketing audience management, as well as targeting optimization and personalization in advertising.
2. The Recommender SE is an extension over the central OPENi architecture that allows developers to build applications enhanced with recommendations natively provided by OPENi, without violating users’ cloudlet privacy policy.
3. The Analytics SE provides insights into developers about transactions and usage of their API-enabled applications, relative to OPENi calls and user demographics.
4. The Timeline SE is to provide a common interface for organising user data, enabling the adding, updating, retrieval, and deleting of user data related to time

5. Conclusions

Web and mobile application users are getting progressively accustomed to the idea of having access to their data, regardless of their location or the device they are using. This has caused the dispersion of valuable user information across multiple cloud storage spaces, but also a proliferation of cloud based services that has revolutionized the way these applications address personal and business needs. OPENi brings to both data owners and application developers the Cloudlet, a unified cloud storage space, designed to hold heterogeneous information, leveraging the power of aggregated data in a privacy-aware manner. OPENi makes the data owner and not the service provider in charge of the permission rules enforced on his data. Service providers, through the application developers, can benefit from a more open and trustful relationship with their users, building more powerful applications with the provided unified Graph API. OPENi greatly depends on the idea of communities, both user and developer-wise, so the challenge lays ahead to bring people into this ecosystem. To that end, training platforms will be deployed, courses will be offered and multiple other disseminations actions, including Hackathons, will be held. As another future step, the proposed OPENi ecosystem, will be deployed and tested with other data sources, targeting alternative stakeholders outside of the mobile applications market; thus governments or federal organisations, even enterprises which may use such solutions to improve and make their IT more open and transparent.

Acknowledgments. This work has been created in the context of the EU-funded project OPENi (Open-Source, Web-Based, Framework for Integrating Applications with Social Media Services and Personal Cloudlets), Contract No: FP7-ICT-317883.

References

- [1] OPENi Project. OpenSourceProjects repository. <https://opensourceprojects.eu/p/openi/>
- [2] I. Alvertis, T.Kastrinogiannis, R.Kleinfeld, E.Robson et. all, "OPENi Graph API Framework in the Social Standards Landscape," *position paper in W3C workshop on "Social Standards: The Future of Business*, August 7th-8th, 201, San Francisco, USA.
- [3] <https://www.ctrl-shift.co.uk/research/product/88>, accessed online April 2015
- [4] www.pewinternet.org/2014/11/12/public-privacy-perceptions/, accessed online April 2015
- [5] Innovation Opportunity. Learnings from the midata Innovation Lab <https://www.ctrl-shift.co.uk/research/product/81>
- [6] Datacoup. A personal data marketplace. <http://datacoup.com/docs#about>
- [7] Experian. <http://www.experian.com/> , accessed online April 2015
- [8] Free is a Lie. Aral Balkan. TNW Europe 2015 presentation. <https://www.youtube.com/watch?v=upu0gwGi4FE>
- [9] Epic Browser. <https://www.epicbrowser.com/> , accessed online April 2015
- [10] DuckDuckGo. <https://duckduckgo.com/>, accessed online April 2015
- [11] BlueKai. <http://www.bluekai.com/> , accessed online April 2015
- [12] Smart Discloser Policy. <https://www.data.gov/consumer/smart-disclosure-policy> , accessed online April 2015
- [13] Alvertis I., Petychakis M., Lampathaki F., Askounis D., Kastrinogiannis T. " A Community-based, Graph API Framework to Integrate and Orchestrate Cloud-Based Services." AICCSA. 2014.
- [14] Petychakis, M., Alvertis I., Biliri E., Tsoouplis R., Lampathaki F., Askounis D. "Enterprise Collaboration Framework for Managing, Advancing and Unifying the Functionality of Multiple Cloud-Based Services with the Help of a Graph API." Collaborative Systems for Smart Networked Environments. Springer Berlin Heidelberg, 2014. 153-160.
- [15] Feuerlicht, George, and Hong Thai Tran. "Enterprise Application Management in Cloud Computing Context." Proceedings of the 16th International Conference on Information Integration and Web-based Applications & Services. ACM, 2014. APA
- [16] Fielding, Roy Thomas. "Architectural styles and the design of network-based software architectures." 2000.
- [17] Tsoouplis R., Petychakis M., Alvertis I., Biliri E., Lampathaki F., Askounis D., "Community-based API Builder to manage APIs and their connections with Cloud-based Services". CAiSE Forum. 2015.
- [18] Abbas, Assad, and Samee U. Khan. "A Review on the State-of-the-Art Privacy-Preserving Approaches in the e-Health Clouds." Biomedical and Health Informatics, IEEE Journal of 18.4 (2014): 1431-1441.
- [19] Rahimi, M. Reza, et al. "Mobile cloud computing: A survey, state of art and future directions." Mobile Networks and Applications 19.2 (2014): 133-143.
- [20] Zhu, Hengshu, et al. "Mobile app recommendations with security and privacy awareness." Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014.
- [21] Schwab, K., Marcus, A., Oyola, J. O., Hoffman, W., & Luzi, M. (2011). Personal data: The emergence of a new asset class. World Economic Forum. Retrieved on May 12th, 2015 from http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf