Privacy by Design and the Emerging Personal Data Ecosystem

Ann Cavoukian, Ph.D.
Information & Privacy Commissioner
Ontario, Canada

Foreword by Shane Green
CEO of Personal

October 2012

Acknowledgements

The Information and Privacy Commissioner of Ontario, Canada, would like to gratefully acknowledge the contributions of the following individuals whose efforts were invaluable in the drafting of this paper: Michelle Chibba, Director of Policy and Special Projects, IPC, and Policy Department staff; Josh Galper, Chief Policy Officer and General Counsel, Personal; Drummond Reed, Respect Network; Alan Mitchell, Strategy Director, Ctrl-Shift; Claire Hopkins, Marketing and Communications Director, Ctrl-Shift; and Liz Brandt, CEO, Ctrl-Shift.

We also appreciate the opportunity to co-launch this paper with the Society for Worldwide Interbank Financial Telecommunication (SWIFT) and acknowledge their contribution to the case study section. We would especially like to thank Peter Vander Auwera, Innovation Leader, SWIFT, and Pierre Blum, Senior Product Manager, SWIFT.

Information and Privacy Commissioner
Ontario, Canada

2 Bloor Street East
Suite 1400
Toronto, Ontario
M4W 1A8
Canada

416-326-3333
1-800-387-0073
Fax: 416-325-9195
TTY (Teletypewriter): 416-325-7539
Website: www.ipc.on.ca
Privacy by Design: www.privacybydesign.ca

TABLE OF CONTENTS

Foreword

Samuel Taylor Coleridge's famous line from "Rime of the Ancient Mariner" – "water, water everywhere … nor any drop to drink" – could easily be updated for the 21st century when it comes to the state of data about people and our lives: "Data, data everywhere, but not a bit (or byte) for me." Simply put, our data is everywhere, yet there is no place where we can easily find or use it.

How is this possible, when we live in a Big Data world where the data we create or that is captured about us grows at an exponential rate that exceeds Moore's Law? Just about everything that we do – from mundane form filling to using the latest mobile devices and apps – generates data. Companies certainly understand the importance of our data and capture it for their own use, mining value from it every day. Even so, they struggle to make sense of it all, hoarding it in silos that, ironically, greatly limit its ability to be used.

Now, imagine breaking down these silos, liberating the data, and bringing it together in a secure place where a person could easily access this information and decide how best to use and share it. Then imagine letting the smartest developers build apps on top of such permission-based data, so that people could harness its power to radically improve their lives, on their own terms. The potential use cases are as unlimited as the data the person might store in his or her vault and the personal, private networks they will wish to create when they push their data into the world beyond their vault.

For example, analyzing information about our own purchases, payments and habits could make us smarter financially. Mashing together health records with dietary and exercise tracking could make us healthier. We could use our information to signal purchase intent directly to retailers, earning us better offers and allowing our favorite stores to create more satisfying and welcoming relationships with us. From a productivity perspective, we could use the data in our vault to literally make form filling obsolete and reclaim tens of billions of wasted hours annually across the globe.

There is no need to imagine such services. Technology makes such opportunities possible. Only inertia and certain business practices stand in the way.

A new personal data sector is taking root and creating user-centric, user-driven tools to give individuals more control over their own data. Start-ups focusing on personal data vaults and the safe, private exchange of data form a vibrant core of this industry. As evidenced by the World Economic Forum's May 2012 Rethinking Personal Data project report, the largest companies in the world as well as governments have become keenly interested in this sector.

Putting users at the center and in control of their own data is not new, but it is an idea whose time has come. This is partly because the current path, as the World Economic Forum report highlights, is inherently unsustainable.

More than ever before, the privacy and security practices of companies and governments are front-page news, and regulators and politicians are scrutinizing them closely. Seeing the well-publicized triumphs and tribulations of the largest search engines and social networks, people are starting to wake up and ask tough questions about privacy, transparency, security, and why they lack the power to use and benefit from their most personal of assets – information about themselves and the people, places, things and activities in their lives. Companies have also begun to realize that today's online advertising model is dysfunctional, inefficient and alienates the customers and dollars they wish to attract and retain.

Personal sits in the center of this emerging personal data ecosystem. We are the first commercially available platform to give individuals the ability to securely import, store, share and reuse all the important data, notes and files in their lives through a vault and personal network connecting them to trusted people, organizations and apps. An end-to-end solution for personal data, Personal is simultaneously the vault and network of nodes that allows people to share and benefit from their data however they choose. Other companies are emerging in this space, focusing on all parts of the personal data spectrum – data vaults, networks, identity management, and other areas. We can barely imagine all of the amazing benefits that will come from this user-centric data world, and we are excited to help make it a reality.

If this new sector is to succeed and gain user trust, the companies in it must adopt *Privacy by Design* (*PbD*) principles in their technology and business practices, as described in this paper. And I can think of no better or more esteemed authority to author this study than Dr. Ann Cavoukian and her team. Dr. Cavoukian's coinage of *PbD* and long record of leadership on privacy as a scholar, advocate and regulator are renowned and well deserved.

The result of their work is a pioneering examination of the opportunities and challenges of the personal data ecosystem. The paper serves to put sharper edges around this emerging category, and I believe it provides a lasting and important intellectual cornerstone for its development. I think you will agree.

Shane Green

Co-Founder and CEO
Personal

Introduction

The collection of personal information in our vastly networked world has grown by several orders of magnitude. Indeed, personal data is being viewed as "the new oil of the Internet and the new currency of the digital world."[1] Much of the data collected is intended to enhance user experience through new services, efficiencies, convenience, etc. Although it is widely recognized that personal data can be used to create economic and social value, some service providers and third parties believe they are the ones who should be controlling our data, rather than serving as its dutiful custodians. Over time, this will contribute to a lack of transparency and erosion of our privacy.[2]

Consumers are very aware of the changing nature of online privacy. In fact the concern that the erosion of privacy causes consumers ranks second only to the fear of another financial crisis.[3] The Internet presently lacks an identity layer and this causes digital identities to be fragmented and difficult to exert control over[4]—this, despite the fact that it is estimated that an average individual releases over 700 items of personal data per day.[5] In turn, organizations extract, store, and profit immensely from this data. Unfortunately, this often means that the relationship between individuals and organizations suffers from distrust largely due to the individual's inability to play an active role in controlling and even profiting from the use of their own information. Although there are long-standing discussions regarding the "ownership" of personal data, this paper avoids that debate and instead focuses on the technologies and initiatives associated with the Personal Data Ecosystem (PDE)—a collection of tools and initiatives aimed at facilitating individual control over personal information.

This paper will describe the systems and initiatives that are driving the development of the PDE, and how they seek to address the challenge of protecting and promoting privacy, while at the same time, encouraging socio-economic opportunities and benefits.

Although the definition of informational privacy will differ among jurisdictions, the essence of privacy relates to the ability of individuals to have control and freedom of choice about the collection, use and disclosure of information about ourselves—what we might call our personal data flows. The lack of transparency and accountability regarding data flows is a major factor contributing to consumer privacy concerns. So much of our data is

1 Meglena Kuneva, European Consumer Commissioner, March 2009 "Personal data is the new oil of the Internet and new currency of the digital world."

2 Mike Swift, "Battle brewing over control of personal data online," *San Jose Mercury News* July 4, 2011.

3 An October 2011 McCann Worldgroup Company report "The Truth About Privacy: What every marketer who handles consumer data should know" notes that 70% of people worry about the erosion of personal privacy, while 78% worry about a further global financial crisis.

4 Information and Privacy Commissioner of Ontario. "7 Laws of Identity: The Case for Privacy-Embedded Laws of Identity in the Digital Age." (2006).

5 Ctrl-Shift, "Personal Data Stores: A Market Review," (2012), 8.

under the control of service providers and other third parties who seek to create value in what is being referred to as 'the age of Big Data'.[6] The challenge we face is protecting and promoting individual privacy while at the same time allowing for the socio-economic opportunities and benefits derived from the permissioned contextual use of our personal information.

The purpose of this paper is to look at these developments from a privacy perspective and identify how *Privacy by Design* (*PbD*) is essential to its success. The expectations of what constitutes a healthy privacy-protective relationship between individuals and organizations is being reset under the umbrella of the emerging PDE. The PDE is being supported by new technologies, such as the Personal Data Vault (PDV), that allow individuals to control and manage their own information. For example, the same item of personal data that individuals currently have to share manually multiple times with different organizations, such as their email address or phone number, could be stored once in their PDV and then shared selectively with organizations of their choice with a single click.

This paper will describe the systems and initiatives that are driving the development of the PDE, and how they seek to address the challenge of protecting and promoting privacy, while at the same time, encouraging socio-economic opportunities and benefits. By giving individuals: a) more explicit control over the sharing of their personal data online, and b) new trust frameworks that raise the collective expectation of how companies and organizations will respect an individual's right to control their personal data, the rise of the PDE can be the biggest leap forward for personal privacy on the Internet since the advent of the privacy policy. To assist with the challenges and possible privacy issues arising from the creation of PDE, the experience garnered through the application of *PbD* to systems relevant to the PDE will be explored as guidance for those working in the PDE field. The paper also explores the implementation of *PbD* within the PDE through two case studies focusing on two different platforms for personal data. The first is about Personal.com, including an analysis of the company's application of *PbD* to its service. The second is about the Respect Network and SWIFT's Digital Asset Grid.

Current approaches to privacy have resulted in a toxic battleground over personal data where organizations want more and more of it and try harder to get it, and where individuals feel increasingly vulnerable, intruded upon, and unwilling to share information. Left unchanged, this will result in an adversarial trust and opportunity-sapping stalemate.

The new PDE resolves these issues to generate a new 'win-win'. When individuals are beneficiaries of the use of their data and confident it will not be abused, they are more willing to contribute to economic and social activities as active information-sharing partners. This is the only way to realize the economic and social potential of the digital age.

In this paper we focus on the positive potential of these developments. We are, of course, acutely aware of the many ways 'a good idea' can be lost in bad implementation, distorted and undermined by vested interests, and generate new problems and abuses. We will return to these issues in future papers.

6 Information and Privacy Commissioner of Ontario, Jeff Jonas. "*Privacy by Design* in the Age of Big Data." (2012).

The Personal Data Ecosystem

The Personal Data Ecosystem (PDE) is the emerging landscape of companies and organizations that believe individuals should control their personal data, and who make available a growing number of tools and technologies to enable this. Aside from legal requirements, the starting premise of the PDE is that individuals control the sharing of their own "official record," (also called a "golden record") and set the rules as to who can access and use their personal information for what purposes. In this way the individual becomes the central point of data integration, and individuals always have the ability to extract their data and take it wherever they wish.

There are three essential elements for the PDE to function effectively:[7]

- **Technology to control personal data:** this technical component includes the storage, access control, security, authentication, and user interface which provides the user with a set of features to manage their data and identity.
- **Supporting factors:** data management standards, communications standards, trust frameworks (rules and processes for the setting, monitoring and enforcement of permissions in permission-based information sharing), interoperability among various stakeholders and third parties, legal interoperability, regulatory requirements and settings – all of these lay the foundation of a reliable and scalable PDE.
- **Stakeholders:** the main stakeholders in the emerging PDE are individuals who are increasingly managing and controlling their own personal data for their own purposes; specialist services such as PDVs helping individuals to do this; governments and other public organizations; and private sector businesses that wish to begin sharing and using information with individuals on a new trust basis.

The PDE trend is in contrast to the status quo where personal information is collected and stored by many different applications or service providers rather than by the individual (Fig. 1). This status quo is largely asymmetrical. Individuals often do not know who, within the service provider, is able to see the data uploaded. In addition, the individual has little, if any, control over the uses of one's personal information by service providers, such as for behavioural advertising purposes. The value exchange is also uneven – the service provider is able to profit from the use of an individual's personal information, while the individual usually does not.

7 See World Economic Forum. "Personal Data: The Emergence of a New Asset Class." 2010; World Economic Forum. "Rethinking Personal Data: Strengthening Trust." 2012.

The PDE stands in contrast with the model used by Facebook, Google, eBay, and many others in which users' data exists in a distributed and silo format. While each of these online services collects, manages, and utilizes user data in a highly organized, systematic, professional and technologically advanced way to meet their organizational objectives, individuals have fallen behind in terms of having access to coordinated, advanced data management technology. The current model is fragmented and inefficient for the following reasons:

1) Users have limited control over the management and usage of their data;

2) Each of the services has different privacy policies and settings;

3) Users have little or no ability to manage privacy settings;

4) Each online service requires users to establish a password and a user name, as well as to share detailed personal information (e.g. address, date of birth) specifically for that service, duplicating sensitive data and increasing exposure risk;

5) Each online service must independently attempt to verify the online identity of the user to their required level of assurance; and,

6) Each online service ends up with only a partial view of each user, leading to guesswork, error and waste.

By contrast, in the PDE, the user is in control as the main player. This means the user has the ability to manage his or her personal information; create a single, integrated view of one's behaviours and activities; provide identity and claims verification; selectively share this view with the organizations of one's choice; better use one's information as a tool; receive personal information handed back from organizations; use analytics applied to one's information to spot trends; communicate and share opinions and views with others (e.g. P2P product reviews); and set priorities and planning for different aspects of one's life (e.g. getting married, planning for retirement).[8] In sum, individuals would be able to analyze and set priorities or constraints based on their own behaviour, life goals and events, rather than having to accept analysis and priorities set by others.

8 Ctrl-Shift (n 5).

Status Quo vs. PDE

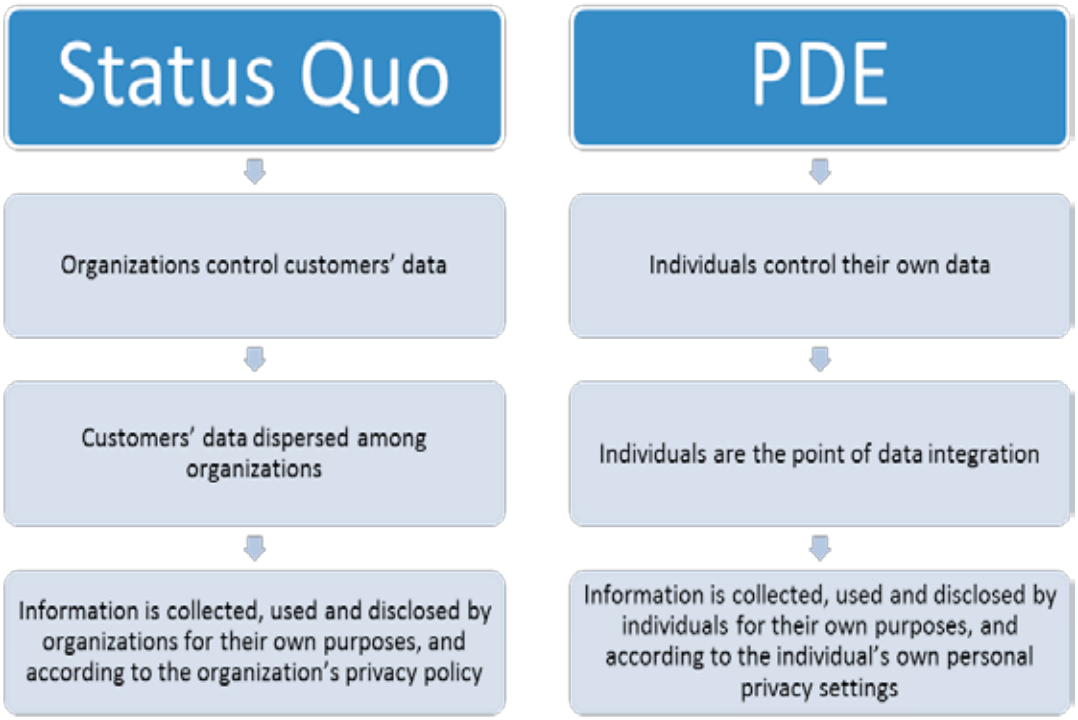

Figure 1 – Status Quo vs. PDE

Technological, commercial, and government factors are also driving the development of the PDE. Technology processing power has increased, while the cost of processing information has decreased. As a result, it is now possible for individuals to have the same sophisticated data generation, management, sharing, and analysis tools as organizations. Since the advent of Web 2.0, information generated and sent out by individuals outstrips the information sent out by organizations to individuals.[9]

There is a rise in the view that personal information is a resource akin to currency, and that individuals and service providers aligned with their interests are now in a position to profit from personal data sharing.[10] Granted, there exists a long-standing discussion surrounding the ownership of personal data.[11] While these discussions are important, our focus in this paper is on the technologies and initiatives associated with PDE rather than proposals for its legal framework.

9 Ctrl-Shift. "The New Personal Data Landscape." 2011, p. 6.

10 E.g. the World Economic Forum has referred to personal information as a "new asset class." See generally WEF Report 2010 (n 7). Also, European Consumer Commissioner Meglena Kuneva dubbed personal data "the new oil of the Internet and the new currency of the digital world." Kuneva (n 1).

11 See generally Ann Cavoukian, and Don Tapscott. "Consumer Privacy: What You Should Know." Chap. 7 In *Who Knows: Safeguarding Your Privacy in a Networked World.* Toronto: Random House of Canada, 1995. See also Laudon, Kenneth C. "Markets and Privacy." *Communications of the ACM* 39, no. 9 (1996): 92, and Samuelson, Pamela. "Privacy as Intellectual Property?". *Stanford Law Review* 52 (2000): 1125.

The commercial impetus for the development of the PDE is captured by the World Economic Forum in two reports issued as part of its "Rethinking Personal Data"project.[12] The first one was issued in 2010 and titled, *Report about Personal Data: The Emergence of a New Asset Class.* It identifies several PDE "key imperatives for action" including the need for the private sector and policy-makers to invest money in trials to establish the necessary trust frameworks, as well the need for greater dialogue between regulators and the private sector. The second report was issued in 2012 and titled, *Rethinking Personal Data: Strengthening Trust.* It makes additional recommendations about developing consensus regarding the rules for obtaining individuals' trusted and permissioned flow of data in different contexts.

The Personal Data Ecosystem Consortium is an industry association which supports a user-centric model across its membership of over 30 companies,[13] including Personal.com, Reputation.com, and the Respect Network.[14] Another organization, Ctrl-Shift, provides market research and intelligence that assists organizations to see the world through the eyes of the customer.[15] They forecast the impact – and opportunities created – by customers for organizations, and believe that empowering consumers is a fast-growing business opportunity. They hope that their work will help businesses to adopt a truly customer-centred perspective so that this shift in control will benefit everyone – not just buyers and sellers, but the economy as a whole.

Respect Network Corporation, based in San Francisco, is establishing the 'Respect Network' as a trusted personal data network.[16] Unlike centralized social networks, the Respect Network is a decentralized, multi-provider network much like today's email or banking networks. Partners in the creation of the network include Kynetx, Gluu, The OpenXDI Project, Project Danube, The Customer's Voice, Planetwork, and Bitworld. Recently, Neustar and Swisscom have also joined as founding partners along with four companies offering user-centric identity verification technologies – Miicard, BioID, Virtrue, and TrustCloud. All members subscribe to the Respect Trust Framework, which is a new model for personal data sharing listed with the Open Identity Exchange.

12 World Economic Forum, "Rethinking Personal Data," http://www.weforum.org/issues/rethinking-personal-data/.

13 Personal Data Ecosystem Consortium, http://pde.cc/. Startup circle members are: (2012) Allfiled, bitWorld, Cloudstore Technologies, Consumer Marketing Rights, COMRADITY, Lifedash, Metaconnectors, PIB-d Ltd, Planetwork, Privowny, Tangled, Virtrue; (2011) Azigo, Buyosphere, Connect.me, archify, Gluu, Kynetx, Mydex, MyINFOSAFE, Peercraft, Personal InfoCloud, Personal, Privo, Project Danube, Qiy, Reputation.com, Singly, Synergetics, SwitchBook, The Customer's Voice.

14 Reputation.com's CEO Michael Fertik had previously predicted an emerging "privacy economy" in which several companies advancing onto the market offer products and services for individuals to better control their data, and early indications show there is much business interest in allowing individuals to control their data.

15 Ctrl-Shift. "A control shift is underway," online: http://ctrl-shift.co.uk/

16 The Respect Network. http://respectnetwork.com.

Governments are also taking steps to hand data back to their citizens, and are encouraging other organizations to do the same. Over the last several years, the U.S. Government has launched an open data initiative to release mass data sets to the public and individual records to individuals about such things as health care, energy consumption and education. For example, the U.S. Department of Veteran Affairs created the Blue Button initiative to allow veterans to download their personal health record as a text file or PDF. It is anticipated that the Blue Button concept will be expanded to additional U.S. departments.[17] The Green Button initiative follows along with the common-sense view that consumers should have access to their own energy usage information in a downloadable, easy-to-use electronic format, offered by their utility or retail energy service provider. In September 2011, U.S. Chief Technology Officer Aneesh Chopra challenged utilities across the country to participate in the "Green Button" initiative.[18] California utilities, for example, have already implemented this functionality.[19]

Under the project title 'midata,' the U.K. government has encouraged organizations to release the personal information they hold back to their customers in a portable, machine-readable, reusable format. The Mydex PDV service assists with the initiative.[20] The U.K. government is also exploring a proposal to introduce a legal requirement that, upon request, organizations must provide to their customers their transaction history and consumption data in an open standard machine-readable format.[21]

17 United States Department of Veterans Affairs, "What is the Blue Button Initiative?," http://www.va.gov/bluebutton/.

18 White House Blog, "Green Button Giving Millions of Americans Better Handle on Energy Costs," http://www.whitehouse.gov/blog/2012/03/22/green-button-giving-millions-americans-better-handle-energy-costs.

19 Jeff St. John, "California Gets the Green Button," *Greentechmedia.com*, January 18, 2012.

20 Mydex, "Mydex and the UK government's new midata policy," http://mydex.org/2011/11/03/mydex-uk-governments-midata-policy.

21 United Kingdom Department for Business Innovation & Skill, "Midata 2012 review and consultation," http://www.bis.gov.uk/Consultations/midata-review-and-consultation?cat=open.

Personal Data Vault

The PDE landscape includes a broad selection of technologies, services and tools. Currently the most mature are Personal Data Vaults[22] (PDV) which help individuals to collect, store, use, share, grant access to, and manage their own personal information in a manner that is completely within their own control. The PDV market is largely made up of startup companies with investments ranging from $1 million to $11 million US.[23] Personal.com, a platform that offers users both a PDV and personal networks for safe data exchange, alone, raised $7.3 million in 2011 and another $3.5 million in 2012.[24] Venture capital investment represents the first stage in the development of a PDE/PDV market. The second stage will be to create wide scale sustainability, meaning a market for such products that would become self-funding and profitable. The PDV market by itself in the United Kingdom is predicted to reach £30 million by 2016. However, if there is a market explosion, the value could grow as high as £1 billion for the same period.[25] In turn, this could dislodge the market in acquiring customer data from third parties, which is currently valued at $2 billion in the United States alone.[26]

PDVs give the user a central point of control for their personal information (e.g. interests, contact information, affiliations, preferences, and friends), including structured or unstructured data, such as text, images, video or sound. The information that a person chooses to put into a PDV may be general in nature, can relate to a specific topic, such as health or education information, or can be information relating to a particular objective, such as managing one's online presence. A key concept behind PDVs is 'controlled push' and 'informed pull.'[27] The customer engages in a controlled pushing out of their personal information out, but they can also pull information in by requesting data from different sources, based on the customer's own criteria (e.g. best price for car insurance). Figure 2 sums up some of the core functions of a PDV.

22 Alternative terms used to describe a Personal Data Vault include: Personal Data Store, Personal Data Locker, Personal Cloud, and Personal Data Service.

23 Ctrl-Shift (n 5) p. 4.

24 Robin, Francesca. "The Emerging Market That Could Kill the iPhone: A Handful of Tech Startups Are Competing for a Foothold in the Nascent Market for Personal Data Control. And That Could Mean Major Changes for the Likes of Apple, Google and Microsoft." *Fortune/CNN.com*, August 1, 2012.

25 Ctrl-Shift (n 5) p. 7.

26 Robin (n 24).

27 KuppingerCole. "Life Management Platforms: Control and Privacy for Personal Data." 2012, p. 10.

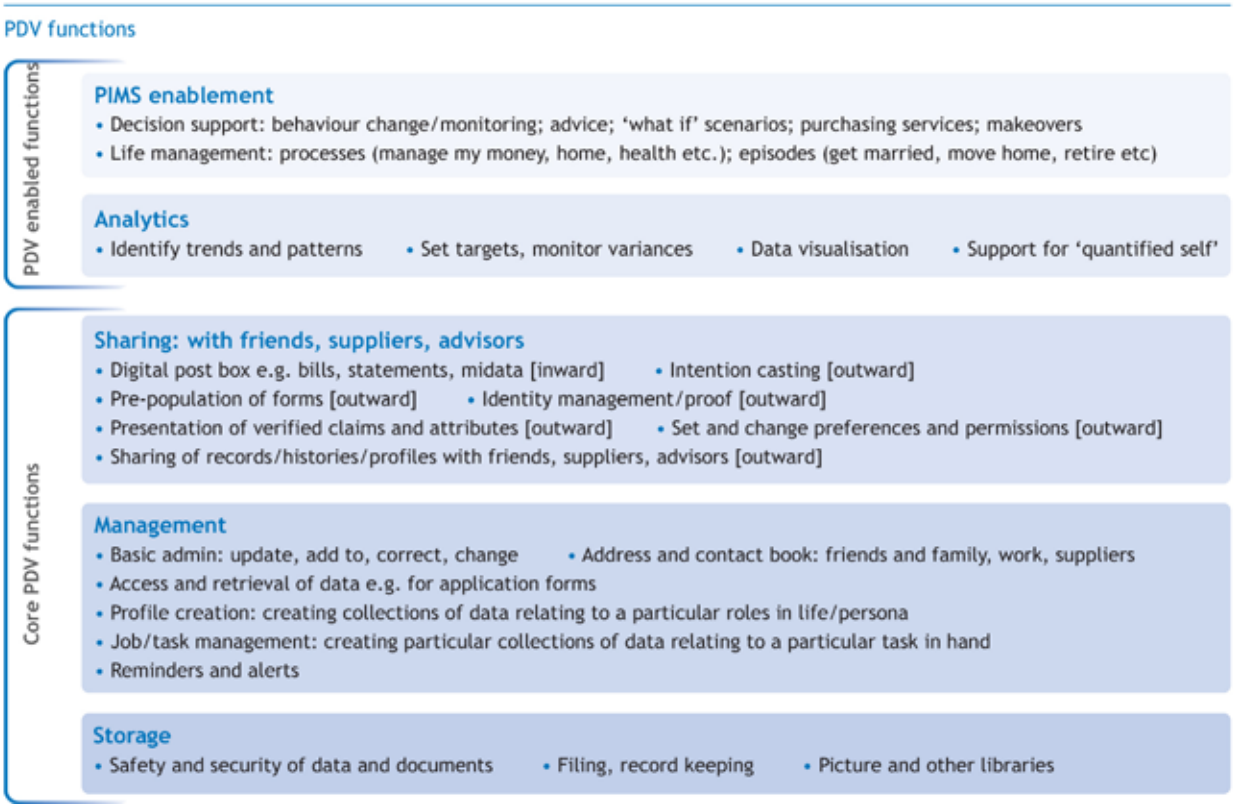

Figure 2 – PDV Functions

Source: Personal Data Stores: A Market Review, Ctrl-Shift 2012

A PDV may enable users to choose whether to be discovered by other users or third party service providers, or to share personal information with the same, according to criteria set by the individual. PDV systems use specialized software and distributed hardware for data storage, authentication, access control mechanisms. Many of them also offer an API (Application Programming Interface) for developer access. A PDV may reside in one location or can be distributed among several federated sources. It can also be self-hosted by the user (i.e. using one's own server or have a third party hosting company that acts in the legal role of personal data agent).[28]

Additional PDE technologies, services and tools include information logistics platforms and services that can be used for efficient information delivery and exchanges. Also, personal information management services (PIMS) can assist individuals in researching and coordinating life processes and episodes, such as getting married, moving or retiring. Analytics tools can monitor one's patterns, examine variances, set personal targets and goals, and visualize data. Targeted personal data services, such as to manage one's online reputation, function to enhance PDVs and could eventually integrate with PDV (see Fig. 3).

28 Drummond Reed, Joe Johnston, Scott David. "The Personal Network: A New Trust Model and Business Model for Personal Data." (2011).

Research by Ctrl-Shift has identified up to 20 PDVs that have launched or will have launched by the end of 2012. Figure 4 shows those that can be publicly identified (there are more launches planned but not yet made public). The figure also highlights the different parts of the emerging personal data ecosystem – infrastructure providers such as the Respect Network, PDV providers such as Mydex in the U.K. and QIY in Holland, and specialist users of personal data whose services link in to individuals' PDVs on a permissioned basis.

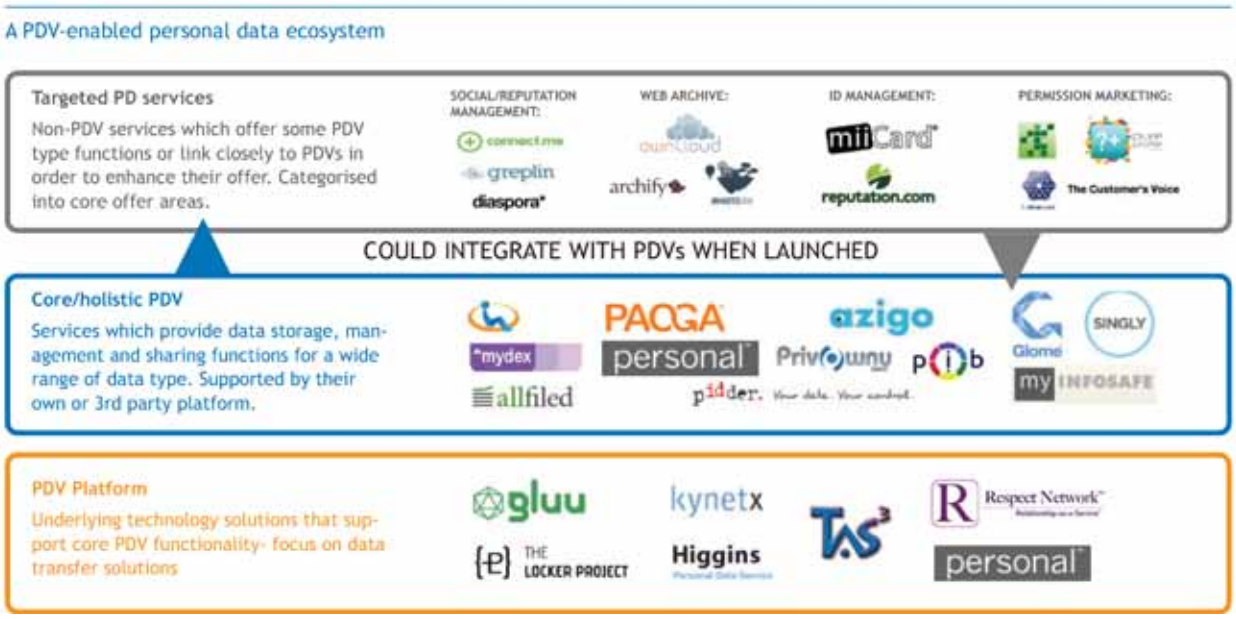

Figure 3 – PDV function stack

Source: Personal Data Stores: A Market Review, Ctrl-Shift 2012

An example of a data sharing platform is the Respect Network, which is building an interoperable network for privacy-protected exchanges of personal information based on the open standard OASIS XDI (Extensible Data Interchange) protocol and the Respect Trust Framework agreement listed with the Open Identity Exchange.[29] The goal is to make multiple PDVs from different providers interoperable with each other and with the businesses that want to connect with them, much like a credit card network makes electronic payments interoperable between different banks and merchants. The Respect Network business model is also similar to that of a credit card network. The key difference is that instead of charging an interchange fee based on the value of a transaction, businesses pay a relationship fee for the value of each customer relationship maintained over the network. The network also features a network-wide peer-to-peer reputation system that applies equally to vendors and customers.

29 The Respect Network (n 16).

Another platform is Personal.com. Available on Web and mobile devices, Personal gives individuals a data vault and tools to control, share and obtain value from their personal information through personal networks with other people, organizations and apps. Using Personal, people can safely and easily aggregate, leverage and exchange structured data, notes and files about their lives and identities. Recently, Personal publicly launched its application programming interface, which allows companies and app developers to use Personal's privacy- and security-enhancing platform features for exchanging data and authenticating users within their own apps.[30] Personal was founded on the principles that individuals should be able to control and gain both economic and non-economic value from their personal information. Privacy and security are embedded into its technology, business practices and enterprise. To underscore this focus and the primacy that an individual has over their own data on Personal, the company refers to its users as "Owners."

An example of a targeted PDV service is Reputation.com. Reputation.com is an online reputation management company that assists individuals to control their online search results due to inaccurate, misleading or outdated material, which can adversely influence how Web searchers view them. In addition, Reputation.com helps to prevent private information, such as where someone lives, their income, or their marital status, from being made public. [31]

PDV Scenario

Alice authenticates herself to her PDV. Having a PDV allows Alice to keep her personal information or "life data" in one place. Alice's PDV will authenticate her to organizations that recognize her data sharing platform with her consent so she only has to log in once rather than logging in to each online account separately. Organizations within the data sharing platform do not have to authenticate Alice or to offer her products and services. Alice has full control over the data in her PDV – she decides what information to share, with whom, and under what conditions. The service she uses also provides "data portability" – the ability for her to easily take the contents of her PDV to another service provider.

Alice's favorite retailer is a member of her data sharing platform which means she was able to download her transactional history from this retailer into her PDV account. Now Alice can share her data anonymously with a personal information management service which analyzes her spending habits and makes suggestions based on her financial goal of purchasing a vehicle.

30 See https://www.personal.com/gemware/pages/personal-launches-personal-platform; and http://developer.personal.com/

31 Reputation.com has provided online reputation management and Internet privacy protection to clients in more than 100 countries since 2006, and is the recipient of the 2011 World Economic Forum Technology Pioneer Award. "Control your online image," http://www.reputation.com.

Alice finds out that a charitable medical research foundation wants access to her health records for research purposes. As part of Alice's data sharing platform, she shares her health records on the condition that the foundation does not pass it along to anybody else and that it remains anonymous – she specifies that she does not want her name and address to be linked to the records. Alice will rely on contractual provisions to ensure that her conditions are met. However, in the future technology solutions such as "SmartData,"[32] may exist where a virtual agent will understand the various contexts in which Alice wishes to share her personal information, and it will automatically follow the rules she sets.

32 Tomko, G. (2012). *SmartData - Privacy Meets Evolutionary Robotics in the Matrix: Protecting Freedom Using Virtual Tools*. Paper presented at the IPSI SmartData International Symposium, Toronto, Ontario. http://www.ipsi.utoronto.ca/sdis/

Privacy Issues and the PDE

While the concept behind the PDE aligns with many aspects of the protection of privacy, there are, nonetheless, some areas which require careful attention to ensure that a PDE successfully protects personal information. Questions of interoperability, interactions and information-sharing mechanisms between PDE actors may have an impact on privacy. In a PDE environment, in which personal data is collected and shared with the permission of the individual, the devil will truly be in the details. In the wrong hands, one's PDV and activities within the PDE could be exploited as a major surveillance tool.

The focus here must be on issues across the overall framework for a PDE because all stakeholders within the PDE will have to address them, not just individual PDV operators. Privacy and trustworthiness are almost always more difficult to establish within an ecosystem of multiple stakeholders than within a single enterprise. Across multiple stakeholders, there may likely be many different policies, different security architectures, and different deployed tools and technologies, all of which need to be both interoperable and consistent in the protections provided for shared data. Strong security measures undertaken by a single stakeholder become meaningless if its data trading partners do not have compatible measures; the policies and technologies of all the ecosystem members must satisfy the requirements of the trusting party (the individual). Weak links in security can also lead to data breaches, compromised identity credentials, and identity theft. There is a need to proactively ensure privacy within any federated identity system, given the privacy issues arising from data-in-motion between multiple stakeholders. Different approaches to privacy within the ecosystem could lead to information being collected, used, and disclosed contrary to the individual's preferences. Unintended consequences could lead to an erosion of trust, harming the PDE as a whole.

While the individual retains control over his or her raw data in the PDV, this does not preclude third parties in the PDE from retaining copies of one's data with the individual's permission. This means all actors in the PDE must strictly follow the permissions for data held by others in the ecosystem, because the retention of copies of that data increase the chances of unintended disclosure and erosion of personal control.[33]

Taking a proactive approach to privacy will be essential to the success of any PDE initiative. The next section provides an overview of how to achieve this through *Privacy by Design*.

33 Harmonizing these high-level policies for the treatment of personal data is the goal of new Internet-scale trust frameworks like the Respect Trust Framework used by the Respect Network.

Privacy by Design and the PDE

Privacy by Design (*PbD*) is a concept developed back in the '90s to address the ever-growing and systemic effects of Information and Communication Technologies and of large-scale networked data systems. *PbD* advances the view that the future of privacy cannot be assured solely by compliance with regulatory frameworks; rather, privacy assurance must ideally become an organization's default mode of operation. *PbD* extends to a "Trilogy" of encompassing applications: 1) IT systems; 2) accountable business practices; and 3) physical design and networked infrastructure. The objectives of *PbD* – ensuring privacy and gaining personal control over one's information and, for organizations, gaining a sustainable competitive advantage – may be accomplished by practicing the following 7 Foundational Principles:[34]

The 7 Foundational Principles of *Privacy by Design*

1. ***Proactive*** not Reactive; ***Preventative*** not Remedial
2. Privacy as the ***Default Setting***
3. Privacy ***Embedded*** into Design
4. Full Functionality – ***Positive-Sum***, not Zero-Sum
5. End-to-End Security – ***Full Lifecycle Protection***
6. ***Visibility*** and ***Transparency*** – Keep it ***Open***
7. ***Respect*** for User Privacy – Keep it ***User-Centric***

Now is the time to apply *PbD* to the PDE, since it is currently in the early stages of development. Trust and governance are key forces that will shape the PDE. Success will depend on establishing the privacy and security of the user's personal information. In this regard, transparency and clarity will be essential. In addition, PDE companies will need to devise easy-to-use features, employ data minimization when sharing data, and ensure that the user is fully aware of the privacy implications of his or her data sharing decisions. While it is envisioned that PDE will also improve information exchanges between individuals, governments, and corporations, a challenge is the development of privacy-protective protocols and platforms that facilitate interoperability between data sets.

34 Information and Privacy Commissioner of Ontario. "*Privacy by Design*: The 7 Foundational Principles." (2009).

For several years, the IPC has examined emerging technologies and best practices that are relevant to PDE, which can assist in developing the PDE in a manner consistent with *PbD*.

1. ***Proactive*** not Reactive; ***Preventative*** not Remedial

The *PbD* approach is characterized by proactive rather than reactive measures. It anticipates and prevents privacy-invasive events. It does not wait for risks to materialize, nor does it offer remedies for resolving infractions once they have occurred – it aims to prevent them from occurring. In short, *PbD* comes before, not after the fact.

As outlined in the previous section, the challenges and risks to privacy in the development of a PDE are much higher than developing information systems at the enterprise level. Although a PDE represents a significant increase in user control, it nonetheless involves the collection, use and disclosure of personal information among several stakeholders.

The IPC has written about undertaking risk assessments when establishing an ecosystem of multiple stakeholders and ecosystems in a cross-layer cloud computing environment (e.g. Federated Privacy Impact Assessment or F-PIA),[35] the risk of unintended consequences in building information architectures which could be used for different purposes in the future,[36] as well as the importance of designing user interfaces with user-centricity in mind.[37]

To proactively ensure privacy within the PDE, and given the multiple stakeholders that could potentially be involved in a PDE scenario, stakeholders should conduct an F-PIA or similar type of assessment that considers issues arising from data-in-motion in which personal information is transferred among various stakeholders. At the engineering level, developers should assess whether the infrastructure to be created could have unintended uses or consequences. For example, could an additional infrastructure be built on the contemplated one, and what would be its implications for privacy?[38]

The PDE should establish its privacy and trustworthiness by ensuring that all stakeholders in the PDE have common policies, technologies and tools that are interoperable and consistent in protecting shared data. Personal information should be handled according to those established rules or trust frameworks, conforming to promises made at the time of collection.

---

35 Information and Privacy Commissioner of Ontario, and Liberty Alliance Project. "The New Federated Privacy Impact Assessment (F-PIA)." 2009; NEC Company Ltd., and Information and Privacy Commissioner of Ontario. "Modelling Cloud Computing Architecture without Compromising Privacy: A *Privacy by Design* Approach." (2010).

36 Information and Privacy Commissioner of Ontario, and Kim Cameron. "Wi-Fi Positioning Systems: Beware of Unintended Consequences." 2011.

37 Information and Privacy Commissioner of Ontario, Justin B. Weiss. "*Privacy by Design* and User Interfaces: Emerging Design Criteria - Keep It User-Centric." (2012).

38 E.g. permanent identifiers were assigned to Wi-Fi networks devices for communication between those devices. Eventually those identifiers were used to create Wi-Fi positioning systems to identify the location of mobile devices.

2. Privacy as the *Default Setting*

We can all be certain of one thing – the default rules! The power of the default condition cannot be underestimated. *PbD* seeks to deliver the maximum degree of privacy by ensuring that personal data are automatically protected in any given IT system or business practice. If an individual does nothing, their privacy remains intact! No action is required on the part of the individual to protect his or her privacy – it is built into the system, automatically, by default.

It is essential that when the user is in control of all of the data, the default condition of 'no action required' be adopted, as enabled by data minimization, especially in a multiple identity claims environment. We have written, for example, how in the case of the Internet, its basic infrastructure was built without an identity layer, meaning that users do not always know who or what they are connecting to online.[39] However, in creating an identity layer, privacy implications must be identified and addressed.[40] In the context of extremely large datasets, i.e. "Big Data," we have written about the importance of building privacy (e.g. data minimization[41]) into technological advances which seek to capture, store, manage, and analyse personal information.[42]

In creating the identity layer for the PDE, stakeholders should ensure that privacy is reflected in the building blocks of the chosen identity metasystem. Privacy requires data minimization at every stage of the information lifecycle, and if personal information is not needed, it should never be collected in the first place. Research in this area is already underway. For example, researchers have embedded privacy into the use of a PDV on mobile devices, in which a set of filters controls and minimizes the flow of data from the PDV to third party applications.[43] In addition, techniques exist to perform analytics on anonymized data, which will serve to enhance trust in information-sharing environments.[44]

39 7 Laws of Identity (n 4).

40 *Ibid.*

41 To be clear, within a PDV, the individual is free to pursue a policy of 'data maximization' about themselves, i.e. to manage and use as much of their personal data as they wish.

42 *PbD* in the Age of Big Data (n 6).

43 Shilton, K., J. Burke, D. Estrin, R. Govindan, M. Hansen, J. Kang, and M. Mun. "Designing the Personal Data Stream: Enabling Participatory Privacy in Mobile Personal Sensing." (2009), p. 7.

44 *PbD* in the Age of Big Data (n 6) p. 11.

3. Privacy *Embedded* into Design

Privacy should be embedded into the design and architecture of IT systems and business practices. It should not be bolted on as an add-on, after the fact. The result is that privacy becomes an essential component of the core functionality being delivered. Privacy becomes integral to the system, without diminishing functionality.

It is an opportune time to build privacy into the architecture and design of a PDE when it is in its nascent stage. We have written extensively about the need to embed privacy as early as possible into the design of systems in complex information-sharing environments dealing with large amounts of information.

Embedding privacy must occur within and across the PDE at all levels, including hardware manufacturing, operating platform, software and thin-client layer, as well as any outsourced or mobile components. Internal control systems based on essential preconditions can ensure that only authorized stakeholders in the PDE can contribute to and access the data. The system design principles of a PDE architecture where individuals can manage their own data must take into account a number of additional considerations:[45]

- *Participant Primacy*: Users should retain control over their raw data and be able to make decisions to share subsets of the data. This does not preclude third parties from keeping copies of the data, if permitted by the individual.
- *Data legibility*: The system must provide high-level tools and guidance on the implications of users' decisions about their data. It allows users to understand the risks of their participation, and helps them to make better decisions about data collection, sharing and retention.
- *Long-term engagement*: The system should encourage the continued engagement of users by allowing them to check if their data is still visible and relevant, and through the use of triggers,[46] make continuing and ongoing decisions about their sharing policies.

The use of intelligent agents within the PDE could help users make informed decisions about their personal information based on, for example, an agent's observations and reports.47 In the future, "SmartData" could move the PDE to a more ubiquitous Privacy by Design PDE. SmartData will consist of Internet-based autonomous, cognitive agents that act as a data subject's online surrogate, securely storing one's personal information, and intelligently disclosing it based upon the context of the data request, in accordance with the user's instructions. [48]

45 Shilton et al. (n 44).

46 Triggers can be by way of new applications that generate new kinds of data, or by occasional or periodic verification of agreements with third parties.

47 Shilton et al. (n 44) p. 7.

48 Developed by Dr. George Tomko, Expert-in-Residence at the University of Toronto. Tomko, George J., Hon Kwan, and Don Borrett. "SmartData: The Need, the Goal, the Challenge." 2012. For further information see http://www.ipsi.utoronto.ca

4. Full Functionality – *Positive-Sum*, not Zero-Sum

PbD seeks to accommodate all legitimate interests and objectives in a positive-sum, or doubly enabling "win-win" manner, not through a dated, zero-sum approach, where unnecessary trade-offs are made. It avoids the pretense of false dichotomies, such as privacy *vs.* security, demonstrating that it is indeed possible to have multiple functionalities.

From a positive-sum perspective, privacy can be seen as a business strategy that brings with it a sustainable competitive advantage, such as enhanced user trust. Privacy delivers shareholder value, builds trust, strengthens the value chain, and gives businesses a competitive advantage.[49] If individuals are confident that they are genuinely in control, they would be more willing to share more information with organizations, such as preferences, plans and intentions. This is fundamental a "win-win," i.e., if personal data is truly going to become "the new oil of the Internet,"[50] then business interests in deriving the economic value of personal information and individual interests in maintaining control over personal information must both be met within the PDE.

5. End-to-End Security – *Full Lifecycle Protection*

Privacy, having been embedded into the system prior to the first element of information being collected, extends throughout the entire lifecycle of the data involved, from start to finish. This ensures that at the end of the process, all data are securely destroyed, and in a timely fashion. Thus, *PbD* ensures cradle-to-grave lifecycle management of personal information, end-to-end.

In the context of systems relevant to the PDE, the IPC has written about security considerations within a Government 2.0 environment where citizens co-manage their personal information on a secure infrastructure supported by trusted back-end systems.[51] We have looked at security and cloud services, in which the entire lifespan of data must be evaluated, including the protection of any significant meta-data that is generated based on patterns of data access and sharing (including identities of the data requesters).[52] The importance of ensuring the secure destruction of personal information is an area that warrants strong consideration but is often overlooked.[53]

---

49 Cavoukian, Ann, and Tyler Hamilton. *Privacy Payoff: How Successful Businesses Build Customer Trust.* Toronto: McGraw-Hill Ryerson, 2002.

50 Kuneva (n 1).

51 Information and Privacy Commissioner of Ontario. "Privacy and Government 2.0: The Implications of an Open World." (2009).

52 Modelling Cloud Computing (n 35).

53 Information and Privacy Commissioner of Ontario, and National Association for Information Destruction Inc. "Get Rid of It Securely to Keep It Private: Best Practices for the Secure Destruction of Personal Health Information." (2009).

It almost goes without saying that security within a PDE must be exemplary. In a multi-stakeholder ecosystem such as the PDE, this means that it is not enough for most actors to be secure – all must be secure. Personal information which has reached the end of its lifecycle in a PDE must be destroyed in a consistently secure and privacy-protected manner. Methods of destruction for paper records include mechanical (crosscut shredding, pulping, pulverizing), to pieces millimeters in dimension, and incineration to white ash. For electronic media, methods of secure destruction include mechanical (degaussing and sanitation / secure erase) to an unusable state. Deleting files or reformatting hard drives is no longer sufficient to ensure that electronic media is securely destroyed.

6. *Visibility* and *Transparency* – Keep it *Open*

PbD seeks to assure all stakeholders that whatever the business practice or technology involved, it is, in fact, operating according to the stated promises and objectives, subject to independent verification. Its component parts and operations remain visible and transparent, to users and providers alike. Remember − trust but verify.

Users of a PDE must have a clear understanding of overall data management, access and sharing by the service provider and third parties. The sense of control over their personal data that users gain over time, along with the increased sensitivity and awareness of that data, requires transparency to be the focal point of any PDE business approach. We have written on the issue of accountability requiring clear organizational rules, with an explicit commitment to the policies that are the basis for those rules.[54]

Stakeholders within the PDE should adopt accountable business practices in the following areas:

1. **Policies and Commitment**: PDE stakeholders should have solid policies, commitments to protect privacy, meaningful transparency, and a willingness to demonstrate their own capacity to uphold promises and obligations.

2. **Implementation Mechanisms**: PDE stakeholders should have robust internal standards and controls that integrate privacy into the design of their systems, as well as processes to guide and support decision-makers. Practices should incorporate ethics and values-based considerations, and should fully consider a broad range of their performance risks.

3. **Assurance Practices**: PDE stakeholders should have an ability to monitor and evaluate how they are doing in terms of accountability, and to make real-time course corrections, where necessary.

54 Information and Privacy Commissioner of Ontario, Hewlett-Packard, and The Centre for Information Policy Leadership. "*Privacy by Design*: Essential for Organizational Accountability and Strong Business Practices." (2009).

7. *Respect* for User Privacy – Keep it *User-Centric*

At its core, respecting the user means that, when designing or deploying an information system, the individual's privacy rights and interests are accommodated, right from the outset. User-centricity is designing for the user, anticipating his or her privacy perceptions, needs, requirements, and default settings.[55] It means putting the interests, needs, and expectations of users first, not those of the organization or its staff. Empowering users to play active roles in the management of their personal data may be the single most effective check against abuses and misuses.

The evolution of a PDE[56] is by its very nature aligned with the seventh foundational principle of *PbD*, precisely because of its focus on user control. The principle of '*Respect* for User Privacy – Keep it *User-Centric*' requires that architects and operators keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options. The PDE makes respect for user privacy a cornerstone of its systems, practices, design and infrastructure. Doing so results in a paradigm shift in the assumptions about privacy from "organization-centric" to fully user-centric, which pushes past the current practice of organizational promise-making regarding data protection and elevates it to ecosystem-wide trust frameworks. PDE transforms this debate by recognizing privacy as a personal setting where the individual is empowered to choose what information he/she wishes to share.

This *PbD* principle has important links to the field of user interface design. User interface is the system by which people (users) interact with a machine. It includes hardware (physical) and software (logical) components. Usability is the degree to which the design of a particular user interface takes into account the human psychology and physiology of its users, and makes the process of using the system effective, efficient and satisfying. Usability is mainly a characteristic of the user interface, but is also associated with the functionalities of the product and the process to design it.

In the context of *PbD* generally, and of this principle in particular, the notion of usability is expanded to include the extent to which a system/interface is usable for the achievement of objectives related to informational self-determination, as well as for communicating expectations and providing opportunities for feedback that help shape and clarify those expectations. The field of user interface design is already well established. From the perspective of user interface design, the privacy challenge is to make consumers aware of privacy practices without getting in the way of their experience. Operational aspects of this principle include measures to assure transparency, attain informed user consent, provide rights of access and correction, and make effective redress mechanisms available. Organizational policies and processes should demonstrate the same degree of consideration for users at all touch points and interactions.

55 McDougall, P. (August 8, 2012). Microsoft IE 10 Makes 'Do Not Track' Default, *InformationWeek*.

56 Including PDVs and data sharing platforms.

Implementing *Privacy by Design* in the PDE

It is said that personal data vaults "build out" and data sharing platforms "build in", and both meet each other to create the PDE. In many cases, services are evolving rapidly to become both. The recommendations for *Privacy by Design* (*PbD*), given above, form a strong foundation for what should be done to protect privacy in the PDE. In this section, we build upon this foundation by demonstrating what can be done to protect privacy, by considering two case studies. The first one examines the personal vault feature offered by Personal.com. While the vault feature is examined for Personal, they offer additional PDE features, including a data sharing platform. The second case study examines two data sharing platforms, the Respect Network, and the Digital Asset Grid.

Case Study: Personal.com's approach to *Privacy by Design*

Although there are a number of PDE platforms, we have selected Personal.com as the illustrative example of how a *PbD* framework has been applied to a commercial PDV and personal networks platform.

Personal offers a Web and mobile service to give individuals a data vault and tools to control, share and gain value from their personal information, including through personal networks. Using Personal, people can safely and easily aggregate, leverage and exchange structured data, notes (i.e., unstructured data), and files about their lives and identities. One key characteristic that distinguishes Personal from other platforms (e.g., online social networking sites) is that the service enables individuals to easily enter, import, consolidate and manage information across their lives in digital form, all while being a user-driven tool, as opposed to one that is centered around an enterprise. Beyond offering non-economic benefits to users such as secure sharing of important information, automatic form filling using the data in one's vault, reuse of data, and apps to make one's data even more useful, one part of Personal's business model is to create a marketplace in which users drive the value proposition through their data, with Personal acting as an agent that only benefits economically if the user does.

On Personal, individuals can add or import their data and share it through bundles of structured and unstructured fields of data and files called Gems (Fig. 4).[57] Gems are organized around specific themes to enable users to efficiently organize and search for information, as well as to grant and request access to targeted, relevant data with other users and organizational partners of their choosing. A person can store data in his or her Gems that cover every area of life, such as passwords, consumer products, loyalty programs, and financial and health information. Within the vault, an individual can choose whether to grant another user access to one of their Gems.

57 Each structured data field represents a different preference, data value or piece of information. Importantly, users (and no one else) always choose the information to store or import into a Gem.

Figure 4 – Gems – Containers used for Data and File Storage and Sharing

In the future, using data stored in his or her own vault, a user will be able to signal an intention to buy a particular article or service through Personal's platform to relevant corporate partners. This information will be shared anonymously (without the email or name of the user), and the data will be used to attract only relevant and customized offers that will be sent to the user's data vault. A user can decide to stop granting access to his or her anonymized data and thus stop receiving offers to his or her data vault, at any time. If the user obtains value in the form of a discount or cash back through this fair exchange of their data, Personal will charge the merchant a transaction fee that is a small percentage of the value realized by the user. In this way, the financial interests of the user and Personal are closely aligned.

Personal has embedded *Privacy by Design* directly into its platform, with examples relating to data and network control, transparency, security, portability, and data minimization, found below.

Data and Network Control

Gems contain fields of structured or unstructured data, in addition to files, and are organized by category. The data fields contain either sensitive or non-sensitive data, and, in some cases, Personal goes beyond these definitions to designate certain fields as "sensitive" that would not otherwise have been designated as such, under applicable laws. Any data marked "sensitive" is encrypted within Personal (noted by a lock next to the field) and can only be viewed with the password that the user chooses. All files and photos uploaded to Gems are also encrypted. Only the user knows the password, which Personal does not store.

Everything a user does on Personal is initiated by a user's affirmative action. Only the user can grant access to his or her data stored on Personal. Access is always permission-based, and happens within a given session. This means that there is no such thing as uncontrolled access to one's information without the user intending and confirming that another person may have such a right of access. The user alone controls the data to import or manually input into the data vault, whether to grant access to an entire Gem or part of one, and deciding who will receive access. Most important, each Gem may be shared with as many or as few users or organizations on the Personal platform as the initiating user chooses. This is, in a real sense, the opposite of the current enterprise-centric model. Further, users can either provide viewing-only access to a Gem or access that allows another user to adopt the information as the person's own; users may revoke access privileges to data granted on a viewing-only access basis, at any time.

On a more technical level, once a user decides to grant access to one or more Gems, the system creates a "valet" key for the user that is the intended recipient of the grant. The user who has been granted access must then use his or her password to access the particular information, providing a second layer of encryption and password protection for the data. Personal cannot use the "valet" key – only the intended user may use it. In addition, Gems are content-specific, ensuring that the personal data residing within them is an accurate, adequate and reliable representation of the user's information.

Transparency

Transparency is a core principle and functionality of Personal. Personal's user interface has been designed to advance the user's control of his or her data and information exchange. Indeed, the system is designed to tell a user who has access to a user's information, at any time. Overall, illustrating who can see what data will help users understand the consequences of data sharing. Some of these functionalities include:

- When sensitive information is saved in a Gem, it is graphically indicated as being encrypted.
- Each user has a "Network" tab in his or her account that shows the list of contacts within a user's network. A user can grow his or her network by granting access to his or her Gems to contacts and by being granted access to others' Gems. Upon clicking on a contact, the system shows the user exactly which Gems have been granted to the contact and which Gems the contact has granted to the user. In this way, the Network features serve as a kind of "personal graph" for a user's own information.
- Along similar lines, Personal also provides each user with a "Feed" showing each instance of the user having requested or having been granted a Gem and the identity of the contact. When a new instance of sharing occurs, the

user is alerted to it in real time by a new message appearing in his or her feed (Fig. 4). The “Feed” is an organic record of who has been granted access to user data and the data that the user has extended to others.

Figure 5 – Personal.com feed shows activity regarding Gems

- Because Personal’s platform is data-driven and the data is stored in structured form, a user can choose to grant someone access to his or her information at an entire or multiple Gem level.
- If one user requests a Gem from another user, the second user has the right to refuse access to it.
- At the present time, users can share information in two ways – either on a viewing/access-only basis or allowing the import of his or her data into the other user’s vault (Fig. 5). If the user has granted viewing or access-only rights to the second user, the first user may always choose to revoke access to his or her Gem.

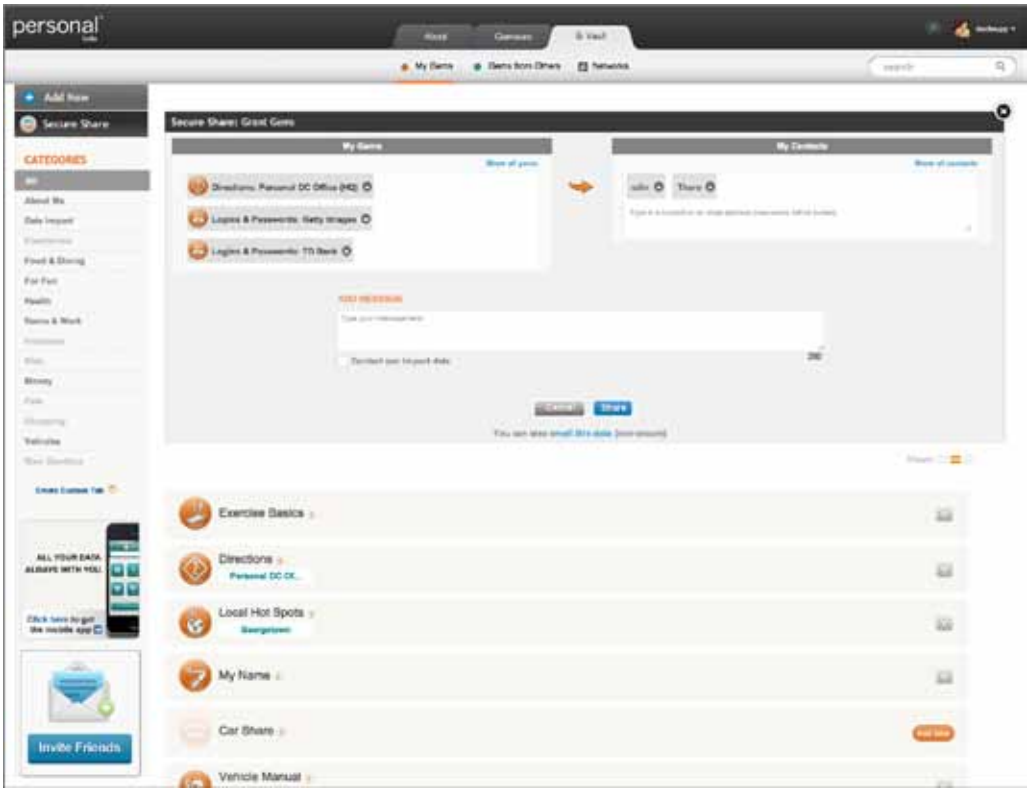

Figure 6 – Personal.com secure sharing user interface

Security

Upon registering, users choose a password, which Personal does not store. Instead, the company uses the bcrypt algorithm. This produces a hash or a long string of letters and numbers that are stored for comparison with the password users enter. As part of bcrypt, the password is made more complex through a process called "salting." The entire hashing process is then repeated many times through a process known as "stretching." This results in strong one-way encryption, meaning that Personal cannot reverse-engineer your password. Each time users login to Personal, the system will simply re-run bcrypt on the newly-entered password. The system then compares the two hashes (the stored one and the computed one) to determine if they match. If so, a user can login. If the passwords do not match, the login attempt is rejected.

If a user forgets the password or, for some other reason, chooses to reset it, all sensitive data will be deleted permanently from his or her data vault and, more generally, from the system. Non-sensitive information will be retained in his or her data vault. Although it might pose an inconvenience, wiping away this sensitive data will provide heightened protection for the user in the case of a user-compromised password or a breach. In addition, it disables personally identifiable data extraction as the linkage may be broken by removing sensitive data.

Beyond password practices, Personal takes many steps to secure users' data and is constantly working to improve them. For example, sensitive data stored in Gems is encrypted and each Gem containing sensitive data is encrypted uniquely. Personal also encrypts all user data while in transit from a user's browser to Personal's database servers, whether designated as sensitive or non-sensitive. This includes when users access their data or grant access to others. Personal also uses secure cookies with HTTPS to further protect user data.

Personal does not use privacy-threatening third party content delivery networks (also known as CDNs) that could potentially track users to further ensure that user data is never exposed. Personal also does not allow third party tracking of users within its product. In addition, in the context of crash reports, the company has, by default, filtered out any user data that could appear in fields so that employees are not inadvertently exposed to any user data. And, Personal does not use data de-duplication methods that can raise security and privacy concerns.

Personal purposely avoided using "free analytics" services because, although arguably free of charge, such tools often "phone home" to the manufacturer, and the price of use can end up being unintended data exchanges for the user. In addition, Personal has built its own analytics system to monitor activity in aggregate in its system, such as number of registrations, number of Gem shares by category, number of Gems completed, and fields completed. It is important to note that Personal's proprietary analytics software focuses only on meta or transactional data within its system at an aggregate level – and never on the actual data values that a user inputs or imports into a particular Gem. In the case of sensitive data, the company has no technical means whatsoever to access it.

Finally, Personal regularly conducts audits of its security, including penetration testing by outside firms, and fosters a culture of security and privacy within its company that touches every employee. Furthermore, the company's databases have a log system that records any employee's entry and access into the databases, ensuring traceability and control of the user's data access.

Portability

Personal users have the option to export their own data from their vaults to an Extensible Markup Language (XML) file, a simple format that is both human and machine-readable. Personal believes users must be empowered not only to benefit from their data while on Personal, but also to leave Personal, should they wish to do so. This option is built into the system, thereby ensuring that, on Personal, people are truly in control of their data, throughout the entire lifecycle of data use.

Data Minimization

Personal has built a "delete button" into its platform to enable users to easily delete their accounts. As noted above, too, unlike most other network services, the system service logs do not contain user data, which is filtered out using Ruby on Rails' parameter filtering before the logs or crash reports are written. In addition, the company uses POST rather than GET requests to keep networking equipment from inadvertently logging user data. Thus, even logs do not contain any user-provided data held in a vault once users close their accounts.

Because sharing through Gems is designed to occur at a granular, permission-based level, data minimization in granting access to information is built into the platform. In addition, Personal will not grant any third party access to user data, except when specifically requested by the user or as required by law. Regarding responses to court authorized processes, such as judicial warrants, Personal will be unable to produce any decrypted sensitive data, as only the user possesses the password to decrypt it. In addition, Personal has purposefully minimized the company's ability to access user data. For example, since no one within the company (or third party) has a user's password, it is not possible for an employee to access sensitive user information.

Personal's Legal Framework

Personal has also created a legal framework for individuals who share information with other people, organizations and apps through their vault. These rules govern the networks that individuals build with their data, and form the legal embodiment of the platform's technical features. It starts with the Owner Data Agreement ("ODA"), which every user agrees to when registering for Personal. The ODA is a legally binding contract between the user and Personal that states, among other things, that the data stored in Personal belongs to the user, only the user controls who gets access to his or her data, and companies and developers (called "Data Users") that are granted access to data by the user are also legally obligated by the ODA. Called unprecedented and modeling a new legal position,[58] the ODA also incorporates covenants that Data Users must meet to work with Personal, including respect for user privacy and security and transparency about the use of shared data.[59] In this way, Personal has built into its platform the legal rules for scaling in terms of users and the businesses, organizations and apps users would like to add to their networks for granting access to their data for specific, voluntary purposes.

58 Doc Searls, The Intention Economy: When Customers Take Charge, (Harvard Business Review Press, 2012), 186.

59 See https://www.personal.com/data-user-covenants

Case Study: Respect Network and SWIFT Digital Asset Grid as Data Sharing Platforms

The Personal Data Ecosystem envisions multiple PDV providers, with individuals being able to choose their preferred provider the same way they currently choose their bank within the global banking network or their email provider within the global email network.[60] PDV providers will be interoperable through *data sharing platforms*. To use an analogy, a data sharing platform is to PDVs and businesses what a credit card network is to banks and merchants. Data sharing platforms address three critical interoperability issues:

1. **Technical interoperability:** how PDVs, apps, and businesses can safely and securely share personal data not just with the necessary permissions and protections, but with the common semantics needed for all parties to understand and make use of the data.

2. **Legal interoperability:** how PDVs, apps, and businesses can enter into data sharing relationships in which the appropriate legal assurances and protections are available to all parties, and which satisfy the regulatory requirements of the different jurisdictions involved, while at the same time removing the frictions that often inhibit individuals and businesses from realizing the full value that can be realized from trusted personal data sharing.

3. **Business interoperability:** how PDVs, apps, and businesses can employ transparent, sustainable business models that align incentives to help individuals protect and realize the value of their personal data, rather than track and exploit it in ways that individuals can neither see nor understand.

This three level "stack" in the PDE is often referred to as "BLT" (Business/Legal/Technical). A diagram of these layers is shown as Figure 7:

60 As with email, individuals should also have the option of self-hosting, i.e., running their own PDV, typically using open source software and hardware designed for this function, such as the Freedom Box. http://en.wikipedia.org/wiki/Freedom_Box

BLT: The Business/Legal/Technical Stack

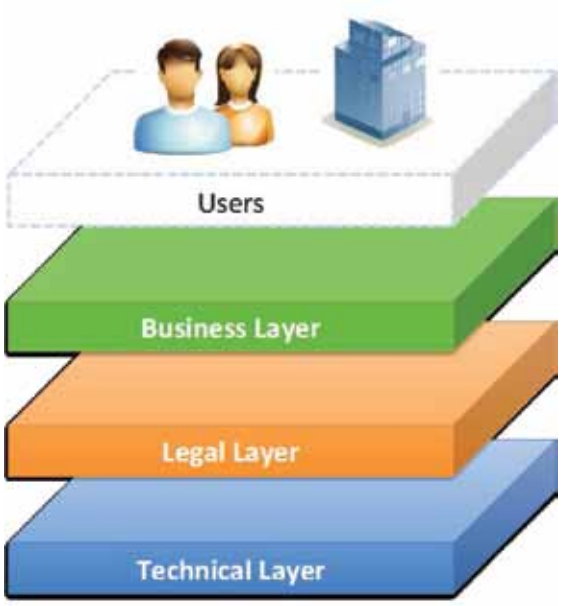

Figure 7 – The BLT layers of the PDE

Examples of PDE data sharing platforms may be found below.

Respect Network

The Respect Network is currently being developed by Respect Network Corporation, a PDE startup based in San Francisco, and 16 Founding Partner companies around the world, including Neustar, a trusted neutral provider of real-time information and analysis internationally, and Swisscom, Switzerland's leading telecoms provider and one of the most trusted brands in Europe.[61]

The foundation of the Respect Network is the Respect Trust Framework,[62] the first global trust framework for personal data listed with the Open Identity Exchange,[63] the international non-profit home for digital trust frameworks. Winner of the Privacy Award at the 2011 European Identity Conference, the Respect Trust Framework is based on five core principles for personal data sharing that cross almost all jurisdictions and Fair Information Practices Principles.[64]

61 http://respectnetwork.com/founding-partners/

62 http://openidentityexchange.org/trust-frameworks/respect-trust-framework/

63 http://www.openidentityexchange.org/

64 http://en.wikipedia.org/wiki/FTC_Fair_Information_Practice. Note that a full analysis of how the five Respect Principles map to FIPPs from around the world is included as an appendix to the Respect Trust Framework.

From this cornerstone, the Respect Network is building out the BLT layers as follows:

1. **Technical interoperability** is based on two core technologies: the XDI (Extensible Data Interchange) open standard format and protocol for semantic data interchange being developed by the OASIS XDI Technical Committee[65] and the KRL (Kinetic Rules Language) open source event-based programming language being developed by Kynetx,[66] one of the Respect Network Founding Partners. In particular, the combination of XDI link contracts—portable, machine-readable permissions for shared data—and KRL rules—portable, machine-readable instructions for how to follow an individual's and businesses policies regarding shared data—provide a powerful platform for permissioned, protected, personally-controlled data sharing on a global scale.

2. **Legal interoperability** is provided through a modular trust framework architecture beginning with the Respect Trust Framework as the overarching set of principles agreed to by all members of the Respect Network. In addition to the five core principles, the Respect Trust Framework also defines a peer-to-peer reputation system covering all members of the network. This reputation system, already implemented by the Connect. Me social vouching and discovery service,[67] provides additional incentives for all participants to observe the Respect Principles for personal data.

3. **Business interoperability** is based on an open and transparent business model in which Respect Network access is free for individual members of the network because business members of the network agree to pay a relationship fee for the value of the personal data sharing relationships facilitated by the network. This model is highly analogous to the business model for the credit card networks, where merchants pay an interchange fee for the value and convenience of the payment transactions facilitated by the network. This business model also aligns the interests of individuals and their PDV providers to protect and realize the value of the personal data they share over the network.[68]

65 http://www.oasis-open.org/committees/xdi/

66 http://en.wikipedia.org/wiki/Kinetic_Rule_Language

67 https://connect.me/about

68 In the terminology of VRM (Vendor Relationship Management), these PDV providers are "fourth parties", i.e., agents who are acting on behalf of buyers, vs. third parties, who are agents acting on behalf of sellers. The VRM concept of a fourth party is similar to the concept of a buyer's agent in real estate.

SWIFT Digital Asset Grid

SWIFT[69] is a member-owned cooperative that provides the communications platform, products and services to connect more than 10,000 banking organizations, securities institutions and corporate customers in 212 countries and territories. SWIFT enables its users to exchange automated, standardized financial information securely and reliably, thereby lowering costs, reducing operational risk and eliminating operational inefficiencies.

Launched in 2009, Innotribe is SWIFT's initiative to enable collaborative innovation in financial services. Innotribe fosters creative thinking in financial services, through debating the options (at Innotribe events) and supporting the creation of innovative new solutions (through the Incubator, Startup Challenge and Proof of Concepts).

The Digital Asset Grid is a research project backed by the Innotribe Incubation Fund, whose purpose is for banks to provide a platform for secure peer-to-peer data sharing between trusted people, businesses, and devices. As a data sharing platform, the SWIFT Digital Asset Grid is a superset of the Respect Network BLT architecture, which is why Innotribe has been sponsoring work with Respect Network Corporation and the Respect Network Founding Partners to help develop prototype infrastructure and applications for the Digital Asset Grid.

The goal of the Digital Asset Grid is to help all participants in the digital economy deal with the exploding complexity of data by doing for data what SWIFT has already done for payments: provide a new scalable global network—one that operates parallel to SWIFT's existing financial infrastructure—for "digital data banking" i.e. trusted transactions of any digital asset between any two parties on the network.

This network is designed to act as a digital map that describes the location of the data, the trust framework(s) under which access to the data is available, the digital identities who have access to that data, and the access and usage rights these identities have under those trust framework(s).

This means that from a technical interoperability perspective, the Digital Asset Grid will use the same XDI and KRL architecture as the Respect Network. However, from a legal and business perspective, it will be a superset because it will encompass other trust frameworks and business models that extend beyond the personal data interoperability requirements of the Respect Network. From the perspective of the legal layer, this superset relationship is illustrated in Figure 8:

69 Society for Worldwide Interbank Financial Telecommunication

Trust Framework Architecture of the Digital Asset Grid

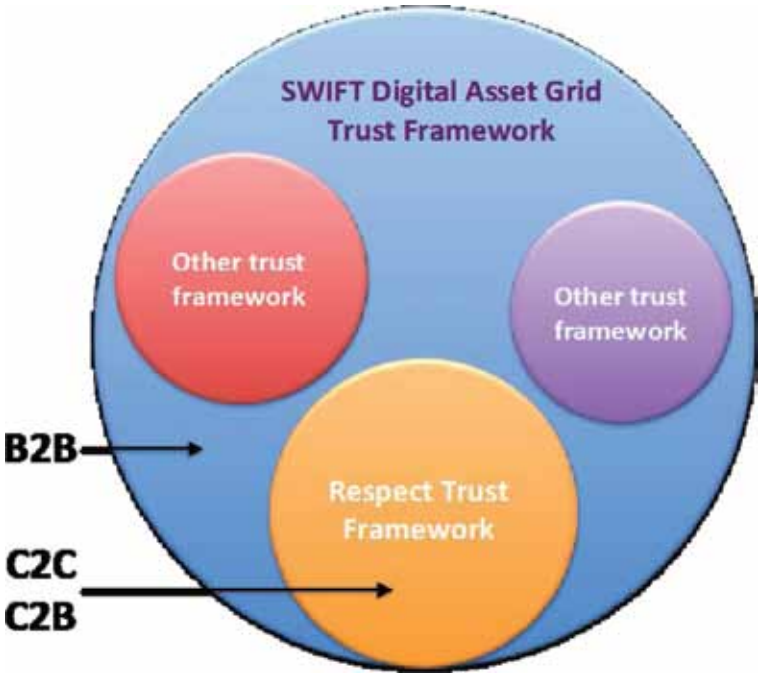

Figure 8 – The trust framework architecture of the Digital Asset Grid

From the perspective of the Personal Data Ecosystem, the development of the Digital Asset Grid will provide enormous impetus for global adoption of a robust PDE because it will bring SWIFT's core skills regarding governance, identity, security and operational excellence to the establishment of a global data sharing platform as ubiquitous and reliable as today's global banking network. This should enable PDVs to flourish and also encourage widespread development of new PDV-powered applications.

Conclusions

Privacy is all about control – exerting control over the uses of your own personal information. If control is in the hands of someone other than yourself, then your privacy may be at risk. Once your data is compromised, the harm has often been done. In the words of then-U.S. Supreme Court Justice William Douglas, "Once privacy is invalid, privacy is gone." To place control of one's personal information into the hands of the data subject is truly a game changer. It moves privacy well beyond the normative sphere of laws and best practices, and into an actual technological and marketplace context, characteristic of *Privacy by Design* (*PbD*).

These developments are in contrast to the often asymmetrical relationship individuals have with organizations that collect, use, and disclose their personal information. Individuals must disclose information every day to different organizations. This results in fragmented and inefficient interactions between individuals and organizations, which can be reversed by making the individual the central point of integration for their personal data.

In this paper we have described the concept of a Personal Data Ecosystem and its first two key technological and services components, the Personal Data Vault and the data sharing platform. We have described how the PDE aligns with the concept of *PbD*, which also has at its core a focus on the user. In this way PDE can complement *PbD*. Given the shift in control occurring in the commercial and public spaces outlined in this paper, it is with great anticipation that the PDE is being realized.

PbD concepts will be helpful to PDE stakeholders wishing to get ahead of the privacy challenges we have identified in this paper. For example, simple and intuitive user interfaces that make it easy and natural for individuals to make decisions about their privacy will be essential. While regulatory frameworks, policies and procedures may be adequate for the status quo, to fully realize this paradigm shift, we need to do further work on the technology side. The idea of individuals having "SmartData" agents should be pursued. These virtual agents would represent our privacy interests online and through intelligent algorithms, they would be able to understand context and inferences about our personal data on the Web held by third party applications or organizations. Lastly, organizations whose goals are consistent with the PDE must proactively protect the privacy of their users by undertaking risk assessments and by making privacy the default setting, through data minimization and other means, at every stage of the information lifecycle.

Privacy must be embedded into the each layer of the PDE, accompanied by exemplary systems security and accountable business practices. By following these measures, the future of privacy will be assured – now, and well into the future.

Resources

Books

Cavoukian, A., & Hamilton, T. (2002). *Privacy Payoff: How successful businesses build customer trust*. Toronto: McGraw-Hill Ryerson.

Cavoukian, A., & Tapscott, D. (1995). *Who Knows: Safeguarding your privacy in a networked world*. Toronto: Random House of Canada.

Mitchell, A. (2002). *Right side up: Building brands in the age of the organized consumer*. London: HarperCollinsBusiness.

Searls, D. (2012). *The Intention Economy: When Customers Take Charge*. Boston: Harvard Business Press Books.

Reports

Ctrl-Shift. (2009). The new Personal Communication Model: The rise of Volunteered Personal Information.

Ctrl-Shift. (2011). How customer friendly are retailers' privacy policies?

Ctrl-Shift. (2011). The new personal data landscape.

Ctrl-Shift. (2012). Personal Data Stores: A Market Review.

KuppingerCole. (2012). Life Management Platforms: Control and Privacy for Personal Data.

Mydex. (2010). The Case for Personal Information Empowerment: The rise of the personal data store.

Mydex. (2011). Mydex and the UK government's new midata policy, from http://mydex.org/2011/11/03/mydex-uk-governments-midata-policy

National Institute of Standards and Technology. (2011). *Special Publication 800-145: The NIST Definition of Cloud Computing*.

World Economic Forum. (2010). Personal Data: The Emergence of a New Asset Class

World Economic Forum. (2012). Rethinking Personal Data: Strengthening Trust.

Papers

Craig Burton, Scott David, Drummond Reed, Doc Searls, & Windley, P. J. (2012). From Personal Computers to Personal Clouds: The Advent of the Cloud OS.

Drummond Reed, Joseph Johnston, Scott David. (2011). The Personal Network: A New Trust Model and Business Model for Personal Data.

Drummond Reed, & Windley, P. J. (2012). The Personal Channel: The Extraordinary Benefits of Communicating Via Personal Clouds *The Live Web Series*.

Information and Privacy Commissioner of Ontario. (2006). 7 Laws of Identity: The Case for Privacy-Embedded Laws of Identity in the Digital Age.

Information and Privacy Commissioner of Ontario. (2008). *Privacy in the Clouds: Privacy and Digital Identity - Implications for the Internet*

Information and Privacy Commissioner of Ontario. (2009). Privacy and Government 2.0: The Implications of an Open World.

Information and Privacy Commissioner of Ontario. (2009). *Privacy by Design*.

Information and Privacy Commissioner of Ontario. (2009). *Privacy by Design*: The 7 Foundational Principles.

Information and Privacy Commissioner of Ontario, Hewlett-Packard, & The Centre for Information Policy Leadership. (2009). Privacy by Design: Essential for Organizational Accountability and Strong Business Practices.

Information and Privacy Commissioner of Ontario, & Kim Cameron. (2011). Wi-Fi Positioning Systems: Beware of Unintended Consequences.

Information and Privacy Commissioner of Ontario, & Liberty Alliance Project. (2009). The New Federated Privacy Impact Assessment (F-PIA).

Information and Privacy Commissioner of Ontario, & National Association for Information Destruction Inc. (2009). Get rid of it Securely to keep it private: Best practices for the secure destruction of personal health information.

Information and Privacy Commissioner of Ontario, J. B. W. (2012). *Privacy by Design* and User Interfaces: Emerging Design Criteria - Keep it User-Centric.

Information and Privacy Commissioner of Ontario, J. J. (2012). *Privacy by Design* in the Age of Big Data.

NEC Company Ltd., & Information and Privacy Commissioner of Ontario. (2010). Modelling Cloud Computing Architecture Without Compromising Privacy: A Privacy by Design Approach.

Tomko, G. J., Hon Kwan, & Don Borrett. (2012). SmartData: The Need, the Goal, the Challenge.

Journal Articles

Laudon, K. C. (1996). Markets and privacy. *Communications of the ACM, 39*(9), 92.

Samuelson, P. (2000). Privacy as Intellectual Property? *Stanford Law Review, 52*, 1125.

Shilton, K., Burke, J., Estrin, D., Govindan, R., Hansen, M., Kang, J., & Mun, M. (2009). Designing the personal data stream: Enabling participatory privacy in mobile personal sensing.

News Articles

Fertik, M. (2012, April 3, 2012). Your Future Employer Is Watching You Online. You Should Be, Too, from http://blogs.hbr.org/cs/2012/04/your_future_employer_is_watchi.html

John, J. S. (January 18, 2012). California Gets the Green Button, *Greentechmedia.com*. from http://www.greentechmedia.com/articles/read/california-gets-the-green-button/

Robin, F. (August 1, 2012). The emerging market that could kill the iPhone: A handful of tech startups are competing for a foothold in the nascent market for personal data control. And that could mean major changes for the likes of Apple, Google and Microsoft. *Fortune/CNN.com*.

Swift, M. (July 4, 2011). Battle brewing over control of personal data online, *San Jose Mercury News*.

Speeches

Kuneva, M. (March 31, 2009). *Keynote Speech: Roundtable on Online Data Collection, Targeting and Profiling*. Brussels.

Studies

McCann Worldgroup Company. (2011). The Truth About Privacy: What every marketer who handles consumer data should know.

Websites

Personal.com. http://www.personal.com

Personal Data Ecosystem Consortium. http://pde.cc/

Reputation.com. http://www.reputation.com

Respect Network. http://respectnetwork.com

United Kingdom Department for Business Innovation & Skill. Midata 2012 review and consultation, http://www.bis.gov.uk/Consultations/midata-review-and-consultation?cat=open

United States Department of Veterans Affairs. What is the Blue Button Initiative?, http://www.va.gov/bluebutton/

White House Blog. (2012). Green Button Giving Millions of Americans Better Handle on Energy Costs, http://www.whitehouse.gov/blog/2012/03/22/green-button-giving-millions-americans-better-handle-energy-costs

World Economic Forum. Rethinking Personal Data, http://www.weforum.org/issues/rethinking-personal-data/

Overview of the Organizations

Office of the Information and Privacy Commissioner of Ontario, Canada (IPC)

The role of the Information and Privacy Commissioner of Ontario, Canada, is set out in three statutes: the *Freedom of Information and Protection of Privacy Act*, the *Municipal Freedom of Information and Protection of Privacy Act* and the *Personal Health Information Protection Act*. The IPC acts independently of government to uphold and promote open government and the protection of personal privacy. Under the three *Acts*, the Information and Privacy Commissioner: resolves access to information appeals and complaints when government or health-care practitioners and organizations refuse to grant requests for access or correction; investigates complaints with respect to personal information held by government or health-care practitioners and organizations; conducts research into access and privacy issues; comments on proposed government legislation and programs; and educates the public about Ontario's access and privacy laws. More at: www.ipc.on.ca and www.privacybydesign.ca.

Personal.com

Personal (www.personal.com) gives individuals a web and mobile vault for securely sharing and storing data, notes, files, and photos so they can better manage their lives at home and work. Built on a privacy- and security-by-design platform, Personal helps individuals to connect with trusted people, organizations and apps through private networks they control, and to properly leverage their data so they can realize the most value from it -- all with the peace of mind that the data remains legally theirs, they know who can access it, and it cannot be shared without their permission. In addition to being on the web, Personal is available for iPhone and Android devices, and recently launched the Personal Platform for developers: http://developer.personal.com/.

Respect Network

Respect Network is building the world's first trusted personal data network. All members, both individuals and businesses, subscribe to the Respect Trust Framework, the new model for personal data sharing that is listed with the Open Identity Exchange and won the Privacy Award at the 2011 European Identity Conference. Individuals can join the Respect Network today through the Connect.Me social business card and social directory service available at http://connect.me/. More information about the Respect Network is available at http://respectnetwork.com/.

Ctrl-Shift

Ctrl-Shift is a specialist research/consultancy organization focused on helping organizations understand and respond positively to the changing nature of consumer empowerment. One of these changes is the trend towards individuals acting as managers and controllers of their own data. More at: http://ctrl-shift.co.uk/.

Office of the Information and Privacy Commissioner, Ontario, Canada

2 Bloor Street East
Suite 1400
Toronto, Ontario
Canada M4W 1A8

Web site: www.ipc.on.ca
Privacy by Design: www.privacybydesign.ca

October 2012

Ctrl-SHIFT

Information and Privacy Commissioner, Ontario, Canada

Respect Network