Association for Information Systems
AIS Electronic Library (AISeL)

Wirtschaftsinformatik Proceedings 2013 | Wirtschaftsinformatik

2013

Personal Information Markets AND Privacy: A New Model to Solve the Controversy

Alexander Novotny
Vienna University of Economics and Business, Vienna, Austria, alexander.novotny@wu.ac.at

Sarah Spiekermann
Vienna University of Economics and Business, Vienna, Austria, sarah.spiekermann@wu.ac.at

Personal Information Markets AND Privacy: A New Model to Solve the Controversy

Alexander Novotny and Sarah Spiekermann

Vienna University of Economics and Business, Vienna, Austria
{alexander.novotny,sarah.spiekermann}@wu.ac.at

Abstract. From the early days of the information economy, personal data has been its most valuable asset. Despite data protection laws, companies trade personal information and often intrude on the privacy of individuals. As a result, consumers feel out of control and lose trust in electronic environments. Technologists and regulators are struggling to develop solutions that meet businesses' demand for more personal information while maintaining privacy. However, no promising proposals seem to be in sight. We propose a 3-tier personal information market model with privacy. In our model, clear roles, rights and obligations for all actors re-establish trust. The 'relationship space' enables data subjects and visible business partners to build trusting relationships. The 'service space' supports customer relationships with distributed information processing. The 'rich information space' enables anonymized information exchange. To transition to this model, we show how existing privacy-enhancing technologies and legal requirements can be integrated.

Keywords: informational privacy, personal data markets, privacy regulation

1 Introduction

The digital economy is in a deep crisis. From the inception of the digital economy, personal information (PI) has emerged as its core asset. PI is "any information relating to an *identified or identifiable* natural person" [1]. Abundantly leveraged as a free commons, PI is at the core of Internet economics and is considered the motor for online innovation. "Personal data is the new oil of the Internet and the new currency of the digital world" [2]. It finances the Internet's free content. It strengthens an Internet company's competitive stance. In many cases, it is even the only ingredient that brings an online service to life (e.g., social networking).

However, PI is also at the core of three facets that constitute humanity: Identity, dignity and privacy. And as PI is increasingly collected, used, packaged, and sold, more conflict arises around how people can retain control of their identities and protect their dignity and privacy. Under the umbrella terms "data protection" and "privacy" - the ability to control both the circulation of PI (out-flowing information) and the access of others to the self (in-flowing information) [3] - a global political debate has

11th International Conference on Wirtschaftsinformatik,
27th February – 01st March 2013, Leipzig, Germany

emerged. This debate centers on whether people shall be enabled to control their PI and which aspects companies shall be allowed to use.

Unfortunately, the economic realities of personal data markets on one side and data protection efforts on the other are drifting apart. Companies capitalize on opportunities to collect and trade PI at an unprecedented scale. Uncontrolled PI trading has evolved [4]. Every time users surf online, an average of 56 parties track their activities on a website, largely without their consent or knowledge [5]. Companies claim "legitimate" business interests in the data they collect. The digital marketing association thinks that "marketing fuels the world" [6]. Major self-regulatory efforts of the industry, such as the Safe Harbor Agreement and the Do-Not-Track initiatives, are failing [6]. As a result, almost every regulatory privacy framework in the world (EU data protection directive 95/46/EC, Convention 108, OECD Data Protection Guidelines, US Bill of Rights Proposal, and more) is now being overhauled with the goal of strengthening consumer rights.

However, will regulation and self-regulation initiatives achieve what they say they aim for? With increasing business interest in personal information and an escalating conflict between privacy rights groups, regulators and industry, we believe that the time is ripe to develop a tenable vision of PI markets. This vision must allow for an innovative information-rich world while maintaining privacy. Fruitful streams of research and innovation depend on data about individuals. However, harm to human dignity and privacy must be avoided, and people must remain masters of their identities. What if we had digital markets that used and traded PI but let people control their information and identities?

Because of incongruous technical, economic and legal assumptions, it seems as if we are far from shaping such a future. Technology scholars have developed valuable privacy-enhancing technologies (PETs) that could put PI management back into consumers' control [7-8]. However, their technical proposals often build on the assumption that people prefer anonymity in transactions with companies [9-10]. Consumers, in contrast, often don't mind being identified in transactions with business partners, and companies are keen to foster 'personal' relationships [11]. While most PET proposals imply that consumers will invest time into privacy management, people simply expect regulators to protect them and companies to behave in an ethical way [12]. Finally, the PET community insists on terms such as "data minimization" [13], which are hardly realistic in times when users submit 95 million tweets on Twitter and send about 47 billion (non-spam) e-mails on an average day. The result is a patchwork of PET solutions that are adopted by neither industry nor governments.

Besides the difficulties to deploy easy to use PETS, economists disagree on the effects of privacy on welfare [14]. Chicago school proponents argue that PI disclosure benefits society because information asymmetries are reduced [15-16]: As companies learn more about their customers, they can serve customer preferences better. In contrast, critics contend that privacy protection generates social welfare [17]. Everyone acknowledges that people need control over the use of their PI [18-19]. But no consensus is reached yet on whether people should legally own their PI and get a property right [1], [19]. Many want to view privacy exclusively as a human rights issue because they are concerned that people could be 'propertized' [3], [20]. But giving peo-

ple control over their PI has driven human-rights based privacy regulation so far [18]. As a result, only a few scholars have theorized about how PI markets could be organized with privacy [3], [19], [21-22]. Where this has been done, models fail to consider the current technological landscape and legal environment.

This article makes an attempt to fill the visionary gap we need to make privacy efforts work in the economic environments we face. Based on insights about consumer behavior, market mechanisms, existing regulation and privacy technologies, we propose a 3-tier model for PI markets. Our model embraces information richness as the future of a digital economy. 'Social data' originating from people will inevitably be an important resource. We acknowledge that many transactions will be identified. However, the market we propose aims to empower people as much as companies. People and companies are assigned a few core rights and obligations resulting in a new and simple market structure. Many of these rights and obligations are already established; however, they are either weakly enforced or their importance is not recognized by policy makers. In our model, company obligations vis-à-vis consumers are enforced by the law and supported through privacy-enhancing technologies. The design of our model is guided by the principle of combining legal and technical enablers which mutually complement each other in asserting market rule enforcement. Our model is limited to the private commercial PI sphere, excluding government activity.

In the next section, we describe our vision of a functioning PI market in which privacy can be preserved and consumer trust in PI handling can be re-established. In the subsequent sections, this hypothetical market model is described in detail, including the derivation of technical and legal requirements to enforce it. The paper closes with a critical discussion of our model's benefits and challenges.

2 A Three-tier Model for PI Markets

The model builds on the existing PI ecosystem. Currently, this system is complex and opaque, and its players engage in many secondary data use activities that undermine consumer privacy and trust [4]. We create transparency and simplicity by assigning existing players to a simple three-tier market structure (see Figure 1). The *first market tier*, which we call 'relationship space', includes the business relationship between data subjects and 1st tier partners. The *second market tier*, 'service space', includes the distributed computing and service infrastructures that enable today's business relationships. It integrates all those processors who need to receive customers' PI to directly enable and enrich 1st tier services. We distinguish between service delivery providers, which are necessary to perform the principal service, and service enhancement providers, which contribute to the 1st tier business relationship. The *third market tier*, 'rich information space' encompasses those players who do not directly support the 1st tier relationship. Participants in this part of the market can process as much data as they want, but the data they work on needs to be anonymized - to a degree that it cannot be linked with reasonable effort to 1st or 2nd tier transactions or data subjects. Each time PI is transferred to 'rich information space', it has to pass what we call the

'anonymity frontier'. When information passes the frontier, it loses its 'personal' nature.

The stakeholders in our model are connected by contractual relationships. For any given relationship, market actors are unambiguously assigned to one of the three tiers. Table 1 summarizes the rights (Right 1-3) and obligations (Obl. 1-9) of all actors in our model. Usually, the data subject and 1st tier partner agree on a contract governing the exchange of service, compensation and PI. 1st tier partners arrange service-level agreements with service delivery and enhancement providers specifying the expected service quality. In exchange, service delivery and enhancement providers receive monetary compensation, or the right to use and sell anonymized information (AI). Market participants in the 3rd tier close sales contracts over AI with other actors.

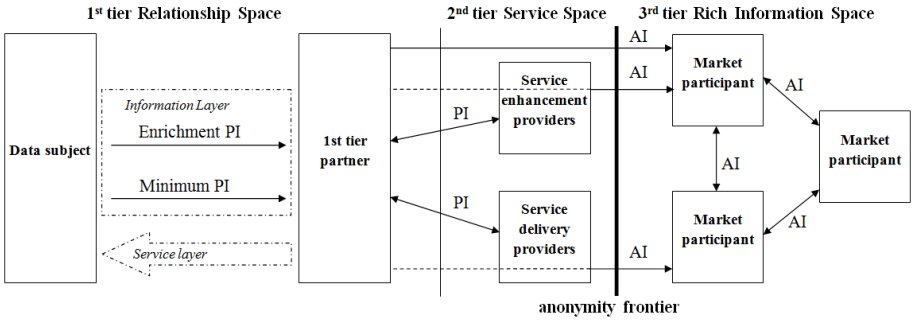

Fig. 1. Three-tier model for PI markets

2.1 The 1st Market Tier: Relationship Space

The 1st market tier is termed a "relationship space": visible 1st tier partners maintain identified one-to-one relationships with their customers. All PI they receive is the recognized property of their customers and can be used only for purposes set down in PI usage policies, which accompany every PI exchange. The 1st tier is characterized by identified business relationships between users and one visible company, a separation of service and information exchange and the right to a privacy-friendly service, legitimized information collection, people's property rights in their personal information and liability of the 1st tier partner for any PI abuse. The next paragraphs justify these characteristics from an economic and human rights perspective and point to their technical and legal implementation.

Identified Business Relationships and a Unique, Visible 1st Tier Partner. Because personalized customer relationships have proved effective, companies have invested in CRM solutions. Companies need and want identified customer relationships [11]. And many customers are willing to provide their PI in the service context if they receive appropriate benefits. Therefore, we depart from traditional data protection visions that promote the idea of total anonymity vis-à-vis companies [9].

However, users want predictable relationships in which they can control the use of their PI [23]. Predictability is supported when users are dealing with only one visible

PI-collecting business partner. We define *partner visibility* as a state in which data subjects visiting a physical or electronically-enabled premise can unambiguously and effortlessly name the commercial entity that they are transacting with. Customers in a physical retail store such as Walldepot see Walldepot as the 1st tier partner (and not, for example, the shelf suppliers).

Table 1. Actors in the three-tier model and their rights and obligations

		Property right in PI	Right to a privacy-friendly service	Right to alienate AI	Obtaining legitimization for PI usage	Handling of PI in accordance with PI usage policy	Offering a privacy-friendly service option with minimum PI	Acknowledgement of standardized PI usage policies	Initiating an accountability system	Request for authorization of 1st tier partner when using PI	Separation of PI from multiple data subjects or 1st tier partners	Demonstrate PI usage rights to authorities	Anonymization of information exchanged in the 3rd tier
		1	2	3	1	2	3	4	5	6	7	8	9
Role	**Definition**	**Rights**			**Obligations**								
Data subject	Natural person disclosing PI in the course of a service transaction in a business relationship with the 1st tier partner.	x	x										
1st tier partner	Visible and primary opposite party in the service transaction and, from the viewpoint of the data subject, the party that is responsible for the PI.			x	x	x	x	x	x		x	x	x
Collector	Party that gathers the PI from the data subject either by interrogation or observation.			(x)		x					x	x	x
Controller	„Natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing" of PI (Art 2 Directive 95/46/EC).			(x)		x				x	x	x	x
Service delivery provider	Entity authorized by the 1st tier partner that is necessary to perform the principal service.			(x)		x				x	x	x	x
Service enhancement provider	Entity authorized by the 1st tier partner that is not the service delivery provider but contributes by sufficiently close enrichment to the relationship between the 1st tier partner and the data subject.			(x)		x				x	x	x	x
Market participant	Any party including businesses, private persons, and governments who exchanges AI with other entities in the marketplace.			x									x

All parties having a contractual agreement with data subjects need to be visible. Otherwise, they are not allowed to collect any PI through mechanisms such as cookies or uploaded software daemons. If data aggregators and brokers want to collect PI from users, they need to establish a distinct and visible relationship with data subjects.

The reason for this one-partner rule is that people lose control when multiple parties invisibly collect their PI at the same time. This loss of control promotes distrust on the web [23]. From a company perspective, the one-partner rule enables companies to regain the monopoly on PI collection in their transactions. This increases the power that they get from competitive information, as unrelated data traders will not have access to identified information anymore.

Selected technical enablers: identity and claim mechanisms; graphical user interface design. ***Legal enablers:*** use of standardized symbols for signaling 1st tier partner; mandatory 'one-visible-partner' rule; legal liability of 1st tier partner for PI use.

Separation of Service and Information Exchange and the Right to a Privacy-Friendly Service. Today, most online transactions are of a composite nature. Information is collected as a service spin-off [24] without making the 'information deal' visible to the customer. In our model, companies are asked to distinguish an information layer and a service layer within a business relationship. The service layer embraces the delivery of the principal service to the data subject, such as the sale and delivery of a book. Within the information layer, PI is split into necessary information for delivering a service ("minimum information") and additional information that is used to enrich and enhance the service experience ("enrichment information"). Minimum information can be defined as the set of PI that is necessary and sufficient to perform the principal service. For the online book retailer, the minimum information is the name, delivery address and payment information. The individual's purchase history, date of birth, and affinity profile, in contrast, is what we consider enrichment information. Additionally, partners are obliged to offer one service option (Obl. 3) that requires data subjects to disclose only the minimum amount of their PI. Thus, people always have the right to a privacy-friendly service (Right 2). This right repackages the existing concept of "data minimization" [13] but limits its scope to users preferring data minimal over information rich services.

Consider, for example, a web search engine, look.com, which offers three service options. Selected by default, the privacy-friendly option requires the individual to pay a subscription fee of, let's say, € X per month; this option neither records search queries of the data subject nor shows any personalized ads. In contrast, the € Y priced second option collects more PI and uses it for an agreed time period to provide a richer service experience, such as individualized search results. The third option commercially leverages users' PI for an agreed time period, such as for the targeted placement of ads. This option may be free. The user trades his or her PI in exchange for the free search service. The 'free' mentality governing online business relationships today would make room for a more realistic view of what digital services are actually worth. The result of such a separation of service options would benefit all market participants: Competition in the market for PI may be improved because the salience of the information transaction increases [24]. In addition to service quality, marketers could compete on PI usage rights and privacy. They could realize new revenue streams from privacy-friendly service options. And people would finally get a true choice of PI disclosure options.

A market challenge is that 1st tier partners could deliberately create opacity by providing myriad options, with variations on factors such as retention times or usage purposes for the PI. We therefore see the need for standardized PI usage policies that are adopted by regulators, at least for the privacy-friendly baseline offer (Obl. 4). ***Technical enabler:*** standards for the presentation of minimum PI service options. ***Legal enabler:*** mandatory separation of the service deal from the PI deal; obligation

to offer one service option with minimum information use at reasonable quality and price (Obl. 3); mandatory compliance with standardized privacy policies.

Legitimized Information Collection and Liability. The legitimization of data collection is probably the most important bridge between US American and European data protection frameworks [25]. Legitimization justifies the collection and use of PI. It can be obtained either through the active consent (Obl. 1) of a data subject or by legal empowerment; for example, mobile operators are legally required to preserve some connection data. Consent is a voluntary, timely agreement of the data subject to the 1st tier partner's PI terms, which should be explicitly communicated to the data subject [26].

Reconsider the search engine example from above: The default option would be the privacy-friendly version of a service. At one click, customers can explicitly opt into the free version. Whatever service option an individual chooses, all parties handling PI shall respect the agreement between data subjects and their 1st tier partners manifested in electronic PI usage policies (Obl. 2). Software agent solutions, such as P3P agents, enable people to initially configure their privacy preferences in their clients once (i.e., in the browser); for example, people might object to data processing for marketing purposes. A client-based architecture choice gives users more control over settings [27]. The user's software agent matches PI usage preferences with companies' standard usage policies (cf. 'Privacy Bird' presented in [7]) and supports the negotiation of an agreed PI usage policy. People are empowered to actively take advantage of their legal rights in every transaction and companies benefit from better data quality and compliance.

Moreover, the 1st tier partner is legally liable for any collection and use of PI. Liability safeguards the data subjects' property right and a contractually agreed PI usage policy. Liability of the 1st tier partner is natural from a customer perspective. The 1st tier partner acts as the single point of contact for the data subject. Most importantly, we envision that the 1st tier partner is responsible for implementing a technical accountability system that ensures that PI usage rights set down in electronic PI usage policies are obeyed (Obl. 5). Accountability ensures that any access, use, disclosure, alteration, and deletion of PI can be traced back to the party who has done so by using technical means. The 1st tier partner shall therefore have a technical infrastructure that can demonstrate PI usage rights to authorities and auditors at any time (Obl. 8).
Technical enablers: standards for the presentation and content of PI usage policies, software agent-supported privacy policy negotiation, use of an accountability system to enable and monitor policy-compliant use of PI (e.g., sticky policies, audit logs).
Legal enablers: 1st tier partners obtain legitimization for PI usage; handling of PI in accordance with electronic PI usage policies (Obl. 2); legal obligation to have and regularly audit an accountability system; 1st tier partner liable for all PI transactions.

Property Rights to Personal Information. A core component of our model is that data subjects have property rights for their PI (Right 1). The property right to PI cannot be alienated [1], [19]. Because of its personal rights character - similar to moral rights in copyright - seizing PI-related rights shall be prohibited. It is the characteristic

of identifiability that inseparably binds PI to an individual. Identifiable information can never be an object separable from a beholder, such as a book can be divided from its owner. However, usage rights to PI can be transferred. From a human rights perspective, data subjects have the biggest interest in the PI asset. Thus, they are the natural holders of this property right.

The main reason for proposing property rights to PI is a psychological one: Property rights would create stronger asset awareness in the minds of all stakeholders. The awareness that PI is an asset of economic value makes data subjects more informed when deciding about disclosing PI [28]. Equally, companies will probably be more cautious and reflective in collecting and using it. To make people aware of this asset, we must label information self-determination rights as "property rights".

Technical enabler: policy repository on the client side. ***Legal enabler:*** recognition of a property right to PI (for an elaborate discussion of this proposal, see [18-19]).

2.2 The 2nd Market Tier: Service Space

Typically, the 1st tier partner is assisted by subcontractors, outsourcers, and strategic alliances to deliver services and products. This complex service web adds to the insecurity of today's personal information markets. In fact, consumers are most concerned about secondary uses of their data by invisible partners [29]. For this reason, we create a 'market chunk' where this web of invisible service providers is organized. The 2nd tier includes all companies that contribute to the services delivered in the 1st tier.

PI abuses arise when parties at greater distance from the initial service perceive less responsibility for the PI they use. To extend the context-based trust between data subjects and 1st tier partners, 2nd tier service providers must be legally tied to the initial business relationships. This tie is created via a chain of accountability that ensures authorization, non-repudiation, separation, and auditability. Since all 2nd tier providers need to serve the 1st tier business relationship with the customer, our model ensures contextual integrity of PI use. PI is used within the boundaries of contextual integrity when the applicable social norms of appropriate PI collection and distribution are upheld in a given situation [30]. The following characteristics enable the 2nd tier:

Tying the Service Space to 1st Tier Relationships. We distinguish between service delivery and service enhancement providers (see Table 1). Service delivery providers such as parcel services delivering book orders are necessary to perform the principal service. They are always immediately involved in the 1st tier relationship and, for instance, include entities supporting the accountability and security of transactions. Service enhancement providers might also need to receive PI. These providers are parties that *directly* or *immediately contribute* to the 1st tier business relationship. For instance, they provide advertisements matching the interests of book purchasers. In case a data subject chooses such an enhanced service option, the service delivery providers can also handle enrichment information and service enhancement providers can process minimum information. ***Technical enablers:*** privacy policy language; accountability system to enable and monitor policy-compliant use of PI. ***Legal enabler:*** legal obligation to have and regularly audit an accountability system.

Authorization, Nonrepudiation, Separation and Auditability. For 2nd tier parties, an accountability system must comply with the requirements of authorization, nonrepudiation, separation, and auditability. First, authorization requires that access to PI by the service provider is approved by the 1st tier partner on an individual transaction basis (Obl. 6). When a customer purchases a book, the online shop must explicitly authorize a credit-scoring agency to use customer data for a credit check. Second, nonrepudiation prevents service providers from falsely denying that they have accessed, used, altered or deleted PI. Third, separation requires that PI units stemming from different service transactions, data subjects, and 1st tier partners are kept in strict isolation unless the legitimized purpose allows for the combination of PI (Obl. 7). This safeguards contextual integrity. Fourth, auditability ensures that compliance can be demonstrated at any time to authorities and auditors (Obl. 8). ***Technical enabler:*** use of an accountability system to monitor policy-compliant use of PI (e.g., sticky policies, audit logs). ***Legal enabler:*** separation of PI from multiple data subjects or 1st tier partners; legal obligation to and auditing of the accountability system.

2.3 The 3rd Market Tier: Rich Information Space

The 3rd tier is a market space where businesses, individuals, governments, and other parties not contributing to an identified business relationship freely exchange and trade information. They, however, need to ensure *anonymity* according to state-of-the-art technical standards. PI may originate from data subjects, but when the anonymity frontier is passed, this information becomes a *freely exchangeable* good. Innovation can be vividly spurred on the basis of this data. We assume that the marginal utility from identification outside of business relationships is so minimal that it does not justify the ensuing privacy risks. Severe *sanctions* should be imposed on 3rd tier market players who distort competition by holding identifiable or re-identifiable data.

Anonymity and Sanctions. Data subjects want to retain control over the distribution of their PI and want to share in good peace of mind. A straightforward way to create control and peace of mind is to legally enforce anonymity of all data except in situations where identification is needed or desired by the customer (1st and 2nd tier). People are granted a privacy commons, a shared space of anonymity [19]. In our model, this space is created by ensuring that PI cannot leave the contextual boundaries of the 1st and 2nd tier. When it does, it must be anonymized. What constitutes sufficient anonymization is a dynamic concept dependent on the current state-of-the-art of technology. Regulators should document and update current standards for anonymization in so called "BREF"s, best available techniques reference documents, which have been applied successfully for integrated pollution prevention and control (IPPC, Directive 2010/75/EU). Currently, the concepts of k-anonymity [10], l-diversity [31] and t-closeness [32] suggest that it is sufficient to have a large anonymity set of individuals, diverse attribute values and similar attribute value distributions. Each market participant in the 3rd tier is obliged to respect these anonymity mechanisms (Obl. 9) and is regularly audited for the fulfillment of this requirement. Finally, damages and

penalties for the illegal acquisition, possession, use or sale of identifiable information are necessary to protect a trustworthy market regime. ***Technical enabler:*** anonymization. ***Legal enabler:*** legal obligation and auditing of anonymity requirement in 3rd tier; sanctions for breaking the anonymity rule.

Free Exchange. Free trade of anonymized information increases the amount of exchanged information. Any market participant shall have free access to the 3rd tier market, including data subjects who may want to sell their anonymized information directly. Compensating for the costs 1st tier partners incur in our model, they have the right to anonymize and sell any PI collected independent of the data subjects' consent. Market participants can resell anonymized data once they acquire it (Right 3). ***Technical enabler:*** privacy-preserving data mining. ***Legal enabler:*** right to alienate AI.

3 Implementing the Three-tier Model

As has been outlined throughout Section 2, technical and legal enablers must support the implementation and enforcement of our model. Many of these technologies and legal enablers already exist. This section outlines how our model builds on these existing enablers and which need to be developed or changed.

3.1 Technical Enablers

Well-established privacy-enhancing and security technologies enable the enforcement of our model [33]. Table 2 gives an overview of selected technologies and assigns them to the relevant market tiers. To implement the requirement of accountability in the 1st and 2nd tiers, different systems based on sticky PI usage policies and audit logs are available [8], [34-35]. Most accountability systems suitable for ensuring contextual integrity are based on cryptographic technologies that can be easily applied in distributed environments [36-37]. Determining the responsible party for a data breach can be achieved by available identity technology. Existing security mechanisms, such as SAML, can identify the 1st tier partner and the data subject [38]. To specify the content of PI usage policies privacy policy languages are necessary [39]. Some privacy policy languages have already been standardized by the W3C consortium (P3P). Since negotiating these policies is a laborious and complex task for the data subject and 1st tier partner, architectures can make the task easier by employing software agents that semantically understand policy content [40]. The usability of privacy functionality and user agents at the interface between human individuals and machines is more and more improved [7]. Although data subjects are possibly identified on the application layer, they might want to be anonymous to third parties on the communication layer; to ensure their anonymity, data subjects can employ existing web anonymity technologies that protect the interaction between data subject and business partner [9]. Anonymity on the web can be supported by the "do not track" functionality that many web browsers will offer; this functionality indicates to the communication partner that no PI shall be collected. Additionally, anonymization technologies

are needed to realize sufficient anonymization of PI in the 3rd tier [10], [31-32]. Privacy preserving data mining technologies can guarantee endured anonymity [41].

Table 2. Assortment of existing technologies to support enforcement in the market tiers

Relationship Space (1st tier)	Service Space (2nd tier)	Rich Information Space (3rd tier)
Accountability system Sticky policy, Privacy injector, Privacy-aware access control, Distributed auditing logs		Anonymization k-anonymity, l-diversity, t-closeness, graph anonymity
Identity mechanisms SAML, OAuth, OpenID		
Contextual integrity-compatible cryptography Identifier-based encryption, NOYB		
Privacy policy languages POL, PrimeLife policy language, E-P3P, EPAL, Rei, EnCoRe, PERFORM, Ponder, Contextual Integrity language		
Privacy policy negotiation P3P, PISA		Privacy-preserving data mining Randomization, Perturbation, Differential privacy, KD cycle-based data mining
Web anonymity and pseudonymity agents LPWA, Crowds, Hordes, Onion Routing, Mixminion		
Do Not Track		
Human-computer interface Privacy pictograms, User privacy agent interface design, Visual tagging		

3.2 Legal Enablers

Our model shall not only be technically feasible, but shall also be meaningful to public policy. Policy makers need to know which of the rights and obligations we propose already exist in the current legal framework. One important idea is to consider PI as the private property of data subjects [18]. A property right to PI (Right 1) is reflected in the principles of informed consent (Art 7 Directive 95/46/EC, Para. 7 OECD, Art 2 FTC Fair Information Practices (FIP)) and the right to object (Art 14 Dir. 95/46/EC). Missing is the recognition of full property rights to PI in civil law, however.

So far, a data subject's right to a privacy-friendly service (Right 2) exists only at a very limited scope. For example, Art 8 Directive 2002/58/EC mandates service providers to offer an option preventing the presentation of calling line identification. All of our model's other obligations already exist in legal frameworks. For instance, the obligation of anonymizing any information exchanged in the 3rd tier (Obl. 9) already exists to some extent in the principle of data quality (Art 6 Directive 95/46/EC). PI should "be kept in a form which permits identification of data subjects for no longer than is necessary […]". To this vague formula, our model adds a clear anonymity frontier that unambiguously determines when anonymization takes place. Best available technique reference documents (BREF), kept current by data protection authorities, prescribe state-of-the-art anonymization technologies. Thus, only minor adaptations to the current legal framework are needed.

4 Discussion

Our vision for a personal information market establishes compromise between players in the current PI ecosystem and data protection proponents. Our model embraces the fact that data richness is the future of a digital economy and creates room for information-rich services and data trading as well as identified customer relationships. At the same time, our technical and legal suggestions empower people to participate in PI markets and protect their privacy. To help people understand their transactions with companies and the value of their PI, we create a new and simple market structure that assigns clear rights and obligations to all market players. Trust built by a clear allocation of rights also aids companies and legal enforcers.

We are aware that many of the rights, obligations and legal and technical enablers we propose are not new. They have been proposed for over two decades by researchers in privacy, identity, security, and legal studies and debated by companies and regulators. We do not need new security mechanisms which can, for instance, identify the 1st tier partner, but build on existing technologies which have been outlined in Section 3.1. However, no one has demonstrated how all of the puzzle pieces could be arranged in a market model to benefit both people *and* companies.

Personal information markets working to the benefit of people and the economy require that the enforcement of market rules is improved. The main design principle of our market model is to combine legal and technical mechanisms which mutually overcome its weaknesses. A legal property right to PI backed up by technical accountability of data usage simplifies law enforcement access for data subjects.

A sour apple that companies have to swallow is to finally provide people with a privacy-friendly default service option. But, as we show in this article, the apple isn't that sour. Companies can finally re-enter competition on the basis of service qualities. Furthermore, our model meets the privacy preferences of different individuals: Access to content at potentially lower cost for those who are willing to 'pay' with their PI and alternative versions for customers that are concerned about their privacy. Privacy rights proponents may argue that this preference-based market structure disadvantages the poor, who may be forced into selling their PI. This argument is true only if marketers choose to have people pay for the privacy-friendly version. Marketers could also make the data-rich version more attractive from a service perspective – with greater functionality and no ads – while offering a baseline service with non-personalized ads in a privacy-friendly way.

Finally, even if individuals opted into the usage of their PI in exchange for the service, our market proposal provides privacy protection: After all, companies would be accountable and liable for how they use PI. Limitless reuse and repackaging out of context would be outlawed. Privacy risks would hence be limited even for those who share. As data subjects will have property rights to their PI, they will also be brought back to the negotiating table. Property rights, a right to privacy-friendly service options and defaults, company accountability and a transparent market structure promise to re-establish the trust we need to see information services flourish.

A core benefit of our model is also its main technical challenge: the creation of a free market space that ensures anonymity. Ensuring anonymity becomes more diffi-

cult as technology becomes more powerful, facilitating identification. Anonymization could reduce the entropy of information to such an extent that the utility for information users would vanish. For multidimensional PI that contains many attributes about data subjects, the 'curse of dimensionality' forces that information is extensively aggregated to guarantee reasonable anonymity [41]. Utility-based privacy preservation, however, guarantees that the utility of anonymized data does not drop by more than a defined threshold ε, known as ε-differential privacy [42-43]. Data protection authorities define the "BAT" (Best Available Techniques) (Directive 2010/75/EU) that guarantee sufficient anonymity. Flourishing service spaces based on "non-identified, social data" instead of "personal data" may be the result. Information buyers want to obtain a representative sample of a population of individuals, not the information of identified single data subjects [41].

Finally, two more fundamental challenges of our model must be considered: concerns of 'monopolizing' information and the international enforceability of our model. The idea that personal data could be recognized as property originated in the US; this idea has been met by the criticism that people shouldn't be 'propertized' [3], [20] as well as a series of other arguments (for an overview see [19]). Ralph Waldo Emerson once remarked, "As long as our civilization is essentially one of property, of fences, of exclusiveness, it will be mocked by delusions." For these reasons, we view the idea of property rights to PI critically. However, because markets already treat PI as property, we ask only that people get the same rights that companies have already claimed for themselves. Moreover, a property right would not substitute, but rather enhance the human rights basis of privacy [18]. In Europe, it would provide people with an additional legal instrument, giving them easy access to existing, well-proven enforcement structures. Data subjects would be enabled to effectively claim their rights to PI on their own instead of calling on data protection authorities.

Another challenge of our model is its international practicability. Recent years have shown how difficult it is to reach international consensus on data protection or privacy. Even more difficult is enforcement. The Safe Harbor Agreement between the US and Europe on data handling practices is a good example of failure. A more effective path could be to implement and enforce binding hard-law for data protection. For example, property rights are enforceable as well-recognized legal instruments within both the European and US legal orders. If Europe and the US applied property rights to PI [25], the rest of the world would potentially follow suit.

Acknowledgement

We would like to thank Julian Cantella for the editing of this paper.

References

1. Bergelson, V.: It's Personal But Is It Mine? Toward Property Rights in Personal Information. UC Davis L. Rev. 37, 379 (2003)
2. Personal Data: The Emergence of a New Asset Class, World Economic Forum, Jan (2011)

3. Noam, E.M.: Privacy and Self-Regulation: Markets for Electronic Privacy. In: Wellbery, B.S. (ed.): Privacy and Self-Regulation in the Information Age, pp. 21-33. NTIA (1997)
4. Rethinking Personal Data: Strengthening Trust, World Economic Forum, May (2012)
5. Angwin, J.: Online Tracking Ramps Up - Popularity of User-Tailored Advertising Fuels Data Gathering on Browsing Habits. WSJ, June 18, B1 (2012)
6. Bott, E.: The Do Not Track Standard has Crossed into Crazy Territory, http://www.zdnet.com/the-do-not-track-standard-has-crossed-into-crazy-territory-7000005502/
7. Cranor, L.F., Guduru, P., Arjula, M.: User Interfaces for Privacy Agents. ACM Transactions on Computer-Human Interaction 13, 135-178 (2006)
8. Karjoth, G., Schunter, M., Waidner, M.: Privacy-Enabled Services for Enterprises. In: 13th International Workshop on DEXA, pp. 483-487, Aix-en-Provence (2002)
9. Gritzalis, S.: Enhancing Web Privacy and Anonymity in the Digital Era. IMCS 12, 255-287 (2004)
10. Sweeney, L.: k-Anonymity: A Model for Protecting Privacy. IJUFKS 10, 557 (2002)
11. Spiekermann, S., Dickinson, I., Günther, O., Reynolds, D.: User Agents in E-commerce Environments: Industry vs. Consumer Perspectives on Data Exchange. In: Eder, J., Missikoff, M. (eds.): LNCS, Vol. 2681, pp. 696-710. Springer, Berlin (2003)
12. Personal Data in the Cloud: A Global Survey of Consumer Attitudes, Fujitsu Res. Inst. (2010)
13. Borcea-Pfitzmann, K., Pfitzmann, A., Berg, M.: Privacy 3.0 := Data Minimization + User Control + Contextual Integrity. IT 53 (1), 34-40 (2011)
14. Acquisti, A.: The Economics of Personal Data and the Economics of Privacy. 30 Years after the OECD Privacy Guidelines. OECD (2010)
15. Posner, R.A.: The Economics of Privacy. Am. Econ. Rev. 71, 405-409 (1981)
16. Calzolari, G., Pavan, A.: On the Optimality of Privacy in Sequential Contracting. Journal of Economic Theory 130, 168-204 (2006)
17. Acquisti, A., Varian, H.R.: Conditioning Prices on Purchase History. Marketing Science 24, 367-381 (2005)
18. Purtova, N.: Property Rights in Personal Data: a European Perspective. Dissertation, Uitgeverij BOXPress, Oistervijk (2011)
19. Schwartz, P.M.: Property, Privacy, and Personal Data. Harv. L. Rev. 117, 2056 (2003)
20. Cohen, J.E.: Examined Lives: Informational Privacy and the Subject as Object. Stanford Law Review 52, 1373-1437 (1999)
21. Laudon, K.C.: Markets and Privacy. Communications of the ACM 39, 92-104 (1996)
22. Aperjis, C., Huberman, B.: A Market for Unbiased Private Data: Paying Individuals According to their Privacy Attitudes. HP Working Paper (2012)
23. Smith, H.J., Milberg, S.J., Burke, S.J.: Information Privacy: Measuring Individuals' Concerns about Organizational Practices. MIS Quarterly 20, 167-196 (1996)
24. Jentzsch, N., Preibusch, S., Harasser, A.: Study on Monetising Privacy: An Economic Model for Pricing Personal Information. ENISA (2012)
25. Purtova, N.: Property Rights in Personal Data: Learning from the American Discourse. CLSR 25 (6), 507-521 (2009)
26. Art29WP: 01197/11/EN WP 187 - Opinion 15/2011 on the Definition of Consent, Article 29 Data Protection Working Party, Adopted on 13 July 2011 (2011)
27. Spiekermann, S., Cranor, L.F.: Engineering Privacy. IEEE Transactions on Software Engineering 35, 67-82 (2009)
28. Spiekermann, S., Korunovska, J., Bauer, C.: Psychology of Ownership and Asset Defense: Why People Value their Personal Information Beyond Privacy. In: International Conference on Information Systems (ICIS 2012), Orlando, FL (2012)

29. Culnan, M.J.: "How Did They Get My Name?": An Exploratory Investigation of Consumer Attitudes toward Secondary Information Use. MISQ 17, 341-363 (1993)
30. Nissenbaum, H.: Privacy as Contextual Integrity. Wash. L. Rev. 79, 119 (2004)
31. Machanavajjhala, A., Kifer, D., Gehrke, J., Venkitasubramaniam, M.: l-diversity: Privacy Beyond k-anonymity. TKDD 1, 3 (2007)
32. Li, N., Li, T., Venkatasubramanian, S.: t-Closeness: Privacy Beyond k-Anonymity and l-Diversity. In: 23rd IEEE ICDE '07, pp. 106-115 (2007)
33. Shen, Y., Pearson, S.: Privacy Enhancing Technologies: A Review. Report HPL-2011-113, Hewlett-Packard Laboratories (2011)
34. Ringelstein, C., Staab, S.: DIALOG: Distributed Auditing Logs. In: IEEE ICWS '09, pp. 429-436 (2009)
35. Mont, M.C., Pearson, S., Bramhall, P.: Towards Accountable Management of Identity and Privacy: Sticky Policies and Enforceable Tracing Services. In: 14th Int. Workshop on DEXA, pp. 377-382, Prague (2003)
36. Mont, M.C., Bramhall, P.: IBE Applied to Privacy and Identity Management. Technical Report HPL-2003-101. Hewlett-Packard Laboratories (2003)
37. Guha, S., Tang, K., Francis, P.: NOYB: Privacy in Online Social Networks. In: 1st Workshop on Online Social Networks, pp. 49-54. ACM, Seattle, WA (2008)
38. Recordon, D., Reed, D.: OpenID 2.0: a Platform for User-Centric Identity Management. In: 2nd ACM Workshop on DIM, pp. 11-16, Alexandria, VA (2006)
39. Berthold, S.: Towards a Formal Language for Privacy Options - Privacy and Identity Management for Life. In: Fischer-Hübner, S., Duquenoy, P., Hansen, M., Leenes, R., Zhang, G. (eds.): Privacy and Identity 2010, Vol. 352, pp. 27-40. Springer, Boston (2011)
40. The Platform for Privacy Preferences 1.1 Spec., W3C, 13 Nov (2006)
41. Aggarwal, C.C., Yu, P.S.: A General Survey of Privacy-Preserving Data Mining Models and Algorithms. In: Aggarwal, C.C., Yu, P.S. (eds.): Privacy-Preserving Data Mining, Vol. 34, pp. 11-52. Springer, New York (2008)
42. Ghosh, A., Roth, A.: Selling Privacy at Auction. In: 12th EC, pp. 199-208. ACM, San Jose (2011)
43. Dwork, C., McSherry, F., Nissim, K., Smith, A.: Calibrating Noise to Sensitivity in Private Data Analysis Theory of Cryptography. In: Halevi, S., Rabin, T. (eds.): TCC 2006. LNCS, Vol. 3876, pp. 265-284. Springer, Berlin (2006)