On data markets as a means to privacy protection
An ethical evaluation of the treatment of personal data as a commodity

Hendrik vom Lehn

Master Thesis

Faculty of Technology, Policy and Management
Delft University of Technology

August 2014

TU Delft Delft University of Technology

Master thesis information

Title:	On data markets as a means to privacy protection
Subtitle:	An ethical evaluation of the treatment of personal data as a commodity
Graduation date:	August 25, 2014
Programme:	MSc Engineering and Policy Analysis (EPA)
Faculty:	Technology, Policy and Management
University:	Delft University of Technology

Author information

Author:	Hendrik vom Lehn
Email:	h.vomlehn-1@student.tudelft.nl
Student number:	4240952

Graduation committee

Prof.dr. Jeroen van den Hoven	(Section Ethics / Philosophy of Technology)
Dr. David Koepsell	(Section Ethics / Philosophy of Technology)
Dr. Scott Cunningham	(Section Policy Analysis)

Contents

Executive summary

A number of technological developments such as cloud computing and big data analysis have affected the way in which personal data are processed. These developments go coupled with the currently prevalent business model of free online services that are financed through advertisements and an analysis of user data. Based on these developments, it seems that the new requirements have exposed deficits in the current approach to data protection in the European Union. In the debate on this topic, one of the solutions that are discussed is to create market structures in which users can sell personal data to businesses, thereby gaining control over the ways in which their data is used. Such an approach would constitute an alternative way to the protection of privacy, which is different from the current form of data protection. In order to better assess the validity of claims about the effectiveness of such an alternative approach, it therefore is of importance to know the possible effects that data markets would have on the privacy of online service users.

This study investigates this question by means of an ethical evaluation. Since the definition of privacy as such is highly contested, it is not straightforward to determine what an impact on privacy would constitute. To this end, a literature review on the different meanings of privacy is conducted first. The conclusion in this regard is that privacy is a cluster of concepts which does not allow for a single definition. However, for an ethical evaluation it is the moral reasons for the protection of privacy that should be in the focus, and the precise definition of privacy is of secondary relevance. Based on this result, an evaluation framework is constructed for use in this study.

Another aspect that requires clarification upfront for an ethical evaluation is the question of what specifically constitutes a data market. An investigation of the relevant literature in this regard shows that first instances of data markets are about to appear in practice, but that most proposals only exist in theoretical form. Only secondary markets for personal data — which are not accessible by users themselves — exist in practice. First approaches of real data markets seem to emerge, but are in a very early phase that is too premature for the sake of a detailed evaluation. Yet, there are a number of interesting approaches in the literature which propose the concept of a data market in abstract form. This study makes a selection of these proposals and uses them for an ethical evaluation.

The outcome of the ethical evaluation shows that there are a number of different effects that could occur if these data market approaches would be implemented. Although some of these effects are indeed positive for the protection of privacy, there are various effects that would be detrimental to privacy. Most importantly, data markets could lead to a loss of individual autonomy and have adverse effects on the societal function of privacy. Striking is also the symbolic change to privacy as a human right that a commodification of personal data might entail. Moreover, it has to be considered that data markets as a regulatory infrastructure would require the collection of additional data for their own functioning. This in turn leads to new questions of privacy protection that would have to be solved. Overall, it can be said that there is not a single and clear impact on privacy, but a wide range of possible effects that are connected in an intricate manner.

To which extent these effects would occur is contingent upon the behaviour of users in such markets and the design parameters of possible data market approaches. Central in this regard is the form in which users would gain access to a data market, and in which way they would be concerned with single market transactions that they engage in. Also, the scope of data markets is of relevance. Although detrimental to allocative efficiency, it would be beneficial for the protection

of privacy if data markets are restricted in their scope concerning the market participants and the type of data that is traded therein. Furthermore, the specific design of data markets is relevant for the behaviour of users, and thereby the consequences of their actions. The data market proposals that this study analyses are not specific enough in order to assess all of these parameters, but provide useful indications for elements that are of relevance in this regard.

Concerning the overall problems with data protection in an age of big data, it is not apparent at all whether data markets would indeed form a better way of protecting the privacy of online service users. Many of the improvements that data markets could bring could likely also be achieved by modifying the existing methods of data protection. Policy makers in the European Union should therefore not focus on the solution of data markets, but strengthen the existing and proven mechanisms for data protection. This entails providing sufficient funds to the supervisory authorities, stimulating research on the existing weaknesses of data protection, and speeding up the process of political decision making concerning data protection issues.

Preface

Coming from the field of computer science and having worked on the development of privacy-enhancing technologies, I got interested in the broader questions concerning the protection of privacy. Especially interesting in my opinion is the development towards a stronger economic importance of data, which is often discussed disconnected from the issues of privacy protection. I started to wonder in which way the economics of data and the protection of privacy are connected. I reflected on specific policy proposals, such as that of a tax on personal data discussed in France early 2013[1], but also got curious about the general implications for privacy if the value of data is increasingly discussed from an economic point of view.

When I started to look for a suitable topic for a master thesis, I therefore knew which field would be of most interest to me. A visit to the *7th International Conference on Computers, Privacy and Data Protection (CPDP)* that took place in Brussels in January 2014 gave me further clarity in this regard. I found it particularly interesting that data markets are also discussed on a European conference on data protection, although the discussion on this form of regulation seems to stem primarily from the United States. Although intuitively appealing in argument, I wondered what the overall privacy implications of data markets would be. I therefore used the opportunity to write my thesis for a masters degree in Engineering and Policy Analysis at Delft University of Technology on this topic.

Acknowledgements

During the process of writing this thesis I have been in contact with a number of people who helped me in defining the research topic and conducting the research. I am particularly grateful to my supervisors Jeroen van den Hoven, David Koepsell and Scott Cunningham. Their feedback during the process of writing my thesis helped me to focus on the relevant questions and improve possible weaknesses in my work. I furthermore would like to thank Thomas Baar and Ilse Oosterlaken, with whom I had important discussions that helped me in defining my research topic. Finally, I am thankful to Marc van Lieshout (TNO Delft) for his input on relevant literature on this topic.

[1] See my blog post from January 21, 2013 on this point: `http://hendrik.vomlehn.de/2013/01/21/the-french-data-tax-and-the-influence-on-privacy/` [accessed August 17, 2014]

1 Introduction

> When market reasoning travels abroad, beyond the domain of televisions and toasters, market values may transform social practices, and not always for the better.
>
> Michael Sandel (2013, p. 127)

Throughout the last years, privacy has been a highly discussed issue in many European countries. Besides the revelations about governmental spying through whistleblower Edward Snowden (The Guardian, 2013), the data collection practices of online service providers like Google or Facebook have been highly discussed. The technological changes surrounding cloud computing and the possibilities emerging through the application of big data analysis (Mayer-Schönberger & Cukier, 2013) have a huge impact on the privacy of online service users. Although the European Union is currently updating its legislation on data protection (Kuner, Burton, & Pateraki, 2014), it is not likely that the debate on privacy will come to a halt in the near future.

Data protection as a regulatory approach to protect privacy is supposed to set boundaries on the extent and the manner in which data collection may take place. Both the concept of privacy as such and the emergence of data protection as a regulatory approach are closely linked to technological developments. It is therefore not surprising that technological change leads to debates and the need for adaptations in these fields as well.

The currently ongoing technological change concerns the way in which information infrastructure is being used. The emergence of technologies bundled in the umbrella term of cloud computing has led to a much more centralized way of information processing. On top of that, storage capacity and processing power became relatively cheap and economic incentives to utilize existing data increased. This led to the emergence of big data and the phenomenon that data is increasingly used for secondary purposes, meaning purposes that are different from the reason data has been collected for in the first place.

This trend led to new business models and a twofold debate. It is by now common practice that an online service is offered to the user free of charge and is instead financed through revenues that are generated by the secondary use of the data, mostly advertisement and marketing purposes. Since this has implications on user privacy, a debate on privacy evolving around the question to which extent service providers may collect user data has been sparked. At the same time, the usage of user data is being discussed from an economic perspective. While the first debate mostly concerns the individual (micro-level), the economic debate is mostly on a macro-level, revolving around terms such as data as the *new oil* (Kroes, 2013) or data as *a new asset class* (World Economic Forum, 2011). Figure 1.1 illustrates the two different levels at which the current debate takes place.

Out of the many reactions to the ongoing developments, one approach being discussed is to assign property rights to personal data and thereby facilitate a market for it (e.g. Novotny & Spiekermann, 2013; Lanier, 2013). Also, first commercial activities in this regard can be observed (e.g. Soenens, 2013; Simonite, 2014). The appealing factor of this approach is that it would include the user into the economical activities surrounding their[1] data and would thereby connect

[1] Note: For the purpose of gender neutrality in my writing I make use of the *singular they*, meaning that I use 'they' and its inflected forms also as a pronoun for the singular.

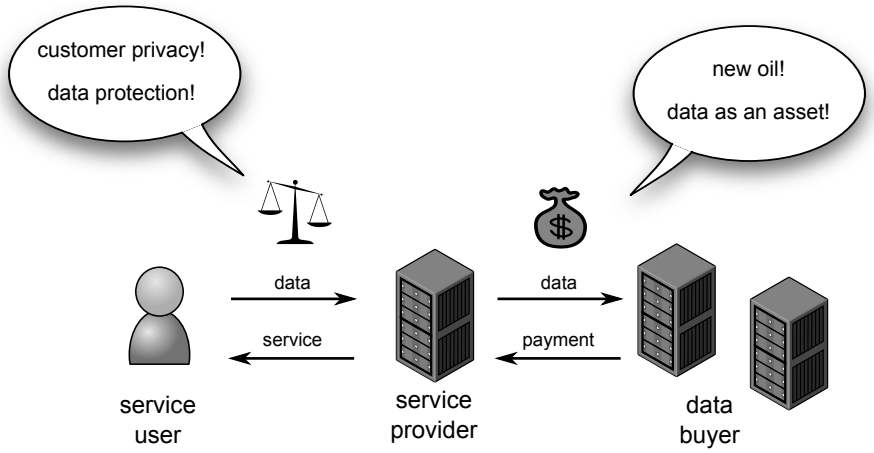

Figure 1.1: Illustration of the current situation and debate concerning the use of personal data.

the two levels of the aforementioned debate. Also, some proponents of this approach claim that it would constitute a form of privacy protection, since the user would be more closely involved in the use of their data and even would have a financial interest in it.

Whether such a data market would lead to a better form of privacy protection for online service users is not immediately apparent. This partially stems from the fact that privacy as such is not a clearly defined concept. Or as Solove puts it, "[p]rivacy is a concept in disarray [and n]obody can articulate what it means." (2006, p. 477) An analysis of approaches to privacy protection therefore has to be critical about this conceptual ambiguity of the notion of privacy as such.

The goal of this study is to research in which way data markets would impact the privacy of online service users in the European Union. I do so from the perspective of an ethical evaluation, taking different conceptions of privacy into account. To this end, I start this study with an investigation of the different meanings of privacy. I then review different approaches of data markets as proposed in the literature. Besides the ethical evaluation as the main focus of this study, these elements thereby form further contributions which could also stand on their own.

As I show in the following, data markets would affect the privacy of online service users in a number of ways. Depending on the specifics how they would be implemented, they could either contribute towards user privacy or also be detrimental in this regard. However, in order to achieve the former, they would have to be implemented in a way which incorporates many precautions and would ultimately be not too different from the existing data protection legislation in Europe. Besides the more direct effects, the transformation of data into a commodity would influence the social norms that surround it and would have to be accounted for. The precise outcome is furthermore strongly contingent upon the behaviour of users, which is a field in which much more research would be necessary in order to arrive at clear results. To design a data market in a way that improves the protection of privacy with regard to the status quo in Europe therefore is not an easy undertaking. Insofar it seems doubtful whether a paradigm shift towards a commodification of personal data in the form of data markets is worth the risk to ultimately be derogatory to user privacy.

Before going into the actual contents of this study, I want to provide some more background and introduce the focus of this study in more detail. In the remainder of this chapter, I therefore first provide an overview over the technological and societal changes that have led to the situation in which data markets are discussed as a solution to problems with privacy protection (Section 1.1). Afterwards, I describe the status quo with regard to data protection in the European Union (Sec-

tion 1.2). The shortcomings thereof have led to a bigger debate on the status of privacy protection, which I will outline in Section 1.3. Before coming to the discussed option of data markets, I briefly discuss the concept of property (Section 1.4). I then introduce the research question of this study in more detail (Section 1.5), before I finally outline the structure of this study (Section 1.6).

1.1 Big Data, secondary data use, and data trading

The current debate on the protection of privacy in the context of online services, is partially fuelled by developments in technology and how technology is put into use. Trends such as *cloud computing* or *big data* help to characterize these developments, but a look at the larger developments around it can also be helpful. Nissenbaum (2010, pp. 38) identified four pivotal transformations. The first is a democratization of information technology, meaning that IT became widely accessible and its use is not limited to a handful of organizations any longer. The second pivotal transformation is the high information mobility, referring to cheap storage and the possibilities that communication networks such as the Internet offer. Facilitated by the first two developments, the easy aggregation of information emerged as the third transformation. Finally, the trust in information as such and the belief to solve societal problems through the use of information constitutes a fourth transformation. Taken together, these transformations are able to characterize many of the more specific technology trends such as cloud computing and big data.

Cloud computing is one of the developments that had a huge influence on the privacy debate over the last years. It is not a very specific development in terms of the technological change, but rather a collection of different developments to outsource IT infrastructure (storage, computing and applications) and access it over the Internet. Cluley brings the essence of cloud computing to the point by suggesting that "every time we say cloud we say 'somebody else's computer'." (Cluley in Palmer, 2013) Already before the NSA revelations, experts warned that privacy and security questions are not considered enough when making use of cloud computing[2], but especially since then the topic has received great attention. Castroz (2013) estimated that the U.S. cloud industry might loose up to 20% of its share in foreign markets based on a loss in trust following the NSA revelations. This shows that privacy-related problems in the context of cloud computing are slowly being taken into account.

Big data is another technological trend that is connected to the privacy debate. It is not so much an advancement in technology itself, but rather a change in the way technology is put into use. It concerns the handling of large data sets, the use of statistical correlations between data fields and the incorporation of data that has not been necessarily collected for that purpose. Mayer-Schönberger and Cukier (2013), who wrote a bestseller book on the big data transformation, characterize big data through three fundamental shifts of mindset. The first one is about the change to process all available data instead of taking smaller samples. The second change is a loss of accuracy of the input data, which can be compensated for through the inclusion of larger data sets. The third one is the switch from causal reasoning to correlations as an explanatory way. These developments might at first sight seem unrelated to privacy, but are not. The strong reliance on data that has been collected for other purposes makes it impossible for individuals to oversee the consequences of data collected about them and makes informed consent impossible. The exclusion of causal reasoning furthermore means that consequences possibly affecting individuals are left to algorithms and are out of reach for humans to decide on. Big data thereby does not affect privacy in any new direct way, but leads to a change of character and quantity in the risks that the use of data brings with it.

[2] For example, Microsoft had to admit in 2011 that they have to provide U.S. authorities access to cloud data, even if it is stored on servers within the European Union (Whittaker, 2011).

A new phenomenon from a data-protection perspective is the *secondary use* of information. Whereas in the past most data was collected for a specific purpose, it becomes more common to utilize data for other purposes as well. The developments of cloud computing and big data are not necessarily the reason for it, but reflect the changes that have led to this development. Through the centralization of data storage and the enhanced analysis capabilities, it became possible to use data for secondary purposes, whereas before it simply would not have been efficient enough to do so. This leads to a higher availability of knowledge that has not been accessible before and can also be an improvement of efficiency if data does not have to be gathered multiple times. While one can argue that there is a welfare benefit in the secondary use of data (Jentzsch, 2010), the value created through it has to be balanced against the data protection rights that individuals have (Article 29 Working Party, 2013a).

An aspect that is not taken into account in many debates evolving around privacy, is that it is not only the explicitly provided data which is of relevance, but also data that has been automatically gathered as part of information systems. This means that, for example, privacy issues around Facebook do not only concern the information that users actively have put into their profiles, but also data that is generated through the interaction with the website. The monitoring of data itself is what Clarke calls 'dataveillance' (Nissenbaum, 2010, p. 23). It is thus not the direct surveillance of an object, but indirect surveillance through data that is readily available. This not only concerns specific platforms, but also the aggregation of different activities users perform on the web. Through advertisements placed on websites, companies such as Google are able to track the online behaviour of users across the web (Craig & Ludloff, 2011, p. 6). Also, there is an increasing range of devices that generate data about their users and thereby allow for new forms of dataveillance. Morozov (2013a) warns in this regard that it is the data stemming from 'smart devices', such as maybe toothbrushes or dustbins, that will allow more and more intelligence gathering in the future. In a classification between *volunteered*, *observed* and *inferred* data employed in a study by the World Economic Forum (2011), it is the former which discussions about privacy usually evolve around. However, it is also observed data which is used as input for big data analysis, which in turn generates the inferred data. Therefore, all three types of data haven to be taken into account when discussing the privacy implications of new information technologies.

Connected to the secondary use of information is the 'free service' business model. Many of the online services that people use are provided free of charge to them. In turn, they are financed through the use of personalized advertisement or the sale of customer profiles on secondary markets. In both cases, service providers employ information about their users in order to generate profits on a market that is different from the one the actual service is provided in. This form of business model constitutes a case of data usage for secondary purposes as the data is gathered for the provision of the actual service in the first place. However prevalent this form of business model may be, there is the critique that it is unlikely that users are fully aware of the actual practices they engage in (United Nations High Commissioner for Human Rights, 2014, p. 6). Without the awareness of users, however, it is not only a topic of relevance for privacy as such, but also constitutes a violation of data protection legislation (Article 29 Working Party, 2014b, p. 44)[3].

As indicated above, the economic value of data is by now widely recognised. This is not only reflected through the value that the use of data has for businesses, but can also be seen through the emergence of *secondary markets*[4] In such markets, firms sell information about users to other

[3] Noteworthy is that even on the legal ground of consent such practices are not legitimate if service providers hide these details in the fineprint without truly informing their users about the practices (Article 29 Working Party, 2014b).

[4] Such markets are often simply called *data markets* (Craig & Ludloff, 2011). The firms that sell information on these markets are often called *data brokers* (Committee On Commerce, Science, And Transportation, 2013) and sometimes also *omnibus information providers* (Nissenbaum, 2010). In order to distinguish these markets from markets in which the user is directly involved, I will refer to such markets as *secondary markets*. This terminology is also used in some other studies (see e.g. Acquisti, 2010; Cave et al., 2011).

firms who are interested in such data in order to study their target group. This can concern information about specific individuals, but also more general information about groups of customers. Interesting about this industry is the veil of secrecy under which it operates. As Nissenbaum (2010, p. 49) notes, many companies in this sector do not disclose the sources from which they gather their information. Also, it is common practice to extend this secrecy into the other direction, by in turn contractually limiting customers of these companies to disclose them as their sources (Committee On Commerce, Science, And Transportation, 2013). This leads to the situation that the existence of secondary markets as such and some of the key actors involved therein (see e.g. Nissenbaum, 2010, pp.45) are known, but that the actual business practices happening in this sector are concealed.

The possible implications of secondary data use and data markets in particular can be best illustrated using a few examples. LexisNexis, which is one of the bigger players in the secondary market in the US, sells their data among others to the U.S. government who uses the data to fight tax fraud (The Economist, 2013). Similarly, there is an increasing trend that also healthcare institutions make use of the data that they can commercially acquire from secondary markets in order to build better risk profiles for their patients (Singer, 2014). Although in both cases the goals might be legitimate ones, it is highly questionable whether users are aware of the fact that their data collected in a commercial context is put into use this way. In the Netherlands, it was revealed that GPS navigation vendor TomTom sold information about road usage to the Dutch police, who used the information to set speed traps (Arthur, 2011). While this concerned aggregated data and therefore did not affect the data of individual persons, another case of secondary data use in the Netherlands was different in that regard. A Dutch super market chain provided the shopping data gathered using its loyalty card program to the Dutch authorities, who in turn used it to prove social security fraud (Schmidt, 2010). What these examples have in common is that the data which has been gathered in a very specific context is put into use in a completely different context. The individuals affected through it are probably unaware of these practices and would not make the link between the collection of data in one context and use in another.

But even if information stays within one context, the secondary use of information within a specific context can lead to privacy issues. For example, it was reported that a travel website charged higher hotel prices to users of Apple computers (Mattioli, 2012). This is an example of price discrimination, where factors probably unknown to the individual concerned are taken into account. Another well known example from the commercial context is that of a teen girl receiving advertisements for baby products (Hill, 2012). Through an analysis of her shopping habits, a U.S. supermarket chain found out that she is probably pregnant and sent her coupons for baby products. Through this, her family learned about her pregnancy before she told them about it. Compared to the price discrimination based on the type of computer used, this is a more complex finding based on inferred instead of simply observed data. In both cases, however, data gathered in the commercial context was used in a way in which individuals are not likely to expect, thereby raising questions of privacy.

In general, the secondary use of data and the employment of big data analysis techniques is likely to raise privacy concerns. More recently, the NSA revelations helped to raise awareness in this regard[5]. Noteworthy is that Mayer-Schönberger, who mainly focussed on the benefits of Big Data in his book on big data (Mayer-Schönberger & Cukier, 2013), more recently expressed concerns about the predictive possibilities of big data (Persson, 2014). He says to have been surprised at which scale intelligence agencies are using big data analysis and urges for more modesty

[5] Although this study does not focus on the activities of intelligence agencies, one more note in this regard might be important to take into account. Through their wide-reaching competences and the secondary use of data that is gathered in other contexts, intelligence agencies can increasingly rely on data that private entities indirectly gather for them (Mantelero & Vaciago, 2013). Thereby, privacy issues which at first sight only play a role in the private sector also become a question of governmental surveillance.

in the use of big data techniques. The extent to which data collection and analysis are allowed to be employed, is within Europe regulated through data protection legislation. In the next section, I therefore go into a small parenthesis on the status quo of data protection within the European Union.

1.2 *Parenthesis:* data protection in the European Union

In order to understand the debate about privacy and the introduction of data markets, it is helpful to have a good understanding of the regulatory status quo. Within Europe, data protection as defined in the EU Data Protection Directive (European Parliament, 1995) is the most important element in this regard. Noteworthy about this directive it that it forms an omnibus approach for the protection of informational privacy, with one directive setting the standard for both the private and public sector for all EU member states. Although it ultimately is implemented in national legislation that slightly differs per member state, it nevertheless achieves its aim to harmonize regulation throughout the EU relatively well. Besides the Data Protection Directive there are also a number of other legal sources of relevance for the protection of privacy, but the Data Protection Directive stays the most important one on this subject.

The EU Data Protection Directive applies to all 'personal data' which is data that directly or indirectly identifies a 'data subject' (the individual concerned). Such data may not be collected and processed without limits, but only within the boundaries that the Data Protection Directive defines. Two articles are of particular importance in more practical terms. Article 6 defines the conditions under which data processing must take place. Regardless of the legal ground on which data processing takes place, these conditions have to be fulfilled. This means that even if a data subject would provide consent for a different way of processing, these provisions may not be circumvented. The legal grounds for data processing in turn are defined in article 7. They define the basis on which data processing may take place and if none of them is applicable in a specific situation, the processing of personal data is prohibited. Taken together, these two articles define if and under which conditions processing may take place.

Out of the general obligations defined in article 6, the *purpose limitation* principle is a very significant one. It defines that data processing has to take place for a beforehand specified purpose and that all further of use of the data has to be compatible with it. This entails that a secondary use of data is generally prohibited. This provision is of especial importance in the discussions around big data and additional criteria to assess what constitutes compatible use might be necessary. Among other things, article 6 furthermore specifies that data processing has to be kept to the minimum necessary in order to achieve the specified purpose and that the identification of data subjects may not be possible any longer than is necessary. This can be read as an obligation to anonymise data as early as possible. Important to keep in mind about these provisions is that they are always applicable and even the consent of the data subject cannot lift them.

The legal grounds for data processing are defined in article 7. Well-known is that consent of the data subject is one way to make the data processing legitimate. However, it is only one out of six different legal grounds, out of which two more are also more widely applicable. One of them is data processing which is necessary in order to perform a contract. For example, it is necessary to store the address of a data subject in order to send an invoice. The other more widely applicable legal ground is that of 'legitimate interest'. It requires a balancing test between the interests of the data subject and those of the data controller (Article 29 Working Party, 2014b), but allows to process data without the consent of the data subject. The ambiguity that this balancing of interests entails, however, makes it less popular than the collection of consent which data subjects are willing to give in many cases anyhow.

Further provisions of the Data Protection Directive entail that data has to be kept up to date,

gives the data subject the right to inspect the data stored about them and demand rectification in case of incorrect data. Also, there are procedural limitations to automated profiling and supervisory authorities have to be notified in some cases of data processing. Furthermore, special categories of personal data, such as the ethnic origin of a data subject, require extra precautions in terms of procedures. Noteworthy is also that the transfer of personal data to non-EU countries is only permitted if an adequate level of protection exists in the third country[6]. In total, the EU Data Protection Directive sets limits to the processing of personal data, but more importantly defines the procedures and protection mechanisms that have to accompany the processing activities. In this regard it is not so much a prohibition on data processing, but rather a procedural safeguard to give individuals a say in it.

The EU Data Protection Directive is currently being phased out. It will be replaced by a regulation on the EU level, which will have direct effect and thereby avoids the differences that occur through the diverging implementations in member states. Besides this change in legal basis, it will also update some of the provisions, but is likely to stay overall similar to the current directive. A good overview over the proposed changes and the state of discussion is provided by Kuner et al. (2014). One of the changes under discussion is to strengthen the requirements of consent, so that it is not only used as a legal ground, but also leads to a real consent of the data subject in practice. Furthermore, it is discussed to weaken the requirements for the processing of pseudonymous data, including weaker requirements for the profiling of such. Most important, however, are changes in the procedures accompanying data processing. Among others, it is planned to increase the cooperation between the supervisory authorities of EU member states and give citizens the possibility to have one point of contact in this regard (the so called 'one-stop shop' mechanism). Although the European Parliament already voted on the amendments that it made (European Parliament, 2013b), the regulation has been with the Council of the European Union for longer time and the adoption of the regulation is not expected before 2015 (Foster, 2014).

Next to the Data Protection Directive, a number of other legal provisions are of relevance on a European level. A good introduction into these and their historical development is given by Izyumenko (2011, pp. 19). Important is that next to the notion of data protection as defined by the aforementioned directive, also a right to privacy as such exists in legal terms. Both at an EU level through article 7 of the EU Charter of Fundamental Rights (European Union, 2010), but also through article 8 of the European Convention of Human Rights (Council of Europe, 2010), privacy has the status of a human right. But also in more practical terms, the Data Protection Directive is not the only legal source in this regard. The e-Privacy Directive (European Parliament, 2002), for example, defines provisions on unsolicited advertisements sent via e-mail and the use of cookies. Despite these accompanying provisions, the Data Protection Directive is the most significant legal provision when it comes to the protection of privacy and will therefore also serve as a frame of reference when discussing the status quo within this study[7].

In summary, the EU Data Protection Directive defines the boundaries and procedures of legitimate data processing. It is supposed to give users a say in the use of the data concerning them, but

[6] For data transfers to the United States, which are of especial importance because of cloud computing, there exists the so-called 'safe harbour' agreement. It entails that the U.S. are not required to have an adequate level of protection through binding legislation, but essentially gives firms the possibility to self-certify that they adhere to a number of principles (Busch, 2006). Especially after the revelations of spying activities of U.S. intelligence services, criticism about this form of agreement increased, thereby impeding negotiations about an extension of the agreement (Reding, 2014).

[7] I will, however somewhat deviate from the terminology used within the Data Protection Directive. I will mostly refer to the 'user' instead of the 'data subject' and talk about the 'service provider' where the directive speaks of 'data controller' and the 'data processor'. In this regard, the terminology used within this study is somewhat less precise, but closer to typical scenarios and therefore less abstract.

also establishes supervisory authorities that protect the interests of users. It is therefore not up to the user to protect themself, but a set of institutions is supposed to act on behalf of the user. In this regard, value conflicts related to privacy are dealt with through the mechanism of *firewalling* which Thacher and Rein (2004) distinguish from other ways to deal with value conflicts in public policy: a set of institutions and procedures specialized on the protection of privacy are set up, thereby opposing other interests that might play a role. However, it has to be noted that the Data Protection Directive and its institutions have been established in 1995 and it seems that the procedures laid down therein have difficulties to deal with the more recent data processing practices.

1.3 The ongoing debate about privacy protection for online service users

The technological developments and the increasing use of information technology in all spheres of life have led to a wide debate on the protection of privacy. The problems that are discussed are not necessarily new. Many of the points and issues that are a matter of debate these days have already been pointed to decades ago (see e.g. Miller, 1971). What is new, however, is the societal relevance of the matter. The NSA revelations have in some regard been a peak of this debate, but independent of these matters there has been a wide-ranging discussion about the adequacy of existing regulations within the academic sphere. Within this section, I want to give an overview over the debate focussing on a European perspective. I start with a summary of the problems that occur and then give a brief overview over the solutions that are discussed.

problems with the protection of privacy As already outlined above (see Section 1.1), there are a number of technological developments that have led to a debate about the privacy of online service users. Due to the breadth of the overall debate, I cannot summarize it in its complete scope at this point. Central for this study is the question in which way data markets would impact privacy in a European context. As data markets might constitute an alternative means to data protection, I will focus within this section on insufficiencies that are of relevance for the EU Data Protection Directive.

Purtova (2011) provides a detailed analysis of the current problems with data protection and how the EU Data Protection Directive addresses them. She concludes that the current data protection mechanisms are still adequate in their substance, but fails in the procedures that are applied (Purtova, 2011, pp. 182). This expresses itself in a failure to deal with the principles of transparency and accountability, which are of central importance for data protection. Due to the high number of data streams and the high number of actors, individuals nowadays are not able to easily track who processes their data in which way. Even if individuals formally have all the rights in this regard, it is difficult to enforce them in practice. But also the supervisory authorities struggle with the high demands that are put onto them and are often limited by the insufficient resources that are available to them. This means that the procedures that are put into place are incapable to deal with the changed situation.

One of the problems that becomes easily apparent in practice is the strong reliance on consent. As explained above, informed consent is one of the legal grounds upon which data processing may take place. But also for cookies, the e-privacy Directive (European Parliament, 2002) prescribes notifications that inform the user about the use cookies on a website. Similarly, the U.S. knows a culture of 'notice and consent' in which users are informed about the use of their data through online services. However, applying the approach of informed consent to the online environment is meaningful on a conceptual level, but fails in practice. Van Alsenoy, Kosta, and Dumortier (2013) warn in this regard between the huge gap between theory and practice. One of the bigger

obstacles are simply the huge requirements in terms of time to read through all the texts that one encounters when using online services. From this perspective, an informed consent approach creates burdensome transaction costs on the side of the user (Martin, 2013). McDonald and Cranor (2008) estimated that a user in the United States reading all of the privacy policies that they are confronted with would spend 20 hours each month if they would seriously read them. Besides the time requirement itself, there is the problem that these texts are difficult to read and often take a legalistic form that is not easy to understand for the user. Nissenbaum (2011, p. 36) in this regard speak of the 'transparency paradox', meaning that the effort to increase transparency leads to an actual decrease of transparency because too much information is provided. The characterization as a paradox also shows that there is no easy solution to this problem if following the route of informed consent.

Another problem that shows up in practice is the emergence of secondary use. It is not a completely new phenomenon, but as outlined above, increased in relevance at a massive scale. The principle of purpose limitation as one of the general obligations for data processing expresses that a secondary use of personal data is in general prohibited. Therefore, purpose limitation and secondary use seem to be mutually exclusive. The Article 29 Working Party (2013a) recently published an opinion on the principle of purpose limitation in which they express that the secondary use of data is possible within strict limits. Important is that the purpose is clearly specified beforehand and compatible with the original reason for data collection. This indeed excludes the use of data for purposes that are completely different and unknown at the time of data collection, but allows to build in some flexibility for the use of data, if it relates to the service provided and is specified beforehand. In any case, it is apparent that the strong demand for secondary use of data cannot be followed upon if the principle of purpose limitation is uphold.

A further obstacle to be mentioned here is the distinction between anonymous and personal data. As the Data Protection Directive is limited to personal data, properly anonymised data are exempted from regulation. This makes it an important distinction on a conceptual level. In practice, however, it is difficult to determine when data are truly anonymous. A well-known example in this regard is a competition by the online video rental service Netflix, for which it released an apparently anonymised dataset of usage data. Through a combination with user profiles from a public movie recommendation site, Narayanan and Shmatikov (2006) could successfully deanonymise the data provided by Netflix. In more general terms, a study by Koot (2012) shows that a small number of data fields is often sufficient to uniquely identify an individual. As outlined in an opinion by the Article 29 Working Party (2014a), it is generally possible to properly anonymise data by using robust mathematical methods. However, applying such methods and still keeping usable data is often not possible and technically complex. Avoiding the applicability of data protection regulation by anonymising data is therefore not as easy as it is sometimes believed. Besides these problems with anonymisation as such, there is also critique about the exemption of anonymised data. Vedder (1999) proposes to extend the scope of data protection in this regard by also including data that can have the same adverse consequences as the use of personal data can have. The strong distinction between anonymous and personal data therefore at first sight seems to be less relevant than the current scope of the EU Data Protection Directive gives to it.

A last issue to bring up at this point is not a specific problem with data protection as such, but concerns the debate about privacy problems. As I will discuss extensively in Chapter 2, the notion of privacy is very ambiguous and can thereby lead to confusion in debates about it. The often inherently used dichotomy between public and private is one of the problems in this regard (Nissenbaum, 2010, pp. 89). To that end, Selinger and Hartzog (2014b) propose to put the concept of *obscurity* more strongly into the focus, because it allows for more nuances how information is accessible, thereby leading to stronger explanatory power. Another problem that often arises is the presentation of privacy problems in terms of a zero-sum game where, for example, privacy interests

have to be balanced against public security (Hoepman & van Lieshout, 2012). As, I discuss further below, the field of privacy-enhancing technologies is an example in case showing that this is not necessarily true. Finally, the discussion about privacy problems often shows characteristics of a 'wicked problem' (Rittel & Webber, 1973). Different actors involved often define the problem differently. While, for example, industry actors might perceive privacy as a barrier for market-adoption, individuals will perceive a privacy-problem as an infringement of their rights. Also, in many cases the solutions discussed serve as a problem-definition. Moreover, privacy problems don't have a definitive solution, which is shown by the always ongoing debate about them. Taking these aspects together, it is not surprising that there is so much debate about problems concerning privacy.

possible solutions to privacy problems There is not only discussion about problems, but also specific solutions are being discussed. Although this study focusses on data markets as one form of solution, I will at points refer to other solutions as well. Therefore, it is helpful to have an overview over these. I have to limit the discussion at this point to a selection of the broad range of solutions discussed in the literature, but I try to indicate the most important directions at least.

As individuals in Europe have many rights that they could exert, it is often assumed that individuals just do not know enough about the actual practices and have to be better informed. In order to raise awareness, there are a number of approaches to take. One is to tackle the problem of unreadable privacy notices. Examples in this regard are the use of layered privacy-notices that start with an overview at a high level and only show details upon request (van den Berg & van der Hof, 2012). The use of more 'visceral' forms of notices, such as the use of anecdotes, is another avenue that can be explored in this regard (Calo, 2012). To nudge the user in specific situations instead of explaining everything upfront in privacy notices is a different direction to take (Acquisti, 2009). An example of this is to display a random sample of the prospective audience of a Facebook-post at the moment the user would submit their post (Wang et al., 2013). More generally, the goal of such approaches it to provide a stimulus about the possible adverse consequences where usually only the benefits of an action are directly visible (Pötzsch, 2009).

Another solution tackling the over-reliance on consent is to automate the process of giving consent. In the case of tracking-cookies, for example, the e-Privacy Directive (European Parliament, 2002) requires websites to inform their users upfront about the use of this technology. As this concerns the same type of situation again and again, a solution to this problem might be to configure the desired behaviour in the web-browser and transmit this information to websites. Interestingly, such a proposal was part of the discussion for the upcoming Data Protection Regulation (European Parliament, 2013a, amendment 105), but did not make it into the final vote of the European Parliament (European Parliament, 2013b). On a more theoretical level, there also exist proposals in the literature which discuss the use of automated consent through software agents for more diverse use-cases (Le M'etayer, 2009). However, there exist enough technical proposals in this regard, but it is up to industry adopt them (Spiekermann, 2014).

A more general strand of solutions that attack privacy problems from a technical perspective are *privacy-enhancing technologies* (PETs). Definitions of PETs vary as much the term privacy itself, but one can say that the common goal of such technologies is "to shield the user from some aspects of on-line surveillance" (Danezis & Gürses, 2010). In more general form, PETs are about the application of data security technologies for the sake of privacy protection (London Economics, 2010, p. ix). For example, anonymisation techniques can be employed to shield the identity of users, or to simply minimize the amount of user data that is transferred across a network. While PETs are without any doubt useful for the protection of privacy, they increase the technical complexity and mean extra development costs for industry. To businesses it is often unclear why they should take these extra efforts, especially when they limit their possibilities for the

economic exploitation of data (Cave et al., 2011, p. 59). From an industry-perspective, the deployment of PETs therefore create a trade-off between the competitive advantage that they might gain by employing them and the reduction in benefits from the use of data (London Economics, 2010, p. x). Nevertheless, such technologies are continously developed and especially pushed through academia. A very recent example in this regard is the 'openPDS' system for the storage of metadata (de Montjoye, Shmueli, Wang, & Pentland, 2014). With this system, it is technically possible to store metadata about the use of information technology on the user's device and only provide the minimal necessary information to service providers who want to offer personalized services. As with all PETs, however, it remains to be seen to which extent market pressures or regulation are strong enough in order to lead to a deployment in practice.

A solution that is very different from the existing data protection legislation is to focus more strongly on the risks of data use. This development seems to be driven by the practice of secondary data use and its incompatibility with purpose limitation. According to proponents of such approaches, one should not regulate the collection of data as such, but rather regulate in which ways data may be used. Examples where such arguments can be found include a recent governmental report to the Dutch parliament (Minister en Staatssecretaris van Veiligheid en Justitie & Minister van Binnenlandse Zaken en Koninkrijksrelaties, 2013, p. 4) and a report to the U.S. president (President's Council of Advisors on Science and Technology, 2014, p. 49). Although regulatory approaches focussing on the actual use of data might be better able to deal with the problematic of secondary use, they move away from the concept of a precautionary principle and provide less transparency to the user how their data are used.

Another approach focussing on the risks of data use are *privacy impact assessments* (PIAs). Similar to impact assessments that are common in other fields (e.g. environmental impact assessments), PIAs aim to assess the possible impacts of information technology on privacy during the development phase. As such, they focus on the risks in specific use cases, but can be applied complementary to existing data protection legislation. A number of approaches in this regard have been developed, but a clear standard in this regard is lacking (Oetzel & Spiekermann, 2013). Also, there is no general requirement to conduct such an assessment yet. Furthermore, there is the critique that such approaches focus too much on the risks of data use, while neglecting the benefits that users might have from the use of their data (Polonetsky & Tene, 2013). However, it is usually the benefits of data disclosure that are in the focus, while privacy-related aspects are easily ignored (Pötzsch, 2009). Insofar, mandatory PIAs could provide a reasonable addition to existing regulation.

The improvement of oversight could be a further mechanism to deal with some of the current problems with data protection. As reported by Purtova (2011, pp. 183), the national supervisory authorities currently lack resources to deal with the tasks they have been ascribed. The provision with sufficient funds could therefore be a first step to enable the authorities to responsibly deal with their duties. Further improvements in oversight that are proposed in the upcoming data protection regulation (European Parliament, 2013b) are higher fines that can be imposed and a better coordination of supervisory authorities within the EU. But also outside of the EU data protection sphere solutions dealing with supervision are discussed. Mayer-Schönberger and Cukier (2013) propose to introduce audits on the use of big data technology. The focus hereby would, however, not be on the collection of data, but the risks that stem from an improper interpretation of results.

As partially outlined above, the upcoming Data Protection Regulation tackles a number of issues of the current directive. For a full overview over the proposed changes I refer to an article by Kuner et al. (2014), but I want to indicate a few of them at this point. One of the changes is the introduction of an icon-based representation of privacy policies, thereby addressing the problem of informed consent. Another change is the aforementioned one-stop shop principle that aims at a better cooperation of the supervisory authorities throughout the EU. Also, the appointment of

data protection officers within companies shall become mandatory in case of activities affecting a higher number of individuals. Overall, there are many changes under discussion, but the slow progress in negotiations with the Council (Foster, 2014) shows the difficulties that the political decisions around data protection have.

Finally, there is the proposal to involve the individual more strongly through the use of data markets. Part of the argument is that individuals who have property rights in their data feel more strongly connected to it and better defend their rights (Lessig, 2006). Proposals in this regard mostly stem from the United States (see e.g. Laudon, 1996; Lanier, 2013), but are also discussed within a European context (see e.g. Soenens, 2013; Novotny & Spiekermann, 2013). As the ways in which data markets would impact on user privacy is the central topic of this study, I do not describe them any further at this point, but instead refer to the discussion ahead[8].

Overall, there is not one specific problem with informational privacy, but a set of different problems. The ambiguity in the problems discussed and also in the notion of privacy as such thereby hamper the discussion of these problems. Accordingly, there also is not *a solution* to privacy problems, but different solutions trying to tackle different problems are under discussion. Proposals for data markets that shall involve the individual more strongly appear to be very different from many of the more partial fixes to data protection. It therefore is of interest to investigate in which ways data markets would impact the privacy of online service users in the European Union.

1.4 *Parenthesis:* the concept of property

An aspect strongly connected to the proposals for data markets is the introduction of property rights for data. However, it is important to take into account that property as a legal concept is different from property in the more colloquial sense and that there are cultural differences to the concept of property. In order to better distinguish the different aspects connected to property throughout the remainder of this study, I give a short introduction to the concept of property at this point.

The most important insight about the concept of property is that it concerns a set of different rights. Purtova (2011, p. 67) explains that the French law distinguishes between three rights in this regard: *usus* as the right to use, *fructus* as the right to use the yields, and *abusus* as the right to dispose of something. Honoré (1961) distinguishes even six different rights and additional characteristics that are connected to ownership. Although the different rights usually come together, it is important to distinguish between these different conceptual aspects when discussing issues of property.

Confusing in this regard can also be the cultural differences that arise due to a discrepancy between the *common law* and *civil law* (Purtova, 2011, pp. 64). The civil law, which has been strongly influenced through the French revolution, usually treats all the different rights as a bundle that comes together. It therefore represents a conception of ownership as a unitary right, which cannot be divided. In common law, on the other hand, property is a more flexible concept. Different forms of property exist, which makes property a fragmented concept. Ownership therefore does not automatically entitle to all of the different rights in this regard. This difference between the civil law and the common law represents a cultural difference which has to be taken into account in any discussion about property rights in Europe and the United States.

Another distinction that is of importance in the case of data markets is that of the special category of *intellectual property*. As Resnik explains, intellectual properties "[...] require special legal protection because they are non-exclusive: two people can possess and use the same item of

[8] For an overview over data markets discussed in the literature see Chapter 4.

intellectual property without preventing each other from possessing or using it." (2003, p. 320) Since data has these characteristics as well, data markets would also have to deal with the special properties of intellectual property.

Connected to the concept of property is the notion of a *commodity*. A commodity is characterized by three distinct features: it is alienable, commensurable, and fungible (Resnik, 2003). An object is alienable if it can be transferred, is commensurable if one can easily compare it with similar objects and fungible if one can replace it without loss. In the case of personal data, all three features are questionable. The alienability, for example, might be limited due to the non-exclusive character of data, the commensurability might be difficult because the data always concern a specific individual and the fungibility might not be given because of the intrinsic value that data might have for the individual concerned. It is therefore questionable whether markets for personal data can function in the same way as markets for other commodities do.

The question of alienability is of further concern for data markets since a full alienability of personal data effectively means that an individual waives the rights that they have to their data. If one establishes the link between data markets as a means to protect privacy and individuals can completely alienate the rights in their data, this would mean that it is possible to waive the right to privacy. It can therefore be reasonable not to grant the full set of property rights for personal data, but limit the rights of individuals in this regard. Furthermore, as Purtova (2013, p. 83) explains, property rights can not only have a market function, but also be of a protective form if they are instantiated in the right way. In every discussion that brings up property rights as a means to privacy protection, it therefore has to be differentiated in which form and to which end property rights are being used.

1.5 Research question and approach

As outlined above, there are a number of problems that the current European approach to data protection cannot deal with in an adequate manner. New practices such as cloud computing, big data techniques and the free service business model have changed the landscape of information practices. One of the responses to this change is the call for data markets in which the individual can engage in. Whether such an approach could deal with the contemporary issues of privacy protection more adequately is not immediately apparent. It therefore is of interest to investigate in which ways data markets have an impact on privacy. Formulated more precisely, the central question of this study is as follows:

> In which ways would data markets impact the privacy of online service users in the European Union?

embedding in 'the problem with privacy protection' The question that this study deals with is therefore different from the issue of privacy protection at a larger scale. As discussed in Section 1.3, there are a number of issues with the current approach of data protection. The questions involved therein show the characteristics of a wicked problem and different stakeholders have different perceptions of the problem. Also there is a broad range of solutions discussed, out of which each tackles a subrange of the issues discussed. However, it is not the goal of this study to deal with the larger question, which one might call 'the problem with privacy protection'.

Nevertheless, it is important to keep in mind how the central question of this study is embedded in this larger debate. The approach of data markets is discussed as one of the possible solutions to the 'problem with privacy protection'. Insofar, this study forms a contribution to the larger debate, but does not attempt to find an immediate solution to the larger problem.

research approach The overall approach taken to answer the answer the central question of this study is that of an ethical evaluation (Van de Poel & Royakkers, 2011, pp. 145), focussing on values related to privacy. In order to conduct the evaluation, two further steps have to be taken beforehand. First, the ambiguity about different conceptions of privacy requires to arrive at more clarity about the notion of privacy. Second, it has to be established in more detail what specifically constitutes a data market.

Therefore, I first conduct a literature review on different conceptions of privacy. I try to provide a synthesis of the different meanings of privacy and arrive at a conclusion in which way the impact on privacy can be evaluated. Based on this, I develop an evaluation framework that is suitable to answer the central question of this study.

Regarding the specifics of data markets, I provide an overview over different data market approaches discussed in the literature. Out of these, I select a handful of approaches that are specific enough in order to be scrutinized in terms of an ethical evaluation and also fulfil the criterion that they involve the user in a direct way. I decided to base the evaluation on abstract proposals discussed in the literature, because real instances of data markets are just about to emerge. Otherwise, an evaluation on a real case might be more beneficial, since more details could be taken into account then.

The ethical evaluation itself is based on the evaluation framework which I develop in Chapter 3. It consists of a structural analysis based on the contextual integrity framework (Nissenbaum, 2010), and an ethical evaluation including values that are of relevance for markets and privacy.

I use an ethical evaluation as the overall method for this study, because it is best-suited to evaluate the impacts of data markets in a general form. To assess the impacts of data markets in the form a full impact assessment is not possible, since too many parameters would be unspecified. This concerns the implementation details of the data market approach as such, the context it is embedded in and, foremost, the behaviour of users in such markets. It therefore is better fitting to assess the more general implications that would be possible in the form of an ethical evaluation.

perspective of this study Hence, the perspective of this study is an ethical one. It is descriptive and evaluative in character, but does not attempt to be of normative form. In order to support the arguments made in the evaluative part, research results from the fields of economics and behavioural economics will be incorporated. The study furthermore incorporates knowledge from the field of law in order to refer to existing and planned legislation that is of relevance. This gives the overall study an interdisciplinary character.

Geographically, the study focusses on the situation within the European Union. However, I also include a section on privacy in the United State (see Section 2.5). This is in order to set policy proposals stemming from there into perspective and to account for cultural differences with regard to privacy.

Through its contribution of the ethical evaluation, this study is not only of relevance for the field of ethics, but also for the field of policy analysis. Within the activities that make up policy analysis defined by Mayer, van Daalen, and Bots (2004), this study falls into the category of *clarification of values and arguments*. Through its evaluative character, this study provides clarification about the values that data markets would affect and can be used to analyse arguments that stakeholders make in this regard. Furthermore, this study can be characterised as an ex-ante policy evaluation (Dunn, 2008), which results from the inclusion of specific proposals for data markets discussed in the literature.

societal relevance of the topic The ongoing debate about user privacy in the context of online services and the current efforts of the European Union to update its data protection legislation (European Commission, 2012) show the need to bring data protection in line with technologi-

cal developments. This need is also reflected by the high number of reports that governments and international bodies have published throughout the last years (see e.g. U.S. Department of Commerce, 2010; World Economic Forum, 2011; Federal Trade Commission, 2012; Committee On Commerce, Science, And Transportation, 2013; Minister en Staatssecretaris van Veiligheid en Justitie & Minister van Binnenlandse Zaken en Koninkrijksrelaties, 2013; President's Council of Advisors on Science and Technology, 2014; United Nations High Commissioner for Human Rights, 2014). However, there is no consensus on the best approach to bring privacy protection in line with technological development. This shows that there is generally need for further research in this regard.

Within the field of solutions that are under discussion, the approaches of data markets stick out, since they are in multiple ways fundamentally different from the existing data protection legislation. At first sight appealing about them is the strong focus of the role on the user and the closure of the gap between the two sides of the debate portrayed in Figure 1.1. However, the divergence from existing approaches asks for extra caution on the possible effects that data markets would have. While it is relatively clear how a differently structured privacy notice will affect the behaviour of a user, the introduction of a data market will likely have more diverse implications. Also the possible shift from privacy as a human right towards a commodification of data asks for precaution. In order to assess different policy options involving data markets, it therefore is important to have a clear picture of the ways in which data markets would impact the privacy of online service users.

With the results of this study, policy analysts who want to analyse different policy options for the protection of privacy gain insights into the different effects how data markets influence privacy. This is necessary as a preliminary step in order to conduct further assessments on the possible implications and suitability of policy options. Also, researchers who work on specific data market solutions obtain valuable insights into the different pitfalls that have to be taken into account when designing data markets that account for privacy. Moreover, the results of this study are useful in order to analyse the validity of claims regarding data markets and their effects on privacy in more general terms. As such, the contributions made by this study can be of use in a number of fields and help to better protect the privacy of online service users[9].

terminology used throughout this study In this document, I use a number of terms which are sometimes ambiguous in their meaning. With the following list I intend to clarify in which way I make use of these terms:

- **market**: a market system in the abstract sense; not necessarily a specific market occurring in practice
- **data market**: a market system in which a service user is being paid for the data that is generated through their participation in online services; unless specified explicitly, it does not refer to secondary markets
- **secondary market**: a market system in which data concerning users is traded, but to which only service providers and not the user themself have access to
- **data/information**: used somewhat interchangeably, with a preference for data[10]; the developments around Big Data show that the traditional distinction between the two is less meaningful than it used to be
- **user**: the user of an online service offered by a service provider
- **service provider**: organisation offering an online service to users
- **privacy**: mostly referring to *informational privacy*

[9] For a discussion on the societal relevance of the protection of privacy itself, please see Section 2.3.
[10] Also the EU Data Protection Directive speaks of personal data and not information.

- **data protection**: regulatory mechanism to protect privacy; in its substance only partially overlapping with privacy (see Section 2.4)

1.6 Structure of this study

As indicated above, the remainder of this study consists of multiple steps. In Chapter 2, I provide an overview over the different conceptions of privacy. I do so by means of a literature review, focussing on definitions, methods and reasons for the protection of privacy. I conclude that privacy is a cluster of related concepts and that it is ultimately the moral reasons to protect privacy which are of importance. Based on these insights, I assemble the evaluation framework that I use for this study in Chapter 3.

Chapter 4 provides an overview over different data market approaches and thereby introduces the object of this study in more detail. Within this chapter, I describe the origins of data markets, refer to existing practices and discuss data market approaches that can be found in the literature. Out of the different approaches found in the literature, I make a selection of a handful of them and describe them in more detail. Together with a characterisation of the status quo of 'free services' in terms of a data market, this selection serves as the object of study for the actual evaluation.

The evaluation itself, and thereby the main objective of this study, can be found in Chapter 5. I provide a structural analysis of the different data market approaches first, thereby pointing to the different elements that are of relevance for the impact on privacy. Afterwards, I perform an ethical evaluation of data markets by focussing on values that are of relevance for markets and privacy. I perform this evaluation in a more general form first and only later-on link the conclusions back to the selected data market approaches.

I finish this study by summarizing the findings of this study in Chapter 6. In this chapter, I furthermore give indications on the contributions of this study and possible directions for future work. Going somewhat beyond the main question of this study, I also provide pointers to the prerequisites under which data markets could be effective for the protection of privacy and compare their usefulness against the status quo of data protection. I close the chapter with policy recommendations in this regard.

2 Conceptions of privacy

As shown in the previous chapter, there is a lot of debate about the ways how privacy can and should be protected. In order to better understand statements which claim that a certain mechanism affects privacy, it therefore is imperative to have a good understanding of the concept of privacy itself. However, as Solove points out, privacy is a concept in disarray, whose meaning cannot be easily articulated (2008, p. 1). In this chapter I introduce different conceptions of privacy in order to show the diversity of meanings that privacy can have. Through a discussion of the most important features, I lay the foundation for the evaluation framework that I develop in Chapter 3.

The goal of this chapter is not to provide a complete overview over the different meanings of privacy in the literature, nor is it to provide a fully sound taxonomy of different meanings of privacy. Others, as Solove (2008) or Smith, Dinev, and Xu (2011), have made corresponding attempts and I refer to the existing literature in this regard. Rather, my goal here is to highlight the different angles from which privacy is being discussed and to show the diversity in focal points being taken[1]. The classifications that I make thereby serve as guidance through the discussion.

In order to view the diversity of the concept of privacy, I discuss its meaning in the literature from three different perspectives. First, I discuss different definitions of the concept of privacy as such (Section 2.1). But even without precisely defining it, ways to protect privacy are often being discussed. As a second perspective, I therefore discuss different ways to account for privacy (Section 2.2). Furthermore, the moral reasons for the protection of privacy (Section 2.3) are of importance as they show why it is that privacy ought to be protected.

For the context of this study, there are two additional elements concerning privacy that deserve consideration. Since data protection is the predominant approach for the protection of privacy within Europe, it is helpful to consider the relation between privacy and data protection (Section 2.4). The fact that much of the discussion concerning data markets stems from the United States makes it necessary to understand the differences between privacy protection in Europe and the United States (Section 2.5). Finally, I close this chapter with conclusions about the concept of privacy that are of importance for the remainder of this study (Section 2.6).

2.1 Definitions of privacy

Definitions of privacy do not stand on their own, but are often closely linked to particular ideas how privacy can be accounted for and why it ought to be protected. For clarity and to better show the diversity in meanings, I tried to separate these aspects from each other and instead discuss them one after the other. In this section I start with different definitions of privacy, thereby giving a first overview over the different meanings it can have. For a better structure of this overview I distinguish three different categories: privacy as a right, privacy as a state and privacy as property.

privacy as a right The probably oldest and best-known definition of privacy belongs to this first category and is the *right to be let alone* by Warren and Brandeis (1890). Driven by the technological developments of their time (instantaneous photographs) and the societal consequences stemming from it, they argued for this right to be recognized in legal terms. In their article they

right to be let alone

[1] NB: The views that I describe within the section are those found in the literature and I do not necessarily agree with all of them.

discuss how such a right should be legally grounded and whether copyright law is applicable in this case. Their definition is therefore one of the earliest attempts to translate the need for privacy into legal terms. However, it is left open what privacy itself is considered to be and what it specifically entails to be let alone. An interesting detail about this definition is that it is almost a prototype to express a right to privacy as a negative freedom, an observation also made by Hildebrandt and Koops (2010, p. 436).

privacy as control

Another well-known definition of privacy is "the claim of individuals, groups or institutions to determine for themselves when, how, and to what extent information about them is communicated to others." (Westin, 1967, p. 7) Similar to Warren and Brandeis, the trigger for the development of this definition was technological change, this time being the uprise of computers and relatively large-scale data collection. In substance it is much more explicit than the right to be forgotten and can be said to be a defintion as a positive freedom. Worth mentioning is that Westin explicitly had the relevance of privacy for society in mind and did not only consider it to be of value for the individual (see Miller, 1971, p. 49).

informational self-determination

The right to *informational self-determination* that the German constitutional court established in 1984 (Hornung & Schnabel, 2009) is in substance similar to Westin's definition of privacy. Also the reasoning of the constitutional court is based on the societal importance of privacy that Westin already mentioned. This right is not only of relevance for the German legal system, but became the foundation of EU data protection legislation as well. Also the OECD guidelines on the protection of privacy (OECD, 1980) can be said to be along the lines of Westin's definition, which gives it a world-wide relevance.

human right

Apart from the substance of definitions, there also is the legal status which makes privacy a right. In Europe, article 8 of the European Convention on Human Rights gives the "right to respect for his private and family life, his home and his correspondence." (Council of Europe, 2010) This makes privacy a human right within Europe, which also entails that privacy cannot be waived by individuals (Izyumenko, 2011). Although this formulation is in its substance not very helpful to define privacy, it is important to consider the legal status that is thereby ascribed to privacy.

privacy as a state Another way to define privacy is to focus on the question whether someone has privacy, thereby considering privacy to be a state. In this way, the definition of Westin (1967) discussed above can not only be seen as a right, but also as a state. This means that one has privacy if one is able to determine for oneself when, how, and to what extent information about one is communicated to others. In this way privacy is the outcome of the applicability of a given method. This is why methods to account for privacy are often closely connected to the definition of Westin. I discuss this in more detail in Section 2.2 below.

privacy as control

An account of privacy which is defined in this way is given by Roessler: "Something counts as private if one can oneself control the access to this 'something'." (2005, p. 8). Based on this definition, she distinguishes between three different dimensions of privacy. *Decisional privacy* concerning the interference in one's decisions, *informational privacy* meaning the protection of personal data, and *local privacy* as the traditional spatial dimension. Accordingly, she speaks of a violation of privacy if one of these three dimensions is being interfered with in an illicit manner.

contextual integrity

An extensive definition of privacy that also takes the form of a state is provided by Nissenbaum through the framework of *contextual integrity* (2010). The central insight of her work is that social contexts differ from each other and that one has to assess questions of privacy with regard to these contexts. Flows of personal information can occur within or in between contexts and context-relative informational norms prescribe how personal information should flow. Nissenbaum proposes her framework of contextual integrity as a benchmark for privacy and defines it as follows: "Contextual integrity is defined in terms of informational norms: it is preserved

when informational norms are respected and violated when informational norms are breached." (Nissenbaum, 2010, p. 140).

Besides conceptually unpacking the concepts of a context and context-relative informational norms, Nissenbaum also provides a decision heuristic that can be used to evaluate whether an information system violates contextual integrity, and if so, whether that violation is justified. (2010, pp. 181). She thereby accounts for the fact that informational norms are not set in stone and might change for good reasons. In combination, these tools provided through the framework give, on a conceptual level, clear advice how to assess questions of informational privacy. However, an important prerequisite is that a very specific situation to be assessed is given, because otherwise the concept of contexts is not applicable.

degree of access

An interesting definition of privacy in the form of a status is provided by Gavison (1980). A good summary of her definiton is provided by Nissenbaum: "Gavison defines privacy as a measure of the access others have to you through information, attention, and physical proximity." (Nissenbaum, 2010, p. 68) Interesting about Gavison's definition is that it does it does not describe privacy as a binary state, but as a degree of access. With such a definition it thus becomes possible to have more or less privacy, whereas most conceptions of privacy are about the question whether or not one has privacy. Another aspect worth highlighting is the inclusion of attention as one of the dimensions, which is an aspect not considered in many definitions.

epistemic account

Another definition of privacy which considers privacy to be a degree concept is the *epistemic account of privacy* provided by Blaauw (2013). In this view, privacy is an epistemic relation between a subject that has privacy, a set of true propositions on the subject and a set of individuals with respect to whom the subject has privacy. The degree concept comes in through the amount of propositions, the number of individuals and the certainty of the epistemic relation (i.e. the strength of the belief).

privacy as property A very different view of privacy can be found in the field of economics, namely to consider privacy as a commodity that can be sold through economic transactions. This view seems to have two different origins. First, the role of property rights as a means to protect privacy, and, second, considerations about the desire for privacy in the context of economic transactions.

property rights

The source of the first reason, namely that of property rights, can be found within the legal system of the United States. Warren and Brandeis discuss in their article on the right to be let alone which legal basis this right should have and find that it already has been acknowledged in previous jurisprudence. They acknowledge that this right is different from existing forms of property protection, such as physical or intellectual property, but nevertheless arrive at the conclusion that it has the characteristics of property (Warren & Brandeis, 1890, p. 205). As reported by Levine, these considerations have highly influenced the thinking about privacy in the Western world (1980, p. 8).

intellectual property

Particularly interesting is the relation of privacy to intellectual property. Warren and Brandeis consider the protection of intellectual property to be a special case of the right to be let alone (1890, p. 205). Others, have investigated whether intellectual property rights could be a means to promote privacy. In an article on the accounts that exist for the existence of intellectual property, Resnik (2003) argues that there are reasons to utilize intellectual property rights as means for privacy protection, but concludes that this justification is only of limited application since many protected works are considered public. It therefore seems that there is an overlap in the concepts, but that neither can be used to fully account for the other.

privacy as a commodity

An example of the second type of consideration can be found in article by Rust, Kannan, and Peng (2002). They initially define privacy itself in terms of degree of access (see above), but only to then describe the increasing data collection practices by businesses and the conflicts

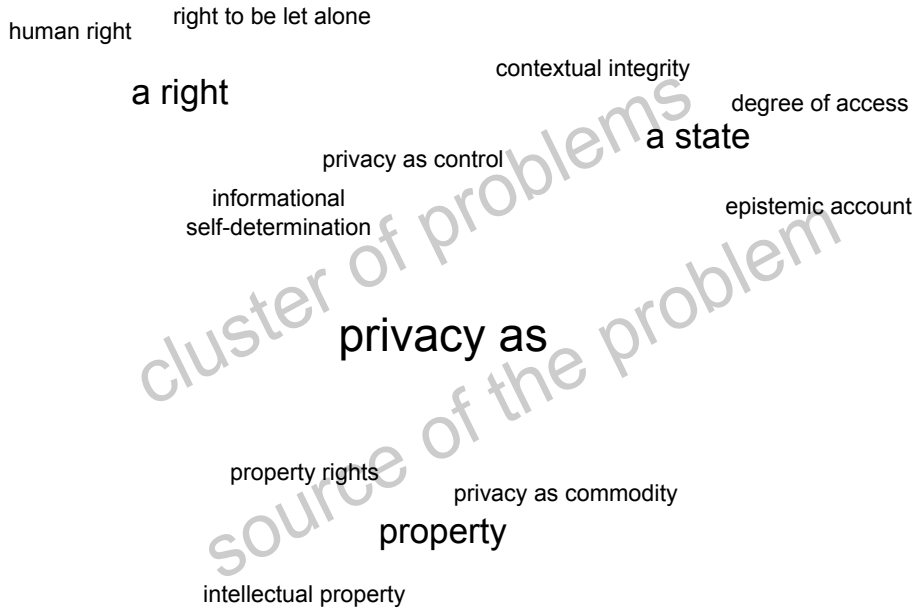

Figure 2.1: Overview over the different definitions of privacy.

which privacy creates. Based on these developments and the desire for privacy that customers have, they derive a model how businesses could sell 'units of privacy' (Rust et al., 2002, p. 458) to customers. Thereby, they effectively consider privacy to be an asset that businesses can sell in order to get compensated for data they might otherwise collect.

other views on privacy A last definition of privacy to be mentioned here does not fit into any of the three categories described above. Solove (2008) argues that there is no single definition of privacy that can account for all the different meanings that it entails. He sees privacy as a "set of protections against a plurality of distinct but related problems" (Solove, 2008, p. 171). From his point of view, the term privacy is useful as a shorthand to refer to this cluster of problems, but trying to find the common denominator of privacy will only lead to confusion. To give greater clarity to this conception, he provides an extensive taxonomy of privacy problems that are common in activities concerning data processing.

cluster of problems

Van Den Hoven (2008) makes the point that the conceptual ambiguity of the term privacy is in itself a source of problems when it comes to practical solutions that are to be found. He argues that we can do without a conceptual definition of the term and should instead focus on the problems that we want to solve. In this regard, the moral reasons to protect personal data are more important than a conceptual analysis of the term privacy itself.

source of the problem

As shown in Figure 2.1 privacy can be defined in very different ways. Often definitions share some characteristics, but still provide another angle to view privacy from. The three dimensions of privacy as a right, privacy as a state and privacy as property that I used to guide the discussion, therefore only serve as one possible approximation to capture common elements. In the next section I show another way how privacy is more implicitly defined, namely through different methods to account for privacy.

2.2 Methods to account for privacy

The previous section has shown that there are numerous ways to define what privacy is and that some scholars claim that privacy is not even a single concept. The goal of this section is to investigate in what different ways privacy can be accounted for. Some of these methods are linked to one of the definitions above, but others are more general and are not immediately connected to a specific definition.

In order to structure the overview of different methods, I make use of a framework introduced by Lessig. He distinguishes between four types of constraints that can be used for regulation: the law, social norms, the market and architecture (Lessig, 2006, p. 123). While the first three should be self-explanatory, the last type of constraint might require some explanation. Architecture refers to constraints that limit an individual in performing certain actions or make it preferable to choose certain actions. Such constraints can, for example, occur in the form of physical limitations (like a door lock) or software code limiting the user to specific actions. Important about this framework is that none of these constraints stand on their own. Instead, these different constraints coexist and influence each other. Nevertheless, it is helpful to conceptually distinguish these types of constraints. In what follows, I group the different methods to account for privacy by the type of constraint that primarily or initially is used to achieve the desired consequences.

the law Within the regulation of privacy governed by law, one can distinguish between two main lines of approaches. The one being data protection legislation and the other intellectual property rights. I already described the data protection legislation in the European Union in Section 1.2 and I am going to provide some more insights about an intellectual property rights approach in Section 2.5. Within this section, I therefore limit the discussion to a rough outline of these two different concepts.

data protection

Data protection legislation matches closely to the definition of privacy as control given above. Through legal provisions individuals are empowered to determine which information about them is known to others. The law thus lays the foundation of this method, but the other constraints also play a role. Without the awareness of individuals about their rights, the protection of privacy as a competitive element in the market, the enforcement of these rights by data protection agencies and software code that complies with these regulations, the system of data protection would not be effective.

intellectual property

As discussed above, the legal doctrine of intellectual property rights is another way to protect privacy. It also provides a form of control over personal data, but in a more indirect and general form. Through the recognition of property rights in data, anyone who wants to process personal data has to acquire the corresponding rights first. This, again, is a form of protection initiated through the law, but also dependent on the other forms of constraints. Especially important within the scope of this study is the question to what extend market mechanisms can be used to stimulate the protection of privacy in this form.

An important consideration with regard to the method of intellectual property rights is the question whether these rights are only seen as a method to account for privacy or also as a grounding for a right to privacy itself. When reading the classic texts by Warren and Brandeis (1890) or Westin (1967), it seems that their considerations about the role of property rights are an expression of the former. Another view, however, is to also ground the right to privacy itself in the ownership of information. This view can be derived from the Lockean concept of *self-ownership* and is for example discussed by Blok (2002, p. 26). I am going to discuss this point in more detail when describing the moral reasons for privacy protection in Section 2.3.

social norms Next to the codified law, social norms play an important role in the protection of privacy. In the case of informational privacy, data protection legislation provides the legal framework, but social norms define how these are filled in in practice. The questions which violations of data protection are brought to the attention of a data protection officer, to what extent the protection of privacy is a competitive advantage and which data processing practices implemented in digital systems are found acceptable is primarily based on the prevailing social norms.

privacy as a separate dimension

Morozov argues that consumers have to take ethical considerations into account when they share personal data (2013a). He draws a parallel to climate change where some decades ago consumers made decisions solely based on the market price of energy. When it became widely known that the market price of energy does not account for environmental damage, consumers started to take ethical considerations next to the energy price into account. Morozov argues that consumers should do likewise when it comes to the consumption of services that entail data sharing as part of the pricing model, because data sharing can entail negative consequences for other individuals as the one who shares the data. This line of thought, namely to include ethical considerations into decisions about the sharing of personal data, is what I will call *privacy as a separate dimension* in the following. It is closely connected to consumer behaviour and the social norms regarding data collection practices.

privacy through accountability

Another aspect closely connected to social norms is the accountability of businesses who utilize personal data. Through the developments of big data and an increase in secondary use of personal data, it becomes difficult to decide at the time of data collection whether it concerns a legitimate activity or not. This has led to an increasing number of calls to not judge based on the legitimacy of data collection, but to focus more on the actual consequences of how the data is used (see e.g. Boyd, 2011; Mayer-Schönberger & Cukier, 2013; Minister en Staatssecretaris van Veiligheid en Justitie & Minister van Binnenlandse Zaken en Koninkrijksrelaties, 2013; President's Council of Advisors on Science and Technology, 2014). As of now, this is primarily a consideration taken into account by the users of online services, but some would like to see this consideration in the center of data protection legislation.

the market As the discussions on the different methods found in the law and social norms has shown, market considerations play an important role when it comes to the effectiveness of mechanisms aiming at the protection of privacy. Currently, this mainly concerns the decision of users which online services to use, but through the trend towards the Internet of Things all kinds of consumer goods will increasingly be affected as well.

data markets

Next to this development, however, stands the idea to create markets for the trade of personal data themselves. As shown in Section 1.1, such markets already exist, but do not involve the user themself yet. The claim that such a market in which the user directly sells their data is a means to protect privacy is the central focus of this study. I therefore do not further elaborate on this idea at this point, but refer to the later chapters. However, it is important to mention that it is this type of regulatory constraint where the idea of a data market fits in.

architecture For the case of informational privacy, architecture might be the strongest form of regulatory constraint. Through architecture in the form of software, or simply *code*, it is determined which information is shared with whom and who has access to it. In his book 'Code', Lessig (2006) argues that the regulatory power of code is underestimated by politics and should be more dominant in discussions about the regulation of information systems. In the remainder of this section, I show in which forms code can play a role in the protection of privacy.

privacy as control

An important role of code in the protection of privacy is to give users the choice that data protection legislation entitles them to. The privacy as control definition puts the focus on the control that users have to determine which information is known about themselves. This means

that information systems have to account for this type of control. Without this form of control being built into the code, users cannot determine which information about them is known and thereby privacy is violated.

privacy by obscurity

Selinger and Hartzog (2014b) argue that the concept of *obscurity* is a better way to think about many of the current debates on privacy. In their view, obscurity provides "probabilistic levels of protection by increasing the transactional cost of finding or understanding information" (Selinger & Hartzog, 2014b). It expresses the thought that the classification of information in either public or private is inadequate. What matters in practice is how difficult it is for someone to obtain a certain piece of information. Understood in this descriptive manner, it can help to better understand some of the debates about privacy. Employed more actively, this insight can also be used to make online communication more obscure on purpose. Although current technological and political changes lead to an overall reduction of protection that obscurity offers, making information more obscure can nevertheless provide some protection. This protection, however, mainly affects day to day activities and is less effective when people invest explicit effort to find information.

The importance of obscurity has also been acknowledged by Solove, who speaks about *increased accessibility* as one of the privacy harms in his taxonomy on privacy (2008, pp. 149). Also, the role of obscurity has recently been recognized by the European Court of Justice (ECJ) in the case C-131/12 on Google vs. Spain (2014). The ruling mandates Google to remove search engine results upon the request of individuals who feel that their privacy is invaded by a particular entry in the Google search engine. This also holds for websites that are still accessible, which means that the ECJ implicitly acknowledges the role of obscurity. Unfortunately, the debate following the court ruling has mainly been along the lines of 'the right to be forgotten', thereby leading to some of the fallacies that a discussion focussing on obscurity could have avoided (Selinger & Hartzog, 2014a).

co-privacy

Another method to account for privacy that is that of *co-privacy* (Domingo-Ferrer, 2011). Domingo-Ferrer proposes to build information systems in such a way that it becomes attractive for users to behave in a manner that promotes privacy. Formalized through game theory and Nash equilibria, co-privacy is implemented in code by building protocols in such a way that it becomes attractive for users to help each other in protecting their privacy. Another option is to build applications in such a way that they give the user some form of utility by protecting their privacy. An important point of this approach lies in the fact that technology not only makes it possible for users to protect their privacy, but even promotes this behaviour.

privacy-enhancing technologies

A more general form of privacy protection through code are privacy-enhancing technologies (PETs)[2]. This umbrella-term covers a broad range of technologies that play a role in the protection of privacy. Elements often used in that regard are methods to achieve the confidentiality of information (e.g. encryption), ensure the anonymity of users, or more generally try to minimize the amount of information that is necessary to achieve a given end. Through the utilization of such methods in the construction of information systems, the privacy of a user is better protected than it would be if the information system would have been built in a straight-forward manner.

A noteworthy aspect about privacy-enhancing technologies is that they can create a narrow focus on what privacy constitutes. In my personal experience acquired through working in this field, I made the observation that scholars who work on PETs often equate a specific method of protection with privacy as such. Scholars working on encryption technologies, for example, would consider privacy as confidentiality of information and accordingly see privacy as protected if all communication is being encrypted. Smith et al. (2011) seem to share this experience, as they include a section on what privacy is not (anonymity, secrecy, confidentiality, security, ethics) in their review on privacy-related research.

[2]For an overview over such technologies see (Danezis & Gürses, 2010).

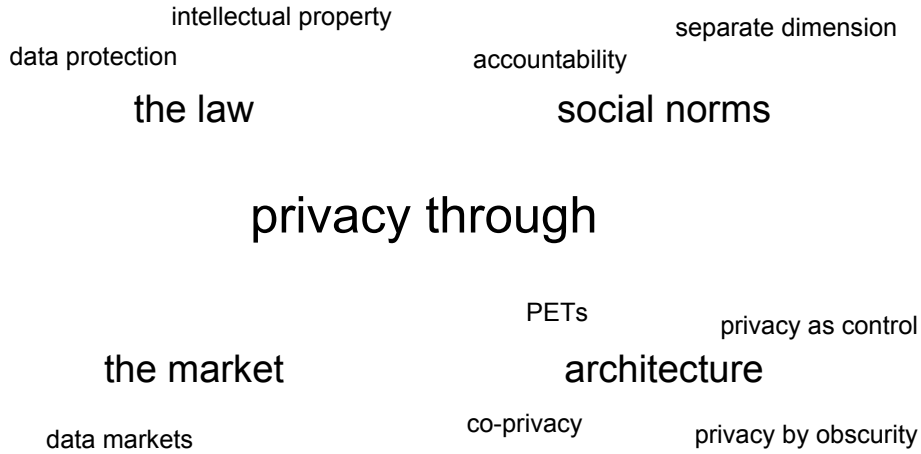

Figure 2.2: Overview over the different methods to account for privacy.

In more general terms, this observation applies to all the methods discussed in this section. The fact that a particular regulatory constraint is being employed with the goal to protect privacy, can lead to the impression that the object of regulation is what constitutes privacy[3]. In this regard, the different methods presented in this section can also be considered to be definitions of privacy.

Figure 2.2 shows an overview over the different methods to account for privacy. As I explained earlier, it is always an interplay of the different regulatory constraints that is needed to achieve an intended outcome. An intellectual property approach implemented in the law seems to fit well with the approach of data markets. Whether data markets are indeed a possible way to account for privacy is the goal for the remainder of this study. However, before turning to this question, I will look into the moral reasons why privacy should be protected.

2.3 Moral reasons to protect privacy

Besides defining privacy in itself and looking at the different methods to account for privacy, another way to view privacy is to focus on the moral reasons to protect privacy. Besides additional clarity about the function of privacy, this can also help to improve discussions about privacy. As Van Den Hoven suggests, the focus should be on the moral reasons for data protection instead of a recurring discussion about definitions of privacy itself (2008, pp. 303). In this section I provide an overview over the different moral reasons to protect privacy that can be found in the literature. I start with accounts related to harm and then go into reasons relating to autonomy. After that, I briefly consider the relation of privacy to ownership of data, before I finally discuss the public good character of privacy.

Connected to the discussion of moral reasons to protect privacy, is the question what type of value privacy is. Considering privacy to be an *intrinsic* value means that the protection of privacy is valuable in itself and can not be explained through other values. Considering privacy to be an *instrumental* value, on the other hand, means to completely seek the moral reasons to protect privacy in the fulfilment of other values such as autonomy. Next to this distinction, Roessler refers to a *functional* explanation which accounts for the intrinsic value of privacy, but acknowledges that privacy also has important functions in the fulfilment of other values (2005, p. 69). This question is important insofar, as accounts of privacy considering it to be a purely intrinsic value can by definition not contribute to this section. I do not want to explicitly choose for one perspective at this point, but use the explanatory power of reductive conceptions to further explore what privacy

[3] See also the remarks by Purtova (2011, p. 41) how the conceptualisation of privacy channels the development of law.

is and why it is valued.

prevention of harm Van Den Hoven mentions *information-based harm* as one of the moral reasons to protect personal data (2008). Examples of such harm include identity fraud on the Internet, or the planning of a robbery through intelligence gathering. Although the protection of personal data in such contexts limits the epistemic freedom to know, it seems a justified limitation based on Mill's harm principle (Van Den Hoven, 2008, p. 311). A good overview over different types of harms inflicted through the violation of privacy is provided by Solove. He names *physical injuries*, *financial losses and property harms*, *reputational harms*, *emotional and psychological harms*, *relationship harms*, *vulnerability harms*, *chilling effects*, and *power imbalances* as possible types of harms (Solove, 2008, pp. 174).

information-based harm

Hildebrandt and Koops discuss the harm that can be inflicted through the use of profiling technologies (2010). Through the use of stochastic processes and the errors that occur as part of it, individuals can be assigned into categories they do not belong to. If this false categorization then leads to real consequences for the individual concerned, the profiling leads to *unjustified discrimination*. But even if the process is free of errors, the use of profiling can still contribute to *unfair discrimination*. A prominent example of this type is price discrimination, a technique where customers are charged different prices based on the assumed willingness to pay[4]. Since in both cases the discrimination becomes possible through the collection of data, the protection of personal data is a measure to prevent the discrimination practices.

unjustified and unfair discrimination

Of particular relevance for unjustified discrimination is the practice of *non-distributive group profiling* (Hildebrandt, 2006). With this technique, a profile to be applied to a particular target group is created. However, the attributes used in this profile are not necessarily those of each individual in the target group, leading to non-universal generalisations. If such a profile is created and incorrectly applied to an individual not exhibiting this attribute, they will nevertheless be treated as if they did. The malignant property of this technique is that the profile can be created from a dataset of individuals who consented to this form of profiling, but then later applied to another group who did not consent to be analysed. This means that the originally individual decision to consent into profiling has possibly adverse consequences for others as well.

A further issue connected to the discrimination through profiling is the availability of a due process. The EU data protection Directive (European Parliament, 1995) and also the OECD principles (OECD, 1980) prescribe transparency about the collection practices and give the individual the right to inspect the data collected about them. Several scholars (see e.g. Mayer-Schönberger & Cukier, 2013; Morozov, 2013b; Solove, 2008; Hildebrandt & Koops, 2010; Purtova, 2011) warn that the increased use of profiling and Big Data make it difficult to enforce this right in practice. This is due to the numerical increase in the use of such techniques, but also the increasingly hidden character of profiling activities as part of ambient technologies. The claim made here is that the way how technology steers human life becomes invisible and can therefore also not be challenged by the individuals concerned[5]. Solove therefore comes to the conclusion that we should turn to Kafka's 'The Trial' and not only focus on Orwells's 'Nineteen Eighty-four' when discussing the negative consequences of data mining (2008, p. 194).

undue process

Another important distinction to make is the question which types of information-based harm one recognizes. The practice of U.S. courts to only recognize privacy harms that can be proven lead Calo to the distinction between objective and subjective privacy harms (2011). While the former category concerns the actual use of personal information in a harmful way, the latter category concerns the perception of unwanted observation. Subjective privacy harm can stem from the

subjective privacy harm

[4]For an introduction to price discrimination and the relation to privacy see (Odlyzko, 2003)

[5]The importance of this development is not to be underestimated, as the current data protection legislation is based on the assumption that individuals know about the personal data relationships and actors (Purtova, 2011, p. 52).

knowledge about the possibility of possible adverse effects in the future, but also includes embarrassment based on the revelation of personal details. As Calo argues, "[i]t is enough to believe that one is being watched to trigger adverse effects" (2011, p. 1154), which closely connects subjective privacy harm to the *panoptic effect*. In order to measure subjective privacy harm, the degree of aversion against the observation and the experienced extent of observation can be used as a metric. Defined in this way, subjective privacy harm can form a useful category to complement the more widely recognized objective privacy harms.

Also other scholars recognize the importance of subjective privacy harm, although under varying labels. Solove in that regards speaks about *emotional and psychological harms* (2008, pp. 175). Reiman introduces the notion of 'internal censorship', that she distinguishes from other types of risk that are also closely connected to the 'informational panopticon' (Reiman in Nissenbaum, 2010, p. 75). Nissenbaum furthermore connects these panopticon-like effects to autonomy, insofar that privacy is a prerequisite for voluntary actions and the ability to freely formulate plans (Nissenbaum, 2010, p. 82).

autonomy The connection between autonomy and privacy has been extensively investigated by Roessler. In her book *The Value of Privacy* (Roessler, 2005), she gives an extensive account in which ways privacy is related to autonomy and thereby also to freedom. Starting from a definition of privacy along the lines of privacy as control, she describes how the absence of control by others is necessary in order to conceive, develop and pursue goals. She further specifies this link for the cases of decisional, informational and local privacy, leading to the three categories of privacy that she distinguishes. Decisional privacy focusses on the sovereignty over the interpretations and decisions of one's own life. Informational privacy concerns the protection of authenticity of one's self-representation, which is connected to the societal value of self-determined citizens and the proper functioning of relations. Local privacy is necessary to provide the room necessary for self-presentation and the ability to invent oneself. As Roessler notes, "[...] privacy is characterized by the inextricability of ethical and moral perspectives and problems." (2005, p. 141) Because of this complex relationship, I can not fully explicate the relation between privacy and autonomy at this point, but refer instead to the work of Roessler (2005).

A more compact account of the connection between privacy and autonomy can be found in (Nissenbaum, 2010, pp. 81). Nissenbaum distinguishes between three different relations, of which the first one sees privacy as one form of autonomy, expressed through the ability of informational self-determination. The second type sees privacy as a contribution to the material conditions necessary for the exercise of autonomy, expressed through the freedom of scrutiny by others. The third relation can be found in the capacity to review principles of action. The former type of relation thus considers privacy to be a distinct type of autonomy, whereas the latter two consider the relation to be causal, with privacy being a precondition for autonomy. This latter type of relationship is also recognized by Hildebrandt and Koops, who see the relation of privacy as precondition for autonomy as an enabler of positive freedom and a public good (2010, p. 436).

ownership A very different type of reason for the protection of privacy can be found in the expression of property rights connected to personal data. Based on the Lockean idea of the *ownership of the self*, several scholars argue that individuals have property rights in the data concerning them. Spiekermann, Korunovska, and Bauer argue that users are likely to develop a feeling of ownership based upon the time that they invest in the creation of user profiles (2012, p. 5). Blok makes a more general link, explaining how the Lockean idea of property as a natural right was historically also of influence in early U.S. court cases on the protection of privacy (2002, p. 167). This historical link is also described by Solove (2008, pp. 26).

ownership of the self

In practice, however, it is often not clear who invested the labour in the creation of data. As

Cohen notes, it often is the case that third parties gather the information about an individual and thereby invest the labour (2000, p. 1381). Therefore, it is not straightforward to infer that only the individual concerned has rights in their data.

common good Privacy is often considered to be a right that is only of relevance for the individual and not of relevance for society as a whole. And even if societal interests are considered, privacy is often seen as conflicting with other societal interests such as security or efficient markets. As Solove argues, such a view underestimates the value of privacy since its benefits for society are neglected (2008, pp. 89). In the following, I review a number of reasons that show in which ways privacy contributes to society and therefore makes it worthwhile to protect for reasons beyond the individual.

public sphere

At first sight contradictory might be the idea that privacy is a necessary requirement for the public sphere to function properly. However, it is the right mix of privacy and transparency that is necessary for public discourse (Simitis, 1987, pp. 729-737). If there is too little transparency, privacy can form an impediment for the democratic decision-making process. If there is too much transparency, individuals lack the freedom that is necessary to develop the ideas with which they can contribute to this very process. This room for seclusion can not only be found in form of the private life, but also in the form of separated societal sub-systems. The German constitutional court recognized the importance of these aspects in its landmark decision on the right to informational self-determination (Hornung & Schnabel, 2009).

societal norms

Next to the room necessary to experiment with deviating ideas, it can also be of societal value to actually deviate from majority norms. As Solove writes, "[p]rivacy is a recognition that in certain circumstances, it is in society's best interest to curtail the power of its norms" (2008, p. 95). Similar as with ideas, the enforcement of norms needs some room to account for the proper development of these. But also for the protection of minorities an imperfect enforcement of norms is of importance. The societal importance thereof has already been recognized by Westin in 1967, who writes "[d]evelopment of society is particularly important in democratic societies, since quality of independent thought diversity of views, and nonconformity are considered desirable traits for individuals" (Westin in Miller, 1971, p. 49).

contextual integrity

Another aspect connected to societal norms is the functional separation of different contexts. As explained in Section 2.1, the central idea of contextual integrity is the insight that different norms govern different societal contexts. Data flows disregarding the boundaries of contexts are therefore likely to violate these norms (Nissenbaum, 2010). Privacy in that regard is therefore a prerequisite for the correct functioning of norms.

social interaction

A similar argument is made by Roessler and Mokrosinska (2013) regarding the correct functioning of social relationships. They argue that privacy plays in important role for the functioning of friendship, professional relation, and also the interaction with strangers. Privacy in that regard controls the intimacy of friendships, is used to distinguish different types of professional relationships, and allows for *civil inattention* in public. The correct functioning of both professional relationships and friendships is of importance for society, as they provide functionality that other more formalized institutions within society do not fulfil.

common pool resource

Regan (2002) provides another argument very different from the ones mentioned before. She argues that personal data is a common pool resource which quality degrades if too many utilize it. She establishes this connection based on the insight that the overly use of personal information creates a climate of mistrust. This makes privacy a collective value and emphasizes the societal importance of the protection thereof.

discrimination

A last point to be mentioned here is the issue of unjustified and unfair discrimination that I discussed above already. In the case of non-distributive group profiling, the unjustified discrimination can not only apply to those whom data have been collected, but can also apply to those who

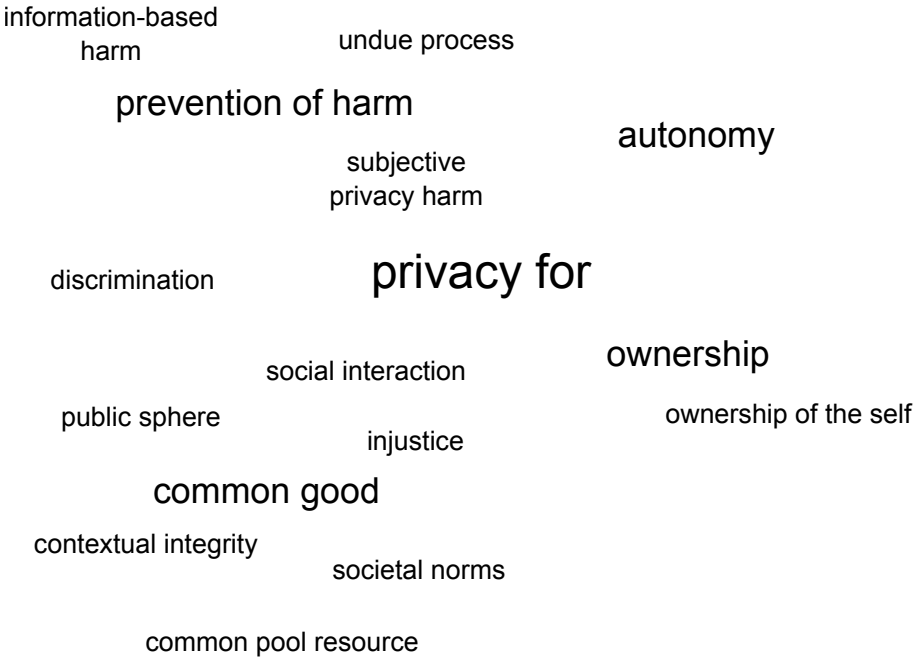

Figure 2.3: Overview over the moral reasons to protect privacy.

on purpose did not do so (see e.g. Selinger & Hartzog, 2014a, p. 8). This characteristic makes the decision whether such form of profiling should occur a decision beyond the individual and therefore of societal relevance. Similarly, unfair discrimination based on collected data can effect others who did not consent to it. For the example of price discrimination, less-frequent shoppers might have to effectively pay a premium (Vanderlippe, 2005). Aggravating this situation is the fact that price discrimination often happens in an overt form (Cave et al., 2011, p. 53), which makes it even more important to decide upon such practices as a societal question.

injustice

Another reason to decide upon issues of data protection as a societal question is the preservation of *complex equality*. Van Den Hoven (1997, 2008) names this as one of the moral reasons to protect privacy. It is based on the concept of *spheres of justice* by Walzer, which entails that goods are allocated by different mechanisms in different social spheres. Van Den Hoven argues that information is also such a type of good and that information exchange between different spheres therefore has to be blocked. This argument has been further analysed by Nagenborg (2009) and is also used by Nissenbaum (2010) as one of the groundings for her notion of contextual integrity.

As this section has shown, there are a number of reasons to protect privacy (see Figure 2.3 for an overview). While there are a number of specific reasons that can be difficult to distinguish, they seem to fall within four broader categories. The main two reasons to protect privacy are the prevention of harm and autonomy. Another not so common reason is the protection of ownership. Furthermore, it is important to realize that there are a number of accounts that explain why privacy should be considered a common good and the protection of it should not be left to the individual. These considerations are the reason why data protection legislation has been installed within Europe. The specific relation of data protection and privacy is what I explore in the next section.

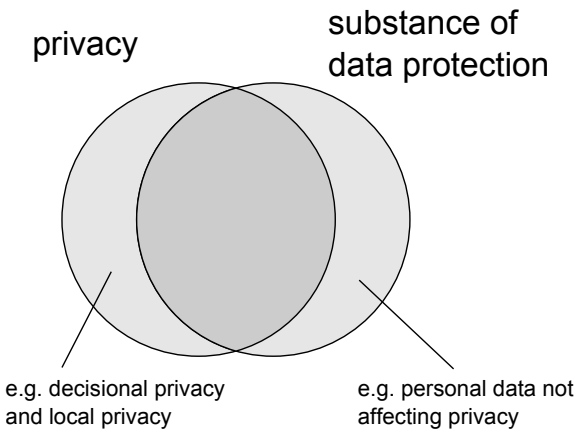

Figure 2.4: Venn diagram showing the overlap between the substance of what is regulated through data protection and privacy.

2.4 Privacy and data protection

In order to better understand the status quo around the protection of informational privacy, it is important to consider what the relation between privacy and data protection[6] is. As discussed in Section 2.2, data protection is the main regulatory means within the European Union to ensure the protection of informational privacy. However, the notion of privacy as such also exists in the juridical sense and it is furthermore of interest how data protection relates to privacy as a more general concept.

As a first approximation it can be said that data protection is a regulatory means to ensure the protection of informational privacy. As Cave et al. (2011, p. 25) note, data protection is a means to ensure informational self-determination. With a conception of privacy as informational self-determination in mind, the link between the protection of privacy and data protection thereby becomes apparent. However, as I discussed above, informational self-determination is only one way to define privacy and is as such also not very precise in operational terms.

When using the three types of privacy described by Roessler (2005), it is clear that data protection mostly concerns informational privacy. It is not intended to protect local privacy and also only marginally touches upon decisional privacy[7]. This shows that with a broad conception of privacy in mind, privacy is more than what it in substance implemented through data protection legislation. Seen like that, one can say that the substance of data protection is a subset of privacy.

The other way round, privacy is also more narrow than what is achieved through data protection. Since data protection applies as soon as it concerns personal data, it affects the processing of all data through which natural persons can be directly or indirectly identified. Such data carries knowledge about a specific individual and thereby potentially bears the risk to infringe on their privacy. However, as Hildebrandt and Koops (2010) note, not all processing of personal data is necessarily relevant for informational privacy. This means that privacy is also a subset of data protection and the two notions only partially overlap in their substance. This relationship is illustrated in Figure 2.4.

An interesting study in this regard has been performed by Oetzel and Spiekermann (2013). Based on the privacy taxonomy devised by Solove (see Section 2.1), they assess which of the

[6] Within this section I am referring to data protection in the juridical sense, i.e. the mechanisms that are currently implemented through the EU Data Protection Directive.

[7] For example, paragraph 15 of the EU Data Protection Directive restricts the ways in which automated decisions that significantly affect the data subject are allowed to be taken. This provision can be seen as touching upon decisional privacy, as it gives procedural guarantees to the way how decisions are being taken.

privacy harms identified by Solove are covered through EU Data Protection legislation. They arrive at the conclusion that data protection is complete in this regard and covers all the relevant privacy harms. Described in the language of the Venn diagram discussed above, this would mean that the subset of privacy that is not covered by data protection should be empty, at least when only discovering informational privacy. However, I find this a somewhat doubtful conclusion. As Solove (2008) describes when building up his taxonomy, he derived the privacy problems largely from cases that appear in the legal domain. This means that his privacy taxonomy mostly covers problems that are already in one form or another addressed in the legal system and it is therefore somewhat self-referential to evaluate legislation based on it. Nevertheless, the study of Oetzel and Spiekermann is valuable as it shows that the EU data protection legislation covers all the aspects of informational privacy that Solove identified based on a legal perspective focussingon the United States.

That there are aspects of informational privacy which are not covered through current EU data protection legislation can also be seen by looking at proposals to extend the scope of data protection. One of these proposals is to extend the scope beyond that of natural persons. As described in Article 2 of the EU Data Protection Directive, the directive only addresses data which directly or indirectly identifies natural persons. As discussed above under the keyword of *non-distributive group profiling*, harm can also occur through data which does not relate to specific individuals. A number of authors (see e.g. Boyd, 2011; Cave et al., 2011; Floridi, 2014) discuss this problem and propose to also recognize the importance of privacy for groups through data protection legislation. However, as it was early recognized by Vedder (1999) privacy rights in the form of collective rights also do not help in the age of profiling, as the groups concerned are not necessarily structured and organized. Contrary to legal persons they would therefore not be able to defend their rights. He therefore proposes to introduce 'categorical privacy', which he defines as an extension of data protection towards anonymous data that can have the same negative consequences as if it was personal data. These two proposals of *collective privacy* and *categorical privacy* show that data protection is more narrow than informational privacy and in this regard not necessarily complete in substance.

A very different distinction between privacy and data protection can be found in the works of Hildebrandt (2006) and Gutwirth and Hert (2008). They differentiate between opacity and transparency tools. In their view, privacy is an opacity tool which means that the goal of privacy is to make the access to the personal sphere more opaque and thereby forms a barrier. Data protection on the other hand is a transparency tool, which means that it provides transparency towards the data subject about the processing of data concerning them. Thereby, it is not prohibitive as such, but rather channels power and provides data subjects with an instrument for control. Although I do not find this distinction very helpful as a means to conceptually distinguish data protection and privacy, it contains an important insight. Data protection does not necessarily have to be about the denial of access to personal data, but rather sets the conditions under which data processing should take place. Understood this way, it can also be seen as an enabler for data processing that would otherwise we be forbidden because of privacy considerations.

The latter distinction is especially meaningful when considering privacy not on a conceptual level, but in the context of European legislation. In the EU, both the Human Rights Convention of the Council of Europe (2010) and the Charter of Fundamental Rights of the European Union (2010) give privacy the status of a human right (see also Section 1.2). Data protection as defined by the EU Data Protection Directive does not define this right to privacy in more detail, but rather adds a complementary notion that exists next to it. From a legal perspective, privacy and data protection therefore exist side-by-side and cannot be simply reduced to one another[8]. This relation in the juridical sense is not only of relevance to distinguish the two notions, but also to assess to

[8] A good introduction to this intricate relation can be found in (Kokott & Sobotta, 2013).

which extent privacy rights can be waived. In this regard, Purtova (2011, p. 235) comes to the conclusion that a waiver of the rights constituted through data protection would be a violation of the fundamental right to privacy and is therefore not possible. Data protection and privacy in the juridical sense are therefore two distinct notions, which are connected in a not necessarily straightforward manner.

Summarizing, one can say that privacy and data protection are two distinct concepts that overlap in the substance of what they address. There is more to privacy than what is covered by data protection and there are aspects of data protection that would not necessarily fall under the concept of privacy. For the part that they overlap, data protection can be seen as a regulatory means to ensure the protection of privacy. This distinction, however, only holds on a conceptual level. In the juridical sense both notions co-exist in an intricate relation established through multiple legal provisions that are defined in different multilateral treaties.

2.5 Privacy in Europe and the U.S.

Although the perspective of this study is a European one, the perspective of privacy in the United States is of interest as the discussion around property rights for data originates there. A more detailed understanding of the differences around privacy protection in Europe and the United States therefore helps to view arguments for property rights in the right light and avoids ill-considered conclusions neglecting the legislative and cultural differences that exist. In this section I first describe the distinctive features of the U.S. approach to privacy protection and afterwards point out some more general differences.

distinctive features of the U.S. approach In the U.S. privacy legislation can be found in two forms. The constitution provides protection from governmental privacy intrusions and tort law deals with privacy intrusions in the private sector. For the area of constitutional privacy protection, the fourth, fifth, and fourteenth amendment are of relevance. They provide a protection from unreasonable searches and seizures, put a bar on self-incrimination and guarantee a due process (Purtova, 2011, p. 102). For the private sector, there are four types of tort that are of relevance: intrusion, disclosure, false light and appropriation (Purtova, 2011, p. 93). Next to the law of tort, there are also a number of laws that regulate the collection of data for specific sectors. Such legislation, for example, exists for health care, the financial industry or the protection of children (Craig & Ludloff, 2011, p. 33).

sectoral approach

Distinctive for the U.S. situation is the strong reliance on self-regulation of the private sector. The only binding privacy legislation in this sector are the four tort types named above. An individual can make use of litigation if they can prove harm that they suffered based on a privacy violation. However, this harm has to be objectively quantifiable, which excludes litigation for subjective privacy harms (Calo, 2011). Next to the possibility of litigation through court, there are also non-binding industry standards, with the *Fair Information Practices* (FIPs) as the most important ones. The FIPs were first published in 1973 (U.S. Dep't. of Health, Education and Welfare, 1973) and cover some of the positive obligations that can also be found in the EU Data Protection Directive. As the FIPs are not binding, an individual therefore has to rely on the regulatory function of the market in this regard.

self-regulation

An element that is less prevalent in the European context is that of a *reasonable expectation* of privacy. Both in court cases concerning the fourth amendment and in tort cases, courts in the U.S. apply a yardstick of what society would find a reasonable expectation of privacy (Solove, 2008, p. 71). What a reasonable expectation of privacy in specific cases can mean seems debatable. In one case, for example, a court found that the police observing houses from above by

reasonable expectation

flying over them does not constitute a violation of privacy since the observation took place from a public point (Solove, 2008, p. 110). In another case, a court declared that people do not have a reasonable expectation of privacy in the phone numbers they dial, since they voluntarily transmit this information to their telephone provider (Solove, 2008, p. 139). Underlying these rulings is the idea that only information that is fully secret can constitute a privacy interest (sometimes called *secrecy paradigm*). Whether this really matches the expectations of society in all cases seems doubtful.

secrecy paradigm

differences between Europe and the U.S. A difference on a more general level is that of privacy as a human right in Europe versus the conception of privacy in a more utilitarian fashion in the U.S (Acquisti, 2010). Connected to this view is the observation that privacy in the U.S. is treated like a commodity (Busch, 2006), whereas the human rights status in Europe makes clear that privacy rights cannot waived. As Zwick and Dholakia (1999) explain, the treatment of privacy as a commodity is the basic assumption underlying the self-regulatory approach that is dominant in the U.S. While the European model with privacy as an unalienable right draws clear boundaries, the U.S. situation that is largely based on self-regulation leads to a commodification of privacy and fits more into a utilitarian approach.

human right vs. utilitarian approach

Another distinction that is more closely related to the governmental context is that of a focus on dignity in Europe compared to the right to be let alone in the U.S (Craig & Ludloff, 2011, pp. 18). As Whitman remarks "American privacy law is a body caught in the gravitational orbit of liberty values, while European law is caught in the orbit of dignity." (Whitman in Solove, 2008, p. 185) This distinction goes together with the divide in regulation for the governmental context and the private sector. As I explained above, the U.S. differentiates between intrusion into privacy by the government and by private businesses an individual establishes a business-relationship with. While the focus in Europe is on the preservation of dignity no matter in face of whom, the U.S. privacy legislation focuses on an absence of interference through the government.

dignity vs. right to be let alone

Connected to this divide is the focus within the U.S. on negative liberty in the form of an absence of government intervention. As Purtova (2011, pp. 100) remarks, the law of tort that is of relevance for the private sector does not provide any positive obligations, which form the essence of data protection in Europe. This is different for the constitutional protection of privacy in the U.S., which guarantees an absence of government intrusion into privacy. For the private sector, markets take the role of enabling factors, which might otherwise be created through law (Cohen, 2000, 1399). Put briefly, the focus on government interference expresses itself through a freedom from privacy intrusion by the government, but goes along with a lack of positive obligations on businesses that might interfere with their customers' privacy.

positive vs. negative liberty

Another difference discussed in the literature is the influence that the works of Hegel and Locke had on the European respectively U.S. perception of privacy. Blok (2002, p. 26) states that the central aspect of personality as devised by Hegel was of great importance in the development of German law[9]. Indeed one can argue that the protection of the personality and possibilities for development and self-expression are central elements of the EU Data Protection Directive. According to Blok (2002, p. 26), Locke was of great importance for the development of U.S. law, which protects many aspects connected to personality through property. Similarly, Solove (2008, p. 26) makes the link between the idea of property rights in information concerning oneself and the works of Locke. Interestingly, the famous article by Warren and Brandeis (1890, p. 204), which is often cited as highly influential for the development privacy rights in the U.S., is rather critical towards a labour-desert theory in the case of privacy. Nevertheless, the perception of privacy as a commodity (see above) and also the more prevalent discussion of property rights fit the link

Hegel vs. Locke

[9] Germany was one of the earliest countries with data protection legislation in place (Blok, 2002, p. 256) and thereby of importance for the development of the European Data Protection Directive.

between Locke's libertarian approach and privacy rights in the U.S.

This section has shown that there are big differences in privacy legislation between Europe and the U.S. Foremost is that the U.S. relies much more strongly on industry self-regulation and do not know an omnibus approach to privacy protection. Since the constitution does not hold for privacy invasions by private parties and sectoral privacy legislation only exists for some sectors, there is a 'regulatory vacuum' when it comes to privacy protection in a more general sense. Together with a more utilitarian approach to privacy and a general preference for market-based solutions, this explains why the idea of property rights for data and the establishment of markets might seem more natural in the U.S. than it is in Europe. This is important to keep in mind when I will introduce the data market approaches in Chapter 4.

2.6 Conclusions on the notion of privacy

The review on literature on privacy in this chapter has shown that privacy is a multifaceted concept and it is difficult to find a single conception of privacy encompassing all of the different aspects. The definitions of privacy (see Section 2.1) each focus on different aspects and there is no single one that is encompassing but still provides a workable definition of privacy. When deciding on a single definition of privacy, I therefore go along with Solove (2008) who defines privacy as a as a cluster of problems. When it comes to the search for an encompassing definition of privacy, however, I agree to Van Den Hoven (2008) in that we can do without a precise definition. What should serve as guidance for the discussion of problems related to informational privacy are the moral reasons for the protection of personal data.

By putting an emphasis on the moral reasons for the protection of personal data, the moral reasons discussed in Section 2.3 thereby become central. As I have shown, there is a wide range of reasons which one can group into four categories. Most apparent is the prevention of actual harm, including subjective privacy harm as a less common but nevertheless important category. A very broad but crucial category is the link between privacy and autonomy. Privacy in this regard provides the spaces that are necessary to make decisions and develop plans. Somewhat less common is the link between privacy and ownership of the self. Finally, it is important to recognize that privacy is not only important for the individual themself, but also crucial as a common good. Many forms of interaction that we have as a society with one another are contingent on privacy. This entails that the protection of privacy is not necessarily for each single individual to decide upon, but to some extent also a collective decision.

I included the methods to account for privacy (see Section 2.2) in order to give an overview over the different ways how questions of privacy can be regulated. Central here is the idea that informational barriers that are put up lead to the rooms that are necessary for privacy. Important to realize in this regard is that it is not only the law alone which has a regulatory function, but also social norms, the market and architecture. The use of architectural constraints is the prevalent element of privacy enhancing technologies, but it is important to acknowledge that architectural constraints often protect only very specific aspects of privacy. Therefore it is always a mix of different regulatory means that should be sought. The specific methods to account for privacy are thus not suitable to define privacy, but are important in order to realize how a protection of privacy works in practice.

With data protection legislation as the main regulatory means how informational privacy is currently protected in the EU, the relation between privacy and data protection (see Section 2.4) becomes important. The central insight here is that data protection is distinct from privacy in two ways. First, data protection is a regulatory means to protect a part of the substance that one can define as privacy. In this regard, data protection is on a different level than privacy. Furthermore,

the substance covered by it is only partially overlapping with what one can consider to be privacy in broader terms. This means that there are privacy-related aspects that have nothing to do with data protection, but also that data protection covers elements that are not necessarily privacy-related. When talking about the protection of privacy, it is therefore important to realize that data protection legislation is an important element, but different from privacy as such and also different from the protection of personal data in the broader sense.

With regard to the fact that much of the discussion about data markets comes from the United States, it is instrumental to have a rough understanding of the cultural and legal differences between Europe and the U.S. in this regard. Concerning the legal differences, it can be noted that the U.S. lacks an overarching system of privacy protection and relies much more on industry self-regulation. This 'regulatory vacuum' important to take into account when discussions about property rights as a more overarching system occur. Furthermore, I have pointed to some cultural differences when it comes to the protection of privacy. Compared to Europe, the protection of privacy is seen as more instrumental and negative liberty plays a bigger role, especially when it comes to governmental intervention. Furthermore, the thoughts of Locke are said to have been of high importance, which explains the strong focus on property rights as a regulatory means. For the discussion of data market approaches below, it is therefore important to take these differences into account when assessing their suitability for a European context.

Summarizing, I consider privacy to be a useful shorthand for discussions about the issues that are related to it. At the same time, however, it can lead to confusion about the specific conceptions that it refers to. For the ethical evaluation that follows further below, I therefore focus on the moral reasons for the protection of privacy. Since this study focusses on informational privacy, this is coupled to the protection of data[10].

[10]The link between the protection of data and informational privacy is not as simple and somewhat more intricate. As the restriction of the EU Data Protection Directive to personal data shows, there exist data that presumably does not necessarily impact on anyone's privacy. The proposed extensions to the scope of data protection legislation and the controversies around Big Data, however, show that the current definitions of personal data might be too narrow either. In the following, I therefore argue based on the hypothesis that all data has the inherent potential to be of relevance when it comes to the protection of privacy. Even if a specific data set does not invade someone's privacy directly, it might nevertheless be of importance when combined with personal data in the more narrow sense. For this reason I base my discussion on a direct coupling between informational privacy and the protection of data.

3 Evaluation framework and methodology

The discussion in the previous chapter has shown that there are many aspects to the concept of privacy that make it difficult to define or measure privacy in a single way. Nevertheless, the goal of this study is to evaluate in which ways the introduction of data markets would impact the privacy of online service users. In order to achieve this goal, a method to evaluate the impact on privacy is needed. The goal of this chapter therefore is to develop an evaluation framework that is of help in this regard.

Since I partially base the evaluation framework on that of contextual integrity (Nissenbaum, 2010), I first give a more detailed description of contextual integrity than what I did in the previous chapter. Afterwards, I describe the evaluation framework that I am going to use. Finally, I point out how this fits into the overall methodology of this study.

3.1 The framework of contextual integrity

One of the important insights captured in the concept of *contextual integrity* is the idea that the social norms governing privacy are not universal, but are dependent on social contexts. Nissenbaum (2010) motivates her approach on the shortcomings of the private/public dichotomy and argues that a richer set of variables has to be taken into account. She names examples like *health care*, *education*, *employment*, *religion*, or the *social marketplace*. She bases the general relevance of context-dependent social norms on a number of different accounts from different academic fields.

Important to note is that Nissenbaum is not the only scholar recognizing the importance of social contexts for informational privacy. As I wrote above (see Section 2.3), Van Den Hoven (1997) provides an account based on Walzer's spheres of justice. In that regard, he sees information as a primary good that needs to be distributed according to the rules of a specific context, implying that information flows between different spheres have to be suppressed. This account is one of those that Nissenbaum based her framework of contextual integrity on.

The importance to distinguish social contexts also found its way into recent policy discussions. Based on the increasing frustration about informed consent in the form of long legalistic texts (see e.g. McDonald & Cranor, 2008; Van Alsenoy et al., 2013), there have been several appearances in which the role of context-dependent social norms has been recognized. The European Parliament proposed to introduce the condition of *reasonable expectations* as part of the legal basis of *legitimate interest*, so that data processors could only invoke this legal basis if the users can expect them to do so (European Parliament, 2013b, article 6(1)). Since reasonable expectations are likely to depend on the context in which the processing takes place[1], this would strengthen the relevance of context-dependent norms. In the Unites States, the Federal Trade Commission proposed to explicitly introduce pre-defined categories of 'commonly accepted data collection', but failed to do so because of industry lobbying efforts (Federal Trade Commission, 2012). Lately, Nissenbaum (2014) also published a reaction to a White House proposal that proposed 'the respect for context' as part of a privacy bill. In her reaction she discusses various interpretations what a context constitutes and concludes that only social contexts provide the distinctions necessary. She thereby

[1] Nissenbaum also emphasizes this connection between her framework of contextual integrity and the concept of reasonable expectations (2010, p. 162).

disapproves of attempts to define contexts in the more simple form of technologies, business models or business sectors.

More specifically, Nissenbaum defines contexts as "structured social settings characterized by canonical activities, roles, relationships, power structures, norms (or rules), and internal values (goals, ends purposes)." (2010, p. 132) Such a context has to be defined for the socio-technical system to be assessed, but the relevant characteristics can mostly be derived from earlier developed less-technical counterparts. This means for example that the use of information systems in health care can be oriented along the practices that were established in health care without information systems. Building on the definition given above, Nissenbaum mentions *roles*, *activities*, *norms*, and *values* as the most important elements for the framework of contextual integrity.

Out of the many norms that govern a context, *informational norms* are of especial importance for contextual integrity. These concern the flow of personal information and are in this regard most important for informational privacy. Nissenbaum (2010) considers four key parameters that characterize informational norms: *contexts*, *actors*, *attributes*, and *transmission principles*. Contexts thereby refer to the contexts as characterised in the previous paragraph. Actors are distinguished through senders, receivers, and subjects that the information transmitted refers to. Attributes refer to the kind of information being transmitted, or *data fields* in more technical terms. Transmission principles determine the norms in which way information ought to be transmitted and include principles such as confidentiality or reciprocity.

This definition of informational norms forms the conceptual core of the contextual integrity framework. Nissenbaum (2010) considers contextual integrity to be preserved if and only if these informational norms are respected. Accordingly, she devises a decision heuristic based on the four components (contexts, actors, attributes, and transmission principles) of an informational norm. If the introduction of a technical system changes any of the four parameters, it constitutes a *prima facie violation* of contextual integrity. Defined in this way, Nissenbaum considers the framework to be of descriptive nature. Thereby it is a tool to describe informational practices and to determine whether a change of these is likely to lead to reactions along the lines of privacy violations.

When only considering prima facie violations, contextual integrity seems to suggest that all changes in informational norms would be prohibited. As soon as a new information practice does not fit into a schema of established norms, contextual integrity is violated. However, also informational norms can change over time and new practices might be legitimate. Therefore, Nissenbaum (2010) considers this part of her framework to be of descriptive nature only. In order to determine whether changes in informational norms are acceptable more considerations are necessary. What Nissenbaum suggests in that regard is to focus on the question whether a new informational practice is better able to support context-relevant values than the status quo. By extending the framework in this way, it is possible to use contextual integrity to determine whether a new informational practice is acceptable although it constitutes a prima facie violation. This extension is what Nissenbaum calls *augmented contextual integrity decision heuristic*.

With the conceptual decomposition of a context and informational norms, the descriptive concept of prima facie violations and the more evaluative considerations added in the augmented decision heuristic, Nissenbaum (2010) offers a valuable package for the evaluation of privacy questions surrounding information practices. She presents the augmented decision heuristic in the form of nine questions (Nissenbaum, 2010, p. 182-183), which, however, I find somewhat ambiguous in their formulation. I therefore stick to the conceptual considerations that she makes, without making use of the specific questions. How I instead assemble these elements into an evaluative framework is what I show in the next section.

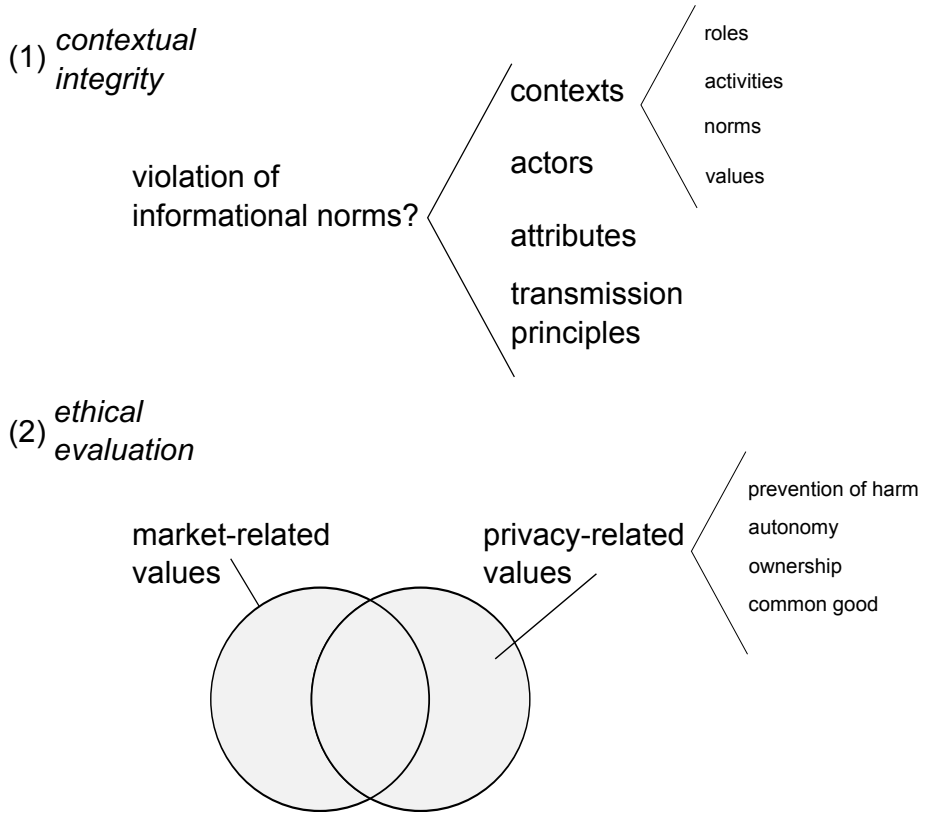

Figure 3.1: Overview over the two steps to be taken in the evaluation framework. The first part of the framework is based on the contextual integrity framework devised by Nissenbaum (2010).

3.2 Evaluation framework used in this study

As the discussion on contextual integrity above has shown, the framework of contextual integrity is strong in the structural components that it adds to an analysis. It offers guidance by conceptually decomposing the relevant aspects of an information practice and puts an emphasis on the informational norms governing a system. In case of new practices for which informational norms have not been established yet, the focus shifts to the values that are of relevance in the given context. This characteristic in turn connects contextual integrity to an ethical evaluation of a more general kind. For the purposes of this study, I therefore combine the structural analysis of contextual integrity with a more general ethical evaluation (Van de Poel & Royakkers, 2011, pp. 145).

Figure 3.1 shows a graphical overview over the evaluation framework. The first step consists of the prima facie assessment devised by Nissenbaum (2010) and leads to a conclusion whether informational norms are violated. In this regard, it has to be assessed whether the context in which information is shared, the actors involved therein, the attributes of the transmitted data and the principles according to which it is transmitted have changed through the introduction of a new information practice. If this is the case, or if informational norms for the context in question have not developed yet[2], there is a prima facie violation of contextual integrity. A prima facie violation does not automatically mean that one should reject the new information practice, but only says that it does not adhere to contextual norms and will thereby likely lead to a discussion on the grounds of privacy concerns.

[2]Nissenbaum mentions this point as part of her augmented contextual integrity decision heuristic (2010, p. 182).

The second step of the evaluation framework is an ethical evaluation. It entails an investigation into the effects of data markets on the values that are of relevance in this regard. For this study, these are the values that are of importance for the trade in markets in general terms and values that concern the moral reasons for the protection of privacy. The latter are of importance because of the central question concerning the effects of data markets on privacy. The former are relevant in order to evaluate the more general societal consequences of data markets. As I have have not discussed yet which values are of relevance for markets in general terms, I will do so before performing the actual evaluation in Chapter 5.

Important to realize is in which way the two components of the evaluation framework are connected. For the purposes of this study I include the component of contextual integrity for its strength in a structural analysis. Instead of looking at the complex situation of a data market at once, the structural analysis of contextual integrity allows to disassemble the information practice into its components and evaluate them one by one. In this regard, I do not include contextual integrity as a single definition of privacy, but for its structural guidance. This understanding of the structure of the problem then helps for the more general evaluation in the second step. Only then is it when the actual evaluation of the effects of data markets takes place. As I concluded in Section 2.6, it is the moral reasons for the protection of data that are of relevance, whereas the exact definition of privacy as such is of secondary importance. By using contextual integrity for its structural strength combined with a more general ethical evaluation based on the moral reasons for the protection of data, I thereby do not limit myself to any specific conception of privacy but stay broad in this regard.

The evaluation framework as described in this section is tailored towards the purposes of this study. However, it is not that specific that it would be completely useless for other evaluations of privacy. By replacing the market-related values with other values that are of relevance for a given problem, it could also be used to evaluate the effects of another problem in which questions of privacy seem to be an issue.

3.3 Methodology of this study

The framework as I introduced it above provides a meaningful way to guide the evaluation of data markets without being restricted to one narrow conception of privacy. It provides conceptual guidance to structure an evaluation and is also generic enough to not disregard any conception of privacy. Through the incorporation of more generally relevant moral reasons for the protection of privacy, it also includes pointers to existing considerations concerning the use of personal data. This aspect makes an evaluation more tangible and allows it to build on the knowledge derived from existing discussions concerning questionable information practices.

As I pointed out earlier, the general character of this study is of descriptive and evaluative form, mainly looking from the field of ethics. This is what the evaluation framework provides. At points I will also refer to considerations from the fields of law, economics and behavioural economics. These are, however, meant to inform the analytic character of this study and not the central perspective of the study itself.

In the next chapter, I discuss the origins of the discussion concerning data markets as a means to privacy protection and introduce several proposals that are discussed in the literature. Out of these proposals I will select those that are on the one hand specific enough to be scrutinized in terms of the ethical evaluation and on the other hand involve the online service user to a sufficient degree. As mentioned earlier, this study does not concern secondary data markets that do not allow the user to participate in them. I will then evaluate the data market approaches identified in this way in Chapter 5. In summarized form, the overall methodology of this study can then be represented as shown in Figure 3.2.

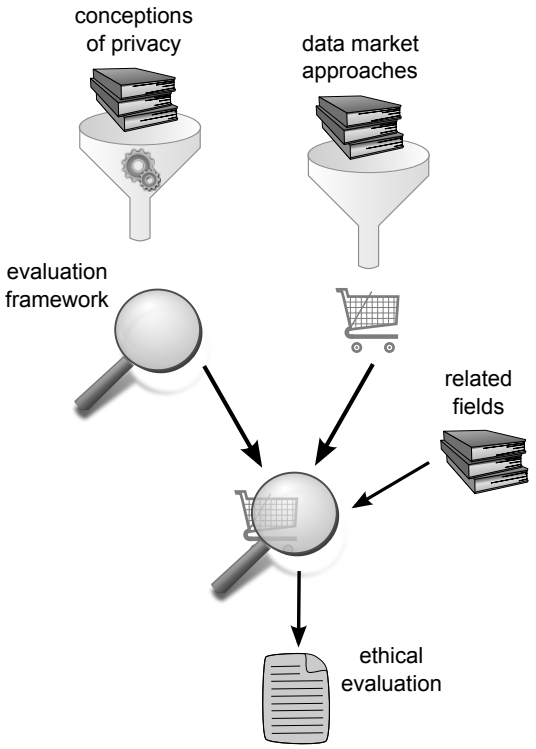

Figure 3.2: Graphical representation of the research methodology applied in this study.

4 Data markets

There is an increasing discussion concerning proposals to enable a trade with personal data, either by assigning property rights or other means to facilitate market structures. The central question of this study is in which ways data markets would impact the the privacy of online service users. In order to evaluate it, an understanding of what constitutes a data market is necessary. The goal of this chapter is to discuss proposals of data markets and identify a number of approaches that are used for the subsequent evaluation.

Most ideal in order to evaluate would of course be data markets that already exist in practice. Only then the complete functioning of such a market can be fully assessed. Every proposal of a data market in the literature will necessarily be incomplete to some extent, meaning that some parameters are left open, more specific implementation details are not described, or it is unclear in which context such a data market would be instantiated. However, for the purposes of this study also descriptions in the literature can be useful if they exhibit a high enough level of detail. Furthermore, it is also the differences between different data market approaches that useful information regarding their impact on privacy can be derived from. For the ethical evaluation that follows, data market approaches proposed in the literature are therefore also sufficient if the level of detail at which they are discussed is high enough.

The remainder of this chapter is structured as follows. I first give some background on the increasing debate about data markets as a means to privacy regulation (Section 4.1). In Section 4.2, I describe the practices around the trade of personal data that already developed. After that, I go into a small parenthesis concerning economic considerations regarding privacy (Section 4.3). Then follows the main purpose of this chapter, namely introducing the data market approaches that can be found in the literature. After giving a more general overview over the literature (Section 4.4), I describe these approaches that I selected for further evaluation in more detail (Section 4.5).

4.1 Origins of the debate

The proposal to create markets for personal data and the link to privacy protection are not new. Already Warren and Brandeis (1890) discussed whether privacy is a form of property right and and Westin (1967, pp. 324) argued that property rights are the best way to achieve privacy in the form of *privacy as control*. In that regard one cannot speak of a recent debate. More recently, however, the emergence of secondary data markets (see Section 1.1) and the increasing recognition of the economic value of personal data (see e.g. World Economic Forum, 2011) gave the discussion new importance. Also new is that the debate which mainly took place in the United States so far swaps over to Europe as well[1].

In general, there are two kinds of origins for the debate on data markets. One is the use of property rights as a regulatory constraint to protect privacy and the other are economic considerations concerning efficiency and the value of data. The legal considerations concerning property rights are mainly of relevance in the discussion in the United States because of the limited privacy regulations that are currently in place (see Section 2.5). As Purtova explains, there a number of arguments in the U.S. debate: property as a means to achieve some form of privacy regulation at all, property as more efficient than current tort law, property as a more general regulatory system and

[1] see e.g. the proposal of Novotny and Spiekermann (2013) and the discussion in the work of Purtova (2011).

even arguments concerning political lock-in at the U.S. federal level (2011, pp. 126). Although these might be important considerations for U.S. policy, these points can be neglected for a discussion of the EU situation. The strong and general protection established through the Data Protection Directive (European Parliament, 1995) already provides within the EU what these arguments are aiming for. Also important to consider in this regard is the preference for self-regulation within the United States (see e.g. Gutwirth & Hert, 2008).

Economic considerations regarding the efficiency of markets are a second origin of the data market debate. As I will explain in Section 4.3 in more detail, there a number of considerations when it comes to an economic point of view on privacy. The Chicago School, with Posner (1981) as one of its most important scholars, considers privacy as a market inefficiency and therefore undesirable. On the contrary, other scholars emphasize the public good nature of privacy and the social costs incurred due to externalities (see e.g. Sholtz, 2001). Also, information asymmetries between service providers and users can lead to a lemon market situation (see e.g. Vila, Greenstadt, & Molnar, 2003). Moreover, it is important to consider the role of privacy regulation in the abilities for price discrimination (see e.g. Odlyzko, 2003). Finally, the desire for privacy protection can also be seen as a way to turn privacy itself into a tradeable good (see e.g. Rust et al., 2002). All these considerations emphasize the economic importance of regulating privacy. The assignment of property rights could thereby create economic certainty, whereby details as the assignment of default entitlements and limits on transferability would still have to be decided on.

A different reason stemming from an economic perspective is the rise of the information economy. Secondary markets for the trade of personal data are growing (Cave et al., 2011), the World Economic Forum (2011) considers data to be an 'asset class' and France discussed to introduce a tax on data concerning their citizens (Pfanner, 2013). The overall importance of data as a driver of the information economy simply cannot be ignored, which makes the question of property rights in this economic resource of relevance. Lanier (2013) even argues that the economy will shrink if we stick to the idea of 'free information' and do not account for the value of information economically. Moreover, service users might develop a feeling of property in information concerning them and therefore would defend their rights better if property rights were assigned (Lessig, 2006). All these points stress the importance of data in the information economy and suggest to apply market principles to data more strongly. Even if there are no formal property rights to data within the EU, one cannot simply ignore the trend of 'commodification' (Purtova, 2011, p. 32) of personal data.

4.2 Existing data market practices

As I outlined above, there is already a development towards the commodification of personal data. Regardless of the property rights status in a legal sense, the trade with personal data and the economic relevance thereof are undeniable. In this section I want to give some background on these practices that already take place. These practices are not necessarily data markets in the strict sense, but are related and therefore of influence for the discussion. The goal of this section is not yet to make a selection of data market approaches, but rather to give some background on existing practices and to distinguish them later from the data market proposals discussed in the literature.

An important distinction is that there are two different types of data markets. On the one hand there are *secondary markets* (also called data brokers or aggregators). They buy personal data or more generic profile information from service providers, insurance companies, telecom operators, and other businesses. The user may indirectly benefit from such practices through reduced service costs or even 'free services', but receives no direct remuneration for the sale of their data. As I pointed out in in Section 1, I distinguish these from data market approaches in which the user is directly involved and receives an explicit remuneration for their data. For the sake of simplicity, I

refer to the latter as *data markets*.

secondary markets I cannot provide an in-depth overview over secondary markets at this point, but instead refer to the descriptions already made in Section 1.1 and the literature[2]. Instead, I want to point out some key characteristics of secondary markets.

A first point of concern is that secondary markets are to a large extent invisible. As a user of a 'free' or data-subsidized service one does not get insight into the practices concerning ones data. The literature discusses some of the bigger actors involved (see e.g. Craig & Ludloff, 2011), but information on the extent of trade and pricing are difficult to find. What from a user perspective might be especially important to know is the amount of subsidy that the service provider includes in their service costs this way.

Another problem concerning the opaque practices of secondary markets and the value of personal data is the valuation of firms. An interesting example in this regard is the micro-blogging service Twitter. Although the company makes only little revenue and is estimated to not be profitable until 2015, it has been valued at the stock market at a price of converted $78 per user (Eavis, 2013). While Twitter might be an extreme example in this regard, the same effect is also visible with other firms who offer 'free services' online, such as Facebook. This shows that there is a value in these data-intense practices, but it is unclear how to economically value the firms behind it. The prevailing explanation for this effect is that the stock market valuation accounts for the expected future profits (see e.g Eavis, 2013). Another explanation, however, is offered by Lanier (2013) who claims that effectively the value of data is taken out of the economy at large and that the economy will shrink with increasingly data-driven services.

Lanier (2011, 2013) also provides an important insight concerning the question who the actual customers of online services are. From an economic point of view, the users are not the customers of online service providers. The real customers instead are those financing the service, which in most cases are advertisers using the online service as a platform for advertising. This is also what the famous quote "When something online is free, you're not the customer, you're the product."[3] expresses. The important insight in this point of view lies in the effect that happens when competing for customers. Companies offering completely disjoint services, such as Google Search[4] or Facebook[5], do not have to compete for their users, but for their advertising customers. These effects are captured in the economic theory of multi-sided markets that I discuss in Section 4.3.

A last point to mention here is the form in which personal data is traded. While it is possible that complete datasets concerning specific individuals are traded, more indirect forms also exist. It is, for example, possible that only generalized profiles of users in anonymous form are being sold. The least privacy-invading form is where a service provider selects the users to which are advertisements are shown themself, thereby not providing any user data to third parties from which monetary compensation is obtained[6]. Yet another form is real-time bidding for advertisements, where advertisers get access to the user after bidding for it in an auction (Olejnik, Tran, & Castelluccia, 2014). Concerning these distinctions, there also exist large differences between the legal domains of the United States and the European Union. While in the U.S. firms are generally allowed to sell complete customer profiles (Solove, 2008), the EU Data Protection Directive only allows anonymised data sets to be given to third parties (Article 29 Working Party, 2014a) without taking additional measures.

[2] See e.g. the work of Purtova (2011, pp. 31), Craig and Ludloff (2011), and the World Economic Forum (2011).

[3] As explained on `http://blogs.law.harvard.edu/futureoftheinternet/2012/03/21/meme-patrol-when-something-online-is-free-youre-not-the-customer-youre-the-product/` [accessed August 17, 2014], the original author of this quote is unknown.

[4] `http://www.google.com/`

[5] `http://www.facebook.com`

[6] This form seems to be employed by Facebook (Chin & Klinefelter, 2012)

data markets As explained above, the focus of this study is on data markets in which the users are involved in a more direct form, meaning to receive remuneration for the sales of their data. First attempts to establish such data markets in practice have been made, but are still in a very early stage.

A company that made big news in the context of data markets is DataCoup[7]. Users of online services can register with DataCoup and allow DataCoup to retrieve information from their social media accounts and credit card transaction stream. DataCoup analyses the data and sells anonymised profile information to third parties. In reward, the users retrieve a monthly payment of $8 (Simonite, 2014). Noteworthy about this approach is that DataCoup generates revenue from the user data on top of the service providers where it was generated. This means that for example Facebook, DataCoup and the users financially profit from this form of collaboration. It remains to be seen whether the service providers where the data originates from accept this practice. Important to realize is that this type of service does not give the users any additional control over the way how service providers handle their data, since the only point where they obtain additional control is DataCoup.

GoodData[8] is another newcomer in this field, but has chosen a technically different approach. Instead of obtaining data from service providers, they ask users to install a web-browser add-on on their computer. This add-on blocks the tracking by service providers and instead captures data streams itself. The user can then choose to provide some of the captured data to GoodData, which analyses it and sells it in anonymised form to data brokers. The profits made from this are not given to the users who contributed their data, but are donated to other nonprofits working for social goods (Schiller, 2014). In this way the users achieve partial control over their data and can choose to donate the monetary value of their data. However, GoodData does not allow its users to receive remuneration for their data themselves.

Another service that made the news were the plans of Reputation.com[9] to launch a product called 'data vault'. The company, which is known for its success in online reputation management for paying customers, announced end of 2012 to start a new product which would undermine the practices of secondary markets. They planned to collect information about users and allow users to share that information in a controlled manner with marketeers seeking access to such data. The hope of their model was that it would be more attractive for marketeers to gain information about consumers in this way than through the data that is involuntarily collected by online service providers. The users of their service would not receive direct remuneration, but other benefits such as coupons (Singer, 2012). The interesting aspect about this approach is that it aims to undermine the success of secondary markets through this way of involvement. However, Reputation.com seems to have discontinued their plans[10]

Plans to establish a more direct form of a data market were pursued by the Belgian entrepreneur Bruno Segers. He co-founded the company IrisPact which should have served as a middlemen between users who are willing to sell data and companies purchasing that information (Soenens, 2013). The service was supposed to be especially privacy-friendly by using asymmetric encryption technologies to restrict the parties having access to data (Blyaert, 2013). Although Segers was successful in raising funds for this endeavour, the company recently stopped its plans because of internal disputes (Snoeck, 2014).

The examples of the existing attempts to establish data markets show that the concept of data markets is tried out in practice, but cannot be easily implemented. Reputation.com and IrisPact

[7] http://www.datacoup.com

[8] https://thegooddata.org/

[9] http://www.reputation.com/

[10] Reputation.com announced to start the service beginning of 2013 (Singer, 2012), but does not list the service on its website. A web search performed on 14 June 2014 only reveals articles announcing the plans to launch the service around that time, but no information that the service actually has been launched.

seem to have failed in realizing their plans. DataCoup and GoodData actually started operation, but are not data markets in the sense that users receive direct remuneration for the profits that are made with their data. In the case of DataCoup, the online service providers continue their operations as before and thereby become competitors of DataCoup on the secondary market. By using GoodData, the user does not receive remuneration, but instead donates profits made from their data. An example of a full data market that started operations does not exist, but the attempts above show that there is some activity leading in this direction.

4.3 *Parenthesis:* the economics of privacy

Although the perspective of economics is not the one I am taking up within this study, it is instructive to understand the characteristics of information and privacy from an economical perspective as well. Many data market proposals include a discussion how they deem to address various economical challenges and it is therefore helpful to have an understanding of these. In this section I want to give a brief introduction into the economics of privacy. I do so, by first discussing the characteristics of privacy in the economical sense. Afterwards, I give a brief introduction into the problems with price-setting of information goods and the theory of multi-sided markets.

characteristics of privacy An early characterisation of privacy from an economical perspective can be found in the works of the Chicago School. Posner (1981) provides a leading contribution in this regard and characterises privacy as an economic inefficiency. He argues that privacy is the equivalent of withholding information from other participants in the marketplace and thereby makes markets inefficient. Against this view argue Hermalin and Katz (2006), who claim that the ex-ante efficiency, market imperfections and intermediate levels of information are disregarded in the argument of the Chicago School. With these factors taken into account it is not self-evident any longer whether privacy does indeed constitute a market inefficiency.

market inefficiency

A very different characterisation of privacy is connected to negative externalities and the social costs they generate. Sholtz (2001) argues that companies who collect large amounts of information about their customers and utilize it for economic benefits, externalize the costs that the use of this information might incur to others. Problems around privacy and information misuse therefore are structurally similar to the problem of social costs. In order to address this problem, he assesses the options of direct regulation, a Pigouvian tax and a reallocation of property rights. He argues for the third option of a property rights reallocation towards customers. This argument, however, is questioned by Hermalin and Katz (2006) who use economic models to assess the effects of such a reallocation. A more general argument against the effectiveness of a simple property rights reallocation is given by Noam, who argues that possible differences in privacy valuation between customers and firms have to be taken into account (Noam in Acquisti, 2010).

social cost

Connected to the social cost argument of privacy violation is the consideration of personal data as a public good. Regan (2002) argues that an overly use of personal data leads to a decrease in trust. Personal data thereby degrades if too many tap into it. Consequently, she argues that the privacy protection is tragedy of the commons problem. Schwartz (2004, p. 2085) contends that not only personal data, but also privacy as such is a public good. He argues that privacy offers the possibility for anonymous interactions and that users who provide information about themselves can still benefit from the overall high level of anonymity. Privacy seen this way is therefore a nonexclusive and nonrivalrous good.

public good

A different characterisation of privacy related to trust is that of a lemons market. Vila et al. (2003) argue that users have high costs to read privacy policies and to check whether service providers violate their privacy rights. Since service providers can economically benefit from the use of information, they have an incentive to defect. Thereby, a market for lemons originates. Vila

market for lemons

et al. model this situation as firms adapting their behaviour depending on the number of customers that actually check how their information is being used. Based on this model, he arrives at the conclusion that in theory there is a Nash-equilibrium, but that overshooting effects can make it difficult to reach it. Regardless of the specific model and its predictions, the recognition of a lemon market structure as such is of importance as also other authors such as Schwartz (2004, p. 2081) recognize it.

price discrimination

Another characteristic connected to the the effects of privacy on economics is that of price discrimination. Since the availability of information about customers is what makes price discrimination possible in the first place, privacy can be seen as a limiting factor for it. Odlyzko (2003) investigates this relation and argues that the desire of companies to perform price discrimination will lead to a decline of privacy. Since customers usually do not like price discrimination, he argues, it will have to take place in concealed forms such as loyalty programs. Indeed, the trend towards a stronger use of such programs and a thereby hidden introduction of price discrimination can be seen in practice (Vanderlippe, 2005). Interesting with regard to the relation between privacy and price discrimination is that the direct identification of a customer and their purchasing habits allows for first degree price discrimination, while the collection of anonymous data can still contribute to third degree price discrimination. The issues around nondistributive group-profiling (see Section 2.3) thereby also become relevant in this context. The characterisation of privacy as a protection of customers against price-discrimination is therefore a more intricate relation, in which the possibility of third-degree price discrimination based on group profiles should be taken into account as well.

privacy as a good

A last characterisation that I want to mention here is the recognition of the fact that the desire for privacy can be interpreted as the option to sell privacy as such as a good. For example, Rust et al. (2002) build a simple economic model on the insight that customers like to have their privacy to be protected and the idea that firms can sell privacy-respecting versions of their products for a premium. Besides the actual model and the maybe too simplistic assumptions that they make, it is the characterisation of privacy as a tradeable good that is of interest here. Based on the tradeoff between the use of information and the protection of privacy, one can therefore consider not only information but also privacy as a good.

difficulties with price-setting A general introduction to the difficulties with price-setting of information goods can be found in the book 'Information Rules' by (Shapiro & Varian, 1999). They explain that for the production of information the sunk costs are dominant, whereas marginal costs for copying and distribution are almost negligible. The situation of multiple competitors supplying the same good to each a share of the total customer base thereby becomes unlikely, meaning that the textbook model of competitive markets is not applicable in the case of information goods. Producers therefore have two main strategies. As a way out, firms can try to be the dominant firm supplying the whole market, or try to differentiate their product. The case for information goods supplied by a firm to multiple customers is thereby different from standard economics.

The situation becomes even more complicated when considering markets for personal data. While in a situation like a movie that is sold on a regular market for consumer goods the group of possible customers is relatively well-known, the possible purchasers of personal data sold on a data market are more diverse. This leads to problems for individuals who want to set a fixed price for their data on a general basis. As Schwartz (2004, p. 2091) discusses, setting a price for data with the possibility of unspecified secondary use in mind is difficult. One can, however, imagine that price-setting becomes easier if it concerns the use of personal data in specific contexts. Similar to real-time bidding for the placement of advertisements (and thereby also access to personal data) that already happens in practice (Olejnik et al., 2014), an auction-based model for the access to personal data could be used.

Another problem that would arise in the case of data markets are the different valuations for privacy that individuals have. As discussed by Aperjis and Huberman (2012), the differences in valuations of privacy could lead to biased samples when personal data is bought on a market. Users who highly value their privacy are likely to demand a higher price for their data and will therefore likely be excluded if a sample of personal data is bought on the market depending on price. Aperjis and Huberman propose to solve this problem through the introduction of an intermediary who bundles the data of a group of individuals with different preferences for privacy-valuation. Through using these intermediaries, other firms could then by a sample of data which is not biased in this regard.

multi-sided markets An issue not directly related to data markets as such, but of high relevance for the status quo of free services is that of multi-sided markets. As Lanier (2013) extensively describes in his book 'Who owns the Future?', the users of service-providers like Google are different from their customers. In the case of Google, for example, it is mainly advertisers who are the customers that Google serves. The fact that two or more different groups are combined through a common market system is not completely new. Newspapers that are partially financed through advertisements, for example, are also a two-sided market. The prevalent model of 'free services' in the online context, however, increases the relevance of economic theories in this regard.

According to Athey (2014, pp. 20) there are two key features that multi-sided markets exhibit. One is the fact that in most of these markets the different groups that participate in it are distinct from one another and are not directly interested in the prices that the other group is charged. The reader of a newspaper, for example, will mostly be interested how much they have to pay for the newspaper and not in the price that an advertiser is being charged. The second feature of a multi-sided market is that the different sides of the market are connected through indirect network effects. The advertiser in a newspaper, for example, will be interested in the number of readers their advertisement is being presented to. The interesting aspect of multi-sided markets therefore is that the different sides are coupled to one another, but not directly through price.

An interesting characteristics of multi-sided markets that are mentioned by Athey (2014, pp. 21) is that behaviour on one side of the market influences the competition on the other sides of the market as well. This has important consequences in the case of markets in which one side of the market is single-homing (i.e. choosing a single platform to engage in), but the other is multi-homing (i.e. choosing multiple platforms to engage in). In this case, there is an incentive for platform providers to gain as much market power as possible on the single-homing side, so that higher prices can be charged to the multi-homing side. An example of this might be that users of social networks mainly use one social network they engage in, i.e. are single-homing. Social network providers such as Facebook are therefore likely to be in a monopoly position for a specific type of social network. Even though advertisers are multi-homing, i.e. place advertisements through multiple channels, a social network provider can exploit their monopoly position in the single-homing side and charge high prices on the multi-homing side. This creates incentives to gain a monopoly position on the single-homing side and explains why it is common to give away the good 'for free' on this side of the market.

4.4 Data markets in the literature

In contrast to the only more recent attempts to establish data markets in practice, numerous proposals for data markets can be found in the literature. As explained in Section 4.1, the academic debate on data markets stems from legal and economic considerations. Correspondingly, there are a number of proposals that focus on specific aspects from one of these two perspectives, but are not detailed enough to point out how such a data market could be established in practice. In this

section, I provide an overview over the literature and identify those approaches that are of interest to be further analysed in this study. I begin with approaches that meet the criteria of this study the least and then work towards the approaches that I consider for further evaluation.

property rights as a means for control One strand of proposals related to data markets can be found in the field of law and discusses to what extent property rights would help to give individuals more control over their data. However, property rights and control over data can take different forms and do not automatically lead to data markets.

As explained in Section 1.4, there are different rights encompassed in the notion of property. Depending on how these are filled in, it is more or less likely that data markets would actually emerge. Harris in that regard considers property to form a scale from 'mere property' to 'full-blooded ownership' (Harris in Purtova, 2011, p. 82). In the case of full-blooded ownership and strong alienability rights, online services users would have maximum control over their data, but could also sell all of their rights on a data market. In the case of mere property, however, users are not able to alienate the property in their data and property rights take a more regulatory character.

An extensive discussion of property rights of this sort can be found in the work of Purtova (2011). She discusses the property rights debate in the United States and assesses whether property rights for personal data would be possible in the European Union as well. Her work is from a legal perspective and mainly focuses on the question if property rights could achieve the legal status quo more effectively. In that regard, she does not consider whether changes in the substance of law would be desirable from a normative perspective, but whether a different implementation of the Data Protection Directive and existing case law could achieve the same ends through more effective means.

EU property rights

Purtova (2011) comes to the conclusion that property rights for personal data would be possible in the EU legal system even under the current regime of the Data Protection Directive. She emphasizes that the current focus on control and consent effectively creates a form of property right and that a formal recognition of property rights would only strengthen these instruments. Regardless of the Data Protection Directive, however, the European Convention of Human Rights (Council of Europe, 2010) and the corresponding jurisprudence by the European Court of Human Rights would set limits on the alienability of property rights. While it would be possible to give away usage rights in exchange for remuneration, it would, for example, not be possible to waive the right to consent. Purtova recommends to establish a property rights regime in order to better channel the rights and obligations that already exist through existing legislation. She proposes to establish a right to sell personal data in the form of usage rights, but to give users control over the details through licenses that are transferred along with the data in the form of sticky policies. In that way, one can say that Purtova is a proponent of data markets, but only in a form that better achieves what the legal status quo already tries to achieve.

EU private law

Other authors in the legal field discuss similar points, but come to slightly different conclusions. Cuijpers (2007) assesses whether the EU Data Protection Directive is mandatory law, meaning whether private parties can work around the obligations set down therein through contractual agreements. She arrives at the conclusion that a private law approach to data protection is possible and that this opens the way for contractual agreements concerning the protection of data. This interpretation of the law has, however, been challenged by Purtova (2011, pp. 199). Regardless of the legal feasibility in that regard, Cuijpers makes an interesting argument concerning the advantages of a private law approach to data protection. She argues that the current Data Protection Directive limits the economic possibilities of users in selling their data and that a private law approach would be desirable from an economic point of view. I that regard, Cuijpers argues for the establishment of data markets through contractual terms, but leaves open what the precise details thereof would be.

The work of Purtova (2011) also refers to a number of scholars who discuss the use of property rights as a regulatory constraint in the legal system of the United States. Among these are Schwartz who argues for a carefully tuned property rights regime to achieve a high level of privacy protection (Purtova, 2011, pp. 135) and Lessig who focusses more on the self-regulation mechanisms that a property rights regime might bring (Purtova, 2011, pp. 130). I discuss the work of these two authors later in this chapter.

Pigouvian tax to reduce externalities A very different proposal related to the potential of data markets can be found in the field of economics. Related to the view on privacy as a social cost (see Section 4.3), a Pigouvian tax can be considered an option to reduce the externalities that the use of data induces.

Sholtz (2001) provides an account for the view on privacy as a problem of social cost. He argues that information asymmetries on the usage of personal data create information asymmetries between users and service providers. This leads to too much disclosure of personal data and thereby creates negative externalities. One of the options to solve this problem is a Pigouvian tax, which would mean that the government introduces a tax which corrects the externalities that are present in the market. While Sholtz describes this option, he contends that such a measure is not relevant in the context of privacy. Instead, he argues that a reallocation of property rights would be more suitable to counterbalance the current market inefficiencies. However, he does not provide details how such a market would look like in more practical terms.

The proposal of a data tax in more general terms is also discussed by Lanier. He acknowledges the benefits of a tax as corrective measure, but provides two arguments against it. From his point of view, a tax would lead to a large bureaucracy that Americans are likely to disapprove. Also, a tax on the use of data might make it impossible for experiments like Wikipedia to get off the ground (Lanier, 2013, pp. 226). In a more recent article focussing on the European debate, he also mentions the possibility of a data tax as a solution to current privacy problems, but disfavours it on the grounds of too much influence by the state and possibilities for corruption (Lanier, 2014).

data tax

Despite these critical voices on a Pigouvian tax, the French government publicly considered idea in early 2013. On the argument that personal data are the raw material of the digital economy, a government report proposes to establish a tax on the collection of personal data. In that way companies like Google could be taxed for the revenues that they make based on the data that French citizens provide. The proposal primarily aims at a possible tax avoidance by service providers on an international level, but also refers to the effects the tax might have on the regulation of privacy (Pfanner, 2013).

Overall, a Pigouvian tax on personal data is only indirectly connected to data markets. Because it would have an influence on the data collection practices of service providers, it can be seen as an alternative to other forms of privacy regulation, among which are data markets. Also, if data markets would be in place, a Pigouvian tax would probably be rejected on the grounds that service providers already pay the users for the use of their data and would have to pay double through the tax. Moreover, it is interesting to consider how a data tax as a regulatory measure would shift the European conception of privacy. While privacy is currently considered an unalienable human right, a Pigouvian tax would convey the message that privacy infringements are acceptable as long as they are paid for[11].

towards real data markets The proposals focussing on property rights and a Pigouvian tax are connected to the debate on data markets, but do not emphasize the creation of data markets as such. Here, I discuss a number of proposals that emphasize the creation of data markets more

[11] A similar argument can be made about the question whether individuals should be able to completely waive their right to privacy in the case of data markets and is considered in Chapter 5.

clearly, but are still not specific enough on the point how these data markets would take form. I then discuss these approaches that I consider tp be specific enough in Section 4.5.

property rights reallocation

A number of proposals explicitly mention the goal to establish data markets that include a remuneration for users, but lack clarity how such a market would take form. As mentioned above, Sholtz (2001) argues in favour of a reallocation of property rights, so that a market for personal data emerges. However, he leaves it at economic considerations based on the Coase theorem and does not provide any details on the specifics of such markets.

full ownership

Another scholar mentioning data markets is Pentland (2009). In the context of an report on data mining, he discusses various options to balance the social value of data mining and privacy concerns by individuals. He proposes data markets as an option to incentivize users to provide their data for data mining purposes. He proposes to give full ownership rights to data, but at the same time install policies that encourage the use of anonymised data for the social good. Besides these initial pointers, he does not provide any more information what the specifics of such policies would be and how data markets itself would take shape.

economic right

A study conducted on behalf of the European Parliament (Cave et al., 2011) discusses the interplay between innovation and privacy protection. In this study, the authors argue that the human right to privacy becomes difficult to enforce when the economic value of data gains further importance. As one of their recommendations, they conclude that an economic right would allow individuals to better control the use of their data. They mention details about possible factors to determine prices on, but do not go into further details on the mechanisms that such a market would have (Cave et al., 2011, p. 97).

complemented property rights

Lessig (2006) is a prominent voice in the debate on Internet regulation and argues for the introduction of property rights in data. He suggests a combination of technological and legal measures to protect privacy. On the technology side, he proposes to set up an identity layer for the Internet that enables reliable identities, but also allows for pseudonymous identities. Furthermore, online services should offer machine-readable privacy-policies so that the user's computer can check these automatically and transaction-costs are thereby lowered. These measure should be complemented on the legal side through a property right for personal data. This property right should be limited in its waivability and supplemented by a legal implementation of the *Fair Information Practices* (U.S. Dep't. of Health, Education and Welfare, 1973). Lessig argues for the use of property rights, because it allows for the fact that users value privacy differently and would help users to engage stronger in defending their rights. The alternative of a liability rule, he argues, would be more suitable if transaction costs were high. This is, however, not the case if machine-readable policies are employed (Lessig, 2006, pp. 225). The proposal of Lessig is quite detailed in the way it encompasses different modes of regulation and also gives details on the form of property rights to be enacted. Still, it leaves open the question whether the property rights in this form would lead to a data market and, if so, in which form it would take shape.

The different approaches discussed so far all argue for the use of property rights or at least a more economic approach to privacy protection. While some only indicate that the trade with personal data would be possible, others also explicitly mention the emergence of a market for personal data. However, none provides a detailed description how such a data market would take shape. Therefore, there are too many open parameters and it is difficult to evaluate how these approaches would influence the protection of privacy.

The proposals made by Laudon (1996), Schwartz (2004), Novotny and Spiekermann (2013), and Lanier (2013) are more detailed in this regard. I discuss them along with a description of the effective status quo in the European Union in the next section.

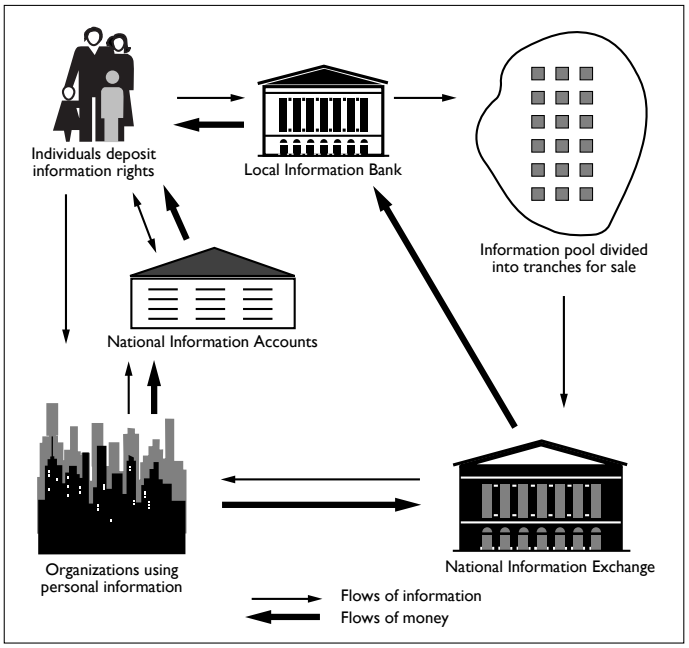

Figure 4.1: Elements and flows of a national information market. **Source:** Laudon, 1996.

4.5 Selected data market approaches

In the previous section I discussed a number of proposals that are close to that of a data market, but that I found not specific enough in order to evaluate the privacy implications thereof. In this section, I discuss four approaches that are more specific in the actual characteristics of the market structure that they propose. These are the *national information market* proposed by Laudon (1996), the proposal of *hybrid inalienability* by Schwartz (2004), the *personal information markets* envisioned by Novotny and Spiekermann (2013) and the more encompassing vision of a *humanistic information economy* by Lanier (2013). In order to have a point of reference for the evaluation following thereafter, I also include a description of the *free service status quo* in Europe interpreted as an implicit form of data market.

national information market A very early proposal for a data market has been made by Laudon (1996). Based on the relatively unregulated situation in the United States, he proposes a *national information market* coupled to a property rights regime for personal data. He grounds the need for such a solution in the upcoming developments of information technology and the failure of classic economic market theory in the case of information. The structure of this market is inspired by the institutions and customer relations in the financial sector, thereby drawing a parallel between financial products and data as an asset. This makes this approach highly centralized in its structure and requires that strong governmental regulation comes along with the institutions to be set up.

Figure 4.1 shows the different elements of a national information market that Laudon envisions. Each individual would have a *national information account* at which the transactions concerning their data are cleared. The data that they would like to sell is deposited at a *local information bank* or can be sold by organisations that run services that the individual uses. The information itself is traded at the *national information exchange* which acts like a stock market. An important detail about the proposed trade is that both organizations and information banks would sell data in *information baskets*, which are tranches of data combining the information of a larger number of

individuals.

The national information market would serve as a replacement of what I defined as secondary markets above. Service providers would still be allowed to collect information about their users as part of the services they offer. Only when the service providers would utilize this information in order to sell it to other organisations, the national information market comes into play. In that regard, the national information market is restricted to the secondary use of information and does not the limit the collection for primary purposes. Another noteworthy aspect described by Laudon is the involvement of the government in the trade with data. The government would be free to collect information about its citizens for its purposes, but would have to buy through the national information market if its wants to obtain information about citizens from service providers. It would also be allowed to sell information about citizens that it collected by itself. These restrictions might be relaxed for law enforcement purposes.

Interesting is how Laudon describes the need for national information markets when grounding it in the legal situation of the United States. He argues that the piecemeal approach of U.S. federal privacy legislation leaves important information systems unprotected, but also criticizes the Fair Information Practices (U.S. Dep't. of Health, Education and Welfare, 1973). He notes that the FIPs were conceived at a time when it was expected that each individual would know about all information systems that contain data about them, but that the widespread use of information technology makes it impossible to apply the FIPs in practice. He concludes that the U.S. legal situation gives individuals only a limited interest in their data and that a property interest would provide a much stronger form or protection. He also mentions that the FIPs fail to recognise the competitive pressures in market structures and the societal harm that information processing can generate (Laudon, 1996, pp. 96). These remarks are remarkable given the time of writing, but also of interest in the light of the European situation. Since the core principles of the Data Protection Directive are similar to that of the FIPs, it means that this critique might partially apply to the data protection regime applicable in the EU. Indeed, Purtova argues that the Data Protection Directive is functioning well in its substance, but fails in its procedures due to the increased number of data streams (Purtova, 2011, pp. 182). This makes it interesting to consider the proposal of a national data market also from a European perspective.

hybrid inalienability Schwartz (2004) proposes a different form of data markets that are not centrally controlled, but instead shall emerge by assigning the proper set of property rights to personal data. He designs the requirements that a property rights regime for personal data should have. He also takes the proposal made by Laudon (1996) into account and argues against the centralized character that the institutions therein would take. In that regard, Schwartz argues for a more distributed version of a data market.

Interesting about the property rights approach introduced by Schwartz is that it does not start from a blank slate, but takes critiques about property rights from the existing debate in the United States into account. He recognises that market failures are likely to emerge and that information asymmetries between users and service providers have to be avoided. Also, the bounded rationality effect has to be taken into account when users enter into contractual agreements with service providers. Furthermore, he acknowledges the public good nature of privacy in an economic sense and emphasizes the importance of privacy for democratic deliberation. He also discusses possible problems due to free alienability of property rights. In the case of free alienability, it would not be possible to limit the secondary use of data and price-setting mechanisms might not work due to the unclear scope how data is used. Finally, he discusses the concern that free alienabilty leads to an increased circulation of personal data and creates additional harm for individuals.

Based on these concerns, Schwartz proposes a model of *hybrid inalienability* (2004, pp. 2094) that forms the core of his contribution. The model evolves around five key characteristics:

1. *Inalienabilities*: Data may be used for primary purposes, but further transfers require the approval of the user. If the user agrees to that, the data may be used for specific secondary purposes only and these conditions are transferred along with the data ('use-transferability restriction'). One-shot alienability of the rights in data is not possible.
2. *Defaults*: The use-transferability restriction is coupled with an opt-in rule, meaning that the user has to receive information about further transfers and has to agree to the transfer.
3. *Right of Exit*: The user has the right to stop further collection through tracking technologies and can cancel their obligations with regard to data collection towards the service provider.
4. *Damages*: Damages are settled via liability rules, meaning that the state determines a fine that is imposed on the service provider. This has the advantage that it is possible to litigate in the case of small and hard to determine damages. It furthermore helps to litigate violations towards a collective without having to identify all individuals concerned.
5. *Institutions*: The property rights are supported through a set of institutions that help creating markets for data, to verify claims about property rights and to oversee the trade with personal data. I describe these institutions in further detail below.

A noteworthy aspect of the proposal by Schwartz are the exemptions on the applicability of the property rights model that he envisages. For the cases of law enforcement, the government access to data in more general terms and also the protection of free speech in the domain of journalism the property rights would not be applicable. Instead, separate rules that control access to data would be necessary (Schwartz, 2004, p. 2096). Furthermore, it is important to note that the use-transferability restriction only targets the secondary use of data. The collection of data for primary purposes would be less restricted and not covered by the data market.

Schwartz does not provide a detailed picture of the market structures that he envisions to emerge. In contrast to the proposals presented in Section 4.4, he does, however, go beyond the claim that property rights would automatically lead to the emergence of a data market. By describing some institutional characteristics as part of the hybrid inalienability model, he acknowledge the need for supporting institutions and provides some detail about these (Schwartz, 2004, p. 2110). With regard to the need for institutions that help to create markets, he disapproves of the national information market idea introduced by Laudon (1996). In contrast to the accounting made through national information accounts, Schwartz proposes to attach metadata about the usage rights and origin to personal data. The trade itself would then happen through different centres for information exchange, instead of one centralized one. With regard to the verification of claims, Schwartz proposes to also establish decentralized institutions that help users in exercising their claims. Finally, he proposes to install a data protection commission similar to the data protection agencies in Europe.

Overall, the proposal of hybrid inalienability lacks details of specific structures and mechanisms. However, it provides important detail with regard to the characteristics of property rights that should be assigned and the kinds of institutions that would be needed to support such rights. In that regard, it is worth to further consider this proposal in the subsequent evaluation.

personal information markets A very recent proposal for data markets has been made by Novotny and Spiekermann (2013). What sets this proposal apart from the other three mentioned in this section is that it does not stem from a United States context, but is explicitly tailored towards the European Union. Furthermore interesting about the proposal is that it addresses the current economic realities with free services financed through secondary markets for personal data. In that regard, it does not disregard the practice of secondary markets in favour of unrealistic regulations, but accounts for the economic realities that govern today's practices.

Novotny and Spiekermann motivate their proposal for *personal information markets* (PI markets) with the increasing gap between the goals of data protection regulation and the economic

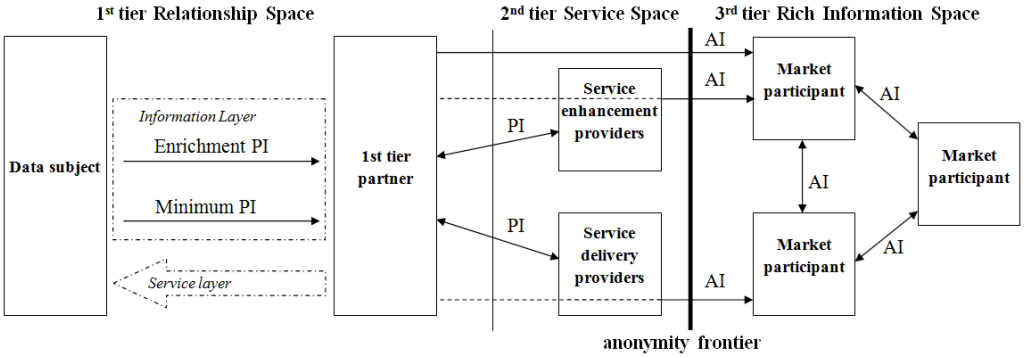

Figure 4.2: The three tiers of a Personal Information market. **Source:** Novotny & Spiekermann, 2013.

realities that govern the use of personal data. With PI markets they want to enable the trade of personal information, while keeping the control about it to users. To this end, they propose a three-tier model which separates between a direct service (first tier), outsourced functionalities that the service provider makes use of (second tier) and a a market for the trade of personal information (third tier). A set of rights and obligations channel the flow of data within these three tiers. Figure 4.2 shows an overview over the data flows in the tiers. I briefly describe each of the tiers in the following paragraphs.

The first tier, also called *relationship space*, covers the relationship between users and service providers. Important in this regard is that a service must be provided by a single clearly identifiable service provider, so that the user has an unambiguous point of contact with regard to his rights. Mash-up services in which it is unclear to the user what the legal responsibilities for each of the service providers are thereby become impossible. Moreover, service providers are required to split up the personal information about the user in minimally required and enriching information. Based on this distinction they have to to offer a privacy-friendly service which only makes use of the minimally required information. This allows to retain more personalized privacy-invading services, but offers the user the option to choose the bare minimum service. It could also be used to offer a free, advertisement-financed service along with a paid version. In either case the collection of personal information requires active consent by the user, which could be automated through the use of machine-readable privacy policies.

The second tier, also called *service space*, includes the service provider and auxiliary services that the service provider makes use of. Again these are split in those services that are necessary for operation (like webhosting) and those that are only necessary in the case of an enriched service (such as an advertisement provider). This allows to keep the flow of personal information to external providers as limited as possible. Furthermore, the additional providers are required to receive authorization from the originating service provider for operations that they perform on the personal information that they have been supplied with.

The third tier, also called *rich information space*, resembles the function that secondary markets for personal data currently have. It allows service providers to trade personal information independent of a specific service contract with each other. Everyone, including users themselves, has access to this market and can sell personal information through it. Consent of the user is not necessary, but information sold through it must have been anonymised in a state-of-the-art manner. Trade with personal information in non-anonymised form is generally prohibited.

Novotny and Spiekermann envision their proposal of PI markets to be realized through a number of technical and legal enablers, out of which the legal ones are of particular interest for this study. Most of the rights and obligations that the different partners of the three tiers would have already

exist within the current legal framework of the European Union. What would have to be added is the right for a privacy-friendly option of services. Furthermore, Novotny and Spiekermann suggest to create property rights in personal data in order to achieve that users have a stronger awareness for their rights. This property right would be unalienable, but allow for the transfer of usage rights in a limited manner.

The proposal of PI markets is of interest for this study because it provides a detailed market structure, including the rights and obligations that different partners in this market would have. It could be largely based on existing EU law and would connect the different parties through contractual relationships. Moreover, an important claim made by Novotny and Spiekermann is that the model of PI markets allows for the use of personal data without violating contextual integrity (2013, p. 1642).

humanistic information economy The proposal for a *humanistic information economy* propagated by Lanier (2013) is very ambitious and is rather a long-term vision than a specific policy-proposal. He argues that online service users have to be paid for their contributions they make in the form of personal data and contends that this is not only for the benefit of privacy, but also necessary to keep the economy at large sustainable. Lanier summarizes his plans in the following way: "[...] a way to conceive the project at hand is to imagine how computer networks could help create a fluid, incremental kind of wealth creation that thrives at a middle-class level and is not zero-sum." (2013, p. 231)

Lanier argues that the current model of 'free services' will shrink the economy over time. He makes the point that the practice of free information means that the value of digital goods is not part of the economy. Nevertheless, digital goods form an increasingly important part of human life and are as such valuable for society. Since in the future more and more of the things that we deem important will be of digital form, this effectively means that the economy will be shrinking. A more vivid to way to put this is as follows: If online services are mostly financed through advertisements and someone is willing to pay for the placement of these advertisements, this means that the users to whom the advertisements are being shown are willing to pay for the products offered. Since online services are financed only through advertisements, this suggests that the products offered in advertisements must for the most part be of non-digital form. However, there will be more and more products in the future that consist purely of information. The debate on file-sharing exemplifies the problems surrounding the economic value of digital goods and new practices such as 3D-printing will only aggravate this situation.

Based on this line of argument, Lanier explains that the current practice of free information will create financial problems for more and more people in the future. At present, it is mostly artists who have problems to make a living from their contribution to society, but in the future this problem will affect an increasingly large part of the population. With the current model, only a few artists at the top can financially sustain their activities. In order to get there, more unknown artists have to take the risks themselves. With more and more societal contributions being of digital form, this will become a problem for society at large. Lanier, therefore, argues that we need a system that sustains a middle class and not only the top.

Lanier furthermore explains that a similar concentration happens at the services that run the information economy. With the current model, it is mostly a few large companies (such as Google and Facebook) that provide information services. These have grown so big because of the network effects that the services they offer create. In the case of Facebook, for example, the service becomes more effective as more users join the service. This positive feedback loop that is present in online services leads to a concentration of power and wealth. Lanier calls the systems that run these services *siren servers*. He argues that siren servers also endanger individual agency and free will, since the Big Data approaches that they employ are only effective if they can predict human

behaviour. Therefore, siren servers constitute a concentration of power even beyond the economic realm. Lanier summarizes this process as "[...] feeding data intro Siren Servers, which lock people in by goading them into free-will-leeching feedback loops so that they become better represented by algorithms." (2013, p. 347)

In order to counteract these developments and create a more equal distribution of wealth and power, Lanier suggests to pay users for data that they contribute. This shall be in the form of nanopayments that are proportional to the degree of contribution and value that results from it (Lanier, 2013, p. 16). He suggests to create the technical infrastructure that would be necessary for such an endeavour according to ideas developed by Ted Nelson in the 1960s (Lanier, 2013, pp. 213). Central to these plans is that no information is copied, but always referred to with two-way links. This means that if B makes use of information that A created, B will create a link to the original copy provided by A. This link can be followed in both ways, so that also A knows who links to the information. Through such a system, it would then be possible to create an universal online market in which it is easily possible to find out who makes use of information that stems from third parties.

The prices in such a market system would be dynamically set through two different components (Lanier, 2013, pp. 259): An instant component represents the price agreed upon between buyer and seller. This component is therefore established through regular market mechanisms, which represent the current value of the data contribution. Additionally, there would be a legacy component which is backwards looking. The legacy component would not be negotiable, but government regulated. Lanier proposes to compose the legacy component out of three different elements: a traditional tax, an acknowledgement of upstream contributions and a correction for lock-in effects. The upstream contributions would be calculated based on the two-way links and make sure that those whose contributions are utilized are financially compensated. The lock-in correction would help to partially adjust for the moral hazard of lock-in situations due to digital services and is calculated according to the price difference to other technical options. It could not undo the technical lock-in effect itself, but would at least mitigate the price differences that occur because of it.

Since such a system would be a considerable change to the current situation, Lanier proposes a number of measures that would help to smoothen the transition. He proposes 'economic avatars' that would allow to pay for services in a different way than the service provider envisions (Lanier, 2013, pp. 268). Multiple users could create risk pools, that for example allow them to have free trial period for services that otherwise require immediate payment. The service provider receives the payment in their preferred form and the users pay in the form they wish. Additionally, Lanier envisions to uphold the option of 'fake free services' for users who are not yet convinced to participate in the new system (2013, p. 321). They would be able to consume services for free, but would also not be compensated for the contributions they make. To furthermore help users with such a system, Lanier proposes to establish optional services that bargain the prices for user contributions on behalf of the users (2013, p. 306). But even with these tools in place, there are many parameters that would have to be filled. Especially difficult would be to set an initial state representing the contributions that were made in the past (Lanier, 2013, p. 322).

Next to the advantages discussed above, Lanier also sees a role for such a system in the protection of privacy (2013, pp. 301). He proposes to support the system by inalienable commercial rights to personal data. This way, not only service providers, but also the government would have to pay for the use of personal data. Lanier goes even that far to propose that law enforcement has to pay for data they gather if does not lead to a conviction. In general, the cost of information would be an incentive to not make use of any information that is not necessary. Through the price-setting function individuals would also be able to control how much information about them is collected.

Overall, the proposal of a humanistic information economy made by (Lanier, 2013) is very ambitious and complex. He extensively discusses the idea in his book 'Who owns the future?'

(Lanier, 2013), but remains vague about many details how such a system would take shape. Moreover, the complex interrelation of the different aspects he envisions make it difficult to extract the key elements from his proposal. Despite these difficulties, the approach is of interest for the further discussion, since it represents the idea of a data market taken to a very extreme form.

free service status quo Yet another form of a data market one might consider is that of the status quo. The predominant deal to use online services free of charge in exchange of personal data represents a very specific form of a market in which users provide as much data as is needed to finance the service. Although this does not fully match the definition of a data market that I use throughout the thesis (see Chapter 1), I nevertheless stretch the definition a bit at this point. I do so for two reasons: First, in order to show the similarities that exist between a data market and the status quo. Second, in order to create a point of reference that can be used for the evaluation in Chapter 5.

I already gave an introduction to the existing practices of secondary markets in Sections 1.1 and 4.2. In summary, one can say that 'free services' are financed through the placement of advertisements and the sale of data in secondary markets. In the case of advertisement, personal data are not required for the placement of advertisements as such, but in order to better select the target audiences to which advertisements are displayed. This process of data collection and selection of advertisements can either be done through the service provider themself (such as in the case of Facebook) or through an intermediary who places the advertisements on behalf of the service provider (such as in the case of ad networks). The second option to generate revenues from personal data is to sell data as such on secondary markets. As explained above, this can happen in different granularities. While in some cases only aggregated or otherwise anonymised profiles are sold, it might also be the case that information about a specific customer is being sold.

With the current legal situation in the European Union, the Data Protection Directive is the most important legal norm concerning the collection of personal data[12]. As pointed out by Purtova (2011), the principles laid down in the Data Protection Directive are in substance still adequate for contemporary information processing practices. A problem, however, occurs through the high number of data flows that nowadays occur in practice. This makes it difficult for the user to assess which of the, sometimes even interwoven, service providers process their data. Likewise, it limits the oversight that data protection agencies can perform. In addition, the practices around secondary use of data lead to ambiguity regarding the application of the Data Protection Directive[13]. To be on the safe side, service providers therefore often ask consent from users and avoid the possible ambiguities of other legal grounds.

With the predominant use of consent as a legal ground for data processing and the requirement of consent for the placement of tracking cookies under the e-Privacy Directive (European Parliament, 2002), the user technically has a choice about the collection of their data. However, service providers often do not offer any meaningful choice[14] and simply require users to give consent if they wish to make use of their service at all. For example, the TU Delft website simply informs its users that it makes use of tracking cookies by Google[15], but does not offer an alternative website without cookies. A prospective student who wishes to inquire about the higher education offered by TU Delft, but does not want to transmit this information to Google, would have to use the telephone instead. Likewise, Facebook offers only type of privacy policy. A user not consenting to their terms, simply cannot make use of the service. Combined with the quasi-monopoly position

[12] In addition, the e-Privacy Directive (European Parliament, 2002) regulates the placement of cookies, which is of high relevance for the tracking technologies employed in the context of advertisement.

[13] See in this regard the opinions of the Article 29 Working Party on purpose limitation (2013b) and legitimate interest (2013b).

[14] See e.g. Novotny and Spiekermann (2013) and Schwartz (2004)

[15] see `http://www.tudelft.nl/nl/cookiebeleid` [accessed August 17, 2014]

of online services and the absence of paid alternative options (see e.g. Lanier, 2013), users do not have a meaningful choice in this regard.

As discussed above under the keyword of multi-sided markets (see Section 4.3), many providers of online services have an incentive to offer their services free of charge. If, as in the example of Facebook, users are single-homing with regard to the social network platform that they use, the service provider has a strong incentive to be monopolist on this side of the market. The framing of a service as a 'free service' is thereby attractive to users and helps in acquiring a stronger market position. The missing revenues from this side of the market can then be compensated through higher prices that are charged on the other sides of the market, which in the case of free services are mostly advertisement and the sale of data on secondary markets. The frame of a 'free service' is therefore no coincidence, but relates to the stronger position on both sides of the market that a service provider acquires by doing so.

In practice, this means that online services are predominantly offered as a 'free service' financed through the collection of personal data, without giving the user any choice about the collection principles. The user has a number of rights along the lines of privacy as control, but can hardly enforce these in practice. This leaves the user with the situation to either make use of a service financed through their data or opt out of digital services altogether. The bottom line of this situation is that the user sells their personal data for the price that they would otherwise have to pay if a paid option of services would exist.

5 Ethical evaluation

The goal of this study is to investigate the question in which way data markets would impact the privacy of online service users. As explained in Section 3.2, I investigate this question in two steps. First, I investigate the data market approaches using the framework of contextual integrity (Section 5.1). This investigation is useful for its structural qualities, thereby pinpointing to the features of data markets that are problematic with regard to privacy. Second, I perform a more general ethical evaluation concerning the values that are of relevance in the case of data markets.

The more general evaluation aims to highlight the problems that the selected data market approaches might create in practice. Before going into the evaluation itself, I first describe the relevant literature concerning the virtues of markets (Section 5.2). Then, I evaluate the influence of data markets regarding market-related values (Section 5.3) and privacy-related values (Section 5.4). I perform this evaluation for data markets in the general sense first and only afterwards link the results back to the specific data market approaches (Section 5.5).

5.1 Data markets seen through the lens of contextual integrity

The framework of contextual integrity is useful in two regards. First, it serves as a benchmark for many of the aspects that are commonly defined as privacy. Although I do not consider a single definition of privacy here, I agree to Nissenbaum that her framework is a good heuristic for the aspects that people find problematic about new information practices. Second, and even more important here, the conceptual decomposition of contexts and informational norms allows to uncover the structure of information practices and to point more specifically to problematic aspects of an information system causing concerns in this regard. Regardless of the question of a definition of privacy, it therefore is helpful to apply the framework in order to gain more insights into the structure of the problem. This is also the reason why contextual integrity is included as part of the overall evaluation framework used in this study.

preliminary considerations The concept of contextual integrity entails an assessment whether the informational norms for a given context are respected. The first question to be answered here is then what a context in the case of data markets is. As Nissenbaum (2014) pointed out, a definition of context in the sense of a business model brings the problem that it would be the business practice itself which defines the informational norms. Interpreted this way, it would be up to service providers to set the expectations for privacy that their users should have. Under this self-referential interpretation of a context, an advancement in privacy-protection would thereby become impossible. For these reasons, Nissenbaum argues that a context has to be interpreted as a social domain. She gives "[...] education, healthcare, politics, religion, family and home life, recreation, commerce, friendship, marketplace, work and more" (Nissenbaum, 2014, p. 23) as examples in this regard.

The context of relevance in the case of data markets, is therefore not the data market as a new business practice, but the existing social context that it operates in. Out of the examples given above, the marketplace as a social domain seems to be a good fit at first sight. However, I argue that *the marketplace* in its traditional reading is different from a market that data markets would constitute. A traditional market is often limited to a specific category of goods and individuals who

engage in trade are not permanently present in these markets. They either enter the marketplace sporadically in the role of a consumer or more regularly as a merchant. In either case, however, the marketplace is a closed-off context which to a large extent does not overlap with other contexts such as education, home life or recreation. This traditional form of a marketplace is likely to be different from the type of market that a data market would constitute.

While a traditional marketplace is functionally separated from other contexts, a data market could develop into an overarching element which connects multiple contexts. If, for example, personal data concerning the free-time activities of a student are offered on a data market and purchased by their school, the data market is not a distinct context any longer, but acts as a bridge between multiple contexts. This effect is likely to be stronger if data that is generated through the provision of a service is offered on the market by automated means. In principle, however, a data market could still be a distinct context similar to traditional marketplaces. If an individual decides to trade a specific data set on a data market and consciously offers the data to a prospective buyer, the trade and the activities involved therein would resemble a traditional marketplace to a large extent. Therefore, it depends on the more specific circumstances whether a data market can be considered to be a context on its own.

In the case that a data market cannot be considered to be a distinct context, it is nevertheless of importance to take its influence on the respective context into account. Nissenbaum provides two insights that might be of relevance in this regard. The first is that the commercial exchange of information can constitute a transmission principle (Nissenbaum, 2010, p. 145). This means that market-related values may play a role not only in the form of contextual values, but also in the way how information ought to be transmitted. The other insight is that it is not necessarily a problem if multiple contexts overlap (Nissenbaum, 2010, pp. 136). It might well be the case that people act within multiple contexts at the same time and thereby perform actions that promote the respective values of both contexts. A problem arises only when the norms of two contexts are in conflict with another. Seen from this perspective, a data market could in principle connect multiple contexts without any problems, as long as the values and norms of the respective contexts stay intact. If a data market connects multiple contexts, the question therefore becomes how the data flow between the contexts and the market-related values influence the respective contexts.

Taking these preliminary remarks into account, I now describe the market approaches selected in Section 4.5 along the lines of the contextual integrity framework. I first argue that all of the five selected approaches would constitute an overarching element and therefore cannot be considered to be a distinct context on their own. After that, I discuss the informational flows in terms of actors, attributes and transmission principles.

context As discussed in Section 3.1, a context can be defined through its roles, activities, norms, and values. However, I argue here that the five selected data market approaches do not constitute a context on their own, but are instead an overarching element interconnecting multiple contexts. These contexts that thereby get connected are also not determined beforehand, but can be different ones. The data market would thereby become a generic bridge breaking up informational barriers between contexts.

In the case of the *national information market* (NIM) proposed by Laudon (1996), individuals could indeed manually sell information on the market. By depositing their personal information at a local information bank, they could actively provide information to be sold on the market. Considering only this component of the market, one could indeed argue that individuals are directly involved in the trade and the NIM thereby could constitute a context on its own. However, it is the local information bank which would sell the information in tranches on behalf of the individual. This means that the individual would only be indirectly participating, but nevertheless the one who

actively provides the information to be sold. Even more important, however, is that also service providers could gather information about individuals and sell them on the market. The individuals concerned would have to give consent and get reimbursed for it, but would have apart from that only little involvement in this form of trade. Since service providers gathering information in this way and other organisations buying the information on the market could be in different social domains, this would clearly give the market an overarching form. Taking the considerations regarding the role of banks and service providers into account, I therefore argue that the NIM as a whole would not constitute a context on its own.

The proposal of *hybrid inalienability* by Schwartz (2004) does not consider a direct involvement of the user in the data market. The user would be required to give consent to all secondary uses of their data, but would not be in direct interaction with the parties whom the data is sold to. The terms of consent and usage policies resulting from it would be the only form of interaction. Further details would depend upon the specific markets that shall arise based on the rights proposed by Schwartz. Since again, the service providers selling data for secondary use are not necessarily part of the same social domain in which the data is used, such a market system would form an overarching element.

The proposal of *personal information markets* (Novotny & Spiekermann, 2013) foresees two forms of involvement of the user in the market transactions. Here, users would make use of online services as it is the case right now. The market aspect only comes into play through the rich information space, which is a market for anonymised information. Since service providers as well as users can sell information on this market, it depends on the form of involvement. If users directly contribute information to be sold, this market could indeed form a social sphere on its own. In the case of information sold by service providers, however, it would constitute an overarching element.

The proposal of Lanier (2013) foresees to compensate users for all forms of digital contributions they make. This can be through more direct contributions in the form of online content, but also more indirect ones as secondary uses of data. The user would have the option to influence one component of the price to be paid, but could also delegate this task to some agent. The explicit involvement of users could thereby become limited. Furthermore, one of the central points in the proposal of the *humanistic information economy* is that every contribution of digital form should be accounted for in economic terms. This means that it would explicitly not be limited to a particular market, but almost by definition be a system spanning different social domains.

The status quo of free online services contains no form of involvement for users in the secondary markets for data that exist. Also, it is to large degree opaque how information is being traded on the secondary markets. The status quo therefore does not constitute a context on its own, but is a system spanning different social domains.

All five data market approaches can be considered to be an overarching element spanning across several contexts. In the case of personal information markets and the national information market, the formation of a specific context would be possible in theory if users would only engage in them through direct means and not through intermediaries. But even then it would depend on the type of information being traded there and how it would be used by the purchasing parties. With the indirect and mixed forms of interaction that all data market approaches provide, it therefore cannot be said that they would constitute a specific context on their own.

The contexts that are of relevance for an assessment of contextual integrity are therefore the ones that would be connected by the different forms of data markets. What kind of social domain that would be is left open by the proposals. I therefore cannot assess the four parameters (roles, activities, norms, and values) in more detail. Important to realize, however, is that the overlap between contexts that data markets constitute can lead to a clash of norms that are prevalent in

different contexts.

Since, an assessment of the specifics of contexts is not possible, I keep the remainder of the discussion limited to the characteristics of information flows that are associated with the different data market approaches.

actors The actors of an information flow are the sender, the receiver and the subjects covered by the information. Since it concerns markets for personal data, data sold will in every case involve an individual. Let me call this individual the *subject X*. Who X is depends slightly on the situation. In most cases, X will be the individual who is supposed to receive remuneration in direct or indirect form through the trade of data. X can also be another person if it concerns a group of people who are covered by the information. However, this is a special case which is not further considered in the following[1]. Another distinction to be made here concerns the anonymity of subject X. In case of anonymity, X is subject of the data, but cannot be identified from it by direct or indirect means. In case of personal information markets, this is always the case for data that is traded in the actual market component, which is the rich information space. For the other data market approaches, an anonymisation of subject X is possible, but cannot be determined upfront.

Another question concerns the sender of information. In case the user is directly involved in the market transaction, the sender is the same as data subject X. As explained above, this is, however, only the case for personal information markets and the national information market. And even with these two data market approaches the direct involvement is only one form of involvement next to the indirect form through service providers. With the indirect form of selling data through service providers, the service provider is the sender of information and thereby different from the subject X.

The last actor to consider is the receiver of information. Since it concerns a data market, the whole purpose is to find potential buyers of the data through the institution of the market. The receiver of information is therefore by design not specified beforehand, which means that an unknown third party will receive the information about subject X. There are two possible restrictions on this general observation. First, it could be possible to limit the data market to a specific group of participants and thereby limit the possible set of receivers. The analysed data market approaches, however, do not foresee this option explicitly. Second, the actor offering data on the market might decline to sell data to a potential buyer. In the case of direct involvement of the user, the user thereby would get the option to refuse selling data to undesired buyers.

In summary, there are two important distinctions to make here. In case of a direct involvement of the user, the sender is the same as the subject of the message. In this case, the subject also has control about the set of recipients. In the case of service providers selling the data, the subject and sender are different and control for the subject about the receivers is at least very limited. The second distinction to be made here is whether the subject of the message can be identified from it. Anonymisation is required for the approach of *personal information markets*, but optional in all other approaches.

attributes The attributes of transmitted information, or more technically the data fields, concern the type of information that is transmitted. Again, the point of a data market in this regard is to be generic, so that there are no clear upfront restrictions in this regard. The attributes could be a single one like a list of email addresses, or a more complex composition of data fields concerning the activities of individual data subjects. The buyer and seller interacting through the data market

[1] The special case in which a single individual tries to sell data concerning a group of people is an interesting case, since it is apparent that consent from the whole group should be obtained. How to organize this process and the distribution of remuneration is an interesting question, that I, however, do not further investigate in this study.

are likely to agree on the attributes before performing the transaction, but the data market itself is a 'neutral' infrastructure in this regard.

The only general observation that can be made in this regard concerns the case of anonymity. In case of anonymised data, it has to be ensured that the data subject cannot be identified within the dataset that is sold on the market. This somewhat limits the possible attributes upfront. Any unique identifiers like the full name of a person, but also combinations of attributes like zip-code, gender and date of birth are likely to uniquely identify a person (Koot, 2012). Such attributes are therefore likely not to appear as part of anonymised data sets. However, the ultimate distinction of whether a dataset is anonymous cannot be made based upon the attributes themselves, but has to be made upon the generality of information within the specific dataset[2].

transmission principles The transmission principles of informational norms concern the way in which information ought to be transmitted. As mentioned above, the characteristic of commercial exchange can constitute a transmission principle, but there also exist more general principles such as confidentiality or reciprocity. Some of the data market approaches provide details that give an implicit indication of transmission principles that are to be established. But also the mere fact that it concerns a commercial market, gives some general indication of the transmission principles that are likely to occur in practice.

One of the central elements of the *national information market* is the national information account. It is not only intended to be used for billing purposes, but also in order to give individual insight into the way how their personal information is used. The establishment of this institutional characteristic would bring with it that the transfer of data towards an organisation buying it requires the organisation to be transparent about the use towards the data subject. Another transmission principle that is mentioned by Laudon (1996) is that the purchase of an information basket explicitly gives the right to use the information for secondary purposes. It therefore would be an inherent characteristic of any data transfer through the market that the data is likely to be used in other contexts than the one which it originated from.

The 'use-transferability' restriction that is part of the *hybrid inalienability* model would constitute a clear transmission principle. The secondary purposes for which data may be used would have to be explicitly specified and transmitted in the form of a policy along with the data. This form of metadata would entail the consent of the data subject, but also explicate for which ends the data may be used.

In case of *personal information markets*, service providers would sell data in the rich information space without the consent of the data subject. This, however, is conditional on the anonymisation of the data. A prevalent transmission principle would therefore be that the data transmitted is decoupled from the data subjects that it originally concerned. This means that the data subject cannot be identified, but also might not even be aware of the transmission. The data is thereby meant to stands on its own and is not connected to any specific individual any more.

A very strong transmission principle in the case of the *humanistic information economy* is that of reciprocity. In this case, this not only concerns data transmissions that are closely connected to market exchange, but all data transmissions that take place. Through the use of two-way links and the requirement to reimburse all individuals for commercial activities that are made possible based on their data, there is a strong principle of reciprocity in each and every data transfer. Explicit consent for the use of data would not be necessary, but the data subject would have some influence through the possibility to demand a too high price.

With the status quo, most data transfers happen based on formal consent provided by the data subject. Due to the practical insufficiencies of the current data protection regime, a voluntary and informed consent, however, becomes questionable. This brings the uncertainty with it that data

[2] See (Article 29 Working Party, 2014a) for a good introduction to the requirements of anonymisation.

which service providers sell on secondary markets might not have been provided by individuals for these purposes.

Although not going into too many details, the descriptions of the data market approaches provided in the literature show differences with regard to reciprocity, the voluntary provision of data, the provision of consent and the awareness of an commercial exchange of data taking place. Important to consider is that the commercial exchange as such might introduce some further implicit transmission principles that are grounded in the market exchange of a good in the more general sense.

Since the buyer of data pays in order to receive the data, this implicitly brings expectations concerning the rightfulness to utilize the data for commercial purposes with it. Even in the case of a data subject who is not aware of the data transfer, the party buying the personal data will be under the impression that there is nothing wrong to commercially use the data. In this case, there would be a mismatch in transmission principles of the two subsequent data transfers between user-service provider and service provider-data buyer. This connects to the danger of a corruption of goods that I discuss in Section 5.2. The commercial exchange context of personal data changes the frame and constitutes an implicit form of communication. The actor purchasing the data might thereby act differently than as if they acquired the data free of charge. The commercial exchange character is a transmission principle that is likely to be apparent in all forms of data markets, but might have consequences that are not desired by all actors.

market metadata In the discussion so far, I focussed on the information flows concerning the actual information that is being sold on the market. However, it has to be taken into account that data markets would not only lead to a transmission of information as a good in itself, but would also generate additional information flows concerning the description of information that is available for trade and the data which is generated as part of transactions taking place. Such data would constitute a form of descriptive metadata that is necessary for the market itself to function. Since this form of *market metadata* is part of the information flows that originate from data markets and has potential implications for privacy, it should be included here as well. Because of the high number of information flows that market metadata would constitute and the lack of clarity of the proposals in this regard, I can only roughly outline how the specific data market approaches would add to an evaluation in terms of contextual integrity in this regard.

In case of the *national information market*, the national information account and the national information exchange are of particular importance. The national information account stores the metadata about all exchanges that already took place in the past and contains this data in a form that is coupled to the individual. Laudon (1996) envisions this element to be run by the government, which adds the government as an important actor receiving the metadata of all transactions. As Laudon envisions this element to lead to increased transparency towards individuals about the use of their personal information, it would also mean that probably more attributes than would be strictly necessary for billing purposes would be transmitted. The national information exchange only deals with baskets of information as a unit of sale, but transactions nevertheless have to include the identifiers of the individuals concerned for accounting purposes. This means that the institution running the national information exchange would have access to the information about the individuals concerned at least in pseudonymous form. Furthermore, it is important to realize that all potential buyers would have access to at least some of the metadata describing an information basket. This would not necessarily have to be the individuals contained in it. Nevertheless, this increases the number of actors receiving metadata drastically. While the metadata contained in the national information account probably would have a transmission principle of confidentiality, the metadata made accessible through the national information exchange are necessary for buyers

and would therefore be of a public character.

The potential privacy implications of such a centralized infrastructure have been recognized by Schwartz (2004), who envisions a more distributed infrastructure as basis for his proposal of *hybrid inalienability*. He suggests that multiple centres for information exchange would take the role of an intermediary bringing buyers and sellers into contact. All data being exchanged through markets would be accompanied by metadata containing the origin of the data, relevant privacy legislation and restrictions on further exchange. Important in this regard, however, is that this metadata would be transferred along with the information that is traded itself. From this perspective, it means that it constitutes additional attributes that are transmitted, but limits the set of recipients to those who also receive the traded information. Also, it should be noted that some of the attributes, such as restrictions on further exchange, form an important part of the transmission principles of the actual information being transferred (see above). Each center for information exchange and the potential buyers furthermore need access to some of the metadata that are relevant to initiate transaction. Overall, this strategy increases the number of actors, but avoids potential problems coupled to a single actor having the complete picture.

The actual market component of the *personal information markets* is the rich information space in which anonymised information would be traded. Novotny and Spiekermann (2013) do not provide much detail about the metadata that this component of their proposed system would generate, but it is apparent that some details about the interacting parties would be necessary in order to initiate a transaction and coordinate the billing process. Particularly interesting is the case in which individuals themselves participate in the market in order to sell anonymised information about themselves. If no further precautions are taken, the individual selling the information would run the risk to be identified through the metadata, although the information being sold has to be anonymous. Technical precautions such as the use of mix networks[3] for the transmission of data and payments could be employed to preserve anonymity also on a metadata level, but some aspects such as the mere fact that a specific set of data is being sold already by itself carries some information. An implementation of personal information markets could thereby carry over the transmission principle of anonymity to the metadata that are necessary in order to establish the market.

Interesting in terms of metadata is the proposal of the *humanistic information economy*. Through the use of two-way links and the inclusion of literally all information as part of a data market, Lanier (2013) effectively proposes a system in which metadata about the use of all information would be generated. Lanier is not explicit about the technical details of his vision, so that it is unclear which specific attributes such metadata would carry and who the recipients would be. Given the purpose of the two-way links as part of the legacy component of the proposed price-setting mechanism, however, it is clear that the metadata would have to be centrally accessible. Also, for the purpose of financial compensation, the parties contributing and using the information would have to be at least indirectly identifiable. This means that an implementation of the humanistic information economy would probably lead to a central repository of metadata containing a list of which information provided by A is used by B. And because of the very goal of the proposal, this would not be limited to information that is explicitly provided as part of a market, but all information that is generated. The instant component of the price mechanism furthermore requires that A and B can negotiate a price, which might make it also necessary for the two parties to be identifiable towards each other. The set of recipients of metadata would therefore not only include the central repository, but also anyone who wants to use information.

The *status quo of free services* is more frugal in terms of market metadata. Since the business

[3] Mix networks are a technical means to create anonymity in a network by combining the data streams of different users. The TOR project (`https://www.torproject.org/`) is probably the most prominent application of this principle in practice.

practices concerning secondary markets are very opaque, it is difficult to say which specific metadata are exchanged in practice. However, since individual users are not directly involved in the markets for secondary information and do not receive financial remuneration, they need at least in principle not to be identifiable as part of the metadata that accompany the exchange of data on secondary markets.

Although the selected data market approaches are not very specific on the market metadata that an implementation would generate, some general observations can be drawn. Most importantly, the mere existence of market metadata is noteworthy. It shows that data markets are not neutral in this regard and lead to the transfer of additional data that would not be necessary otherwise. The market as a system requires additional information in order to work and thereby creates a demand for this type of data. How this affects the privacy of users participating in the market depends on the specifics of the corresponding information flow. The mere existence of this data by itself, however, already constitutes a potential risk, as it might lead to harm if it is accessed in an undesired way.

The differences between the discussed data market approaches in this regard are also remarkable. While implementations of the *national information market* and the *humanistic information economy* would lead to the generation of high quantities of market metadata that are centrally accessible, the *personal information markets* and the proposal of *hybrid inalienability* are more scarce in terms of their data requirements and could also be implemented in a distributed form. A distributed implementation avoids the problem of a single point having complete knowledge, which can be an advantage when it comes to the potential risk of misuse of data.

The differences between the approaches also point to the different purposes of market metadata. It can be necessary for billing purposes, as information for potential buyers and for transparency towards the data subject. The last point is an interesting paradox: For privacy along the lines of informational self-determination transparency towards the data subject is an important requirement. This makes it necessary that information about the use of data is made available to the data subject. Depending how the transfer of this information flow is implemented, this can lead to additional privacy concerns as third parties might gain access to this information.

A similar paradox arises due to the availability of information that is provided towards potential buyers. Since potential buyers have to evaluate the data they are about to buy, a part of the market metadata have to be made available towards a wide public. Even if an individual is directly involved in the data market and might deny the transaction at a later point in time, the mere existence of data has to be widely announced at first. This paradox is structurally similar to that of the *information paradox of inventions*, where an inventor who wishes to sell an invention has to disclose a part of his invention in order to find potential buyers (Arrow, 1962). Besides the billing requirements and the more optional transparency functionality towards the data subject, this requirement of information is an unavoidable part of any market and will therefore exist no matter how the market is implemented.

conclusion The investigation into the informational norms of data markets shows that a number of parameters would be left open. From the perspective of a data subject, in many cases it would be unclear who the receiver of personal data is, what specific attributes about them are transmitted and which contexts would thereby get connected. Also, it might be partially left open which service providers are involved in sending the information and what the transmission principles constituting the transaction are. The degrees to which these parameters are left open vary per data market approach, but are to some extent also an inherent feature of data markets as such.

Concerning the sender of information, the proposal of personal information markets would leave it open which service providers engage in the trade of anonymised personal data. In all other cases,

the data subject has in principle knowledge about this, but might have difficulties to keep track of it in practice. Only in the case of direct provision of information to data markets, it is ensured that the data subject knows about the data that is being traded about them. This in principle possible with the approaches of personal data markets and the national data market, but also is only one way of accessing the market within these approaches. Also, the parameter of transmission principles is specified to a different degree depending on the data market approach. Most defined in this regard are the approaches of hybrid inalienability, where the transmission principles would have to be clearly specified in the form of usage policies, and the humanistic information economy, where the principle of reciprocity plays an essential role. The other approaches somewhat more open about the transmission principles that would be established in practice.

While some of the parameters are specified beforehand through the specific conceptions of a data market, other parameters as the receiver of information and the attributes are always left open. This is, however, not an insufficiency of the respective data market proposals, but inherent to the logic of a market as such. It is the very point of establishing a data market, that various forms of personal data can be sold to upfront unknown buyers. If the attributes and recipients of information would be clear beforehand, it would not be necessary to set up a market in the first place. It is therefore an inherent characteristic of data markets that they leave some of the parameters that constitute an informational norm unspecified.

The same observation holds for the specific contexts that would be connected through the establishment of data markets. A data market approach could in principle be limited for the trade of data within one or more specific social domains. Also it would be possible to design the abstract concept in a generic manner, but still establish them in a more sectorized form in practice. However, such a restriction would to some extent be counter-intuitive to the idea of a market as such, which aims to bring all potential buyers and sellers together. It can be foreseen that without the use of institutions that aim at a separation of markets in this regard, data trade between different contexts would occur. Since the data market approaches discussed in the literature do not name any restrictions in that regard, I also expect the markets that would be established through them to be of overarching form.

While this partial openness about the characteristics of a data transfer is embodied in a data market as such, this does not necessarily hold for individual transactions. The user gets to know some of these details once a specific transaction takes place, but it is questionable to what extent users would be involved in this process. Where a direct access to the data market is being performed, it is likely that the user will get involved in the details of trade. But in the case of intermediaries in the form of service providers, it is questionable whether users will track the details of all transactions in practice[4].

With regard to the notion of contextual integrity, this partial openness in design and practice constitutes a prima facie breach of contextual integrity[5]. But even if all of the parameters were specified, the fact that the data market approaches connect different contexts which used to be disconnected would constitute a prima facie breach. Only a very specific data market that is being used within a single domain would be able to fulfil contextual integrity. Therefore, it is the very design of a generic data markets that constitutes a prima facie violation of contextual integrity.

Besides the information flow of the data that is sold through the market itself, the additional need of market metadata is of relevance for contextual integrity in two ways. First, market metadata can partially constitute a transmission principle for the data that is sold through the market. In the case of *hybrid inalienability* for example, the metadata would be transferred along with the sold data itself and contain information about the use limitations concerning the data. Second, the mere

[4] See the discussion about the failure of notice and consent in Section 1.3.

[5] As Nissenbaum writes, contextual norms that are incomplete with regard to a new information practice, constitute a prima facie breach (2010, p. 182).

existence of market metadata would lead to a prima facie violation of contextual integrity. Since data markets currently do not exist, there are also no informational norms for market metadata established yet. This is another reason why the introduction of data markets would constitute a prima facie violation of contextual integrity.

As Nissenbaum remarks in the context of her augmented contextual integrity heuristic, a more detailed investigation concerning the contextual values is necessary to decide upon the desirability of information practices that constitute a prima facie breach. For an overarching data market this means that the contextual values of both contexts that are connected through the data market have to be taken into account. Since the contexts concerned depend on the specific situation, such an investigation is not possible in general terms[6]. The only conclusion that can be reached is that the selected data market approaches constitute a prima facie violation of contextual integrity.

Important to keep in mind is that a prima facie violation of contextual integrity does not mean necessarily that an information practice is undesirable. It can well be that a change in informational norms or an information transfer between contexts contributes to the relevant values. As Nissenbaum explains, data aggregation for example is not necessarily bad, but can be used for good causes (2010, p. 201). It depends on the specific situation whether a new information practice of this form is desirable.

Although a full assessment in terms of the augmented contextual integrity decision heuristic is not possible in generic terms, there are a number of values that are important regardless of the contexts concerned. These are the values that are of importance for markets and the moral reasons for the protection of privacy. I discuss the influence of data markets on these values in the following sections. This discussion does not lead to a full assessment in the terms of contextual integrity, but would be an important part of any investigation using the framework of contextual integrity for the analysis of more specific situations.

5.2 *Parenthesis:* market virtues

In order to perform an ethical evaluation of data markets, it is not only the values closely connected to the protection of privacy (see Section 2.3) that are of relevance, but also values related to markets in the more general sense. Even if it should turn out that data markets are neutral towards privacy-related values, it is still of importance how they affect other values through their function as a market. In this section, I therefore give an introduction into market-related values based on the literature.

efficiency and freedom

Satz (2012) provides a detailed introduction into the ethics of markets. She points out that markets are mainly recognised for their positive contributions to efficiency and freedom if they function well (Satz, 2012, pp. 17). Efficiency is closely connected to markets as it is the allocative efficiency in terms of matching supply and demand which often makes them attractive in the first place. The value of freedom is closely connected to markets as they provide choice to the individual and decrease the dependency on others and authorities. Next to the freedom that originates in markets themselves, they are also seen as a place where the capacity to choose as such can be developed, thereby forming an instrument which promotes freedom. However, the positive influence of markets on these two values is contingent upon a number of factors and market failures occur where these are not fulfilled.

equality

A central argument by Satz is that "we must expand our evaluation of markets, along with the concept of market failure, to include the effects of such markets on the structure of our relationships with one another, on our democracy, and on human motivation." (2012, pp. 34) She notes that the transfer of income and wealth is not necessarily sufficient for the equal status of citizens

[6] Nissenbaum makes a similar argument concerning one of the examples that she uses to illustrate contextual integrity (2010, p. 171).

in a democracy, but that additionally the access to certain goods such as education, health care, or rights has to be guaranteed, even if this entails to block certain market exchanges altogether (2012, pp. 100). In this regard, Satz puts a strong emphasis on the possible negative consequences that markets can have on the value of equality.

Besides the more general considerations regarding the value of equality, Satz (2012, pp. 91) provides a very specific contribution through her framework of *noxious markets*. In this framework she identifies four common values that are affected in many of the markets that are considered objectionable. Two of the values, extreme harms for the individual and extreme harms for society, are connected to the consequences that markets can have. The other two, weak agency and vulnerability, are related to the existing conditions of participants in the market. If any of these four values is especially pertinent for a given market, the market is likely to be considered objectionable and makes it appear noxious.

individual harm

Extremely harmful outcomes for the individual are one of the noxious market parameters considered by Satz (2012). Markets that negatively affect the well-being or agency of individuals in an extreme form are of this type. Important to realize is also that the negative effect can occur for the individual participating in the market themself or for a third party. Next to harm affecting an individual themself, markets can also produce harmful outcomes for society. This relates to the way how markets shape social interaction and is of especial importance for the equal standing that citizens need to have in a democracy.

societal harm

weak agency

With regard to the conditions that can make market noxious, weak agency is one of the two values considered by Satz (2012). Weak agency can relate to extreme asymmetries in information concerning the characteristics of the exchange or its consequences. Cases in which this is likely to occur include transactions which take a long time to complete or transactions that other people take on behalf of an individual. Important in this regard is that even in the absence of harmful outcomes, weak agency can constitute a problem. The other source condition for noxious markets are extreme vulnerabilities. This occurs when market participants differ in the resources they can bring in, but also in the case of differences in the capacity to understand the transaction they engage in. Besides the risk of exploitation in such cases, some transactions can furthermore actively exacerbate the vulnerabilities that already existed before the transaction. Both, in case of weak agency and vulnerabilities, it is the asymmetry between market participants that makes the market noxious.

vulnerability

severe inequalities

A different characterisation of 'objectionable markets' is provided by Sandel (2013). He distinguishes between two reasons, out of which one are severe inequalities that undermine the voluntary character of exchange. This aspect is similar to the two parameters of weak agency and vulnerability discussed by Satz (2012) and I will therefore not consider it any further.

crowding out

The second reason for objectionable markets considered by Sandel (2013) is that certain types of exchanges can lead to a crowding out of nonmarket values. A well-known example in this regard is that of a daycare center which discovered the effect that parents picked up their children even later after they introduced a fine for late pickups. Sandel explains this behaviour with a change in norms. While parents felt guilty to pick up their children late before the fine as introduced, they considered the fine as a service fee they pay and had not more need to feel guilty. Such forms of crowding out are a form of corruption, whereas corruption not only refers to wrongful gains, but to a degradation of value in the more general sense.

Sandel (2013) points to the differences between intrinsic and external motivations as a possible explanation for such a corruption of values. Interesting in this regard is an article by Bowles (2008) who points to the phenomenon that incentives and altruistic behaviour influence each other. He discusses four different ways in which such influences occur. First, incentives can lead to a change in the frame under which a transaction occurs. Second, incentives can be learned and lead to a change in endogenous preferences over time. Third, incentives can threaten the recognition of self-

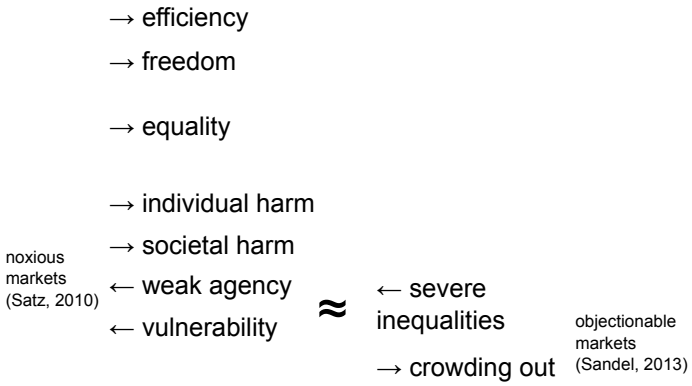

Figure 5.1: Overview over the market-related values discussed in the literature.

('←'=source, '→'=outcome)

determination, leading to defensive reactions. Fourth, the existence of incentives as such conveys information about the beliefs that a principal has about an agent. In each of these cases, incentives can possibly undermine moral values by changing the intrinsic motivations that people have.

In summary, markets are intended to be of positive influence on efficiency and freedom, but can also negatively affect a number of values. In this regard, markets can have negative consequences on equality, lead to individual harm and societal harm. Also, markets can lead to a crowding out of values by changing the intrinsic motivations that people have. Finally, markets are likely to be objectionable in presence of weak agency and vulnerability. Figure 5.1 shows a summary of these values.

5.3 Data markets and market-related values

As discussed in the previous section, there are two main values that are closely connected to the function of markets. These are efficiency and freedom. Besides that, the effects of markets on equality are of general importance. The notions of *noxious markets* (Satz, 2012) and *objectionable markets* (Sandel, 2013) provide more specific insights into the inherent dangers of market transactions. In the following, I evaluate the influence of transactions in the context of data markets on these values.

efficiency One of the main functions of markets is the allocative efficiency in terms of demand and supply that they can create. As Nissenbaum remarks, a common claim is that the availability of more detailed information improves business efficiency by making it possible to identify a desirable customer base (2010, pp. 206). This is basically a corollary of the claim made by the Chicago school that privacy creates a market inefficiency. While there might be some validity in this claim[7], it is important not to ignore the other functions that privacy has for the individual and society. Also important to realize is that the allocative efficiency referred to here is the one in the markets for other goods and services and not the efficiency of the data market itself.

The allocative efficiency within the data market itself is another question. As pointed out by Shapiro and Varian (1999), the price-setting mechanisms for information are not straightforward since the marginal cost of information is almost zero. From the degree of detail provided in the descriptions of the selected data market approaches it is difficult to determine how specifically the

[7] As discussed in Section 4.3, Hermalin and Katz (2006) question the general validity of this claim.

price-setting would work. In every case, however, it is clear that a simple price-setting mechanism which equals supply and demand is likely to fail. Also in substance the allocative efficiency of data markets itself seems questionable. In a simplified view, a higher supply in information through the market can be interpreted as a decreased level of data protection for the users providing the information. Often, the claim is made that users can control their desired degree of privacy through the price they demand for their data (see e.g. Rust et al., 2002; Lanier, 2013). This would mean that an optimal supply in data for businesses has to be matched to the optimal level of protection of privacy of users. The influence of individual preferences for privacy-protection, the difficulty for users to assess their need for privacy protection (see e.g. Acquisti, 2004) and the question how the established price relates to the costs of information services are only some of the questions that make this claim seem doubtful. In every case, it is important to realize that the allocative efficiency within the data market is different from the allocative efficiency within the markets for goods and services that make use of this information.

freedom The other main function of markets is their contribution to individual freedom. Markets are supposed to enhance the capability of humans to choose between alternatives and free them out of coercive relationships. Through the introduction of data markets, users are supposed to be put in the position to select who can make use of their data at which price. This seemingly increase in freedom to choose, however, is doubtful for at least three reasons. First, it is highly questionable whether individuals are able to take meaningful decisions when it comes to the use of their personal data. Second, the influence of individual decisions on others and society at large has to be taken into account. Third, choice may be limited in the way that preferred options are simply not available.

Regarding the first point, there is a massive body of literature indicating that the model of the individual as *rational agent* is flawed[8]. Psychological research has provided insights which indicate that the assumptions of the *homo economicus* are flawed in a number of points (Schneider, 2010). In the context of data markets, it is noteworthy that the effect of *bounded rationality* will decrease the likelihood that an informed decision is being taken. The discussions around the failure of the current notice and consent model (see e.g. Martin, 2013; Van Alsenoy et al., 2013) indicate that users are not willing to take all information that is provided to them into account. Another important insight is that users are likely to discount the probabilistic harm that privacy infringement may induce when comparing it to the immediate benefits that a sale of information might incur (Acquisti, 2004). The discrepancy between a self-assessment of privacy-preferences and the behaviour that users show when it comes to information disclosure, the so-called *privacy paradox* (Norberg, Horne, & Horne, 2007), is a further warning to be careful in this regard. Another paradoxical effect is that of the *control paradox* (Brandimarte, Acquisti, & Loewenstein, 2013), which refers to the effect that individuals tend to disclose more information if fine-grained privacy controls are available. Finally, the insights that in many cases preferences are endogenous and depend on the frame provided (Bowles, 2008) make it doubtful whether users selling their data can be assumed rational agents.

The second point which has to be taken into account concerning the freedom of individuals, is the effect that their decisions have on others. As I explain in more detail below, there a number of ways in which the decision of one individual to provide personal data may affect others. Furthermore, the importance of privacy for society should be considered. For these reasons it is questionable whether it is desirable to provide the maximum individual freedom to sell personal data to users. When the possible harm for others, the effects of privacy on equality and the societal importance of privacy are taken into account, a more paternalistic approach to privacy protection

[8] I can only provide a very rough overview at this point, but will discuss possible effects regarding user behaviour in more detail further below (see Section 5.5).

might be justified.

The availability of options is a third point to be considered here. Choice in the context of a market is only meaningful to the degree that different options reflecting the individual preferences of market participants are available. As the status quo shows, the monopoly position of many online service providers (see e.g. Lanier, 2013) makes it impossible for users to choose a service which reflects their preferences. As Cohen (2000) points out, this brings the danger that service providers unilaterally set the terms under which their services may be used. Without competition reflecting the different preferences of users, it is therefore questionable to speak of choice that markets enable. This primarily concerns the market for services that users make use of. However, all data market proposals foresee the possibility to sell data originating at services for secondary purposes on behalf of the user. This bears the danger of a tight coupling between these markets, leaving the situation of limited options that are available untouched.

These three considerations, the limits of rationality, the influences on others and the availability of preferred options, make it questionable whether the introduction of data markets leads to increased freedom. Under the right conditions, data markets could lead to an increase in individual freedom, but this is not necessarily the case. Important in this regard is also the question to whom freedom is provided: all individuals, individuals in a certain position or with certain preferences, or just the service providers and buyers of data. Furthermore, the distinction between positive and negative freedom seems reasonable at this point. While data markets are likely to increase negative freedom in the sense of less constraints in form of regulations, it is more questionable whether they contribute to positive freedom actually enabling individuals to enjoy a self-determined life.

equality Another value closely connected to markets is equality. While the Libertarian view on the function of markets for *equality of opportunity* is well-established, there are diverging views on the desirability of *equality of condition* (Gosepath, 2011). One concept of interest in the light of data market approaches is the concept of *complex equality* (Walzer, 1983), which recognizes that goods in different social spheres are distributed through different mechanisms. Van Den Hoven (1997, 2008) argues that also information is a good which has to be distributed according to complex equality as well. This consideration connects the above discussion on contextual integrity and highlights the dangers that data markets might have on equality in this regard.

Another point connected to equality is the question how data market approaches handle initial ownership and issues of distribution. When considering market transactions for data to be an element of privacy regulation, the Coase theorem suggests that the initial allocation of rights in data does not matter as it is the relative valuations that ultimately determine whether a transaction takes place (see e.g. Acquisti, 2010). In a similar way, Rust et al. (2002) argue that the two situations of service providers paying customers for information and service providers selling privacy protection to customers are formally identical. While this might be true on a per-transaction basis, it leaves the question of initial allocation untouched. This might according to the Coase theorem be irrelevant regarding the regulatory effect in a economic sense, but should nevertheless be taken into account. Especially if one follows the argumentation of Lessig (2006) and others saying that ownership rights would incentivize individuals to stronger apprehend their rights, the question of initial distribution should matter regardless of the question of economic efficiency. Furthermore, the discussion around the EU Data Protection Regulation shows, the question of default entitlements also plays a role in practice (Purtova, 2013).

Besides the question of initial allocation, the relative change that an introduction of data markets might introduce is a point of discussion. Members of the Chicago School as Posner (1981) warn that privacy rights are redistributive and criticize the market inefficiencies that privacy rights create. However, as Nissenbaum (2010, p. 111) remarks, it is important not to ignore the question whether benefits are evenly distributed in the status quo. If the current situation favours service

providers and the introduction of new privacy rights leads to a redistribution making them worse off, it does not necessarily mean that this change is undesirable. Therefore, when considering the introduction of data markets, it is the resulting situation as a whole that should be assessed and not a possible redistribution in isolation.

noxious markets Satz (2012) concept of noxious markets does not address equality in terms of goods only, but also in terms of equal status in the relationship between people. She argues that it is important for democratic societies that citizens treat each other as equals and that some forms of markets can undermine this form of equality. In her framework, she distinguishes between two sources (weak agency and vulnerability) and two outcomes (individual harm and societal harm) that can make a market noxious. In the context of data markets, the harmful outcomes of data markets are closely connected to the reasons for the protection of informational privacy. Since I discuss these in more detail further below, I only discuss the two sources of weak agency and vulnerability at this point.

Weak agency is indeed of concern when it comes to data markets. Satz (2012, pp. 96) names significant time lags in a transaction and indirect involvement in a transaction as two common sources of agency problems. Both of these are of relevance in the case of data markets. The fact that the decision to sell data happens at one point in time, but possible adverse consequences due to its use only occur at a later point in time constitutes such a time lag[9]. The possibly long time-span between the decision to sell and the consequences occurring based on this decision weaken the agency of a user. In the case of indirect involvement in data markets through service providers selling data on behalf of the user, the problem of indirect involvement is also present. As a direct form of involvement is only seen as an additional way of access in some of the data market approaches, a weak agency based on the form of market access is likely to be present. In a more general sense, the points discussed above under the topic of bounded rationality apply here as well. Overall, the user is likely to be in a position of weak agency if the condition under which the transaction takes place are not designed carefully. With the current experience of service providers setting the terms (Cohen, 2000) and partially also framing their products, it is likely that a situation of weak agency will occur.

Vulnerability in terms of different resources or capacities is another source of noxious markets (Satz, 2012). This is of relevance for data markets in two ways. First, it concerns the circumstances in which data markets are situated. If a data market is accessed by populations with extreme financial disparities, vulnerabilities might occur. A data market set up at a national level might thereby be different from a global data market. Second, data markets themselves might also actively contribute to vulnerabilities. In a TV interview Lanier mentions the option that the worth of personal information might become enough to bring people out of poverty[10]. If indeed the worth of personal information becomes so much that it could constitute a form of basic income, the voluntary character of data market transactions might become endangered. In the absence of other forms of social security, the reliance on the income from data markets would put individuals in a extreme form of vulnerability. In this way, data markets might not only be a problem in cases of vulnerability, but could even aggravate the situation by making other forms of social security seemingly obsolete and unavailable.

In general, it is important to take such more indirect effects into account. It is not only the preferences that are often endogenous (see above), but also the set of options that are available can be endogenous (Satz, 2012, p. 180). Satz makes the argument that a market for kidneys can lead

[9] This is connected to the point of discounting of probabilistic harm discussed above.

[10] The full quote is "But at the point where just your information is worth enough to bring you out of poverty, then we have an option to think about a society in a new way. And this is what I am trying to get people to think about." (Lanier at 48:45 in vpro, 2013)

to a situation in which people who are not willing to sell their organs effectively have to pay the cost for exercising this choice. Such *pecuniary externalities* have to be taken into account when evaluating the effects of data markets. As the example of data income as a basic income illustrates, data markets could shape the background conditions in such a way that users who are not willing to sell their information have to find another source of income in order to be able to make this decision. Thereby, privacy protection would effectively come at a cost, whereas right now it is a right that can be enjoyed free of charge.

objectionable markets Sandel (2013) names two different reasons why markets may be objectionable on moral grounds: severe inequality and a crowding out of nonmarket norms. The first reason is closely connected to the parameters of weak agency and vulnerability discussed above. In the following, I therefore only discuss the effects of data markets on the crowding out of morals. Sandel names the differences between intrinsic and extrinsic motivations as a possible explanation of a crowding out effect. On a similar note, Bowles (2008) names four ways in which incentive mechanisms and altruistic behaviour influence each other: framing, endogenous preferences, overdetermination, and information content. All of these mechanisms are likely to play a role in the case of data markets.

First, financial remuneration is likely to change the frame of a transaction and thereby change the way how it is perceived. As discussed above under the transmission principles of contextual integrity (see Section 5.1), the fact that data is provided in exchange for financial remuneration might communicate to the organisation paying for it that they may use the data for any purpose. This changes the frame of data being provided for specific purposes, to that of a commercial transaction transferring the right to use data. But also the other way round a framing effect is likely to occur. Through the possibility to sell personal data for financial remuneration, individuals are put under the impression that their personal data is a commodity that they can dispose off. This is a threat to the status of privacy as a human right within Europe[11]. Under the current legal regime of Data Protection within the EU, institutions as the data protection authorities and the fact that fines instead of litigation are used to settle disputes clearly communicate the human rights status of privacy. If instead personal data can be sold for remuneration, this implicitly frames privacy as a commodity that can be negotiated upon. This not only endangers the perception of a right to privacy as such, but also is of importance for democracy. As Roessler remarks, the possibility for people to negotiate about privacy might change people's self-understanding and valuation of being self-determined individuals. This is problematic, as democracy is dependent upon individuals "[...] who are aware of and who value their autonomy." (Roessler, 2005, p. 120)

Second, the point of endogenous preferences might also play a role. Data markets would provide individuals with financial incentives to provide information. Users of online service might get used to the fact that they receive financial remuneration for the provision of data or at least are shown the financial benefits that this entails. Once this mechanism of data in exchange for financial benefits becomes widely established, it might seem unreasonable to provide data out of altruistic reasons. This partially is the effect that Lessig (2006) hopes for when he argues for the introduction of property rights to personal data, which shall give individuals a stronger interest in their data. The expectation of financial remuneration shall make individuals alert about the importance of their personal data and instruct them to share it sparingly. However, this might also influence the collection of data for research purposes or the collection of data in the context of governmental services. The preference of a provision of data for legitimate purposes might shift towards a provision of data for a high enough financial remuneration.

[11] There also is the juridical question whether a data market approach would be compatible with the human rights status of privacy in Europe (see e.g. Purtova, 2011), but here it is the perception as a right and not the legal status that I refer to.

Third, the awareness about this commodification might also lead to the effect of overdetermination. Being aware about the effects of financial incentives, people may perceive their self-determination being under attack. As Prins (2006, p. 272) remarks, it ultimately is not the data itself, but the behaviour and identity of individuals that becomes commodified. If individuals realize that they are incentivized to sell this integral part of themself through a market, they might object this commodification on the grounds that they are the ones who should exercise their right of self-determination.

Fourth, data markets would also communicate information. The economic relevance of personal data would not only be discussed at a macro scale (see e.g. World Economic Forum, 2011), but also become clear to each individual user. This is the reason why Novotny and Spiekermann (2013) propose the option of a paid privacy-friendly service as part of their *personal information markets* proposal. Increased transparency about the economic worth of data broken down to the individual, would give online service users a better understanding of the worth that the personal data provided by them entail. This is in stark contrast to the status quo of *free services* where the economic worth of data is not known to individuals.

In addition, it is important to realize that a more common argument in favour of data markets overlaps with the four points made here. It is often argued that data markets would give individuals the autonomy to freely decide whether and how they value their privacy and that strict privacy regulations would be paternalistic (see e.g. Kang in Purtova, 2011, pp. 134). This argument ties in with the effects of frame and overdetermination and might form a counterbalance to the points made above.

Also, the four effects work jointly together and cannot be clearly separated from one another (Bowles, 2008). It is therefore hard to predict the overall effect that a data market approach has in this regard. Important, however, is to recognize that there are a number of reasons suggesting that data markets would have an effect on nonmarket norms in this way.

In summary, data markets are likely to have multiple effects on the different values that are connected to markets. Important is to keep in mind that there are different levels at which these effects take place. The effects on the user have to be distinguished from those on other individuals and those on the service providers. Also, there is not only the data market that is to be established, but also markets for other goods and services might be affected. Furthermore, the point of time at which personal data is sold is not necessarily the point in time at which possible consequences connected thereto occur. I analyse these distinctions in more detail further below. But before doing so, I first discuss the effects of data markets on privacy-related values.

5.4 Data markets and privacy-related values

As the central question of this thesis concerns the effects of data markets on the protection of privacy, it is not only of relevance to evaluate the values that are closely connected to markets, but even more important to take the values closely connected to privacy into account. As I concluded in Section 2, I do not follow any specific definition of privacy and rather regard privacy as a cluster of concepts. Therefore, it is difficult to evaluate the effect of data markets on privacy as such. Regardless of the definition, however, there is more agreement about the moral reasons why we ought to protect privacy, which makes them a good starting point for an evaluation of the possible effects of data markets. In the following, I discuss these along the four key values identified in Section 2.3: the prevention of harm, autonomy, ownership and the common good.

prevention of harm The prevention of harm is probably the most prominent reason in discussions about the protection of data. Failures in which actual harm based on the misuse of

information occur make the reasons for data protection observable and thereby explicate in a very direct form why access to data should be limited. In the context of data markets, the possibility of harm based on data that is sold through the market is a reason that should be taken into account. This consideration also links to the point of individual harm as one of the outcome parameters of the *noxious markets* framework.

The ways in which data markets could lead to information-based harm are manifold. Since data markets are supposed to be a general point of interaction and do not aim at the establishment of a specific data flow[12], it is not possible to demarcate in which specific ways informational harm could occur. The characteristic of an a priori unspecified set of recipients, however, indicates that data markets conceivably lead to an enhanced risk of informational harm. If information is up for grabs on the market and can possibly be used by anyone for any end, this incurs the risk that information is also used in undesired ways. If not prevented by regulatory or technical means, there is at least no reason to assume that a data market would decrease the likelihood of informational harm.

A specific kind of informational harm that is more closely connected to data markets is discrimination. Since the commercial value of data largely stems from profiling for the purpose of targeted advertisement and increasingly also price differentiation, it seems safe to assume that data markets will for the most part also serve this purpose. Which forms of discrimination occur depends on a number of factors. On the one hand, one can argue that a greater availability of personal data will decrease the likelihood of unjustified discrimination. If individuals could be directly involved in the markets that sells information about them, they would at least have the chance to correct erroneous information, so that the quality of information sold on the market increases. This requires however that data markets indeed lead to an increased level of transparency and user involvement, which depends on the specific details of the implementation. Unfair discrimination, on the other hand, might also increase due to the greater availability of data. Also, organizations utilizing data bought on the market would have a stronger claim to use that information for discrimination purposes, as they bought it from individuals potentially even providing it for these purposes. Whether or not discrimination takes place, therefore depends on the usage restrictions of data and regulations that are in place. In every case, these considerations show that the requirement of due process becomes even more important if data markets open up the acquisition of personal data to a large customer base. If implemented with these considerations in mind, a data market can by itself become a part of the procedures for due process. If an implementation does not pay attention to this point, however, discrimination might also increase and become legitimized through the market mechanism.

Connected to the issue of due process is the possibility for subjective privacy harm. If the personal data of individuals is for sale on a data market, it is important that individuals are aware about the ways in which their information is used. The lack of information in this regard or the presentation in ways that cannot be easily comprehended might lead to uncertainty about the usage of data, which in turn will translate into subjective privacy harm. If, however, information about the use of data is supplied in sufficient and understandable form, data markets could also lead to a decrease in subjective privacy harm. Aside from this aspect, the pure existence of data markets can also influence subjective privacy harm. With the status quo of free services, the mere existence of secondary markets is known, but the details of actors who are active therein is largely unknown (Craig & Ludloff, 2011). An increased transparency through more direct data markets might therefore have two effects. On the one hand, it could supply those who are concerned about secondary markets with more details and thereby decrease subjective privacy harm. On the other hand, it might also make the practices of data trade and secondary use more widely known and thereby actually increase subjective privacy harm. Which of these two effects would occur,

[12] See the discussion on recipients and attributes in Section 5.1.

consequently depends on the way how data markets inform individuals about the use of their data.

autonomy The discussion on weak agency in the context of noxious markets already touched upon the connection between data markets and autonomy. As autonomy is a very multi-faceted concept, it is good to focus on one conception of it. Roessler speaks of autonomy as "[...] personal autonomy in the sense of general personal self-determination concerning how I want to lead my life." (2005, p. 51) She sees autonomy as part of freedom in the sense that freedom is about more than mere choice and autonomous decisions ought to be part of it. Roessler furthermore emphasizes that the decisions taken by individuals have to be authentic, meaning that one can identify with the decisions one is taking. Another requirement for autonomy is that one must be able to pursue goals and projects not only in theory, but also on a practical level. For that, in turn, one needs symbolic and literal spaces in which one is free from the access of others, which makes privacy a requirement for autonomy.

Data markets are connected to autonomy in mainly two ways. First, data markets as such provide the user with the option to sell personal data in a market environment. This gives users the choice to keep their personal data or dispose of it for a financial benefit that they deem adequate. While this without doubt provides the user with options that a stricter data protection regime might not provide, it does not necessarily lead to increased autonomy. As Roessler (2005) emphasizes, it is not the mere options that matter, but the question whether these reflect the authentic desires of the user. In a broader context including the services to which data markets might be connected, it is important that the services are offered in a way that do not force a specific decision regarding the sale of data upon the user. Data markets can therefore increase autonomy, but only if they do not impede on any other options and reflect those choices that individuals wish to make.

The second way in which data markets are connected to autonomy is more indirect. It does not concern the data market itself, but the way in which information sold on data markets might be used in connection with the individual concerned. As explained above, much of the data bought in data markets might be used to tailor offers in other markets towards the individual. If information about the individual is for example used for targeted advertisement, the individual concerned might not be aware of other alternatives next to those offered in the targeted advertisements, thereby limiting the autonomy of the individual (Morozov, 2014). Important to realize in this regard is that this limitation is in the first instance invisible, since the options that are not offered are not visible. This effect is what Morozov (2013b) describes as the 'invisible barbed wire' controlling our life. And even if users are in principle aware about these manipulations taking place, they might not be aware about the extent to which profiling enables others to gain knowledge about them (Schallaböck, 2014, p. 40). Intricate about such forms of manipulation is that either way they lead to a limitation of autonomy. If the user is aware about the extent to which these manipulations takes place, this can lead to subjective privacy harm (see above). This means that in expectancy of the possible misuse of data, the user will perceive a limitation in their room to pursue projects which constitutes a violation of autonomy. If the user is not aware about these manipulations, this limitation is not present, but the decision they is taking are not authentic, meaning that again autonomous decisions are endangered. What makes these effects even more intricate is that they concern different marketplaces and different points in time.

Figure 5.2 provides an overview over the different effects that data markets have on autonomy. If at time t1, the user makes the decision to utilize a data market to sell data from a specific service, this will influence knowingly or not the options in a market for goods that are offered to them. If the user does not know about this effect, they cannot act autonomous at time t3 in the market for goods, as it is not an authentic decision being taken. If the user is aware about this effect, subjective privacy harm will lead to a limitation of autonomy at time t2 when data about the user is gathered through utilizing a service. While the option that data markets offer for a decision

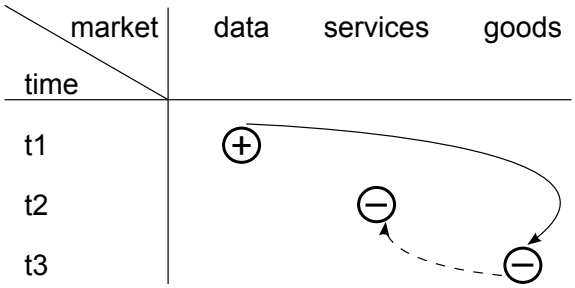

Figure 5.2: Interrelation of the influence on autonomy in different markets.

at time t1 constitute a possible increase in autonomy for the individual, this leads to a possible decrease in autonomy at time t2 or t3.

ownership Although ownership is a maybe less-prevalent reason for the protection of data, it is one that is of relevance in the context of this study. If data markets were introduced, they would emphasize the aspect of ownership much stronger than the current data protection legislation does. The coupling of financial aspects to data would spell out this aspect of privacy that currently is only implicitly perceived by some. As discussed above, a data market approach would require it to be explicit on the question of default entitlements and thereby touch upon the question of data ownership.

If users are the initial owners of the data concerning them, this reflects the idea that data concerning one is part of oneself. Undesired access to this data therefore constitutes a violation of privacy based on this connection to the self. Besides this symbolic relation, data markets would add a financial dimension to this form of ownership over one's data. Undesired access to data would then also constitute a dispossession in financial terms, thereby intensifying the notion of ownership. The ability to demand financial compensation for access to one's data furthermore provides a potential method of data protection. This requires, however, that the individuals concerning the data are allocated the initial ownership of the data.

The question, however, is whether users ought to deserve the full financial benefits of data concerning them. As was early recognised by Miller (1971, p. 213), the commercial value of personal data often originates at the service provider and not the user themself. It is thus a collaboration of the user using a service and the service provider who invested in the infrastructure which creates the value of personal data. This problem becomes apparent in the discussion concerning the right for *data portability* in the proposed Data Protection Regulation (European Parliament, 2013b). As currently proposed, the data subject would have the right to obtain a copy of the data concerning them, but might not be able to use it on a competing platform, based on the intellectual property rights that the originating provider might invoke. Altogether, the current and also proposed EU data protection legislation recognises that there are shared interests between users and service providers. There are no clear default entitlements for either of the two parties (Purtova, 2013). The introduction of a data market would require to make the default entitlements to data explicit in financial terms, which would make it more difficult to acknowledge the shared interests in personal data between user and service provider.

Another change that data markets would bring is the transfer of property interests that the payment of a financial remuneration would bring. As discussed in Section 5.1, the fact that personal data would be transferred under a transmission principle of commercial exchange creates expectations about the rights that the purchase of data entails. Although property always entails multiple components [13], the introduction of a financial remuneration in return for the provision of personal

[13] See the discussion about property in Section 1.4.

data is likely to move the expectations of property interests closer to that of 'full-blooded ownership'. This would create the potential for conflicts with regard to the privacy as property interests that individuals might have in the personal data concerning them. A solution to this problem might be the inclusion of usage policies in the form of metadata that is transferred along with the data itself, as proposed in the case of *hybrid alienability* by Schwartz (2004). If, however, personal data is transferred in isolation through a data market in return for financial remuneration, this might create an expectation of property interests at the receiver side which does not acknowledge the privacy-related interests of the user.

Generally speaking, data markets are likely to enhance the visibility of property interests that are more implicit under the current data protection legislation. As Roessler (2005, p. 142) writes, even in the case of local privacy full ownership is not necessary, but it is the ability to use private spaces that matters. Likewise, informational privacy is also possible if only some features of ownership are available, which creates the possibility of a symbiotic relation between user and service provider. The aspect of financial remuneration in case of data markets bears the danger to disturb this relationship and might have to be counterbalanced with additional measures such as sticky policies containing the usage rights. Positive is the increased transparency that data markets would bring in this regard. This might help to disentangle some of the more hidden privacy conflicts that are grounded in a violation of property interests.

common good An aspect that is often ignored in debates around privacy is the importance of privacy for the common good. Although this is not a problem specific to data markets, it is nevertheless insightful how data markets would take influence on the societal function of privacy. In this regard, the discussion following hereafter also connects to the point of societal harm in the framework of noxious markets. Data markets would influence privacy as a common good is in two ways: through the importance of privacy for the common good as such and through the influence of individual decisions on the privacy of others.

Through the importance of privacy for the functioning of the public sphere, possible influences of data markets would also influence the common good. If the individual is able to sell their personal data and thereby potentially renouncing their autonomy, this would not be limited to consequences for the individual themself. The lack of room for the development of projects would not only be lacking for individual projects, but also for projects that are of societal nature. Autonomy in this regard does not only concern the individual, but also groups within society. Although the current data protection legislation is also not perfect in this regard (Floridi, 2014), it seems likely that data markets which incentivize individuals to sell their personal data and thereby possibly renounce their autonomy will also lead to a decrease of autonomy that is necessary for the functioning of the public sphere.

This possible lack of room for experimentation comes in two forms. First, there might be a lack of privacy and thereby autonomy within a societal sub-system in order to have the room necessary to experiment. Second, the boundaries of societal sub-systems as such could be lifted through the overarching character of data markets. This separation is necessary for the public sphere function as such, but also an important requirement for the development of societal norms and the functioning of social interaction (see Section 2.3). As the discussion on contextual integrity above has shown, it is likely that data markets will lift the boundaries between different social domains, thereby endangering the separation of societal sub-systems. This endangers the functioning of the public sphere, hinders the development of societal norms and leads to problems in social interaction.

Next to this influence on the common good as such, data markets could also have adverse effects on others than the ones who make the decision to sell their data on a market. As discussed under the topic of discrimination (see Section 2.3), non-distributive group-profiling bears the danger that

others than the ones from whom the data originated are affected. If user A decides to sell their data on a data market, this data might be used in a context which influences user B, even if user B decided not to sell their data. Thereby the autonomy with which user A decides to participate in the market, can have an influence on the autonomy of user B in another market and context. This is a general problem that already exists with the status quo of secondary markets (Schallaböck, 2014)[14], but is likely to increase in the case of data markets.

As I have shown in the preceding paragraphs, the decision to sell personal data through a data market can potentially have negative consequences for other individuals and society at large. This is a general characteristic of data sharing, but is of particular importance in the context of data markets. Through data markets, users would gain the option to sell their personal data for financial benefits. It is likely that users will make this decision contingent upon the negative consequences that they expect for themselves. It is, however, important to take also the consequences for others into account when the decision to sell is being taken. The often-made argument that everyone should be able to decide for themselves whether or not to sell their data is therefore flawed.

Data markets are likely to have effects on the moral reasons why we ought to preserve privacy. Central is the somewhat intricate relation of data markets to the autonomy of users at different points in time. Also important is the insight that the possible adverse consequences thereof are not limited to individual harm for the individual selling their data, but can also affect other individuals and society at large. Therefore, caution should be taken that the possibility to sell data is taken responsibly and possible adverse consequences are minimized. To what extent this is possible depends on the specific details of a data market. I investigate these differences together with the conclusions to draw from the discussion on market-related values in the following.

5.5 Discussion and conclusions for data market approaches

As the evaluation in the previous sections has shown, data markets have the potential to influence market-related and privacy-related values in a number of ways. The mechanisms in which this takes place are manifold and it depends on many parameters how they would work out in practice. Although the selected data market approaches are not specific enough to assess all of these parameters, the differences between them give some indication of possible differences in the outcomes. In this section, I try to provide an overview over these differences and thereby link the specific data market approaches to the more general evaluation of the preceding sections.

I start by providing a summary of the findings concerning the impact of data markets on market-related and privacy-related values. As the emergence of these effects is contingent on the way in which users would engage in data markets, I also provide a short overview over findings from relevant studies on user behaviour. Finally, I draw conclusions for each of the selected data market approaches based on these insights.

influences on market-related and privacy-related values As the preceding evaluation has shown, there are a number of ways how data markets could influence values related to markets and privacy. For a better overview, Table 5.1 (on page 81) summarizes the possible effects in a more compact form[15]. Important to keep in mind is that these effects would only possibly occur

[14] The introduction of what Vedder (1999) calls 'categorical privacy' could be a solution in the context of the current data protection legislation. With categorical privacy, data protection would be extended from personal data in the strict sense (see Section 1.2) to data that can have the same negative consequences if it is re-applied to specific individuals.

[15] NB: The possible effects listed in the table are intertwined and cannot be broken down into distinct parts. The table therefore only serves as a rough overview, but is not necessarily complete on its own.

and are not in any case necessary consequences of data markets. Whether they would occur is contingent upon the details in which a data market is implemented and the way in which users would make use of the data market.

the importance of user behaviour In which way users would be involved in a data market depends partially on the specifics of a a data market implementation, but is to some extent also the decision of the user themself. Existing research on the behaviour of individuals in markets and the attitudes of users towards the protection of their privacy can be helpful to understand how users would behave in data markets if they would be set up. A complete anticipation of human behaviour is of course not possible, but insights from these fields can help to identify these elements of data markets that would be of influence in this regard.

There have been a number of empirical studies trying to investigate how users value their privacy in monetary terms (see e.g. Acquisti, John, & Loewenstein, 2009; Tsai, Egelman, Cranor, & Acquisti, 2011; Jentzsch, Preibusch, & Harasser, 2012; Bauer, Korunovska, & Spiekermann, 2012; Savage & Waldman, 2013). However, it is known that users act differently from the privacy preferences that they report (Norberg et al., 2007). This effect, called *privacy paradox*, complicates studies about the valuation of privacy by users. Also, it seems that the evaluation of user preferences in this regard is highly context dependent and that it is difficult to establish the valuation of privacy in general terms (Morando, Iemma, & Raiteri, 2014). Although these studies about privacy valuation provide useful insights, the preliminary conclusion in this regard is that there is no generally applicable model to determine at which price a user would be willing to sell their information. It is therefore also not possible to estimate to what extent users would be willing to give up their privacy in exchange for monetary remuneration.

Some studies make the assumption that the user performs a simple cost-benefit analysis regarding the advantages that a service provides, monetary incentives and the privacy 'costs' they incur through using the service (see e.g. Rust et al., 2002; Awad & Krishnan, 2006). However, as is known from the field of behavioural economics, the model of the *homo economicus* is flawed in many regards (Schneider, 2010). Out of the more general findings from this field, there are a number of psychological effects that seem especially relevant in the context of data markets. These are the influence of *framing*, the *endowment effect* and *loss aversion*. Indeed, framing has been found to be of influence for the willingness to disclose information (Acquisti et al., 2009; Adjerid, Acquisti, Brandimarte, & Loewenstein, 2013). And also the endowment effect is at work when users make decisions about monetary incentives in the context of information disclosure (Acquisti et al., 2009; Morando et al., 2014).

Loss aversion plays a role in the difference between the *willingness to pay* (WTP) and the *willingness to accept* (WTA). Acquisti et al. (2009) found that user are not as willing to pay for the protection of their information as they are willing to accept monetary incentives for information disclosure (WTP $<$ WTA). Similar differences between the WTP and the WTA have been found in another study as well (McDonald & Cranor, 2010). However, there is evidence that there is a willingness to pay for the protection of privacy. Savage and Waldman (2013) have shown that some users are willing to pay a premium for smartphone apps that access less of their data. The differences between the WTP and the WTA, however, emphasize the importance how a service is framed in monetary terms. Therefore, being able to receive a payment for data with which one pays for the provision of a service might in the end not be the same as a 'free service', even if the two scenarios are equivalent from a financial point of view.

Also, it is important to realize that privacy concerns are not the only consideration that users take into account when determining whether to provide access to data concerning them. Spiekermann et al. (2012) have shown that users develop a *psychology of ownership* towards their data. Once users are made aware about the economic worth and the interest in their data, they estimate the

#	value	time	market	actor(s)	possible effect
1	efficiency	t3	goods	business	improvement in allocative efficiency
2	efficiency	t1	data	user+SP	difficulties with price-setting
3	freedom	t1	data	user	user obtains choice to sell data
4	freedom	t1	data	user	user is not able to make rational decision
5	freedom	t1	data	user	effects on others have to be included
6	freedom	t1	data	user	desired options might not be available
7	equality	t1	data	user	complex equality is endangered
8	equality	t0	data	user+SP	default entitlements are difficult to determine
9	equality	t0	data	user+SP	possible redistribution of default entitlements
10	weak agency	t1	data	user	time lag between t1 and t3
11	weak agency	t1	data	user	SP selling on behalf of user
12	vulnerability	t1	data	user	extreme income disparities
13	vulnerability	t1	data	user	data market as basic income
14	'crowding out'	t1	data	buyer	framing: paying for data means owning it
15	'crowding out'	t1	data	user	framing: selling data means it is not important
16	'crowding out'	t1	data	user	endogenous preferences: accustomed to be paid for data
17	'crowding out'	t1	data	user	overdetermination: incentives undermine self-determination
18	'crowding out'	t1	data	user	communication: economic relevance of data
19	'crowding out'	t1	data	user	communication: user is autonomous
20	harm	t3	goods	user	user is harmed in some way
21	harm	t3	goods	user	unjustified discrimination
22	harm	t3	goods	user	unfair discrimination
23	harm	t2	services	user	subjective privacy harm
24	autonomy	t1	data	user	option to sell data
25	autonomy	t2	services	user	subjective privacy harm
26	autonomy	t3	goods	user	non-authentic decision
27	ownership	t1	data	user	strengthened perception of ownership
28	ownership	t1	data	user+SP	clear default entitlements
29	ownership	t1	data	buyer	paying for data means owning it
30	ownership	t1	data	user+SP	distortion of symbiotic relation
31	common good	t2+t3	data	society	distortion of public sphere
32	common good	t2+t3	data	society	distortion of social norms and interaction
33	common good	t3	goods	someone	possible discrimination

Table 5.1: Summary of the different ways in which data markets could influence market-related and privacy-related values.

(t0 = introduction of data markets, t1 = decision to sell, t2 = data is gathered, t3 = data is used,
SP = service provider)

value of their data to be higher. According to the study by Spiekermann et al., privacy interests form a part of the value that users attach to their data, but it is outweighed by the psychology of ownership. This insight connects to strengthened perception of ownership that data markets might create. Data markets might therefore increase the level of care that users have towards their data, even if this means that they defend it on grounds that one might consider to be different from privacy.

Another factor that is of importance for the decision whether to share data is the level of trust towards the party that the data is shared with. The influence of trust as a factor has been early identified (see e.g. Chellappa & Sin, 2005; Dinev & Hart, 2006). A more recent and related insight is the role of the *control paradox*. Brandimarte et al. (2013) have found that more fine-grained control in terms of privacy-settings makes users feel at ease, leading them to disclose more information than they would without these controls. This insight is of importance for the paradigm of privacy as informational-self determination as a whole, but also applies in the context of data markets. If data markets are supposed to put users into control about the question which data to make accessible to whom at which price, this might constitute a more fine-grained control than the one that exists in the status quo.

Furthermore of relevance is the level of understanding that users have about the working of technology and the possible consequences of sharing their data. The general problem of the notice and consent approach (see Section 1.3) that users cannot be expected to fully read and comprehend long legalistic texts might apply in the context of data markets as well. It therefore is of importance that a data market design allows the user to make decisions in an intuitive manner. However, it is not straightforward to achieve this. As the studies on the influence of framing (see above) show, it is subtle differences that can be decisive for the decision to share data. Moreover, a recent study by Hoofnagle and Urban (2014) questions whether users are able to make meaningful decisions in this regard. They found that the largest share of users who are often believed to be simply pragmatic about the protection of their privacy, are in fact uninformed about actual business practices concerning their data. It therefore is crucial to take the understanding of users about their decisions into account when providing them with choices in the context of data markets.

Guidance how to involve users in data markets can also be found in the experiences with tools that are intended to raise privacy-awareness in a more general form. Pötzsch (2009) provides a good overview in this regard. Among other points, she recommends to remind the user about their intentions with regard to privacy, but keep interruptions of the user to a minimum. This may require to measure the privacy attitude of the user beforehand, which can be difficult because of the privacy paradox (see above). An example of such an approach which requires initial learning can be found in a machine-learning algorithm for the privacy-settings of social networks proposed by Fang, Kim, LeFevre, and Tami (2010). A different way to keep the interruptions for the user minimal is to use cues as part of the workflow that function more as a nudge instead of an interruption. Such techniques can for example be found in the studies of Pötzsch, Wolkerstorfer, and Graf (2010) and Wang et al. (2013). For the case of data markets, however, such softer forms of interruptions are probably difficult to achieve, since data will most likely be shared for secondary uses which by definition are not part of the regular workflow of a user.

Overall, there is no easy answer to the question how users would engage in data markets. The extent to which users would share data based on monetary incentives is contingent upon the valuation of privacy, which is difficult to measure in a more general form. Also, existing research indicates that it will probably not be possible to model the behaviour of users in a simple economic model in the form of a cost-benefit analysis. The 'utility function' of a user in this regard is much more complex and dependent upon a number of subtle factors that are influenced through a specific implementation of a data market approach. These factors include the precise wording that

is used, how monetary gains and losses are presented in relation to each other, the extent to which a feeling of ownership for data is created, the trust relationship with the institutions involved, the level of pre-existing knowledge that is presupposed and how the decision to sell is integrated in the workflow of primary applications. Considering the complexity that existing research shows in this regard, it seems unlikely that a general utility function to model the willingness to sell can be established. The aforementioned research results, however, provide insights which aspects to pay attention to if a data market implementation should be designed.

conclusions for the selected data market approaches The level of detail provided in the literature on the selected data market approaches is not high enough in order to assess the more delicate characteristics connected to user behaviour. Yet, the approaches differ through the features that they incorporate. This makes it possible to draw some more general conclusions which of the possible effects summarized in Table 5.1 are likely to occur. In the following, I draw such conclusions for each of the data market approaches selected in Section 4.5 and thereby describe how they deviate from the more general evaluation provided above.

A particular detail about the *national information market* is the fact that data would be sold in baskets. Since the price would not have to be determined per individual, but for a whole basket, this might help to overcome some of the difficulties concerning price-setting (#2). The efficiency within the data market itself might therefore be higher with the national information market than with other approaches.

Another noteworthy aspect concerning the national information market is its planned scope. As the name suggests, it is supposed to be of national scope, which is the United States in the context of the writing of Laudon (1996). The limitation of trade within a specific nation would somewhat alleviate possible problems due to income disparities between market participants (#12).

The proposal of *hybrid inalienability* is special with regard to the restrictions on alienability that it would impose. A one-shot alienability of rights would not be possible and the purposes for which the data is used would have to be clearly specified. Therefore, the price would also have to be set for a specific use and not for the data as such. This might somewhat alleviate the difficulties with price-setting (#2). The limitation to a specific use of data and the fact that the user themself would have to be involved in the decision would also support the user in overseeing the consequences of their decision (#4) and somewhat mitigate the problems with weak agency (#11).

Furthermore, the use of sticky policies that would have to be transferred with the data constitutes a transmission principle. This would change the frame under which the data are sold, thereby somewhat lessening the problem that the organisation buying it is under the impression to have absolute rights to the data (#14, #29). Also, the user selling the data would be able to make explicit that they still values the data and only provides it for a specific purpose (#15).

In the case of *personal information markets* three features are of particular importance. One is the requirement that service providers would have to offer a privacy-friendly, and possibly paid, service. This would give users a meaningful choice compared to a take-it-or-leave-it situation, thereby alleviating the problem that desired options might not be available (#6). That the existence of a paid, but privacy-friendly option would be preferred by at least some users is shown through the experiment by Savage and Waldman (2013) discussed above.

Also the fact that trade in the rich information space would be restricted to anonymous information and all other trade would be prohibited is of influence. On the one hand, this would soften the frame of unimportance that the selling of data creates (#15). On the other hand, users are not even

provided with the possibility to sell non-anonymous data concerning themself. This is a restriction in freedom and autonomy compared to the other data market approaches (#3, #24).

Finally, the possibility of service providers to sell data without permission of the user is remarkable[16]. Combined with the fact that users would only be paid indirectly for data that the service provider sells, this would lead to a number of consequences. Users would not get used to always being paid for their data (#16, #17). Also, the perception of ownership would not be strengthened (#27) as with other approaches. On the other hand, this would better support the symbiotic relation between user and service provider (#30).

The proposal for a *humanistic information economy* is special in a number of points. Central is the idea that the economy should include all value that is created based on data and that the users from whom the data originates are part of it. This has multiple consequences in the way that an economic exchange of data would communicate. Instead of the frame that selling data means that it is not important, the aspect that the user has to be paid for every data contribution would be prevalent (#15). Also the frame that users would get accustomed to be paid for data (#16) might be acceptable since it is an intended consequence and the whole system would be designed for this very aspect. Finally, the economic relevance of of personal data would be emphasized in a very clear way (#18).

The two-component price-setting mechanism would in part alleviate the problems with price-setting (#2), but still leave it up to the user to determine the instant component. The automatic tracking of links for the establishment of prices as part of the legacy component would be a recognition of symbiotic relations in the creation of data mash-ups (#30).

A final point to consider is the possibility of a basic income through data (#13). The problem that the reliance on an income from data constitutes has to do with the vulnerability in market exchanges that this could create. In the case of the humanistic information economy, however, this is again the very idea that would constitute the approach itself. Lanier (2013) argues that in his vision of the humanistic information economy everyone should be able to make a living from their contributions to society in digital form. If this project would be implemented in this all-embracing form, a basic income from data would be the standard and thereby not constitute a vulnerability anymore[17].

The most deviations from the above more general analysis would occur under the status quo. The prevalent arrangement that service providers sell user data in order to offer their services free of charge has some resemblances, but still differs from a full data market in a number of points. Paramount is the point that users do not have direct access to the data market and it is only the service providers who can sell data on behalf of the user. This means that users do not have a choice in this regard (#3, #19, #24), but also somewhat changes the frame that data can be easily disposed of (#15). Furthermore, this makes the price-setting easier as it concerns larger transaction volumes (#2).

Besides the missing access to the data market, users are also not paid in direct form for the data they contribute. The only payment they receive in exchange is that of the free service, for which it is not clear whether it reflects the worth of their data. This means that a data market cannot create vulnerabilities in the form of a basic income (#13). Also, users are not getting used to be paid

[16] One might argue here that this is not of relevance since all data has to be anonymised and users are not identifiable any longer. However, empirical research indicates that users are also concerned about the use of anonymised data (Chellappa & Sin, 2005).

[17] NB: I introduced the point of a basic income as vulnerability above based on a TV-interview with Lanier (vpro, 2013). If put in context for his vision of the humanistic information economy, the idea becomes less objectionable. If, however, a basic income based on data would become possible in the status quo, this might lead to vulnerability.

for data (#16), but are instead used that everything online is available for 'free'. Furthermore, the problem of overdetermination through financial incentives does not exist (#17), but instead exists in the form of incentives for data sharing through free services. Moreover, the lack of payment fails to communicate the economic importance of data (#18) and does not strengthen the perception of ownership (#27).

Moreover, the status quo has a different emphasis on the symbiotic relationship between user and service provider (#30). As ongoing debates about usage rights and default entitlements show, also in this case not all questions in this regard are settled. However, it supports a symbiosis much better than strong default entitlements towards one of the two parties would do. A corollary of the status quo is that it already exists and cannot be introduced, meaning that also no redistribution of default entitlements would take place (#9).

The short discussion concerning the specific data market approaches has shown that the different consequences identified in Table 5.1 would occur to a different extent. However, none of the data market approaches would tackle all of the problems identified in the more general discussion above. Also, it would depend on further implementation details and surrounding conditions in which way these consequences would arise. A more specific analysis would therefore only be possible for a very specific case and cannot be performed for an only mediocre well-defined proposal from the literature. But also the more general insights identified in this chapter are very beneficial as such and I draw general conclusions based on them in the following chapter.

6 Conclusions

The goal of this study was to investigate in which ways data markets would impact the privacy of online service users in the European Union. As the ethical evaluation in the preceding chapter has shown, there are a number of possible effects that the introduction of data markets would have. Most importantly, data markets would be of influence for individual autonomy in a number of ways and precautions would have to be taken that the negative effects in this regard do not occur. Moreover, data markets would affect the symbolic conditions upon which data is transferred, thereby influencing the norms with which personal data is transferred and used. To which extent these effects occur is largely contingent upon specific implementation details and a number of parameters of the context they are embedded in. For better clarity, I provide a summary of these findings in Section 6.1.

It is needless to say that the results of this study do not stand on their own, but integrate into the existing body of research. I therefore provide reflection on the contributions that this study has made and an outlook on possible directions for future research in Section 6.2.

As outlined in Chapter 1, the central question of this study is of interest because of a larger debate on the protection of user privacy. Within this larger debate, data markets are proposed as a possible solution to problems with privacy protection. As I have shown, data markets can lead to a number of effects that are detrimental in this regard. In this light, it therefore is of interest under which conditions data markets would be beneficial to the protection of user privacy. I briefly sketch these prerequisites in Section 6.3.

It furthermore is of interest in which way data markets differ from the current situation in Europe. As pointed out in in Chapter 1, there exists the status quo of free services, combined with data protection partially failing on a procedural level. Since it is this context where the discussion about data markets as a solution to privacy problems takes place, it is of relevance to compare the effectiveness of privacy protection through data markets with this situation. A definitive answer on this point would require further research, but I aim to arrive at a preliminary conclusion in this regard in Section 6.4.

Based on these considerations, I close the chapter with policy recommendations in Section 6.5.

6.1 Summary of the findings

As the ethical evaluation has shown, there is no easy answer to the question in which ways data markets would impact the privacy of online service users. There are a number of different effects that play a role and different parameters of the institutional design that have to be taken into account. As users would have a strong role within data markets, the behaviour of users and design parameters influencing them are of relevance as well. Although research findings about human psychology are helpful in this regard, it is not possible to completely predict the behaviour of users. Most interesting in terms of immediate findings that this study provides are therefore the possible effects of data markets and parameters that are of influence.

possible effects of data markets As I have shown in Sections 5.3 and 5.4, there are a number of different effects that could occur through the introduction of data markets. I identified these effects based on an ethical evaluation focussing on values that are of relevance for the function of

markets and moral reasons for the protection of privacy. For a complete overview, please refer to Table 5.1 on page 81.

The most important findings concerning the possible effects of data markets are:

- Data markets might lead to an increase in *efficiency* in the market for goods. The allocative efficiency in the data market itself, however, is not straightforward to achieve. Data markets might therefore have a mixed influence on the allocative efficiency throughout different markets.
- Data markets would influence individual *freedom* in multiple ways: The option to sell data in a market constitutes an enhanced freedom of choice and is less paternalistic than current forms of data protection. However, it is difficult to ensure that these decisions account for possible adverse consequences and might corrupt *authenticity*. In this regard, the effect on freedom depends on the question whether a positive or negative conception of freedom is considered.
- Connected to this aspect is the influence on *autonomy*. The decision to sell personal data might go coupled with a loss of autonomy in the market for online services and the market for goods. This aspect is somewhat intricate as it concerns different points in time and different markets in which the user engages.
- Data markets would strengthen the perception of *ownership* in data. This might be beneficial for privacy in case of default entitlements towards users. On the other hand, it might distort the symbiotic relationship between users and service providers. In every case, it would lead to distributional questions, thereby affecting *equality*.
- *Equality* is furthermore of importance, since existing discrepancies can lead to a *vulnerability* of users in data markets.
- Data markets could furthermore lead to a *crowding out* of morals. The symbolic change to the provision of data for monetary remuneration alters the frame connected to it, thereby affecting the status of privacy as a human right.
- Through the actual use of data which has been sold, data markets can lead to *harm*. This problem is not inherent to data markets, but might increase in relevance if users provide data more easily because of monetary incentives. Even in the absence of actual harm, data markets can lead to an increase of *subjective privacy harm*, if the possible use of data is not known to users.
- Finally, it is important to recognize the importance of privacy as a *common good*. This means that data markets would not only affect the individual, but also society at large.

relevant parameters These effects are all only possible effects of data markets and would not necessarily occur in every case. While they are strongly contingent upon the actual behaviour of users in these markets, they also depend on a number of design parameters. The structural analysis using the framework of contextual integrity (Nissenbaum, 2010) in Section 5.1 was helpful to identify some of these parameters.

The most relevant parameters affecting the influence of data markets on privacy are:

- The extent to which the recipient, i.e. the buyer, of personal data is known to users. Although it is an inherent characteristic of markets to be open on this point, restrictions in this regard can be beneficial for privacy.
- The scope of the data market as a whole. Data markets which serve as a connection between multiple social contexts are much more likely to negatively affect privacy. This is connected to the more general problematic of secondary use. A limitation of use removes some of the value that data might generate, but is beneficial for privacy.
- The degree to which individuals are involved in specific transactions. Completely delegating the sell of data to a service provider collecting it keeps the user out of the loop, thereby neg-

atively affecting autonomy. An involvement of the user on a per-transaction basis, however, might not be feasible in practice.
- The technical design of a data market influences the extent to which privacy-relevant market metadata are generated. Such data are to some extent necessary for a market to function, but lead to additional privacy concerns that would not be existent without the market.

existence of market metadata The mere existence of market metadata deserves consideration. In order for data markets to function as a social institution and as a technical artefacts, additional information streams are necessary. For the handling of market transactions, a centrally accessible data storage is necessary and also the prospective buyers of data have to be supplied with information. This shows that data markets as a tool to achieve a specific purpose are not neutral in this regard. If they are deployed as a tool for privacy protection, the possibly detrimental effects on privacy through metadata have to be taken into account as well.

differences between the selected data market approaches The data market approaches that I selected for further analysis (see Section 4.5) were helpful in order to determine some of the relevant parameters and to gain more insights into the existence of market metadata. To this end, it were especially the differences between data market approaches that were helpful for this study. However, none of the data market approaches was specific enough in order to fully assess the ethical implications thereof. In this regard, the selected data market approaches have to be seen as a first step towards the creation of market structures for personal data, but not as a complete policy proposal.

6.2 Contributions and future work

In this section, I want to outline in which ways this study contributes to the existing body of research. I do so by listing the contributions that this study made and giving indications in which ways these contributions are of use for further action. However, more research is necessary in order to arrive at more robust conclusions and to better integrate the results of this study with the work in different fields. I therefore also provide an overview over possible directions for future work.

contributions of this study Foremost, this study provides important insights into the possible effects of data markets and the parameters that are of importance in this regard:
- The structural analysis using the framework of contextual integrity (Nissenbaum, 2010) in Section 5.1 provides pointers to these parameters of data markets that affect the protection of privacy. These are of relevance for researchers working on the institutional design of data market approaches and the development of technical solutions for data market infrastructure.
- The possible effects of data markets on the privacy of online service users are the main outcome of the ethical evaluation in Chapter 5. They serve as a caution for policy makers regarding premature conclusions about the effectiveness of data markets for the protection of privacy. The different effects as such are of importance in order to better understand controversies about the protection of privacy. The overview can furthermore serve as starting point for more detailed ethical evaluations on the effects of data markets.

Furthermore, this study advanced the literature on privacy and the economics of data trade in a number of ways:
- The literature review on the different conceptions of privacy (see Chapter 2) is characterized by its discussion of the notion of privacy from three different angles (definitions, methods,

and reasons). It provides an overview over the different meanings of privacy and explains why we ought to protect privacy. As such, it is useful for other researchers who are confronted with the diverse meanings that privacy can have.

- The literature review on data markets (see Chapter 4) includes an overview over data market approaches as discussed in the literature and those starting to appear in practice. This is of use for researchers who want to start similar investigations into the field of data markets.
- The description of the 'free service status quo' in Section 4.5, and more generally the integration of economic aspects with privacy concerns throughout this study, show that the existing situation is to some extent similar to that of a data market. This insight is of importance for researchers who investigate data markets and want to contrast them with the existing situation.

Overall, the results of this study serve as input for policy discussions on the protection of privacy. Policy makers who are confronted with different policy options concerning mechanisms for the protection of privacy receive useful pointers for further investigations that are necessary. Together with knowledge from existing studies on the legal perspective in Europe (see e.g. Purtova, 2011) an the economics of personal data (see e.g. Acquisti, 2010), this study helps to understand in which ways data markets can serve as a solution to current problems with the protection of privacy.

pointers for future work First, it seems as if first instances of data markets are about to emerge in practice (see Section 4.2). It therefore is of interest to study the effects that these *real instances of data markets have* and compare the results to the theoretically grounded findings of this study.

Second, in order to better integrate the role of the user into the results of this study, more research from the fields of *psychology* and *behavioural economics* should be integrated in the analysis of data markets:

- Through the integration of more results from these fields in a systematic manner, it might be possible to arrive at better predictions concerning the decisions that users would make in a data market setting.
- Where existing research findings are not sufficient in order to anticipate user behaviour, new experiments might help to further the knowledge in this regard. However, it has to be taken into account that it will not be possible to completely predict human decisions and data market approaches should be able to deal with this uncertainty.

Third, this study focussed on the most general aspects of data markets. However, there are a number of intricate *corner cases* that should be further studied by means of an ethical evaluation:

- The *role of minors* in online environments is in many cases somewhat precarious. If data markets should be established as a means of privacy protection, it would have to be researched to which extent and in which form minors should be allowed to participate in these markets.
- The balance between the *interest of the government to collect data* for its purposes and the right to privacy is already a difficult issue. In the case of data markets, this balance would have to be expressed in monetary terms and default entitlements would decide on the question for which purposes the government would have to pay for access to data.
- This study assumed that it is specific individuals who participate in data markets. As the discussions about rights for *collective privacy* show, there are cases in which groups of individuals are affected. This raises the question how groups should be involved in data markets regarding the collective decisions to take and the distribution of financial remuneration.

Fourth, this study was conducted from an ethical perspective and only incorporated some of the available knowledge from the field of economics. However, it would be helpful to better integrate

the findings of this study back into the *perspective of economics*. Two different angles would be particularly interesting in this regard:

- One way to characterize the uncertainty with which users would have to sell their data in data markets is that of *incomplete contracting* (Hart, 1995). It would therefore be relevant to express the activities within data markets in terms of incomplete contracting and thereby integrate it into existing research from the field of economics.
- Within the status quo of free services, the relation between users and service providers can be characterized as an exploitation of prosumer labour (Fuchs, 2011). Therefore, a comparison between the free service status quo and data markets through the lens of *Marxian economics* might be insightful in order to analyse in which ways data markets would change this way of exploitation.

6.3 Prerequisites for the protection of privacy

This study has shown that data markets can lead to a number of effects that are detrimental to the privacy of online service users engaging in them. If data markets should be employed in order to help protecting privacy, it is important to set them up in such a way that these effects are controlled for. Although it seems doubtful whether it is possible to design data markets in such a way that none of the negative effects occur, they should at least be designed in such a way that minimizes these effects. In the following, I give an overview over the most important prerequisites that can be derived from the results of this study.

An important aspect is the degree and form in which users are involved in the market. In order for a data market to protect privacy, users should be aware about the transactions concerning their data and the ways in which their data is used. This requires users to be directly involved in the market or at least to clearly specify the use of their data beforehand if the service providers sells data on behalf of the user. Users should also be able to limit the group of recipients of their data, meaning to be able to sell in specialized markets or at least reject transactions with buyers they do not want to share data with. From the perspective of privacy, the data market would most optimally constitute a context on its own. However, this is unrealistic to be economically interesting, since the economic value will mostly come from other contexts. It therefore is important that data markets are installed in such a way that contexts get interconnected without violating their respective norms.

If data are not provided manually but gathered as part of a service, it is important that users are still in control whether the data is sold on the market. This means that users should be able to select which data is sold on markets and are offered an option to use the service, against payment, without the requirement to sell data. As the service provider sells information on behalf of the user, it is important that the user consents into it and, as just mentioned above, can make a meaningful choice to whom to sell the data.

In every case, it is important that transparency about the use of data exists. Service providers have to offer transparency to users about the ways in which they trade data on behalf of the user. Also, users should be able to communicate the ways in which data may be used. The use of sticky policies has the advantage that also the buyers of data are aware of these constraints. Transparency is important for the user to take authentic decisions, but also to prevent subjective privacy harm on the side of the user. It furthermore helps to frame the transaction in the way that the user wishes it to be constituted.

Moreover problematic is the form in which transparency is provided. Important is to involve the user in a way that accounts for the effect of bounded rationality and other relevant psychological distortions that exist. Best practices in this regard do not exist yet and method to do so are a matter

of ongoing research[1]. Of particular importance in this regard is the time lag between the decision to sell data and the points of time at which it is gathered or used. If possible, these points of time should be identical, meaning to involve the user at all points of time and not only to ask for a blanket permission upfront.

The technical design of the market infrastructure is also of relevance, as it can help to minimize the market metadata that are necessary in order for the market to function. Market metadata as such are an inevitable requirement of a data market, but the number of recipients and the attributes necessary can be limited as much as possible. This means that the application of privacy-enhancing technologies is also of relevance for data markets as such. In every case, however, it has to be taken into account that this form of metadata exists and the possibly detrimental effect on privacy has to be considered.

Finally, it is important to not only consider the data market design itself, but also the context that the data market is supposed to be employed in. Extreme disparities in income and the participation out of financial need should be avoided, which means that data markets should be employed on a more local scale and appropriate social security measures should be available.

6.4 Comparison of data markets with the status quo

As discussed in Chapter 1, the EU Data Protection Directive (European Parliament, 1995) partially fails in the application of the procedures that it provides for. Additionally, the strong demand for the secondary use of data leads to tensions with the current approach to data protection. Coupled with the prevalent business model of free services which aggravate these problems, there is the demand to take up work on alternative forms of privacy protection. Since data markets are one of the solutions which are discussed in this regard, it is of relevance to compare the effectiveness of data markets against the status quo of data protection.

Many of the possibly adverse effects of data markets are connected to issues of transparency and the involvement of users in the decision which data may be used for which purposes. The points and problems connected hereto are similar to the elements that the current EU Data Protection legislation also struggles with. If designed in the right way, data markets could in principle help to improve these points and thereby make a contribution to the protection of privacy. However, this would for the most part also be possible with improvements to the current data protection legislation, so that at least the need for data markets in this regard would be questionable.

An exception to this last point is the awareness for the economic value of data that data markets would create. This is in every case a unique advantage of data markets. How this awareness translates into user decisions to sell data is contingent upon other implementation details. Although the economic awareness as such is to be seen as positive, the further effects that it creates might be beneficial or detrimental towards the protection of privacy. Moreover, the financial frame that would be coupled to data transactions could lead to a number of 'crowding out' problems that might disturb the perception of privacy as a human right.

Furthermore, the societal consequences of data markets have to be accounted for. The current approach to data protection is somewhat paternalistic in its form and can more easily deal with this requirement. The increased freedom of choice that data markets would give to individuals might therefore be more troublesome in this regard. With data markets, users are effectively asked to weigh the financial inducement they receive against the probabilistic harm that others or society at large will suffer. This is in structure similar to a common pool resource problem and it is therefore unlikely that users will act in favour of the common good in this matter.

For these reasons, it seems questionable whether data markets are an adequate solution to the problems with the status quo. If the right precautions are being taken and data markets would be

[1] See e.g. the studies of Pötzsch et al. (2010), van den Berg and van der Hof (2012), or Wang et al. (2013).

employed in the right setting, they might be able to solve some of the current problems. However, it is likely that most of the advantages that data markets might bring in this regard could also be achieved with improvements to the existing data protection legislation. This makes it questionable why such a complex institution introducing its own problems should be set up in the first place.

In order to move forward with the current problems concerning privacy, improvements to the existing data protection mechanisms should be taken into account. Out of the proposed solutions that I presented in Section 1.5, the ones addressing the existing problems with consent are of particular importance. A better representation of privacy notices and the use of automated consent are especially promising in this regard. On a technical level, the further development of privacy enhancing technologies is helpful, but research should focus more strongly on economic incentives for business to employ them. Regarding the problems of secondary data use, the introduction of commonly accepted forms of data collection as proposed by the Federal Trade Commission (2012) in the U.S. context constitutes an interesting concept. With such approaches that could be implemented in line with existing data protection mechanisms, the current problems with privacy protection could at least be partially addressed. The introduction of data markets, on the other hand, would constitute a complete paradigm shift with many more unclear consequences.

6.5 Policy recommendations

This study has shown that data markets are not necessarily a suitable solution to current problems with the protection of privacy in the context of online services. The introduction of data markets would create a new set of problems and there is high uncertainty about possible adverse effects. Based on these findings, I therefore cannot give a recommendation for the introduction of data markets.

In order to address the current problems with the protection of privacy as introduced in Section 1.3, I instead give the following policy recommendations:

1. If data markets are considered as a policy option, further research on the adverse consequences of the specific data market approach under consideration is required.
2. Efforts to integrate the debates on privacy protection in the U.S. and the EU should be undertaken[2]. This is necessary in order to avoid friction due to cultural and regulatory differences.
3. The supervisory authorities set up through the EU Data Protection Directive should be equipped with sufficient funds in order for them to fulfil their obligations. The increase in data processing activities has led to a situation in which they have difficulties to carry out their tasks in an adequate manner.
4. A clear political decision on the default entitlements in the ownership of data is necessary. This would not only be of relevance for the introduction of data markets, but is also required for some of the contemporary issues with data protection such as data portability.
5. Further research on solutions addressing the problems with current data protection mechanisms should be conducted. This includes, for example, solutions towards problems with transparency about data processing practices, solutions to deal with the compatibility of secondary use of data, and research on incentive structures for a wider deployment of PETs.

[2] Recently, an expert group of researchers from the Netherlands and the U.S. has taken up work to bridge the gaps between the regulatory approaches of the two jurisdictions (College Bescherming Persoonsgegevens, 2014). This is a good step in this direction, but a more general integration of the overall debate might be necessary.

References

Acquisti, A. (2004). Privacy in electronic commerce and the economics of immediate gratification. In *Proceedings of the 5th acm conference on electronic commerce* (pp. 21–29).

Acquisti, A. (2009). Nudging privacy: The behavioral economics of personal information. *Security & Privacy, IEEE*, *7*(6), 82–85.

Acquisti, A. (2010, December 1). *The economics of personal data and the economics of privacy* (Vol. 1). Background Paper for OECD Joint WPISP-WPIE Roundtable. Retrieved from http://www.oecd.org/sti/ieconomy/46968784.pdf

Acquisti, A., John, L., & Loewenstein, G. (2009). What is privacy worth. In *Twenty first workshop on information systems and economics (wise)* (pp. 14–15).

Adjerid, I., Acquisti, A., Brandimarte, L., & Loewenstein, G. (2013). Sleights of privacy: framing, disclosures, and the limits of transparency. In *Proceedings of the ninth symposium on usable privacy and security* (p. 9).

Aperjis, C., & Huberman, B. A. (2012). A market for unbiased private data: Paying individuals according to their privacy attitudes. *arXiv preprint arXiv:1205.0030*. Retrieved from http://arxiv.org/pdf/1205.0030v1

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In *The rate and direction of inventive activity: Economic and social factors* (pp. 609–626). UMI. Retrieved from http://www.nber.org/chapters/c2144.pdf

Arthur, C. (2011). Tomtom satnav data used to set police speed traps. *The Guardian*, *April 28*. Retrieved from http://www.theguardian.com/technology/2011/apr/28/tomtom-satnav-data-police-speed-traps

Article 29 Working Party. (2013a, April 2). *Opinion 03/2013 on purpose limitation.* Retrieved from http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf

Article 29 Working Party. (2013b, December 11). *Working party comments to the vote of 21 october 2013 by the european parliamentŠs libe committee.* Retrieved from http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2013/20131211_annex_letter_to_greek_presidency_wp29_comments_outcome_vote_libe_final_en.pdf

Article 29 Working Party. (2014a, April 10). *Opinion 05/2014 on anonymisation techniques.* Retrieved from http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

Article 29 Working Party. (2014b, April 9). *Opinion 06/2014 on the notion of legitimate interests of the data controller under article 7 of directive 95/46/ec.* Retrieved from http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf

Athey, S. (2014). *Information, privacy, and the internet*. CPB Netherlands Bureau for Economic Policy Analysis. Retrieved from http://www.cpb.nl/sites/default/files/CPB-Lecture-2014-Information-Privacy-and-the-Internet-an-economic-perspective.pdf

Awad, N. F., & Krishnan, M. S. (2006). The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for person-

alization. *MIS Quarterly*, *30*(1), pp. 13-28.

Bauer, C., Korunovska, J., & Spiekermann, S. (2012). On the value of information - what facebook users are willing to pay. In *Ecis 2012 proceedings.*

Blaauw, M. (2013). the epistemic account of privacy. *Episteme*, *10*(02), 167–177.

Blok, P. (2002). *Het recht op privacy : een onderzoek naar de betekenis van het begrip 'privacy' in het nederlandse en amerikaanse recht*. Boom Juridische uitgevers.

Blyaert, L. (2013). Met irispact wordt bruno segers echt entrepreneur. *datanews*, *May 24*. Retrieved from http://datanews.knack.be/ict/nieuws/met-irispact-wordt-bruno-segers-echt-entrepreneur/article-4000310991663.htm

Bowles, S. (2008). Policies designed for self-interested citizens may undermine" the moral sentiments": Evidence from economic experiments. *science*, *320*(5883), 1605–1609.

Boyd, D. (2011, June 6). *Networked privacy.* Personal Democracy Forum (New York, NY). Retrieved from http://www.danah.org/papers/talks/2011/PDF2011.html

Brandimarte, L., Acquisti, A., & Loewenstein, G. (2013). Misplaced confidences privacy and the control paradox. *Social Psychological and Personality Science*, *4*(3), 340–347.

Busch, A. (2006). From safe harbour to the rough sea? privacy disputes across the atlantic. *SCRIPT-ed*, *3*(4), 304–21.

Calo, R. (2011). The boundaries of privacy harm. *Indiana Law Journal*, *86*(3), 1131-1162.

Calo, R. (2012). Against notice skepticism in privacy (and elsewhere). *Notre Dame Law Review*, *87*, 1027-1072.

Castroz, D. (2013). *How much will prism cost the u.s. cloud computing industry?* (Vol. August; Tech. Rep.). The Information Technology & Innovation Foundation. Retrieved from http://www2.itif.org/2013-cloud-computing-costs.pdf

Cave, J., Robinson, N., Schindler, R., Bodea, G., Kool, L., & van Lieshout, M. (2011). *Does it help or hinder? promotion of innovation on the internet and citizensŠ right to privacy* (Tech. Rep. No. IP/A/ITRE/ST/2011-10). European Parliament Directorate General For Internal Policies. Retrieved from http://www.europarl.europa.eu/document/activities/cont/201112/20111220ATT34644/20111220ATT34644EN.pdf

Chellappa, R. K., & Sin, R. G. (2005). Personalization versus privacy: An empirical examination of the online consumerŠs dilemma. *Information Technology and Management*, *6*(2-3), 181–202.

Chin, A., & Klinefelter, A. (2012). Differential privacy as a response to the reidentification threat: The facebook advertiser case study. *North Carolina Law Review*, *90*, 1417-1456.

Cohen, J. E. (2000). Examined lives: Informational privacy and the subject as object. *Stan. L. Rev.*, *52*, 1373-1438.

College Bescherming Persoonsgegevens. (2014). Amerikaanse en europese privacy-experts zoeken oplossingen om verschil in privacyregels te overbruggen. *press releases of the CBP*, *May 6*. Retrieved from http://www.cbpweb.nl/Pages/med_20140506_meeting-privacykloof-ivir-mit.aspx (accessed August 17, 2014)

Committee On Commerce, Science, And Transportation. (2013, December 18). *A review of the data broker industry: Collection, use, and sale of consumer data for marketing purposes.* Staff Report For Chairman Rockefeller. Retrieved from http://www.commerce.senate.gov/public/?a=Files.Serve&File_id=bd5dad8b-a9e8-4fe9-a2a7-b17f4798ee5a

Council of Europe. (2010, June). *Convention for the protection of human rights and fundamental freedoms (european convention on human rights).* Retrieved from http://www.echr.coe.int/Documents/Convention_ENG.pdf

Craig, T., & Ludloff, M. E. (2011). *Privacy and big data*. O'Reilly Media, Inc.

Cuijpers, C. (2007). A private law approach to privacy: Mandatory law obliged? *SCRIPT-ed*, *4*(4), 304–318.

Danezis, G., & Gürses, S. (2010, April). A critical review of 10 years of privacy technology. In *Proceedings of surveillance cultures: A global surveillance society?* UK.

de Montjoye, Y.-A., Shmueli, E., Wang, S. S., & Pentland, A. S. (2014). openpds: Protecting the privacy of metadata through safeanswers. *PloS one*, *9*(7). Retrieved from http://dx.plos.org/10.1371/journal.pone.0098790

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research*, *17*(1), 61–80.

Domingo-Ferrer, J. (2011). Coprivacy: towards a theory of sustainable privacy. In *Privacy in statistical databases* (pp. 258–268).

Dunn, W. N. (2008). *Public policy analysis: An introduction* (4th ed.). Pearson.

Eavis, P. (2013). Twitter's market valuation suggests wall st. sees huge growth potential. *DealB%k*, *November 6*. Retrieved from http://dealbook.nytimes.com/2013/11/06/twitters-market-valuation-suggests-wall-st-sees-huge-growth-potential/

European Commission. (2012, January). *Proposal for a regulation of the european parliament and of the council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (general data protection regulation).* Retrieved from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0011:FIN:EN:PDF (2012/0011 (COD))

European Court of Justice (ECJ). (2014, May 13). *Case C 131/12*. Retrieved from http://curia.europa.eu/juris/document/document.jsf?text=&docid=152065&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=208758

European Parliament. (1995, October). *Directive 95/46/EC*. Retrieved from http://ec.europa.eu/justice/policies/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf

European Parliament. (2002, July). *Directive 2002/58/EC*. Retrieved from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:en:HTML

European Parliament. (2013a, January 16). *Draft report on the proposal for a regulation of the european parliament and of the council on the protection of individual with regard to the processing of personal data and on the free movement of such data (general data protection regulation).* Retrieved from http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+COMPARL+PE-501.927+04+DOC+PDF+V0//EN (Rapporteur: Jan Philipp Albrecht)

European Parliament. (2013b, October 22). *Regulation of the european parliament and of the council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (general data protection regulation).* Retrieved from http://www.janalbrecht.eu/fileadmin/material/Dokumente/DPR-Regulation-inofficial-consolidated-LIBE.pdf (inofficial consolidated version after libe committee vote provided by the rapporteur)

European Union. (2010). *EU Charter of Fundamental Rights*. Official Journal of the European Union. Retrieved from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2010:083:0389:0403:en:PDF (C 83/02)

Fang, L., Kim, H., LeFevre, K., & Tami, A. (2010). A privacy recommendation wizard for users of

social networking sites. In *Proceedings of the 17th acm conference on computer and communications security* (pp. 630–632). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/1866307.1866378 doi: 10.1145/1866307.1866378

Federal Trade Commission. (2012, 3). *Protecting consumer privacy in an era of rapid change.* Retrieved from http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf

Floridi, L. (2014). Open data, data protection, and group privacy. *Philosophy & Technology*, *27*(1), 1-3.

Foster, S. L. (2014). Timing update for the eu data protection regulation: No news doesn't mean it's gone away. *The National Law Review*, *July 24*. Retrieved from http://www.natlawreview.com/article/timing-update-eu-data-protection-regulation-no-news-doesn-t-mean-it-s-gone-away

Fuchs, C. (2011, January 13). *The political economy of privacy on facebook* (The Internet & Surveillance - Research Paper Series No. 9). Vienna, Austria: Unified Theory of Information Research Group. Retrieved from http://www.sns3.uti.at/wp-content/uploads/2010/09/The-Internet-Surveillance-Research-Paper-Series-9-Christian-Fuchs-The-Political-Economy-of-Privacy-on-Facebook.pdf

Gavison, R. (1980). Privacy and the limits of law. *Yale law journal*, *89*(3), 421–471.

Gosepath, S. (2011). Equality. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Spring 2011 ed.). Retrieved from http://plato.stanford.edu/archives/spr2011/entries/equality/

Gutwirth, S., & Hert, P. (2008). Regulating profiling in a democratic constitutional state. In M. Hildebrandt & S. Gutwirth (Eds.), *Profiling the european citizen* (p. 271-302). Springer Netherlands.

Hart, O. (1995). *Firms, contracts, and financial structure*. Oxford university press.

Hermalin, B. E., & Katz, M. L. (2006). Privacy, property rights and efficiency: The economics of privacy as secrecy. *Quantitative Marketing and Economics*, *4*(3), 209–239.

Hildebrandt, M. (2006). Profiling: From data to knowledge - the challenges of a crucial technology. *Datenschutz und Datensicherheit*, *30*, 548-552. Retrieved from http://www.fidis-project.eu/fileadmin/fidis/publications/2006/DuD09_2006_548.pdf

Hildebrandt, M., & Koops, B.-J. (2010). The challenges of ambient law and legal protection in the profiling era. *The Modern Law Review*, *73*(3), 428–460.

Hill, K. (2012). How target figured out a teen girl was pregnant before her father did. *Forbes*, *Februrary, 16*. Retrieved from http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/

Hoepman, J.-H., & van Lieshout, M. (2012). Privacy. In E. R. Leukfeldt & W. P. Stol (Eds.), *Cyber safety: An introduction* (p. 75-87). Boom Lemma.

Honoré, A. M. (1961). Ownership. In A. G. Guest (Ed.), *Oxford essays in jurisprudence* (p. 107-147). Oxford University Press.

Hoofnagle, C. J., & Urban, J. M. (2014). Alan westin's privacy homo economicus. *Wake Forest Law Review*, *261*. Retrieved from http://ssrn.com/abstract=2434800

Hornung, G., & Schnabel, C. (2009). Data protection in germany i: The population census decision and the right to informational self-determination. *Computer Law & Security Review*, *25*(1), 84–88.

Izyumenko, E. (2011). *Think before you share: Personal data on the social networking sites in europe; article 8 echr as a tool of privacy protection* (Master's thesis, Lund University - Faculty of Law). Retrieved from http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=2027887&fileOId=2027888

Jentzsch, N. (2010). A welfare analysis of secondary use of personal data. In *Workshop on the economics of information security* (Vol. 2010). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.165.2477&rep=rep1&type=pdf

Jentzsch, N., Preibusch, S., & Harasser, A. (2012, 2). *Study on monetising privacy: An economic model for pricing personal information* (Tech. Rep.). European Union Agency for Network and Information Security. Retrieved from http://www.enisa.europa.eu/activities/identity-and-trust/library/deliverables/monetising-privacy

Kokott, J., & Sobotta, C. (2013). The distinction between privacy and data protection in the jurisprudence of the cjeu and the ecthr. *International Data Privacy Law*, *3*, 222-228.

Koot, M. R. (2012). *Measuring and predicting anonymity* (Doctoral dissertation, Universiteit van Amsterdam). Retrieved from http://dare.uva.nl/document/444044

Kroes, N. (2013, 3). *The big data revolution.* Speech. Retrieved from http://europa.eu/rapid/press-release_SPEECH-13-261_en.htm (European Commission - SPEECH/13/261)

Kuner, C., Burton, C., & Pateraki, A. (2014). The proposed eu data protection regulation two years later. *Privacy & Security Law Report*. Retrieved from http://www.wsgr.com/eudataregulation/pdf/kuner-010614.pdf

Lanier, J. (2011). *You are not a gadget*. Penguin.

Lanier, J. (2013). *Who owns the future?* Allen Lane.

Lanier, J. (2014). Wer die daten hat, bestimmt unser schicksal. *Frankfurter Allgemeine Zeitung*, *April 24*. Retrieved from http://www.faz.net/aktuell/feuilleton/debatten/googles-datenmacht-wer-die-daten-hat-bestimmt-unser-schicksal-12907065.html

Laudon, K. C. (1996, September). Markets and privacy. *Commun. ACM*, *39*(9), 92–104. Retrieved from http://doi.acm.org/10.1145/234215.234476 doi: 10.1145/234215.234476

Le M'etayer, D. (2009). A formal privacy management framework. In P. Degano, J. Guttman, & F. Martinelli (Eds.), *Formal aspects in security and trust* (Vol. 5491, p. 162-176). Springer Berlin / Heidelberg.

Lessig, L. (2006). *Code: Version 2.0*. Basic Books.

Levine, M. H. (1980). Privacy in the tradition of the western world. In W. C. Bier (Ed.), *Privacy: A vanishing value* (pp. 3–21). Fordham University Press. Retrieved from http://books.google.nl/books?id=rofFPZc0nooC

London Economics. (2010, July). *Study on the economic benefits of privacy-enhancing technologies (pets).* Retrieved from http://ec.europa.eu/justice/policies/privacy/docs/studies/final_report_pets_16_07_10_en.pdf

Mantelero, A., & Vaciago, G. (2013). The dark side of big data: Private and public interaction in social surveillance: How data collections by private entities affect governmental social control and how the eu reform on data protection responds. *Computer law review international*, *6/2013*(6), 161–169.

Martin, K. (2013). Transaction costs, privacy, and trust: The laudable goals and ultimate failure of notice and choice to respect privacy online. *First Monday*, *18*(12). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/4838

Mattioli, D. (2012). On orbitz, mac users steered to pricier hotels. *The Wall Street Journal*, *August 23*. Retrieved from http://online.wsj.com/news/articles/SB10001424052702304458604577488822667325882

Mayer, I. S., van Daalen, C. E., & Bots, P. W. (2004). Perspectives on policy analyses: a framework for understanding and design. *International Journal of Technology, Policy and Management*, *4*(2), 169–191.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Hodder and Stoughton.

McDonald, A. M., & Cranor, L. F. (2008). Cost of reading privacy policies, the. *ISJLP*, *4*, 543.

McDonald, A. M., & Cranor, L. F. (2010). Americans' attitudes about internet behavioral advertising practices. In *Proceedings of the 9th annual acm workshop on privacy in the electronic society* (pp. 63–72). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/1866919.1866929 doi: http://doi.acm.org/10.1145/1866919.1866929

Miller, A. R. (1971). *The assault on privacy: Computers, data banks, and dossiers*. University of Michigan Press.

Minister en Staatssecretaris van Veiligheid en Justitie, & Minister van Binnenlandse Zaken en Koninkrijksrelaties. (2013, December 13). *Vrijheid en veiligheid in de digitale samenlevings.* Tweede Kamer, vergaderjaar 2013-2014, 26 643, nr. 298. Retrieved from https://zoek.officielebekendmakingen.nl/kst-26643-298.pdf

Morando, F., Iemma, R., & Raiteri, E. (2014). Privacy evaluation: what empirical research on users' valuation of personal data tells us. *Internet Policy Review*, *3*(2). doi: 10.14763/2014.2.283

Morozov, E. (2013a). Information consumerism: The price of hypocrisy. *Frankfurter Allgemeine Zeitung*, *July 24*. Retrieved from http://www.faz.net/aktuell/feuilleton/debatten/ueberwachung/information-consumerism-the-price-of-hypocrisy-12292374.html

Morozov, E. (2013b). The real privacy problem. *MIT Technology Review*, *11*. Retrieved from http://www.technologyreview.com/featuredstory/520426/the-real-privacy-problem/

Morozov, E. (2014). Selling your bulk online data really means selling your autonomy. *New Republic*, *May 13*. Retrieved from http://www.newrepublic.com/article/117703/selling-personal-data-big-techs-war-meaning-life

Nagenborg, M. (2009). Designing spheres of informational justice. *Ethics and Information Technology*, *11*(3), 175-179.

Narayanan, A., & Shmatikov, V. (2006). How to break anonymity of the netflix prize dataset. *CoRR*, *abs/cs/0610105*. Retrieved from http://arxiv.org/abs/cs/0610105

Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.

Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus*, *140*(4), 32–48.

Nissenbaum, H. (2014). Respect for context as a benchmark for privacy online: What it is and isn't. In C. Dartiguepeyrou (Ed.), *Cahier de prospective - the futures of privacy* (p. 19-30). Think Tank Futur Numérique.

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, *41*(1), 100–126.

Novotny, A., & Spiekermann, S. (2013). Personal information markets and privacy: A new model to solve the controversy. In *Wirtschaftsinformatik* (p. 102).

Odlyzko, A. (2003). Privacy, economics, and price discrimination on the internet. In *Proceedings of the 5th international conference on electronic commerce* (pp. 355–366). New York, NY,

USA: ACM. Retrieved from http://doi.acm.org/10.1145/948005.948051 doi: 10.1145/948005.948051

OECD. (1980, September). *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data.* Retrieved from http://www.oecd.org/sti/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm

Oetzel, M. C., & Spiekermann, S. (2013). A systematic methodology for privacy impact assessments: a design science approach. *European Journal of Information Systems*, *2013*, 1-25.

Olejnik, L., Tran, M.-D., & Castelluccia, C. (2014). Selling off privacy at auction. In *Ndss.* San Diego, CA.

Palmer, D. (2013). We should replace the word 'cloud' with 'somebody else's computer', says security expert. *computing*, *December 2*. Retrieved from http://www.computing.co.uk/ctg/news/2316368/we-should-replace-the-word-cloud-with-somebody-elses-computer-says-security-expert

Pentland, A. (2009). Reality mining of mobile communications: Toward a new deal on data. In *The global information technology report 2008-2009* (p. 75-80). World Economic Forum. Retrieved from http://hd.media.mit.edu/wef_globalit.pdf

Persson, M. (2014). Controle over de toekomst. *de Volkskrant*, *March 15*.

Pfanner, E. (2013). France proposes an internet tax. *The New York Times*, *January 20*. Retrieved from http://www.nytimes.com/2013/01/21/business/global/21iht-datatax21.html

Polonetsky, J., & Tene, O. (2013). Privacy and big data - making ends meet. *Stan. L. Rev. Online*, *66*(25).

Posner, R. A. (1981). The economics of privacy. *The American economic review*, *71*(2), 405–409.

Pötzsch, S. (2009). Privacy awareness: A means to solve the privacy paradox? In V. Matyáš, S. Fischer-Hübner, D. Cvrcek, & P. Švenda (Eds.), *The future of identity in the information society* (Vol. 298, p. 226-236). Springer Boston.

Pötzsch, S., Wolkerstorfer, P., & Graf, C. (2010). Privacy-awareness information for web forums: results from an empirical study. In *Proceedings of the 6th nordic conference on human-computer interaction: Extending boundaries* (pp. 363–372). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/1868914.1868957 doi: http://doi.acm.org/10.1145/1868914.1868957

President's Council of Advisors on Science and Technology. (2014, May). *Big data and privacy: A technological perspective.* Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf

Prins, C. (2006). When personal data, behavior and virtual identities become a commodity: Would a property rights approach matter? *SCRIPT-ed*, *3*(4). Retrieved from http://www2.law.ed.ac.uk/ahrc/script-ed/vol3-4/prins.pdf

Purtova, N. (2011). *Property rights in personal data: A european perspective* (Doctoral dissertation, Tilburg University). Retrieved from http://arno.uvt.nl/show.cgi?fid=114387

Purtova, N. (2013). *Default entitlements in personal data in the proposed regulation: Informational self-determination off the table ... and back on again?* (Tech. Rep. No. 016/2013). Tilburg Law School. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2323643

Reding, V. (2014, 1). *Future of the safe harbour agreement in the light of the nsa affair.* Speech. Retrieved from http://europa.eu/rapid/press-release_SPEECH-14-27_en.htm (European Commission - SPEECH/14/27)

Regan, P. M. (2002). Privacy as a common good in the digital world. *Information, Communication & Society*, *5*(3), 382–405.

Resnik, D. (2003). A pluralistic account of intellectual property. *Journal of Business Ethics*, *46*(4), 319–335.

Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy sciences*, *4*(2), 155–169.

Roessler, B. (2005). *The value of privacy*. Oxford: Polity Press.

Roessler, B., & Mokrosinska, D. (2013). Privacy and social interaction. *Philosophy & Social Criticism*, *39*(8), 771–791.

Rust, R. T., Kannan, P., & Peng, N. (2002). The customer economics of internet privacy. *Journal of the Academy of Marketing Science*, *30*(4), 455–464.

Sandel, M. J. (2013). Market reasoning as moral reasoning: Why economists should re-engage with political philosophy. *Journal of Economic Perspectives*, *27*(4), 121-140.

Satz, D. (2012). *Why some things should not be for sale: the moral limits of markets* (Vol. 2). Oxford University Press.

Savage, S., & Waldman, D. M. (2013, October 16). *The value of online privacy.* Available at SSRN: http://ssrn.com/abstract=2341311. doi: http://dx.doi.org/10.2139/ssrn.2341311

Schallaböck, J. (2014, March 13). *Verbraucher-tracking* (Tech. Rep.). iRights law. Retrieved from `http://www.gruene-bundestag.de/fileadmin/media/gruenebundestag_de/themen_az/digitale_buergerrechte/Tracking-Bilder/Verbraucher-Tracking.pdf` (accessed August 17, 2014)

Schiller, B. (2014). Instead of giving google your data for free, now you can donate it for good. *Co.Exist*, *March 28*. Retrieved from `http://www.fastcoexist.com/3026452/instead-of-giving-google-your-data-for-free-now-you-can-donate-it-for-good`

Schmidt, M. (2010). Iedereen is verdachte. *Binnenlands Bestuur*, *September 24*. Retrieved from `http://www.binnenlandsbestuur.nl/iedereen-is-verdachte.306013.lynkx`

Schneider, S. (2010). *Homo economicus–or more like homer simpson?* (Tech. Rep.). Frankfurt: Deutsche Bank Research. Retrieved from `https://www.dbresearch.com/PROD/DBR_INTERNET_EN-PROD/PROD0000000000259291/Homo+economicus+%E2%80%93+or+more+like+Homer+Simpson%3F.pdf`

Schwartz, P. M. (2004). Property, privacy and personal data. *Harvard Law Review*, *117*(7), 2055-2125.

Selinger, E., & Hartzog, W. (2014a). Google can't forget you, but it should make you hard to find. *Wired*, *May 20*. Retrieved from `http://www.wired.com/2014/05/google-cant-forget-you-but-it-should-make-you-hard-to-find/`

Selinger, E., & Hartzog, W. (2014b). Obscurity and privacy. *Routledge Companion to Philosophy of Technology*, *forthcoming*.

Shapiro, C., & Varian, H. (1999). *Information rules: A strategic guide*. Harvard Business Press.

Sholtz, P. (2001). Transaction costs and the social cost of online privacy. *First Monday*, *6*(5). Retrieved from `http://firstmonday.org/ojs/index.php/fm/article/view/859`

Simitis, S. (1987). Reviewing privacy in an information society. *University of Pennsylvania Law Review*, *135*(3), pp. 707-746. Retrieved from `http://www.jstor.org/stable/3312079`

Simonite, T. (2014). Sell your personal data for $8 a month. *MIT Technology Review*, *Februrary 12*. Retrieved from `http://www.technologyreview.com/news/524621/sell-your-personal-data-for-8-a-month/`

Singer, N. (2012, December 8). You for sale: A vault for taking charge of your online life. *The New York Times*. Retrieved from http://www.nytimes.com/2012/12/09/business/company-envisions-vaults-for-personal-data.html?pagewanted=all

Singer, N. (2014). When a health plan knows how you shop. *The New York Times*, *June 28*. Retrieved from http://www.nytimes.com/2014/06/29/technology/when-a-health-plan-knows-how-you-shop.html

Smith, H. J., Dinev, T., & Xu, H. (2011, December). Information privacy research: An interdisciplinary review. *MIS Q.*, *35*(4), 989–1016. Retrieved from http://dl.acm.org/citation.cfm?id=2208940.2208950

Snoeck, D. (2014). Irispact legt boeken neer. *De Tijd*, *May 14*. Retrieved from http://www.tijd.be/nieuws/ondernemingen_technologie/Irispact_legt_boeken_neer.9501570-3130.art

Soenens, D. (2013). 20 miljoen dollar voor belgische start-up van bruno segers. *DMorgen.be*, *December 24*. Retrieved from http://www.demorgen.be/dm/nl/996/Economie/article/detail/1763437/2013/12/24/20-miljoen-dollar-voor-Belgische-start-up-van-Bruno-Segers.dhtml

Solove, D. (2006). A taxonomy of privacy. *University of Pennsylvania Law Review*, 477–564.

Solove, D. (2008). *Understanding privacy*. Harvard University Press.

Spiekermann, S. (2014). Turing test: Why informed consent is now more important than ever. *Der Standard*, *June 10*. Retrieved from http://derstandard.at/2000001888899/Turing-Test-Why-Informed-Consent-is-Now-more-important-than

Spiekermann, S., Korunovska, J., & Bauer, C. (2012). Psychology of ownership and asset defence: Why people value their personal information beyond privacy. In *Proceedings of the international conference on information systems (icis 2012)*. Orlando, Florida.

Thacher, D., & Rein, M. (2004). Managing value conflict in public policy. *Governance*, *17*(4), 457–486.

The Economist. (2013). Stealing from the government. *The Economist*, *November 30*. Retrieved from http://www.economist.com/news/united-states/21590912-cleverer-use-data-and-investigative-collaboration-can-help-cut-fraud-sirfs-up

The Guardian. (2013, 12). *The nsa files*. Retrieved from http://www.theguardian.com/world/the-nsa-files (Retrieved August 17, 2014)

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

United Nations High Commissioner for Human Rights. (2014, June 30). *The right to privacy in the digital age*. Report of the Office of the United Nations High Commissioner for Human Rights. Retrieved from http://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session27/Documents/A.HRC.27.37_en.pdf

U.S. Department of Commerce. (2010, 12). *Commercial data privacy and innovation in the internet economy: A dynamic policy framework*. Retrieved from http://www.ntia.doc.gov/reports/2010/IPTF_Privacy_GreenPaper_12162010.pdf

U.S. Dep't. of Health, Education and Welfare. (1973). *Records computers and the rights of citizens*. (excerpt of the Fair Information Practices available at http://epic.org/privacy/consumer/code_fair_info.html)

van den Berg, B., & van der Hof, S. (2012). What happens to my data? a novel approach to informing users of data processing practices. *First Monday*, *17*(7). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/4010

Van Alsenoy, B., Kosta, E., & Dumortier, J. (2013). Privacy notices versus informational self-determination: Minding the gap. *International Review of Law, Computers & Technology*, *2013*, 1–19.

Van Den Hoven, J. (1997, September). Privacy and the varieties of moral wrong-doing in an information age. *SIGCAS Comput. Soc.*, *27*(3), 33–37. Retrieved from http://doi.acm.org/10.1145/270858.270868 doi: 10.1145/270858.270868

Van Den Hoven, J. (2008). Information technology, privacy, and the protection of personal data. In V. den Hoven & Weckert (Eds.), *Information technology and moral philosophy* (pp. 301–321). Cambridge University Press.

Van de Poel, I., & Royakkers, L. (2011). *Ethics, technology, and engineering: An introduction*. John Wiley & Sons.

Vanderlippe, J. (2005). *Supermarket cards: An overview of the pricing issues.* http://www.nocards.org/overview/. (accessed August 17, 2014)

Vedder, A. (1999). Kdd: The challenge to individualism. *Ethics and Information Technology*, *1*(4), 275–281.

Vila, T., Greenstadt, R., & Molnar, D. (2003). Why we can't be bothered to read privacy policies models of privacy economics as a lemons market. In *Proceedings of the 5th international conference on electronic commerce* (pp. 403–407). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/948005.948057 doi: 10.1145/948005.948057

vpro. (2013, Oktober 28). *Uw persoonlijke data zijn goud waard.* TV documentary broadcasted on Dutch television. Retrieved from http://tegenlicht.vpro.nl/afleveringen/2013-2014/persoonlijke-data.html (accessed August 17, 2014)

Walzer, M. (1983). *Spheres of justice: A defense of pluralism and equality*. Basic Books.

Wang, Y., Leon, P. G., Scott, K., Chen, X., Acquisti, A., & Cranor, L. F. (2013). Privacy nudges for social media: an exploratory facebook study. In *Proceedings of the 22nd international conference on world wide web companion* (pp. 763–770).

Warren, S. D., & Brandeis, L. D. (1890). The right to privacy. *Harvard law review*, *4*(5), 193–220.

Westin, A. F. (1967). *Privacy and freedom*. Atheneum New York.

Whittaker, Z. (2011). Microsoft admits patriot act can access eu-based cloud data. *ZDNet*, *June 28*. Retrieved from http://www.zdnet.com/blog/igeneration/microsoft-admits-patriot-act-can-access-eu-based-cloud-data/11225

World Economic Forum. (2011). *Personal data: The emergence of a new asset class* (Tech. Rep.). World Economic Forum. Retrieved from http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf

Zwick, D., & Dholakia, N. (1999). *Models of privacy in the digital age: Implications for marketing and e-commerce* (Tech. Rep.). Research Institute for Telecommunications and Information Marketing, University of Rhode Island. Retrieved from http://ritim.cba.uri.edu/Working%20Papers/Privacy-Models-Paper%5B1%5D.pdf

Index