Beyond Informed Consent—Investigating Ethical Justifications for Disclosing, Donating or Sharing Personal Data in Research

Christen, Markus ; Domingo-Ferrer, J ; Herrmann, Dominik ; van den Hoven, Jeroen

DOI: https://doi.org/10.1007/978-3-319-61043-6_10

Submission for IACAP 2015 Proceedings

Beyond Informed Consent – Investigating Ethical Justifications for Disclosing, Donating or Sharing Personal Data in Research

Markus Christen, Josep Domingo-Ferrer, Dominik Herrmann, Jeroen van den Hoven

Affiliations

- Markus Christen, Centre for Ethics, University of Zurich, Switzerland, christen@ethik.uzh.ch
- Josep Domingo-Ferrer, UNESCO Chair in Data Privacy, Universitat Rovira i Virgili, Catalonia, josep.domingo@urv.cat
- Dominik Herrmann, Security in Distributed Systems Group, University of Hamburg, Germany, herrmann@informatik.uni-hamburg.de
- Jeroen van den Hoven, Philosophy Section, Delft University of Technology, Netherlands, M.J.vandenHoven@tudelft.nl

Abstract

In the last two decades, we have experienced a tremendous growth of the digital infrastructure, leading to an emerging web ecosystem that involves a variety of new types of services. A characteristic element of this web ecosystem is the massive increase of the amount, availability and interpretability of digitalized information – a development for which the buzzword "big data" has been coined. For research, this offers opportunities that just 20 years ago were believed to be impossible. Researchers now can access large participant pools directly using services like Amazon Mechanical Turk, they can collaborate with companies like Facebook to analyze their massive data sets, they can create own research infrastructures by, e.g., providing data-collecting Apps for smartphones, or they can enter new types of collaborations with citizens that donate personal data. Traditional research ethics with its focus of informed consent is challenged by such developments: How can informed consent be given when big data research seeks for unknown patterns? How can people control their data? How can unintended effects (e.g., discrimination) be prevented when a person donates personal data? In this contribution, we will discuss the ethical justification for big data research and we will argue that a focus on informed consent is insufficient for providing its moral basis. We propose that the ethical issues cluster along three core values – autonomy, fairness and responsibility – that need to be addressed. Finally, we outline how a possible research infrastructure could look like that would allow for ethical big data research.

1. Introduction

Consider the following scenario:

> *Jane is a social science researcher with a broad spectrum of interests. In her study, she wants to understand the connection between the health status of persons and their political preferences across a variety of social and cultural contexts. To do this, she wants access to participants from several European countries for a survey study using a single entry point in a way that she complies with national regulations, given the ethical sensitivity of the data involved. She does not want to pay a fortune for doing this, but she also wants to be sure that all practical issues related to participant payment up to taxation issues are resolved. As the study is multi-disciplinary, she would like to get suggestions from fellow researchers on how the many tricky*

practical details can be resolved; e.g., regarding translation of survey items. Furthermore, she wants to establish a trusted relationship with a sub-group of participants such that these people are willing to engage in follow-up web experiments and donate personal data and access to personal text written on social networks. She wants to be sure that these participants contribute to this study based on an informed decision and that they are enabled to donate data in a privacy-respecting way. Finally, after having finalized her study, she wants to make accessible the collected data in a way that their access and re-use is easy and complies with the European data protection regulations.

This fictitious scenario outlines the many challenges that researchers are confronted with when using the rapidly evolving digital ecosystem for research purposes. This type of research involves issues like participant recruitment, data donation, research community building and sharing of methodologies and results among researchers. It concerns a growing number of disciplines from medicine to psychology, social sciences and even humanities that increasingly use digital means for generating data. Digital research has profound effects on the ways research is organized and conducted nationally and internationally, as well as on the dissemination of skills, research information, and know-how by way of training and network building within their constituent communities (Farago 2014).

Thus, research infrastructures – durable institutions, technical tools & platforms, and/or services that are put into place for supporting and enhancing research – are increasingly set up as Virtual Research Environments (VRE): web portals providing services to users that are connected to underlying databases and repositories of various kinds. VREs are built to carry out scientific research in a community and are used as platforms for exchange between different disciplines or countries (Allan 2009; Carusi and Reimer 2010). A considerable number of such infrastructures already exist in the social sciences and increasingly also in psychology and the humanities (the latter under the label "digital humanities"). Most of these infrastructures solely rely on opening *access* to growing volumes of existing data and facilitating their use by forging common documentation standards and technical platforms across which data can move quickly. A good example for such data services is the European social science data archives consortium CESSDA (www.cessda.org). Also for participant recruitment, several services have been established. Some of them (most prominently *Amazon Mechanical Turk*) were developed for general, commercial crowdsourcing purposes; others (like *FindParticipants.com*) started with the intention to offer services for scientific research. Services that allow "donating" data for research purposes are another growing field. So far, these services have mainly been established in the medical domain, where platforms like *PatientsLikeMe.com* or *Genomes Unzipped* offer the opportunity to patients and citizens to exchange data and knowledge and to make them available for researchers.

All those are examples of research infrastructures that collect data mainly through digital means. Generating such infrastructures is associated with several challenges (Duşa et al. 2014):

1. Ensuring sustainability and establishing permanent/sustainable institutions. This problem mainly refers to financing the VRE, up to now mostly by public agencies, such as national science foundations, government institutions, universities, and European research programs.
2. Facilitating research cooperation and interdisciplinarity. This problem includes establishing common standards regarding data management across disciplines, which is particularly difficult in the broad disciplinary spectrum of social sciences and related fields.
3. Tapping new sources of (big) data. Beside others, this requires motivating citizens to contribute in an informed way to scientific research.
4. Safeguarding data protection. Here, one has to find the right balance between data acquisition and data protection, taking into account that research infrastructures play an important role in establishing best practice of data protection and research ethics.

5. Increasing the visibility of research infrastructures in their respective fields and for the general public. This requires trustworthy, easy-to-use systems that the scientific community embraces.

This broad spectrum of challenges, however, should not mask the more fundamental ethical issues associated with this type of research, namely that the individual should have control over his or her personal data. In 2012, the European Commission proposed a new legislation in the form of a regulation that will replace the Data Protection Directive (European Data Protection Regulation 2012). The General Data Protection Regulation was approved by the EU-Parliament on April 14th 2016, published in the EU Official Journal on May 4th 2016 and entered into force on May 24th 2016. It is applicable on May 24th 2018. The key changes include increased responsibility and accountability for those processing personal data and a requirement for explicit consent for processing activities. Key provisions in this regulation – such as the Right to be Forgotten and the Right to Data Portability – clearly illustrate the goal to put citizens back in control of their data. However, many of the new or modified provisions in the Regulation have been criticized in the course of developing this regulation; in particular, regarding their practical implementation, or whether they are even technically possible at all (Druschel et al. 2012).

Beyond these issues remains the question whether this approach that focuses on control and consent is adequate to the deeper changes that result from big data and the associated digital technologies. After all, one of the novel ideas found in big data research is to work with data that have been collected for a different purpose in order to uncover surprising or valuable information. As Tene & Polonetsky (2012) observe, it can be very difficult to anticipate at the time of collection for what kind of analyses some data will be used in the future.

The following considerations are based on the assumption that one of the most profound effects of this digitalization of information in all spheres of life is that the boundaries around which human beings used to conceptualize and organize their social, institutional, legal and moral world have been torn down, compromised or relativized. While the social online world tends to mirror the offline world, the traditional offline distinctions and demarcations of separate social realms (family and friendship, work, politics, education, commercial activity and production, health care, scientific research, etc.), each governed by context-relative norms, policies and rules, are threatened by the enhanced reproducibility and transmissibility of online data. What we had reasons to care about from a moral point of view in the offline world in these domains cannot be simply sustained and reproduced in a straightforward way in a digital age, which comprises online, offline, and emergent interactions between both. Individual users of digital platforms are only partially aware of these effects, but they begin to appreciate the erosion of social meanings and the frailty of traditional social norms in the digital domain. Affected are core notions like ‘informed consent’, ‘personal information’, ‘anonymity’ or ‘privacy’ as well as their underlying foundational values like ‘autonomy’, ‘fairness’ and ‘responsibility’.

The goal of this contribution is to briefly outline the possibilities and limitations of the classic idea of individual control and consent regarding the use of personal data in the big data context, and to investigate ethical justifications that may support disclosing, donating or sharing personal data, with a focus on using such data in research. This will be done in three steps: First (Section 2), it is assumed – following several other scholars – that the practice of the ‘art of separation’ or the maintenance of ‘contextual integrity’ is a key moral issue that is at stake due to the recent developments in the field of big data. Second (Section 3), it is argued that the core value of autonomy (which provides the moral foundation of control and consent) cannot support the defense of privacy by itself, but must be complemented with two other core values – fairness and responsibility – in order to sufficiently describe the moral landscape of the problem under investigation. Third (Section 4), it is drafted how research relying on (potential) personal data could proceed in order to comply with these values.

2. Contextual integrity and its undermining

In 1983, the political philosopher Michael Walzer introduced the idea of *spheres of justice*, which proposes that societies consist of different social spheres (e.g., medical, political, market, family and educational) each defined by a different type of good that is central to that particular sphere. These different types of goods (e.g., medical treatment in the medical sphere, political responsibility and public office in the political sphere) and the meaning and significance they have in each of these spheres, have their own associated criteria, principles and mechanisms concerning their distribution and allocation. In order to prevent mixing up of distributional criteria and goods from different spheres (and prevent, e.g., allocation of seats in parliament on the basis of financial assets or family relationships or health condition, or making one's ranking on a waiting list in health care dependent on family relationships or college degrees) these spheres have to be kept separated. Walzer refers to the situation where advantages and positions regarding the distribution of a good in one sphere cannot be automatically converted in advantages in another sphere. In each sphere, internal moral considerations are given their due weight, which is denoted with the term *Complex Equality*. This idea of complex equality captures an important aspect of what we mean by 'fairness' and it implies amongst other things that the distribution of access to particular goods tracks the sphere's specific normative considerations (e.g., 'need' in the medical sphere, 'democratic election' in the political sphere). Goods have to be distributed along the mechanisms of the corresponding sphere and goods from different spheres ought not to influence each other in terms of distribution. Put differently, this means that the exchange of goods between spheres has to be "blocked" in order to preserve complex equality. Walzer talks about "blocked exchanges" and the "art of separation". The same ideas regarding social differentiation and quasi-autonomy of social realms with their own internal goals, values and allocation schemes can be found in the work of many other political and social theorists.

Walzer's work has been applied to the realm of information systems by Van den Hoven (1997, 2008) and Nissenbaum (2004). Nissenbaum coined the term *contextual integrity* to refer to this idea, which she considers an "alternative benchmark for privacy, to capture the nature of challenges posed by information technologies" (Nissenbaum, 2004). Contextual integrity thus comprises a wider range of social spheres than the often-applied dichotomy of public and private. Instead, spheres are defined through the expectations and behavior of actors that differ per sphere. In order for contextual integrity and sphere separation to be achieved, the type of information that is revealed and the flows between different parties have to be appropriate for the context. Van den Hoven (2008) considers four different moral reasons to constrain flows of information. Next to the prevention of inequality based on Walzer, he points to information-based harm (e.g., through discrimination), the exploitation in markets, and moral autonomy.

The general challenge of big data is that since information produced within these spheres (health, politics, criminal justice, market) travels much faster (and to greater distances) and is more difficult to control than in the traditional offline world, we face a set of phenomena that threaten the integrity of social spheres and the cultural and social meanings expressed in them, including our values. Of course the boundaries between spheres are to a certain extent relative to time and culture, and not carved in stone forever, but it is important to note that every age, society and culture does in fact draw and treat these boundaries – construed as sets of constraints on the flow of information – as of high normative relevance. This implies that changes to them need to be morally justifiable.

From a purely technological perspective, it becomes more and more obvious that the integration of heterogeneous data describing the activity of individuals in different social spheres enable detailed inferences on the individual. As it is possible to merge different sources of data (e.g., this is the core business of data brokers, among others, see Anthes 2015), this requires studying new methodologies for privacy risk evaluation and the definition of privacy transformations suitable for addressing the multidimensional character of the data. In the literature, there exist some works on the identification

of privacy risks in social network data. Examples include the problem of linking users across different platforms, e.g., Liu and Terzi (2014) who computed the similarity among users by analyzing both generated content and top-k friends. Kosinski et al. (2013) demonstrate that it is possible to infer demographic properties and traits from the set of pages a user "likes" on Facebook. Malhotra et al. (2012) studied a way to construct digital footprints using information retrieval for name disambiguation. Vosecky and colleagues (2009) proposed a method to identify users based on profile matching (either exact or partial). Nunes and colleagues (2012) collected user profiles and, for each dimension of the profile field (e.g., username, picture, location, occupation, etc.), they reduced the problem of user identification to a binary classification task. Jain and colleagues (2013) proposed identity search algorithms to find a user's identity on Facebook, given her identity on Twitter.

Based on such "reconstructions" of individuals, discrimination may occur, which refers to an unjustified distinction of individuals based on their membership, or perceived membership, in a certain group or category disregarding individual merits. Unfair decisions have been observed in a number of settings, including credit, housing, insurance, personnel selection and worker wages, web advertising and recommendation (Romei & Ruggieri 2014). Here, a first crucial problem is *discrimination discovery*, i.e., defining methods capable of providing evidence of discriminatory behavior in activities such as the ones listed above. The legal principle of under-representation has inspired existing approaches for discrimination discovery based on frequent pattern mining (Ruggieri et al. 2010). A number of approaches have been recently proposed to tackle both privacy and non-discrimination risks in disclosing data and models (Hajian et al. 2014). Another source of complexity is when data do not explicitly contain an attribute denoting possibly discriminated groups. This case is known as indirect discrimination analysis (Hajian & Domingo-Ferrer 2013). A well-known example is redlining discrimination analysis, occurring when the ZIP code of residence is correlated with the ethnicity of individuals in highly segregated regions. The second crucial problem is *discrimination prevention*, preventing discriminatory decisions by automatic decision-making algorithms based on data mining. Discrimination prevention consists of extracting predictive data mining models, profiles, or recommendations that trade off accuracy with non-discrimination. There is a blooming research on this problem in the field of data mining, see e.g., the collection edited by Custers et al. (2013). A recent paper by Berendt and Preibusch (2014) has conducted a usability test methodology based on Amazon Mechanical Turk to assess the effectiveness of discrimination-aware approaches. These developments show that the technical capabilities for undermining the contextual integrity of data as well as detecting such integrity breaches are growing, although the former probably to a faster extent than the latter.

Both the new possibilities to merge data that originate from different spheres as well as the associated risks like discrimination point to difficult problems related to informed consent when providing data: First, informed consent is always tied to information in context, characterized by a specified purpose and associated with implicit use limitations. For example, information provided in a health research context is usually associated to disease categories and implies a certain moral impetus, namely that it will result in helping people – either the affected person or persons that in future may be affected by the condition. big data research, however, may obliterate both the information framework (like the disease space) as well as the associated moral intuitions (Christen et al. 2016). Second, if an individual provides informed consent to use data emerging from sphere A as well as to use other data emerging from a separate sphere B this does not imply that the individual provides informed consent to what is logically entailed by A & B. Informed consent is not closed under implication. Third, informed consent is tied to the "personal data paradigm" – but a lot of the data processed in a big data context are not personal data in a straightforward referential sense. This referential sense is the sense that is central to data protection legislation. "Referential" means that information can be related (via some potentially long causal chain) to a natural person. Much of the data is not of this type. At the moment it is processed it does not refer in this sense to any one in particular. This does not imply in the age of big data that that information or the actions involving that information ought not to be constrained on the basis of the moral consideration of the principles we propose.

3. Values affected by big data research

These problems associated with informed consent and discrimination outline that the notion of contextual integrity involves the idea that spheres also differ with respect to the emphasis of certain values. For example, equality plays a particularly important role in the health sphere (everybody should have equal access to health services), fairness is an overarching value in the business domain (the exchange of goods should be fair) and freedom is a guiding value in the political sphere (citizens should, e.g., be able to freely express their opinions). Certainly, all these values (and additional values not mentioned here) are to some extent relevant for each of these spheres – and even within a single sphere people can disagree on the emphasis and interpretation of (possibly conflicting) values. Therefore, due to ethical pluralism, autonomy has become a "meta value" in the sense that it justifies the acceptance of ethical pluralism (within some boundaries, i.e., people are allowed to disagree upon ethical issues) and the right of the individual to act according to own (interpretations of) moral values within the social spheres. Autonomy furthermore provides the moral foundation of the idea that an individual executes control over relevant decisions, actions etc. within social spheres. This goes along with abilities to execute autonomy – and missing abilities to be an autonomous agent, e.g., due to mental illness, may justify bypassing decisions made by the individual).

Therefore, the ideal of autonomy (a.k.a. informational self-determination, that is, the ability of persons to use digital technology in a self-determined and informed way) is often quoted as the indispensable precondition for personal data management. Closely associated to this value is thus the ideal of informed consent, in particular when disclosing information due to using some digital services or when sharing data with third persons. However, as outlined further in the previous section, the recent developments make it questionable that the consent route is a sufficient and meaningful expression of autonomy in the context of big data, in which the amount of information extracted from data (including the elaboration of meta-data) might exceed ex-ante expectations of both users and platform administrators. Furthermore, when individuals use digital platforms, they are often in a position of informational asymmetry: they are not aware of the various informational links between social spheres that are generated in this way and that allow for unexpected benefits and control possibilities by platform providers. The orientation on autonomy puts the focus on the individual and disregards the moral obligations of the other players involved in big data.

In summary, a "minimal ethics" focusing on autonomy and informed consent disregards the "empirical undermining" of autonomy and consent capacity and neglects other morally relevant values. In the following, we propose that the following three values provide a better outline of the moral landscape:

1. **Autonomy:** Users ought to be aware of how their data records are used in order to promote their values and gain control over privacy-related choices.
2. **Fairness:** The benefits of knowledge and information ought to be fairly apportioned to all participants in interactions, so as to rule out inequality of opportunity and exploitation by some at the expense of others.
3. **Responsibility:** Users (both researchers and data providing research subjects) should be held responsible and accountable for the ways in which they use their personal information and the information about other people. If some subjects are wronged, it must be possible to attribute personal responsibility for the wrongs in question.

These guiding values provide a broader in-depth analysis of the main types of moral concerns in the domain of data protection: informational harm, economic disadvantage, discrimination, and threats of self-presentation & identity (Van Den Hoven 2008; Van den Hoven et al. 2012).

Let us explain this point by some examples. Online behavior of users is tracked by advertisement agencies, in order to display more relevant ads. This so-called "behavioral targeting" is commonplace on the Internet today (Hoofnagle et al. 2012). Suppose that this service comes along with immediate benefits in non-material form (recommendations). One concern is that – based on consumer behavior –, the agencies learn habits and personal traits of users that can be used for price discrimination or "price gauging", or that some items might even not be offered (Turow 2011). For example, certain types of users, but not others, are offered special discounts for ordinary consumer products. Or in another case, it could be that an online health insurance provider offers a contract at a higher price.

This is a form of discrimination and relates to the value of *fairness*. Forms of discrimination are not necessarily unethical *per se*, but have to be addressed and analyzed with respect to their justification and counteracted if unjustified. It could be that if a consumer is facing price discrimination in ordinary consumer products, it is up to the user, considered as an autonomous agent, to strike a balance between the potential benefits and the harms of informational exposure. This ethical analysis emphasizing *autonomy* can be matched by a technology that enhances *awareness*, by measuring the informational exposure of the consumer, and other ways to help him or her understand the way his or her information might be used to predict potential harm that he or she faces. These are all necessary steps for promoting more informed decisions, related to the value of *responsibility*.

However, in considering the case in which a health insurance provider is involved, the ethical analysis might take a different course, since the *(contextual) integrity* of two spheres – shopping and healthcare – is violated. In this case the evaluation of the appropriate ethical response may be a form of *empowerment,* which could be promoted by a technology for anonymization and de-linking, or, alternatively, through a policy proposal, such as *extending* the rights of citizens in the digital domain, or by ensuring *accountability* of data mining by advertising agencies.

The recent developments in data protection law in Europe are in accordance with such a broader moral foundation. As mentioned before, the General Data Protection Regulation of the European Union that will replace the Data Protection Directive, include several key changes such as increased responsibility and accountability for those processing personal data and a requirement for explicit consent in cases when it is required for processing activities. However, significant changes have been introduced in order to facilitate processing data inside the internal market as well (e.g., one-stop-shop; one law for the whole of the EU; etc.).

From the legal point of view, when rights are limited by institutional agencies due to legitimate reasons of national security or public safety, a mechanism of assessment (commonly deployed in criminal law under the due process and judicial review procedures) must be enacted by means of public accountability for digital data. This theme is embedded in the current agenda in European and American legislative reform. In particular, it develops the reform activity of the European Data Protection Supervisor (EDPS), by suggesting how a common legal framework in data protection may foster the creation of a "level playing field" (EDPS 2013) and the proposal for the institution of a Public Interest Advocate, as recently suggested by the Report to President Obama by the Review Group on Intelligence and Communications Technologies (Review Group 2013).

For a research context, it is important to mention that the current law prevents to use collected samples in a database for future research projects if not stated specifically in the informed consent form that they can be used for future projects – which is actually the case in most of data collected, e.g. in a healthcare research context. Anonymization has been proposed as a means to bypass missing informed consent in historical data (which is also the solution proposed in Switzerland in the current Federal Act on Research involving Human Beings, Article 32-35). However, we remind that anonymization in a big data context is associated with difficult challenges (Soria-Comas and Domingo-Ferrer 2015). Of course, mere de-identification, i.e., solely removing all the directly identifying

attributes from a dataset is insufficient: identities may be inferred from the remaining attributes or by leveraging context knowledge, resulting in the re-identification of individuals at a later time.

Taken together, also in the case of historical data an ethical focus on informed consent seems to be insufficient due to rather the same reasons as in the more general case of collecting new data. Our next focus, however, is not on historical data and the informed consent issues associated to this problem. Rather, we would like to present a suggestion on how an infrastructure for generating data for research could look like that would comply with the three moral dimensions we have proposed.

4. Ethical handling of data in research – a proposal

An in-depth ethical analysis based on this roughly drafted framework certainly strongly depends on the type of problem under investigation. In the following, the focus will be on research that relies on personal information emerging from individuals – either gained directly (e.g., through surveys or offering possibilities to donate data) or indirectly (e.g., by data mining in social networks). As research often aims to combine data emerging from different social spheres in order to answer specific research questions (e.g., the interrelation of social status and health), the issue of contextual integrity is of particular relevance for researchers that handle such data.

Using the framework above, it is claimed that a research infrastructure that harvests and manages personal data should provide the following functionalities:

- **Autonomy:** Enable research participants to gain awareness on what guides their choices (privacy preferences) and on what they potentially may disclose when providing certain types of data. Shift away the focus from (mere) informed consent towards empowering research participants and data donators.
- **Fairness:** Provide a broader set of utilities (not only monetary compensation) like visualizing the contribution of research participants, e.g., through donated data, to certain scientific results. Create novel types of interactions (using, e.g., co-private protocols, Domingo-Ferrer 2011, and, more generally, co-utile protocols, Domingo-Ferrer et al. 2016) that allow collaborative contribution to a common good (like ensuring each other's privacy). Provide anti-discrimination tools, i.e., models and protocols of data acquisition and analysis for quantifying the risk of discriminatory decisions as a (possibly unwanted) consequence of data profiling and data mining. The goal is to demonstrate that contributing to research is based on a fair exchange and mutual respect of the involved parties.
- **Responsibility:** Ensure longer-term relations between participants and researchers through an infrastructure (social network) that allows for bidirectional relations (e.g., for suggesting new research questions by participants, participant-driven research). Empower the researcher both regarding legal / ethical requirements and technical instruments (e.g., for data anonymization) for doing responsible research with personal data; this may include profile anonymization tools, including masking and synthetic data methods used in statistical disclosure control (micro-aggregation, noise addition, etc.). Empower the participant with the ability to verify how safe is the anonymization performed by the data collector/researcher (Domingo-Ferrer and Muralidhar, 2016).

The goal should be to create a platform that entails technologies that enable user-centric management of personal data covering the whole information cycle: generation, publication, control, exploitation, and self-protection measures. The technological development should include three main axes:

1) The first axis concerns technologies to allow for efficient participant recruitment including all added services (e.g., regarding payment) and at the same time to improve the awareness of research participants about their degree of exposure with regard to their personal data and

the quantification of privacy risks inherent to such exposure. The goal here is to support informed consent by giving participants a clear notion of the risk inherent when providing concrete pieces of information on the platform – in particular if they want to donate data (e.g., emerging from Social Networks the participants are involved in) – and to balance the information asymmetry inherent to this environment.

2) The second axis concerns technologies to protect the data shared by researchers and other users on the platform. To this end a toolbox with anonymization techniques could be provided to support researchers involved in data acquisition; these techniques should have the novel feature that their protection will be verifiable by each data subject (participant contributing data) and that it will be possible to safely disclose their parameters to the data users (researchers). Subject verifiability will guarantee informational self-determination to participants, whereas anonymization transparency towards the researchers will maximize the inferential utility of the anonymized data. Moreover, new privacy-enabled protocols for user-to-provider, provider-to-provider and provider-to-researcher interactions should be designed so that players of such protocols will be self-motivated to embrace them and, thus, protocols can be effortlessly applied in real scenarios. Protocol design could be based on the notion of co-privacy, that is, the property that the best way for a protocol participant to preserve her own privacy is to help other participants in preserving their privacy. In such scenarios in which other relevant utilities (e.g., related to functionality, visibility, availability, security, awareness, analytical utility, etc.) are involved, the more general notion of co-utility could be applied, by which the best way for a player (e.g., users, providers, researchers) to serve her own interests is to help other players towards their own interests.
3) The third axis consists of technologies that facilitate efficient and usable data management on the platform. This involves issues like voluntary data donation, secure data storage, sharing and referencing via data repositories, as well as techniques for visualization.

In contrast to a traditional Internet marketplace, where users are attracted solely by the promise of economic compensation, a research platform should aim to create and maintain an active community that is educated through and involved in research over time. For example, participants may share their personal informational exposure profile with other participants, can create their data control preference profile, can join discussions with other participants as well as with researchers, and even provide genuine ideas as inputs to research. By participating in research, citizens contribute to improve the technology that serves their own empowerment.

The research social network should enable researchers to create, configure and test scenarios of critical data exchanges among specific population targets. The scenarios could be based on a configurable subset of data objects and properties. The researchers will be able to specify the desired criteria for their population (e.g., by giving demographic attributes such as age and gender distribution) as well as the desired privacy attitudes. Participants will be invited to participate and be allocated to the population of a study based on the information in their profile and (if provided) their privacy preferences as obtained by awareness self-assessment tools offered on the platform. The platform should protect the privacy of participants against the researchers and – to some degree – against the provider of the platform itself via the use of anonymization and pseudonymous attestation techniques (such as blinded attribute certificates).

5. Conclusion

In this contribution, we argue that the growing digital infrastructure with its emerging web ecosystem provides research with unprecedented possibilities for accessing data that generate new ethical challenges. A mere focus on personal data control and informed consent does not adequately reflect these challenges. Rather, a variety of issues running from participant recruitment, data donation, data protection, research community building up to sharing of methodologies and results are raised

that need adequate ethical consideration. We propose that the values of autonomy, fairness and responsibility provide a more complete moral grounding of future digital research infrastructures – in particular for disciplines like psychology, social sciences, and public health, where integrated online infrastructures, methodologies and policies of cross-disciplinary data interoperability and sharing are lacking. From such an infrastructure, researchers should expect a cross-cultural, multi-lingual access to participants that is trustworthy, practical, and complies with ethical standards; methods and tools for data anonymization, synthetic data generation, and big data management; access to a research social network to share data, insights and tips when conducting online research (surveys, web-experiments and the like). Participants should expect an infrastructure that provides an easy way to contribute to research and get a fair compensation for it; the possibility to donate personal data for research according to own privacy preferences; access to a research social network that allows for commenting and inspiring cross-cutting research in various fields. The current changes in research involving possibilities for massive data generation and access should be seen as an opportunity to establish new relationships between researchers and their "research object" – human beings as sources of data that is relevant for understanding and improving society.

References

Allan, Robert. 2009. *Virtual Research Environments: From Portals to Science Gateways*. Oxford: Chandos Publishing.

Anthes, Gary. 2015. Data brokers are watching you. *Communications of the ACM* 58(1):28-30.

Berendt, Bettina, and Sören Preibusch. 2014. Better decision support through exploratory discrimination-aware data mining: foundations and empirical evidence. *Artificial Intelligence and Law* 22(2): 175-209.

Carusi, Annamaria, and Thomas Reimer. 2010. *Virtual Research Environment – Collaborative Landscape Study*. A JISC funded project. Available at: http://www.jisc.ac.uk/media/documents/publications/vrelandscapereport.pdf (last access: December 9 2016).

Christen, Markus, Josep Domingo-Ferrer, Bodan Draganski, Tade Spranger, and Henrik Walter. 2016. On the compatibility of Big Data driven research and informed consent based on traditional disease categories – the example of the Human Brain Project. In *Ethics of Biomedical Big Data*, eds. Luciano Floridi, Brent Mittelstadt, pp. 199-218.

Custers, Bart, Toon Calders, Bart Schermer, and Tal Zarsky (eds.). 2013. *Discrimination and Privacy in the Information Society*, vol. 3 of Studies in Applied Philosophy, Epistemology and Rational Ethics. Berlin/London: Springer.

Domingo-Ferrer, Josep. 2011. Coprivacy: an introduction to the theory and applications of co-operative privacy. *SORT-Statistics and Operations Research Transactions* 35: 25-40.

Domingo-Ferrer, Josep, and Krishnamurty Muralidhar. 2016. New directions in anonymization: permutation paradigm, verifiability by subjects and intruders, transparency to users. *Information Sciences* 337-338:11-24.

Domingo-Ferrer, Josep, David Sánchez, and Jordi Soria-Comas. 2016. Co-utility: self-enforcing collaborative protocols with mutual help. *Progress in Artificial Intelligence* 5(2):105-110.

Druschel, Peter, Michael Backes, and Rodica Tirtea. 2012. The Right to Be Forgotten – between Expectations and Practice. ENI SA. Available at: https://www.enisa.europa.eu/publications/the-right-to-be-forgotten (last access: December 9 2016)

Duşa, Adrian, Claudia Oellers, and Simon Wolff. 2014. A Common agenda for the European research infrastructures in the social sciences and humanities. In *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences*, ed. Adrian Duşa, Dietrich Nelle, Günter Stock, Gert G. Wagner, 225-234. SCIVERO Verlag, Berlin.

EDPS (2013): European Data Protection Supervisor, see: https://europa.eu/european-union/about-eu/institutions-bodies/european-data-protection-supervisor_en (last access: December 9 2016)

European Data Protection Regulation (2012): The approved version is available at the following website: http://ec.europa.eu/justice/data-protection/index_en.htm (last access: December 9 2016)

Farago, Peter (2014). Understanding how research infrastructures shape the social sciences: impact, challenges, and outlook. In *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences*, ed. Adrian Duşa, Dietrich Nelle, Günter Stock, Gert G. Wagner, 21-34. SCIVERO Verlag, Berlin.

Hajian, Sara, and Josep Domingo-Ferrer. 2013. A methodology for direct and indirect discrimination prevention in data mining. *IEEE Transactions on Knowledge and Data Engineering* 25(7): 1445-1459.

Hajian, Sara, Josep Domingo-Ferrer, and Oriol Farràs. 2014. Generalization-based privacy preservation and discrimination prevention in data publishing and mining. *Data Mining and Knowledge Discovery* 28: 1158-1188.

Hoofnagle, Chris Jay, and Ashkan Soltani, Nathaniel Good, Dietrich J. Wambach, and Mika D. Ayenson. 2012. Behavioral advertising: the offer you cannot refuse. *Harvard Law & Policy Review* 6: 273–296.

Jain, Paridhi, Ponnurangam Kumaraguru, and Anupam Joshi. 2013. @i seek 'fb.me': identifying users across multiple online social networks. WWW (Companion Volume): 1259-1268. Available at: http://ebiquity.umbc.edu/paper/html/id/624/-I-seek-fb-me-Identifying-Users-across-Multiple-Online-Social-Networks (last access: December 9 2016)

Kosinski, Michal, David Stillwell, and Thore Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences* 110(15): 5802–5805.

Liu, Kun, and Evimaria Terzi. 2009. A framework for computing the privacy scores of users in online social networks. In *Proceedings of ICDM 2009*, The 9th IEEE International Conference on Data Mining: 288-297.

Malhotra, Anshu, Luam Totti, Wagner Meira Jr., Ponnurangam Kumaraguru, and Virgilio Almeida. 2012. Studying user footprints in different online social networks. In *Proceedings of ASONAM 2012*, 1065-1070. arXiv:1301.6870

Nissenbaum, Helen. 2004. Privacy as contextual integrity. *Washington Law Review* 79: 119-157.

Nunes. André, Pável Calado, and Bruno Martins. 2012. Resolving user identities over social networks through supervised learning and rich similarity features. In *Proceedings of the 27th Annual ACM Symposium on Applied Computing*, 728-729.

Review Group (2013): see https://www.dni.gov/index.php/intelligence-community/review-group (last access: December 9 2016)

Romei, Andrea, and Salvatore Ruggieri. 2013. A multidisciplinary survey on discrimination analysis. *The Knowledge Engineering Review* 29(5): 1-57.

Ruggieri, Salvatore, Dino Pedreschi, and Franco Turini. 2010. Data mining for discrimination discovery. *ACM Transactions on Knowledge Discovery from Data* 4(2): Article 9.

Soria-Comas, Jordi, and Josep Domingo-Ferrer. 2015. Big data privacy: challenges to privacy principles and models. *Data Science and Engineering* 1(1):21-28.

Tene, Omer, and Jules Polonetsky. 2012. Privacy in the age of big data: a time for big decisions. *Stanford Law Review Online* 64 (63): 63–69.

Turow, Joseph. 2011. *The Daily You: How the New Advertising Industry is Defining Your Identity and Your World*. New Haven: Yale University Press.

Van den Hoven, Jeroen. 1997. Computer ethics and moral methodology. *Metaphilosophy* 28(3): 234-248.

Van den Hoven, Jeroen. 2008. Information technology, privacy, and the protection of personal data. In *Information technology and moral philosophy*, Jeroen van den Hoven, and John Weckert, eds., 301-321. Cambridge, New York: Cambridge University Press:

Van den Hoven, Jeroen, Dirk Helbing, Dino Pedreschi, Josep Domingo-Ferrer, Fosca Gianotti, and Markus Christen. 2012. FuturICT - The Road towards Ethical ICT. *European Physical Journal - Special Topics* 214: 153–181.

Vosecky, Jan, Dan Hong, and Vincent Shen. 2009. User identification across multiple social networks. In *Proceedings of the First International Conference on Networked Digital Technologies, NDT '09, IEEE*, 360-365.

Walzer, Michael. 1983. *Spheres Of Justice: A Defense Of Pluralism And Equality*. New York City: Basic Books.