



Norwegian
Business School

This file was downloaded from BI Open Archive, the institutional repository (open access) at BI Norwegian Business School <http://biopen.bi.no>.

It contains the accepted and peer reviewed manuscript to the article cited below. It may contain minor differences from the journal's pdf version.

Newlands, G., Lutz, C. & Fieseler, C. (2019). Trading on the Unknown: Scenarios for the Future Value of Data . Sharing Economy Markets and Human Rights. *The Law & Ethics of Human Rights*, 13(1), pp. 97-114. <https://doi.org/10.1515/lehr-2019-0004>

Copyright policy of *De Gruyter*, the publisher of this journal:

De Gruyter allows authors the use of the final published version of an article (publisher pdf) for self-archiving (author's personal website) and/or archiving in an institutional repository (on a non-profit server) after an embargo period of 12 months after publication.

<https://www.degruyter.com/dg/page/576/repository-policy>

Trading on the Unknown: Scenarios for the Future Value of Data

Gemma Newlands, Christoph Lutz, and Christian Fieseler*

In this article, we explore the practices of extensive data collection among sharing economy platforms, highlighting how the unknown future value of big data creates an ethical problem for a fair exchange relationship between companies and users. Specifically, we present a typology with four scenarios related to the future value of data. In the remainder of the article, we first describe the status quo of data collection practices in the sharing economy, followed by a discussion of the value-generating affordances of big data. We then introduce the typology of four scenarios for the future value of data. Finally, the paper concludes with a short discussion on the implications of information asymmetries for a fair exchange process.

Keywords: Big Data; Informed Consent; Sharing Economy; Platforms; Ethics; Information Asymmetry

Introduction

With an estimated 17% of EU consumers having used some form of sharing platform so far, the growth of the sharing economy in Europe has been seen by some as an empowering economic transformation.¹ By enabling private individuals to share their assets in exchange for money, the sharing economy has opened up many new markets on both a local and global scale. However, users of the sharing economy are also involved in an additional and often invisible market: A market for their data. Firmly entrenched within what the United Nations has called a "data revolution," consumer-facing Internet companies have become data hungry, directing their business processes towards the collection and monetization of user data *en masse*.² Such data, in addition to having epistemological, affective, and ethical value, can have a distinct

* Gemma Newlands, BI Norwegian Business School, Nordic Centre for Internet and Society
Nydalsveien 37, 0484 Oslo – Norway, gemma.e.newlands@bi.no.

Christoph Lutz, BI Norwegian Business School, Nordic Centre for Internet and Society
Nydalsveien 37, 0484 Oslo – Norway, christoph.lutz@bi.no.

Christian Fieseler, BI Norwegian Business School, Nordic Centre for Internet and Society
Nydalsveien 37, 0484 Oslo – Norway, christian.fieseler@bi.no.

¹ FLASH EUROBAROMETER 438, THE USE OF COLLABORATIVE PLATFORMS (2016), available at https://data.europa.eu/euodp/data/dataset/S2112_438_ENG; EUROPEAN COMMISSION, A EUROPEAN AGENDA FOR THE COLLABORATIVE ECONOMY (2016), available at <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2016:0356:FIN>.

² UNITED NATIONS, A WORLD THAT COUNTS: MOBILISING THE DATA REVOLUTION FOR SUSTAINABLE DEVELOPMENT (2014), <http://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf>; Marijn Sax, *Big data: Finders Keepers, Losers Weepers?*, 18 ETHICS INF. TECH. 25 (2016).

economic value, whereby current and future value is generated through new technological capacities to integrate and cross-reference multiple large data sets.³ This economic value can be realized by use both within-platform for service improvements and tighter consumer targeting or outside-platform for sale on the broader data market.

However, the potential value of data is unknown before collection. Contemporary data are simply “too dynamic and unpredictable to determine” if and when particular information will become valuable.⁴ Moreover, there is no guarantee during collection that the value inherent within data will be realized, making widespread data collection a form of speculative investment. Since, regardless of outcome, the encapsulated economic value remains the sole property of platforms due to pre-emptive contractual transfer or data collection of passive “digital footprints,” ethical and legal issues are raised as to the appropriate distribution of big data’s potential value.⁵ Although research to date has begun to consider aspects of data rights and data justice, the implications of data hoarding on behalf of platforms demand more attention.⁶ This is particularly the case since the value of data is often unknown and, indeed, unknowable at the time of collection, leading to questions of whether consumers are being adequately informed about the intended purpose at time of collection.⁷

I. Data Collection and the Sharing Economy

According to frequently expressed company rhetoric, sharing economy platforms function as merely the technological intermediaries that enable peer-to-peer transactions.⁸ However, their role as mere “intermediaries” has not hindered platforms from engaging in widespread data

³ Mark Andrejevic, *The Big Data Divide*, 8 INT. J. COMM. 1673 (2014); Ryan Burns, *Rethinking Big Data in Digital Humanitarianism: Practices, Epistemologies, and Social Relations*, 80 GEOJ. 477 (2014); Daniel G. Cockayne, *Affect and Value in Critical Examinations of the Production and ‘Prosumption’ of Big Data*, 3 BIG DAT. SOC. 1 (2016); danah boyd & Kate Crawford, *Critical Questions for Big Data*, 15 INFO. COMM. SOC. 662 (2012); Sheila Jasanoff, *Virtual, Visible, and Actionable: Data Assemblages and the Sightlines of Justice*, 4 BIG DAT. SOC. 1 (2017).

⁴ Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B. C. L. REV. 93 (2014).

⁵ Jim Thatcher, *Living on Fumes: Digital Footprints, Data Fumes, and the Limitations of Spatial Big Data*, 8 INT. J. COMM. 1765 (2014).

⁶ Linnet Taylor, *What is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally*, 4 BIG DAT. SOC. 1 (2017).

⁷ Crawford & Schultz, *supra* note 4; Katina Michael & Keith, W. Miller, *Big Data: New Opportunities and New Challenges [guest editors’ introduction]*, 46 COMPUTER 22 (2013).

⁸ Tarleton Gillespie, *The Politics of Platforms*, 12 NEW MEDIA SOC. 347 (2010); Tarleton Gillespie et al. eds., *The Relevance of Algorithms*, in MEDIA TECHNOLOGIES 167 (Tarleton Gillespie et al. eds., 2014); REBECCA SMITH & SARAH LEBERSTEIN, RIGHTS ON DEMAND: ENSURING WORKPLACE STANDARDS AND WORKER SECURITY IN THE ON-DEMAND ECONOMY, NATIONAL EMPLOYMENT LAW PROJECT (2015), <https://s27147.pcdn.co/wp-content/uploads/Rights-On-Demand-Report.pdf>.

collection. The peer-to-peer nature of the sharing economy means that surrendering private information is considered to be a gateway to access, as information must be shared to generate trust between providers and consumers.⁹ This data, handed over by the user to create such trust, covers basic aspects such as profile information, photographic identification, identification verification documents, and financial information. Offering access as a provider and gaining access as a consumer thus requires the transfer of personal, sensitive information.¹⁰ The practical reality of sharing assets, such as homes, money, or personal possessions also necessitates the collection of accurate data to facilitate effective transfer of ownership. Similarly, the locative nature of ride-hailing, a significant sub-section of the sharing economy, requires accurate location-based data to function. However, sharing platforms also collect additional data of the type referred to as a passive "digital footprint."¹¹ This data, rather than being essential for the operation of the core service, is produced and collected automatically as a by-product.

This distinction between data provided by the user directly and data collected automatically by the platform is enshrined within platform policies. Airbnb's privacy policy, for instance, distinguishes the two categories of information as: "Information You Give Us" and "Information We Automatically Collect From Your Use of the Airbnb Platform and Payment Services."¹² The latter includes data such as usage information (i.e., which pages are viewed or searches are made), and location information (including IP address or precise GPS location if accessed via mobile). Regardless of whether an individual holds an account, Airbnb also collects device information, links to third party applications, and the details of the webpages viewed both before and after visiting the Airbnb website. Following a similar pattern to Airbnb, Uber's privacy policy also indicates a distinction between information provided and information collected automatically, whereby the data collected automatically includes GPS location

⁹ Amitai Etzioni, *Cyber Trust*, volume online first J. BUS. ETHICS. 1 (2017), https://www.deepdyve.com/lp/springer_journal/cyber-trust-MVo4AVPG0P; Asle Fagerstrøm et al., *That Personal Profile Image Might Jeopardize Your Rental Opportunity! On the Relative Impact of the Seller's Facial Expressions Upon Buying Behavior on Airbnb*, 72 COMPUT. HUM. BEHAV. 123 (2017).

¹⁰ Christoph Lutz et al., *The Role of Privacy Concerns in the Sharing Economy*, 21 INFO. COM. SOC. 1472 (2018).

¹¹ Thatcher, *supra* note 5; Digital footprints refer to the aggregate of individuals' personal data traces, accumulated through online activities such as searching, browsing, posting, texting, and liking. They include passive or background tracking of user data, where platforms or other users create data about an individual, as well active participation through sharing content on the Internet. Mary Madden et al., *Digital Footprints: Online Identity Management and Search in the Age of Transparency*, Pew Research Center: Internet & Technology, 16 December (2007), available at <http://www.pewinternet.org/2007/12/16/digital-footprints/>; Marina Micheli et al., *Digital Footprints: An Emerging Dimension of Digital Inequality*, 16 J. INFORM. COMM. ETHICS SOC. 1 (2018); Christoph Lutz & Christian Pieter Hoffmann, *The Dark Side of Online Participation: Exploring Non-, Passive and Negative Participation*, 20 INFO. COMM. SOC. 876 (2017); Stefania Milan, *Political Agency, Digital Traces, and Bottom-Up Data Practices*, 12 INT. J. COMM. 507 (2018).

¹² Airbnb, *Terms of Service* (2019), https://www.airbnb.com/terms/privacy_policy (last updated Jan. 21, 2019).

information, transaction details, device information (including serial number and specific motion tracking), as well as log information such as the IP address, browser information, and information about websites visited prior to interacting with Uber services.¹³

With regard to the purpose of such data collection, platform terms of service provide a brief insight into their plans for using such aggregate data. Airbnb, for instance, simply details that it “may also share aggregated information...and non-personally identifiable information for industry and market analysis, demographic profiling, marketing and advertising, and other business purposes.”¹⁴ Uber’s outline of their intended uses of data similarly follows Airbnb’s pattern, suggesting a platform focused use of data to “provide, maintain, and improve our services.”¹⁵ However, Uber’s attempt to outline its use of aggregate data provides a limited and blanket statement that “[w]e may share your information... [i]n an aggregated and/or anonymized form which cannot reasonably be used to identify you.” Both major platforms in the sharing economy, whose approach to data collection and transparency is largely matched by smaller European-based platforms such as BlaBlaCar, have thus equipped themselves with full license to use their collected data for both internal and external purposes, either in the form of service improvements or more widespread sale to third-parties. The principle of purpose limitation, as enshrined in EU data protection law, consists of purpose specification, namely that data must be collected only for specified, explicit, and legitimate purposes, and of compatible use, namely that data must not be processed further in a way incompatible with those stated purposes. By offering a broad yet unspecific description of the purposes of data collection, sharing economy platforms are attempting to navigate the law for their own self-interest and to maximise the growth potential of their companies. Indeed, it has been well covered in the media that sharing economy platforms have big ambitions: Uber intends to leverage their data for future services such as self-driving cars, whereas Airbnb is intending to move into property design.¹⁶ Greater attention on ensuring purpose limitation, and particularly purpose specification, as a fundamental aspect of EU data protection law, may nevertheless destabilise the entrenched power asymmetries currently in action surrounding user data.

¹³ Uber, *Terms of Service: Uber Privacy Policy* (2018), available at <https://privacy.uber.com/policy/> (last updated May 25, 2018).

¹⁴ Airbnb, *supra* note 12.

¹⁵ Uber, *supra* note 13.

¹⁶ Devin Coldewey, *Uber Shows Off Its Autonomous Driving Program’s Snazzy Visualization Tools*, TECH CRUNCH (Aug. 28 2017), <https://techcrunch.com/2017/08/28/uber-shows-off-its-autonomous-driving-programs-snazzy-visualization-tools/>; Jessica Leber, *Your Future Home Will Be Designed For Sharing (By Airbnb)*, FAST COMPANY (Nov. 4, 2016), <https://www.fastcompany.com/3058520/your-future-home-will-be-designed-for-sharing-by-airbnb..>

This forward-looking approach to data collection, where information is collected, stored, and processed for the purpose of future revenue generation, exists within a broader ecosystem which prioritizes the ubiquitous collection and storage of data.¹⁷ As data mining has become an increasingly essential business process, the potential value from future analysis and utilization has meant that companies have become “data hungry.”¹⁸ Companies collect, analyze, and utilize user data to create new insights, improve algorithms, refine services, enhance business models, and critically inform decision making processes.¹⁹ For this purpose, Internet technology companies have adapted the technical architecture of their online platforms to collect and aggregate large-scale data from varied sources.²⁰ Internet technology companies can also ensure additional value generation by sharing data with third parties, where data is passed from one firm to the next within an “information supply chain.”²¹ Value generation through big data can thus occur among multiple parties, as data can also be sold to other companies, aggregators, or researchers.²² Deighton and Johnson noted that 71% of the market in consumer data was directly or indirectly dependent on data traded among firms.²³

However, an important element in discussions regarding future data monetization is the cost of protecting data and ensuring regulatory compliance. Non-compliance with data protection legislation, for instance, can impose heavy fines and the costs of losing user data to hackers, such as occurred with Uber’s 2016 data breach affecting 57 million users, can be prohibitively high. With the introduction of the European Parliament’s General Data Protection Regulation (GDPR) (Regulation 2106/679) in 2018, even greater protections are assigned to personal data, with accordingly higher fines for its mishandling. In an advance of earlier regulation, Chapter

¹⁷ Kate Shilton, *Four Billion Little Brothers? Privacy, Mobile Phones, and Ubiquitous Data Collection*, 52 COMMUN. ACM. 48 (2009).

¹⁸ Anthony Danna & Oscar H. Gandy Jr., *All That Glitters is Not Gold: Digging Beneath the Surface of Data Mining*, 40 J. BUS. ETHICS. 373 (2002); Sax, *supra* note 2, at 25.

¹⁹ Jonathan S. Ward & Adam Barker, *Undefined by Data: A Survey of Big Data Definitions*, ARXIV (2013), <https://arxiv.org/abs/1309.5821> (submitted on Sept. 20, 2013); TERENCE CRAIG & MARY E. LUDLOFF, *PRIVACY AND BIG DATA* (2011); Steve Lohr, *The Age of Big Data*, N.Y. TIMES (Feb. 11, 2012), <http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html>; David Lazer et al., *Life in the Network: The Coming of Age of Computational Social Science*, 323 SCIENCE 721 (2009); Sax, *supra* note 2; Omer Tene & Jules Polonetsky, *Big Data for All: Privacy and User Control in the Age of Analytics*, 11 NW. J. TECH. & INTEL. PROP. 239 (2013).

²⁰ Sarah M. West, *Data Capitalism: Redefining the Logics of Surveillance and Privacy*, 58 (1) BUS. SOC. 20 (2019); Glen Whelan, *Born Political: A Dispositive Analysis of Google and Copyright*, 58 (1) BUS. SOC. 42 (2019).

²¹ Kirsten E. Martin, *Ethical Issues in the Big Data Industry*, 14 MIS Q. EXEC. 67, 70 (2015).

²² Carolin Gerlitz & Anne Helmond, *The Like Economy: Social Buttons and the Data-Intensive Web*, 15 NEW MEDIA SOC. 1348 (2013); Kirsten Martin, *Data Aggregators, Consumer Data, and Responsibility Online: Who is Tracking Consumers Online and Should They Stop?*, 32 INFORM. SOC. 51 (2016).

²³ John Deighton & Peter A. Johnson, *The Value of Data: Consequences for Insight, Innovation, and Efficiency in the US Economy*, 14 DATA-DRIVEN MARKETING INST. 1 (2013).

8 of the GDPR follows the interest theory of rights and focuses on the duty of data controllers, regardless of whether an individual has stood up to defend their personal rights. It should thus be emphasized that while the benefits of long-term data collection and storage result in benefits for a company, the company also faces financial risks in handling such data.

II. Issues of Data Collection and Informed Consent

Considering the value that can be generated through excessive data collection, we propose that a fair exchange relationship between individual users and companies is substantially challenged. The question of how to operate a fair balance between corporations and users, with regard to the exchange of personal data, has been widely addressed by privacy scholars, who univocally stress the normative ideal of a balance between benefits and possible harms.²⁴ However, most privacy scholars also problematize the information and power asymmetry between corporations and individuals. First, and as a result of the market-dominance of key sharing economy platforms, consumers are subject to lock-in effects, meaning they have few alternatives to choose between. Consumers also have limited portability of their data, in that they can rarely access or transfer the data held about them on company servers.

As a further key point, individuals lack the information or legal literacy to make informed decisions about whether to provide their individual data for a certain benefit. For instance, by referring to the concept of informed consent, scholars have highlighted that informed consent by individuals to provide their data for certain benefits is flawed by default.²⁵ This has led to a scholarly consensus that notice and choice as drivers of informed consent have effectively failed.²⁶ Not only are standard terms and contracts, through which individuals give their consent, likely to be one-sided and generally of low quality, but individuals are also often not able to assess standard terms sufficiently enough to make rational decisions in their own interest.²⁷ This is particularly the case, as contracts for digital platforms are becoming increasingly complex.²⁸ Cognitive limitations lead consumers to neglect non-salient issues in a standard

²⁴ Daniel J. Solove, *Introduction: Privacy Self-Management and the Consent Dilemma*, 126 HARV. L. REV. 1880 (2013).

²⁵ Helen Nissenbaum, *A Contextual Approach to Privacy Online*, 140 DAEDALUS 32 (2011).

²⁶ Mihaela Popescu & Lemi Baruh, *Captive but Mobile: Privacy Concerns and Remedies for the Mobile Environment*, 29 INFORM. SOC. 272 (2013); Paul Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86, N.Y.I. L.Q. REV. 1814 (2011).

²⁷ Floretta Marotta-Wurgler, *What's in a Standard Form Contract? An Empirical Analysis of Software License Agreements*, 4 J. EMPIRICAL LEGAL STUD. 677 (2007); Omri Ben-Shahar, *The Myth of the 'Opportunity to Read' in Contract Law*, 5 EUR. REV. CONTRACT L. 1 (2009); Thomas Wilhelmson, *Cooperation and Competition regarding Standard Contract Terms in Consumer Contracts*, 17 EUR. BUS. L. REV. 49 (2006).

²⁸ MARGARET J. RADIN, *BOILERPLATE: THE FINE PRINT, VANISHING RIGHTS, AND THE RULE OF LAW* (2012).

form contract, or to neglect reading it entirely.²⁹ The dilemma is discussed in Nissenbaum's concept of the transparency paradox, whereby the more information that is shared, the less understandable the statements are for genuine consent.³⁰

The issue of informed consent is further complicated given the complex network effects of value generation through big data analysis. Martin argues that "even organizations with the best of intentions to notify users struggle to communicate complicated and changing policies which, given the large network of actors in the online space, may conflict with the policies of their online partners such as ad networks, third-party organizations, and user-generated applications."³¹ This debate also applies to consent in the sharing economy. As Calo and Rosenblat discuss, the frequency of the contractual changes on some sharing economy platforms result in providers having to agree to new terms of service every couple of days.³² Bar-Gill and Horton went further in arguing that this contractual complexity is a purposeful attempt to exploit the human limitations of processing information, so as to maintain a powerful information asymmetry between platforms and providers.³³ These information asymmetries of sharing economy platforms are largely enforced and perpetuated by the use of algorithmic intermediation, where algorithms act as the "middle man."³⁴ However, algorithms tend to be invisible, so are resistant to scrutiny; they are, according to Pasquale, "black boxed."³⁵ A lack of transparency regarding algorithms and data collection thus complicates the question of fair exchange between companies and users.³⁶

²⁹ Robert A. Hillman & Jeffrey J. Rachlinski, *Standard-Form Contracting in the Electronic Age*, 77 N. Y. U. L. REV. 429 (2002).

³⁰ Nissenbaum, *supra* note 25.

³¹ Kirsten Martin, *Transaction Costs, Privacy, and Trust: The Laudable Goals and Ultimate Failure of Notice and Choice to Respect Privacy Online*, 18 FIRST MONDAY 1, 4 (2013).

³² Ryan Calo & Alex Rosenblat, *The Taking Economy: Uber, Information, and Power*, 117 COLUM. L. REV. 1623 (2017).

³³ OREN BAR-GILL, *SEDUCTION BY CONTRACT: LAW, ECONOMICS, AND PSYCHOLOGY IN CONSUMER MARKETS* (2012); John J. Horton, *Online Labor Markets*, 515-22 (WINE 2010 Proceedings of the 6th International Workshop on Internet and Network Economics, 2010).

³⁴ Gillespie, *The Relevance of Algorithms*, *supra* note 8; Min Kyung Lee et al., *Working with Machines: The Impact of Algorithmic and Data-Driven Management on Human Workers*, 1603 (CHI'15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2015). The notion of algorithmic management describes software-based organization, where workers are dispatched and controlled automatically through – often black-boxed – algorithms rather than through human managers. This type of algorithmic management is particularly common for gig economy services, for example in ride-hailing (e.g., Uber, Lyft) or food delivery (e.g., Foodora, Deliveroo).

³⁵ FRANK PASQUALE, *THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION* (2015); Jenna Burrell, *How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms*, 3 BIG DATA SOC. 1 (2016). The black box problem is particularly pronounced with deep learning algorithms, as widely applied in current web applications such as spam filters, search classification, and newsfeed algorithms.

³⁶ Brent D. Mittelstadt et al., *The Ethics of Algorithms: Mapping the Debate*, 3 BIG DATA SOC. 1 (2016).

By referring to the privacy calculus, privacy scholars have highlighted the trade-off between costs and benefits of data provision as a rational decision. According to the privacy calculus idea, individuals are expected to rationally anticipate and weigh privacy related risks and benefits, and to consequently disclose private information if the benefits outweigh the risks.³⁷ However, scholars have argued that individuals run into difficulties when trying to weigh the costs and benefits of the collection and use of their information.³⁸ The privacy calculus perspective has been criticized for its lack of consideration of emotional elements and situational factors.³⁹ We would argue that in the sharing economy, such emotional and situational elements are particularly salient, as users often use sharing economy services for more than just monetary reasons.⁴⁰

Furthermore, scholars have highlighted that even well-informed and rational individuals cannot appropriately self-manage their privacy due to “structural problems,” which means that future use of data by organizations is not known at the time of data collection. As the creation of future value, as well as privacy harms, are the result of an aggregation of pieces of data over a period of time by different entities, it is “virtually impossible for people to weigh the costs and benefits of revealing information or permitting its use or transfer without an understanding of the potential downstream uses.”⁴¹ Moreover, many uses of data have benefits in addition to costs, and individuals could rationally reach opposite conclusions regarding whether the benefits outweigh the costs.

In the age of big data, as outlined above, much of the value of personal information is not apparent at the time of collection, when notice and consent are normally given. Indeed, scholars have critically highlighted that “given the rapid changes in big data and data analysis, consent in terms of user expectations may easily become outdated.”⁴² Accordingly, consent

³⁷ Mary Culnan & Pamela Armstrong, *Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation*, 10 *ORGAN. SCI.* 104 (1999); Haein Lee et al., *Why Do People Share Their Context Information on Social Network Services? A Qualitative Study and an Experimental Study on Users' Behavior of Balancing Perceived Benefit and Risk*, 71 *INT. J. HUM-COMPUT. ST.* 826 (2013); Han Li et al., *Understanding Situational Online Information Disclosure as a Privacy Calculus*, 51 *J. COMPUT. INFORM. SYST.* 62 (2010); Alan F. Westin, *Social and Political Dimensions of Privacy*, 59 *J. SOC. ISSUES.* 431 (2003).

³⁸ Solove, *supra* note 24.

³⁹ Flavius Kehr et al., *Blissfully Ignorant: The Effects of General Privacy Concerns, General Institutional Trust, and Affect in the Privacy Calculus*, 25 *INFORM. SYST. J.* 607 (2015).

⁴⁰ Eliane Bucher et al., *What's Mine is Yours (For a Nominal Fee) – Exploring the Spectrum of Utilitarian to Altruistic Motives for Internet-Mediated Sharing*, 62 *COMPUT. HUM. BEHAV.* 316 (2016).

⁴¹ Solove, *supra* note 24, 1881.

⁴² Bart Custers, *Click Here to Consent Forever: Expiry Dates for Informed Consent*, 3 *BIG DATA SOC.* 1, 2 (2016).

may be considered outdated when it no longer matches the initial preference of a user, for instance, because they changed their mind or because data processing practices have changed significantly. As Custers discusses, the expectation is that consent is provided forever, despite terms and conditions facing continuous changes.⁴³ By suggesting the inclusion of time-limits on consent, Custers opens a line of discourse on how data should be handled post-use.⁴⁴

III. A Typology of Data's Future Value Scenarios

Since platforms in the sharing economy collect data for the purposes of leveraging it for future profit, this creates a dilemma for a fair exchange relationship, as users are not able to provide legitimate informed consent. Indeed, as Martin notes, “not only are consumers not fully aware, but they also lack the power to choose otherwise.”⁴⁵ In sum, similar to privacy scholars, we argue that a fair exchange relationship between corporations and users is challenged by the unknown future value of big data. Since the fair exchange relationship operates between two parties, platforms and users, an initial conceptualization of the different relationship modes can be theorized by considering the locus of information about the future value of data. Specifically, a fair exchange relationship is reliant on a fair balance of information, between that which is known and that which is unknown. In the following, we propose a typology of four scenarios where the balance of information diverges between users and platforms (Figure 1).

⁴³ *Id.*

⁴⁴ *Id.*

⁴⁵ Martin, *supra* note 31, at 2.

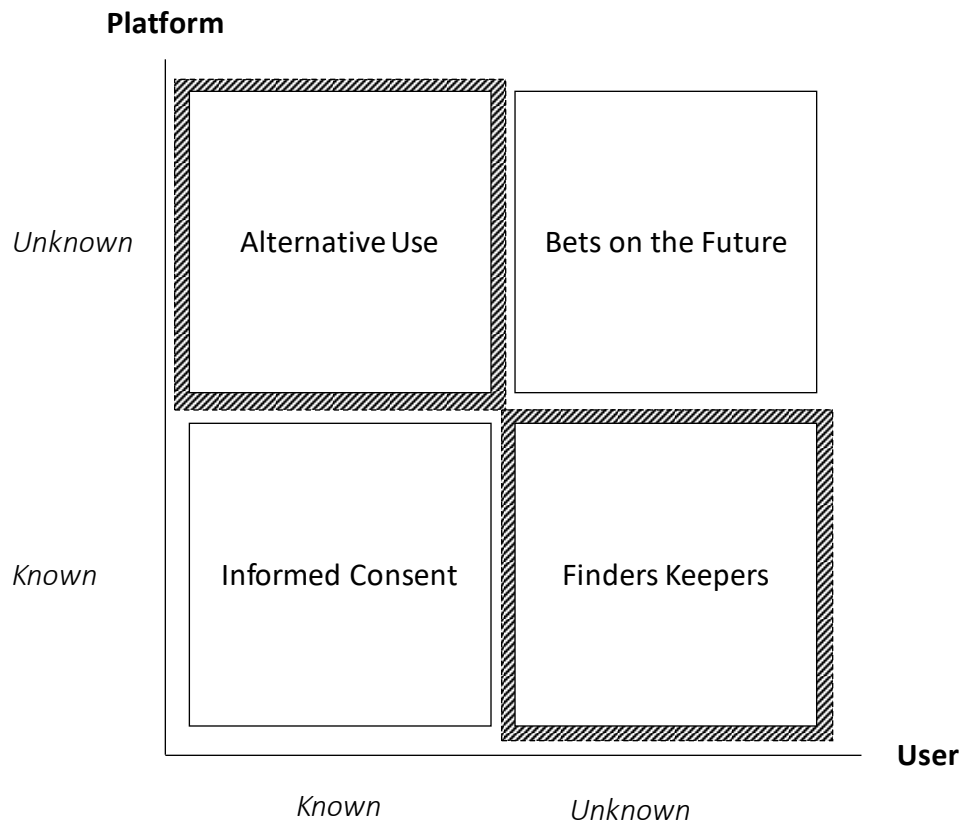


Figure 1 – Four Data Exchange Scenarios

Scenario 1: Informed Consent

In the first scenario, categorized as "informed consent," we presume the existence of true informed consent, where users and platforms are both fully cognizant of the future value of data. In such a scenario, where there is minimal information asymmetry, users either hand over their data or do not hand over their data, but the decision is made based on knowledge of the future value of the data. In particular, this scenario gives users more agency to refuse to hand over their data and retain the ability to formally end their consent at a later date. This would vastly reduce the power of platforms to utilize data based on their own motivations, but be instead aware of the interests and rights of the users.

Such a scenario is, however, idealistic rather than realistic. The limitations of informed consent, as discussed above, are considerable and this scenario would only continue as long as data was not handed over to third parties, where the relationships and combinations of data would make true informed consent difficult if not impossible. This scenario is further negated

by the unknown nature of data value. While a burden lies on consumers to become "truly informed" on the data's value, a broader burden still lies on platforms to make their data practices transparent enough so that the value of the data is apparent.

Scenario 2: Bets on the Future

The second scenario acts as a mirror to the first scenario, positing the existence of a true unknown-unknown relationship, where both the platform and the user are unaware of the future value of the data. This scenario is most likely since, as described above, the value of data is often unknown at the time of collection. To make an example, rating and review data collected by Airbnb, Uber, and similar sharing platforms will become more valuable if citizen scoring is widely introduced.⁴⁶ If this is the case, secondary businesses might be founded that want to interface with rating and review data for their own purposes. However, whether and where citizen scoring will be introduced is, among others, a political question and can thus not be perfectly predicted, limiting the predictability of the future value of rating and review data.

As elaborated, platforms, and especially large ones, collect data excessively as a bet on the future, where certain scenarios might result in a gold rush, even if they are not very likely, while other scenarios might result in a net loss. At the same time, some data is already monetized and thus clearly assigned known future value as these monetization streams are unlikely to stop any time soon. Moreover, the sharing economy presents a heterogeneous context where smaller platforms might lack the motivation, capability, or business philosophy to bet on the future value of data through excessive data collection. With the increasing commercialization of the sharing economy, we might see more sharing platforms become big enough and ready to

⁴⁶ Simon Denyer, *China Wants to Give All of its Citizens a Score – and Their Rating Could Affect Every Area of Their Lives*, THE INDEPENDENT (Oct. 22, 2016, 7:55 PM), <https://www.independent.co.uk/news/world/asia/china-surveillance-big-data-score-censorship-a7375221.html>; Rachel Botsman, *Big Data Meets Big Brother as China Moves to Rate Its Citizens*, Wired (Oct. 21, 2017), <https://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion>; Rogier Creemers, *China's Social Credit System: An Evolving Practice of Control* (May 9, 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3175792. Citizen scoring or social credit system describes national reputation systems that attempt to rate individual citizens' trustworthiness based on large-scale surveillance and big data. The most prominent example is the social credit system currently implemented by the Chinese government.

leverage data mining for bets on the future.⁴⁷ Emerging platforms, due to institutional isomorphism⁴⁸ and market pressure, are likely to invest heavily in data science, exploring ways to amass and monetize user data that has not been exploited for current business activities.

Finally, the bets on the future scenario (as well as the informed consent scenario) obfuscates the power dynamic present between users and platforms, as one party still owns the data. In other words, the real issues for deliberation lie in those two scenarios where one party is in a better situation to the other – in terms of an information advantage. We thus turn to the asymmetric scenarios 3 and 4.

Scenario 3: Alternative Use

The remaining two scenarios depend on the existence of information asymmetries between users and platforms with regard to the value of data. The scenario of alternative use presents a situation where the future value of data is known to the users, but not to the platforms. It is at least conceivable that users may challenge the distribution and appropriation of their data. Instances of reverse engineering⁴⁹, developing alternative platforms, or value creating plugins,⁵⁰ for instance, which the platforms might not necessarily be privy to, all speak to this. Recent literature on big data and data mining has looked at such practices, stressing users' agency instead of understanding them as passive entities.⁵¹ As such, users may have a better understanding of how data can be valued and thus take more agency. Practices related to this are refusing to hand over the data without obfuscation, or taking action to retrieve the data later

⁴⁷ Leandro DalleMule & Thomas H. Davenport, *What's Your Data Strategy?* May-June HARV. BUS. REV. (2017), <https://hbr.org/2017/05/whats-your-data-strategy>; According to industry surveys, a large proportion of organizations fails to leverage big data for decision-making and organizational development. See also NICOLAUS HENKE ET AL., *THE AGE OF ANALYTICS: COMPETING IN A DATA-DRIVEN WORLD*, MCKINSEY GLOBAL INSTITUTE (2016), <https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Analytics/Our%20Insights/The%20age%20of%20analytics%20Competing%20in%20a%20data%20driven%20world/MGI-The-Age-of-Analytics-Full-report.ashx>.

⁴⁸ Paul DiMaggio & Walter W. Powell, *The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields*, 48 AM. J. SOC. 146 (1983).

⁴⁹ Nicholas Diakopoulos, *Algorithmic Accountability: Journalistic Investigation of Computational Power Structures*, 3 DIGITAL J. 398 (2015). The practice of reverse engineering is described as “the process of articulating the specifications of a system through a rigorous examination drawing on domain knowledge, observation, and deduction to unearth a model of how that system works” (*id.* at 404). In the context of reputation systems, for example, a reverse engineering approach would try to uncover why certain users have the ratings they have, thus opening up algorithmic black boxes.

⁵⁰ Neil Selwyn & Luci Pangrazio, *Doing Data Differently? Developing Personal Data Tactics and Strategies amongst Young Mobile Media Users*, 5 BIG DATA SOC. 1 (2018).

⁵¹ Nick Couldry & Alison Powell, *Big Data from the Bottom Up*, 1 BIG DATA SOC. 1 (2014); Helen Kennedy & Giles Moss, *Known or Knowing Publics? Social Media Data Mining and the Question of Public Agency*, 2 BIG DATA SOC. 1 (2015); Deborah Lupton, *Feeling your Data: Touch and Making Sense of Personal Digital Data*, 19 New Media Soc. 1599 (2017); Milan, *supra* note 11.

for themselves.⁵² While an optimistic scenario in terms of user data agency, this scenario raises its own questions in terms of privacy and data protection. For instance, if only a sub-group of users repurpose platform data, then a privacy imbalance remains as each individual user would not know who was using and monetizing their data. Moreover, navigating platforms to extract value from data on the user side requires advanced data science skills and knowledge. On the user side, few people are able or motivated to strategically leverage these skills. Thus, important questions of digital inequalities emerge for this scenario, particularly the question which users benefit and which users suffer from users "knowing" the future value of the data.⁵³

Since platforms want to keep ownership of their data, they are very quick to prohibit data collection through restrictive terms and conditions. Facebook, to take an important example outside the sharing economy, states "you may not access or collect data from our Products using automated means (without our prior permission) or attempt to access data you do not have permission to access."⁵⁴ Further complicating this scenario is the fact that most larger platforms, both in the sharing economy and in other online settings, rely on machine learning algorithms⁵⁵ for many of their operations. These algorithms are opaque, automated, and not accessible to outsiders.⁵⁶ This severely limits even technically savvy users from knowing and leveraging the future value of data, let alone acting upon it.

The argument, suggested by this scenario, is that alternative data uses may create additional value for everybody. The data that is currently monopolized by platforms could be put to good additional use. Hence, a similar transfer into the public domain is conceivable. Indeed, data is not a finite resource and use by one party does not prohibit use elsewhere. The gains would be obvious. It would allow user innovation at the edges and innovations whose added value, assuming they enter the commons, could benefit everyone. However, this is a scenario which is not enabled through the current setup and made further unlikely through power imbalances. Platforms have a prerogative to encroach on alternative use, and for all intents and purposes to act strongly to prevent reverse engineering or independent control over the data accumulated. User creativity is feeble and can always be crushed, whereas the data resides in a

⁵² Finn Brunton & Helen Nissenbaum, *Vernacular Resistance to Data Collection and Analysis: A Political Theory of Obfuscation*, 16 FIRST MONDAY 1 (2011).

⁵³ Andrejevic, *supra* note 3; Matthew McCarthy, *The Big Data Divide and Its Consequences*, 10 SOC. COMPASS. 1121 (2016).

⁵⁴ Facebook, *Terms of Service* (2019), <https://www.facebook.com/terms.php> (last visited Feb. 24, 2019).

⁵⁵ Burrell, *supra* note 35.

⁵⁶ PASQUALE, *supra* note 35.

walled garden within the reach of the platforms. It would also be difficult to find a good justification for this argument. The platforms themselves could risk violating laws and are therefore expected to shy away from giving more access and agency to users that goes beyond the mere control over personal data.

For sharing economy platforms, such as Airbnb and Uber, the empirical reality has been that they tend to collaborate with outside researchers only on very limited terms and on select basis. Several working papers using Uber data involved the collaboration of at least one current or former Uber employee.⁵⁷ Similarly, Airbnb's release of data about New York listings in December 2015 seemed to have been driven by strategic branding, presenting the platform as open-minded and collaborative, rather than by true user empowerment. Immediately prior to the release, Airbnb had removed about 1000 illegal listings from the data set, as revealed by two data activists who had independently scraped Airbnb listings data.⁵⁸ The data activists had found that the removed listings were from hosts with multiple properties, which, according to New York state law, is illegal. In subsequent months, Airbnb faced increasingly tight regulation in New York. Airbnb's data manipulation, as uncovered by the data activists, signalled an untrustworthy image of the sharing platform to the authorities.⁵⁹ The example shows the risk of platforms opening up. Thus, it is likely that the scenario might apply more strongly for smaller, less tightly observed platforms.

Scenario 4: Finders Keepers

The final scenario, termed "finders keepers," most accurately reflects the status quo of data value information. Platforms have greater awareness of the current and future value of data, whereas users are not aware of the value the data holds. If users have some awareness, it is still less than platforms. Yet, the biggest challenge is not that platforms know the value of data, it is that platforms exclusively command all the data and in practice use this power imbalance to create a norm that platforms are the sole owners of the data, and thus sole purveyors of innovation with said data.

⁵⁷ See, e.g., Siddhartha Banerjee et al., *Dynamic Pricing in Ridesharing Platforms*, 15 (1) ACM SIGECOM EXCHANGES 65 (July 2016); M. Keith Chen & Michael Sheldon, *Dynamic Pricing in a Labor Market: Surge Pricing and Flexible Work on the Uber Platform*, Mimeo UCLA (2015).

⁵⁸ Miranda Katz, *A Lone Data Whiz is Fighting Airbnb and Winning*, WIRED (Feb. 10, 2017, 1200 AM), <https://www.wired.com/2017/02/a-lone-data-whiz-is-fighting-airbnb-and-winning/>.

⁵⁹ *Id.*

Moving toward a fairer exchange would require thorough restraint on the platforms' side, and oversight to guarantee predictability for all sides involved. The problem resides that, left to their own devices, platforms have a natural tendency to leave everything open to possibility. This, we would argue, is inefficient, as they lack the capabilities to realize all of the conceivable services that could be developed from data. Thus, an argument could be made that platforms need to limit the scope of how they will use the data, or at least not claim ownership over every infinite use of the data. At the same time, as mentioned, platforms deploy sophisticated machine learning technology and might not always be in full control of what is happening themselves.⁶⁰ This complicates the limitation of scope. Instead of *ex ante* predictability, whereby users would know in advance what result will occur from an algorithm, *ex post* explanation might therefore be more viable and useful. Here, it may be sensible to push for greater explicability about why decisions were made and upon which data. In that regard, the GDPR seemingly introduces a right to explanation for automatic and algorithmic decision-making. However, the feasibility and precision of this right to explanation have been called into question, so that it might be more appropriate to talk about a right to be informed.⁶¹

Nevertheless, we would argue that this argument for at least co-innovation⁶² (if not co-ownership) could particularly be made for the sharing economy. In essence, sharing economy platforms depend more on user goodwill than other technology ventures, and their value creation is marked by significant user involvement. Furthermore, the platforms that are dominant today are so only partly because of their own superior design and technology, but partly also due to a community that grew together with them. Likewise, even today platforms need their users to share on a constant basis, hence this imbalance in intent and language remains ironic. Our argument thus is that a nuanced approach that allows for the mutual interests of companies and users would create a more equitable data exchange.

⁶⁰ Burrell, *supra* note 35.

⁶¹ Sandra Wachter et al., *Why a Right to an Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, 7 I. D. P. L. 76 (2017). It would be beyond the scope of this Article to discuss the right to an explanation (or the lack thereof) in the context of the sharing economy. However, given the scenario of Finders Keepers, the call would be for sharing platforms to offer more transparent communication how and why personal data is used. This is in line with the transparency principles of the GDPR. See Heike Felzmann et al., *How Transparent Is Your AI? Ethical, Legal, and Societal Issues of the Transparency Principle in Cyber-Physical Systems* (Amsterdam Privacy Conference, Working Paper, October 2018).

⁶² We use the term co-innovation here in analogy to the concept of co-creation. Co-creation describes a joint value creation process between companies and customers, where companies and customers join efforts to develop new products and services or enhance existing ones. In the context of the sharing economy, co-innovation would occur if platforms and users engage in co-creation, for example through improving the sharing experience based on user feedback. Coimbatore Prahalad & Venkat Ramaswamy, *Co-Creation Experiences: The Next Practice in Value Creation*, 18 J. INTERACT. MARK. 5 (2004).

Conclusion

In this article, we described the data collection and mining activities of commercial platforms in the sharing economy. More specifically, we analyzed such practices in terms of the unclear future value of data, where big data becomes a speculative investment. We discussed the ethical implications of the future value of data based on four scenarios that denote different levels of platform and user power. The symmetric scenario of informed consent, where both companies and users know the future value of data, seemed unrealistic due to various technical constraints and information asymmetries inherent within the platform ecosystem. The symmetric scenario of bets on the future, where both companies and users do not know the future value of data, seemed more realistic but is undesirable from a user- and platform-perspective. We could argue that this scenario is also undesirable from a societal perspective as users lack the power to gain more agency of their data and platforms lack the power to create beneficial and responsible innovations. The third and fourth scenarios are both characterized by power asymmetries between platforms and users. From a user point-of-view, the alternative use scenario seems to be the most promising and beneficial. In this scenario, users have awareness and control of their data and are able to leverage it for personal benefits. However, our discussion showed the barriers and issues with this scenario that is paralleled in movements that attempt to strengthen user agency, for example, the open data, open source, copyleft and open hardware communities. A true alternative use scenario would probably require a non-commercial set-up of sharing platforms, for example through platform cooperatives.⁶³ Such platforms, however, might be hard to scale. Finally, the finders keepers scenario is arguably the best one for commercial sharing platforms since, in this scenario, the power asymmetries are in their favor. Particularly innovative platforms could increasingly leverage their market power for digital market manipulation and ever more sophisticated surveillance.⁶⁴

The ethics of data value exchange might come down to two conceivable futures. One is where platforms, enabled through their exclusive ownership of data, may place bets on everything and will succeed in some instances, and fail in others. They will generate value for their shareholders in the process, but most likely at the expense of user innovation which, assuming

⁶³ Trebor Scholz, *Platform Cooperativism. Challenging the Corporate Sharing Economy*, (2016).

⁶⁴ Ryan Calo, *Digital Market Manipulation*, 82 GEO. WASH. L. REV. 995 (2013); CALO & ROSENBLAT, *supra* note 32.

some restraints of domain on the platforms' part, might be far greater in sum. Another conceivable future would seem more desirable: One where platforms act more cooperatively, and innovation is enabled through data sharing and collaboration on society's edges. As an outlook, technological developments and innovations on the platform side are likely to shift the sharing economy data eco-system increasingly from bets on the future to Finders Keepers. Different stakeholders, including policy makers, the platforms themselves, the users, the media, and educational institutions, need to work together to instead allow for more user-friendly scenarios, either alternative use or – even better – informed consent. The GDPR is a first step in that direction but it needs to be complemented by educational initiatives that teach data awareness, literacy and agency, by platform self-restraint, and by critical media coverage about problematic data practices by platforms and users.