

Personal data markets

Sarah Spiekermann¹ · Rainer Böhme³ · Alessandro Acquisti² · Kai-Lung Hui⁴

Published online: 30 April 2015

© Institute of Information Management, University of St. Gallen 2015

On every ordinary day, BlueKai transacts over 75 million online auctions for personal information.¹ The company, which belongs to Oracle, says it owns 750 million user profiles² of people who regularly surf the web, and it processes more than 30,000 attributes³ about these users. BlueKai claims to run the world's largest third-party data market place, but it is just one player in a huge web of over a thousand firms that have established themselves in the business that some call “the new oil”: personal data.

Personal data markets thrive, driving online companies' valuation and fueling Internet economics. At the same time this data is not just an ordinary tradable asset. Personal data can be highly sensitive and revealing about a person's identity; processing it is legally restricted by data protection and privacy laws. In many countries, privacy and the right to information self-determination are recognized as a human right. And even

among major high-tech companies, privacy protection now starts to be recognized as essential. While in 1999 Sun founder Scott McNealy claimed that privacy is dead and we should get over it, 2015 saw Apple's CEO Tim Cook say that “information can make the difference between life and death. If those of us in positions of responsibility fail to do everything in our power to protect the right of privacy, we risk something far more valuable than money; we risk our way of life. Fortunately, technology gives us the tools to avoid these risks and it is my sincere hope that by using them and by working together, we will.”⁴

This special issue is placed at the intersection of these seemingly opposing poles of privacy and personal data markets. What are the economic, technical, legal, and business challenges faced by business models of companies like BlueKai? And what about the legitimacy and ethicality of those business models? By presenting a series of papers from global industry players and high-profile academics, spanning rigorous empirical, theoretical and conceptual work, we attempt to provide insight into the complexities of personal data markets and ways to manage and protect privacy within those markets.

We start with an industry perspective provided by Björn Roeber, Olaf Rehse, Robert Knorrek and Benjamin Thomsen of the Boston Consulting Group (Roeber et al. 2015). As one of world's largest consulting houses, BCG has been watching the development of personal data markets closely. It recently predicted that the economic use of personal data can deliver up to EUR 330 billion in annual economic benefit to organizations in Europe alone by 2020.⁵ To understand whether people would be willing to participate in such markets despite privacy concerns the company surveyed 3000 European

¹ OECD, 2013, Supporting Investment in Knowledge Capital, Growth and Innovation

² https://docs.oracle.com/cloud/latest/daasmarketing_gs/DSMKT/GUID-418EDA59-1BD9-40F6-9D57-DD7C266555FF.htm#DSMKT3616

³ <http://bluekai.com/audience-data-marketplace.php>

✉ Sarah Spiekermann
spiek@wu.ac.at

Rainer Böhme
rainer.boehme@icsi.berkeley.edu

¹ Vienna University of Economics and Business (WU Wien), Vienna, Austria

² Heinz College, Carnegie Mellon University, Pittsburgh, PA 15213, USA

³ Institut für Wirtschaftsinformatik, Westfälische Wilhelms Universität Münster, Münster 48149, Germany

⁴ School of Business and Management, The Hong Kong University of Science and Technology, Hong Kong, China

⁴ <http://uk.businessinsider.com/tim-cook-on-online-privacy-2015-2?r=US>

⁵ BCG Report published in the Liberty Global Policy Series; available at: <http://www.libertyglobal.com/PDF/public-policy/The-Value-of-Our-Digital-Identity.pdf> (last visited March, 20th, 2015)

citizens. In this special issue, they report on the scientific background of their study. In line with earlier privacy research they find that nearly all consumers are generally willing to share personal data with organizations, but this sharing depends on the benefits and terms of the exchange as well as the context. Context of data exchange is even more important than the data itself. Consumers tend to clearly discriminate between organizations from various industry sectors, ranking online shops, retailers, and loyalty card providers highest in trustworthiness and social networking services or banks lowest. Thus, some traditional industry players seem to be better positioned than others in the new market arena. People accept active sharing where they are consciously involved in the exchange, but are much less positive about passive collection of information. Data use is fine for them if it is part of an ongoing relationship. For service delivery and marketing purposes companies seem to be allowed to use data, but third-party sharing of data is not appreciated; not in anonymous form and least so in an identified manner. These latter results suggest that personal data markets in their current form will have difficulties to find acceptance among people. In particular identified data use by third parties leads to a strong utility drop for consumers.

Vasilis Gkatzelis, Christina Aperjis and Bernardo Huberman from Stanford have anticipated exactly this identifiability concern and risk aversion and envision a personal data market environment where profiles are brokered anonymously and on a voluntary basis (Gkatzelis et al. 2015). They present a novel pricing mechanism for personal data in scenarios where the data buyer is interested in accurate aggregate information, such as estimates of population means, and needs to pay sellers for contributing their private information. Finding the right price is difficult because too low offers may exclude potential sellers who do not feel fairly compensated and thus opt out from sharing their personal data. This would lead to biased samples and poor estimates. Too high offers are uneconomical for the buyer. The proposed method uses a bundling mechanism to determine the lowest price for unbiased samples given some knowledge about the privacy risk attitudes in the population. Their theoretical model adds to the growing literature on incentives for sharing personal data, which falls right into the scope of personal data markets and privacy.

Another empirical contribution by Irina Heimbach, Jörg Gottschlich and Oliver Hinz from TU Darmstadt leverages a so-far untapped industry dataset (Heimbach et al. 2015). The authors explore the value of third-party use of profile data of online social networking services in e-commerce. Specifically, Facebook profile data can significantly improve the quality of product recommendations. This applies in particular to users with short purchase histories at the specific vendor who would otherwise receive random recommendations. This innovative approach to address the well-known bootstrapping problem

for recommender systems illustrates how to tap the value of personal data through more targeted recommendations in at least two ways. Better recommendations promise additional sales and they increase customer satisfaction, as supported with evidence in the contributed article. It also sheds some light on the question which type of data among the wealth of information in a typical Facebook profile is most useful for this specific purpose. Knowledge about group membership is the most stable predictor, dwarfing more sensitive items such as demographics. This result adds a piece to the mosaic of evidence suggesting that there exist viable ways to align business interest with privacy protection.

Thierry Rayna, John Darlington, and Ludmila Striukova from the ESG School of Management in Paris study the personalized pricing made possible by personal data (Rayna et al. 2015). Consumers are notoriously wary about price discrimination, for fear of being charged higher prices for a given product or service. However, the authors show that it is possible to achieve a situation in which price discrimination is mutually advantageous by rewarding consumers for disclosing personal information. The article examines the conditions under which both buyers and sellers will gain by adopting this pricing model, and show the impact on social welfare. Crucially, the feasibility of mutually advantageous personalized prices relies on firms' ability to monitor consumers. If consumers' actions remain partially hidden, their self-interested behavior may prevent the establishment of this forms of price discrimination.

The scientific contributions from Stanford and ESG envision mechanisms for personal data markets where they presume that people will share personal data with market players for appropriate returns and under mutually agreed conditions (Maguire et al. 2015). But how can people be ensured that data recipients will really treat the data they receive in the way they promise to? How can people's trust in personal data markets be strengthened to a degree that they might become active participants in them? The author team from Microsoft, Sean Maguire, Jeffrey Friedberg, Carolyn Nguyen and Peter Haynes make a technical proposal on how to embed more trust and accountability in the data-sharing ecosystem. They describe a metadata based architecture for user-centered data accountability. At the core of their proposal is to bind policies and permissions negotiated with users to the data that is being collected from them. These permissions travel with the personal data as metadata. Before any entity can process the personal data it must consult the permissions and then act accordingly. A record of interactions is being established, which users and authorities may be able to consult to control proper data handling. If regulators made it mandatory for data market players to systematically negotiate, collect and respect data exchange policies that are bound to the data, a new degree of accountability would be created in personal data markets.

The crucial role of regulators is also highlighted in the final piece of this special issue; a contribution from us, the guest editors of this special issue (Spiekermann et al. 2015). One of the most salient liabilities of holding personal data we see arising is the legal uncertainty surrounding its management today. Privacy regulation is an evolving and among the least globally harmonized fields of law. Many companies today, that process personal data, operate in legal grey zones. Most importantly, most customers are probably not aware of the extent to which their personal data is now being processed by companies. The empirical research on people's privacy expectations suggests that they might be badly surprised when finding out. The position we lie down in our contribution to this special issue is therefore that companies, which hold customer relationships, should go back to more trustworthy relationships with their customers. This implies that they should respect peoples' data protection expectations and consider more carefully whether and how to engage with third parties. At the moment we observe both promise and hype around the idea of building new markets with personal data "oil". But we caution that hypes typically go through cycles, and that we may soon face a period

of disillusionment in which the economic and societal value of personal data assets will need to be carved out.

References

- Gkatzelis, V., Aperjis, C., & Huberman, B. A. (2015). Pricing private data. *Electronic Markets*, 25(2). doi:10.1007/s12525-015-0188-8.
- Heimbach, I., Gottschlich, J., & Hinz, O. (2015). The value of user's Facebook profile data for product recommendation generation. *Electronic Markets*, 25(2). doi:10.1007/s12525-015-0187-9.
- Maguire, S., Friedberg, J., Nguyen, M.-H. C., & Haynes, P. (2015). A metadata-based architecture for user-centered data accountability. *Electronic Markets*, 25(2). doi:10.1007/s12525-015-0184-z.
- Rayna, T., Darlington, J., & Striukova, L. (2015). Pricing music using personal data: mutually advantageous first-degree price discrimination. *Electronic Markets*, 25(2). doi:10.1007/s12525-014-0165-7.
- Roeber, B., Rehse, O., Knorrek, R., & Thomsen, B. (2015). Personal data: how context shapes consumers' data sharing with organizations from various sectors. *Electronic Markets*, 25(2). doi:10.1007/s12525-015-0183-0.
- Spiekermann, S., Acquisti, A., Böhme, R., & Hui, K.-L. (2015). Challenges for personal data markets and an advice for its stakeholders. *Electronic Markets*, 25(2). doi:10.1007/s12525-015-0191-0.