HOW DO ENTREPRENEURIAL FIRMS APPROPRIATE VALUE IN OPEN BIO DATA INFRASTRUCTURES: AN EXPLORATORY QUALITATIVE STUDY

Hannes Rothe[1], Sirkka Jarvenpaa[2], Anna Penninger[1], and Katharina Lauer[3]

[1] Freie Universität Berlin, Berlin, Germany, [hannes.rothe, anna.penninger]@fu-berlin.de
[2] The University of Texas at Austin, USA, sirkka.jarvenpaa@mccombs.utexas.edu
[3] ELIXIR Hub, United Kingdom, katharina.lauer@elixir-europe.org

1 Introduction

With the rise of digital infrastructures and platforms, the question of how value is created and appropriated by complementors in digital ecosystems becomes ever more important (Constantinides, Henfridsson, & Parker, 2018; de Reuver, Sørensen, & Basole, 2017). Advances in digital technologies led to increasing efficiencies of data transfer, storage, and processing. This enabled scalable architectural designs and finally gave rise to digital platforms across a variety of domains (Sanchez & Mahoney, 1996; Tiwana, 2008, 2014). While product-based (Evans & Gawer, 2016; Piezunka, 2011) and industry platforms (Adner & Kapoor, 2010; Cusumano & Gawer, 2002; Gawer & Cusumano, 2008, 2014) are well established categories in management research, we became aware that digital technologies change the mediating role of these 'two-sided markets' (de Reuver et al., 2017). Today, platform providers offer massive amounts of data to complementors in real-time and around the globe (Evans & Basole, 2016; Evans & Gawer, 2016). Complementors integrate this data as boundary objects into their daily practices and activities in order to create added value for their customers (Eaton, Elaluf-Calderwood, Sorensen, & Yoo, 2015; Ghazawneh & Henfridsson, 2013). Together, platform providers and their complementors co-create value as they form a platform ecosystem (Ceccagnoli, Forman, Huang, & Wu, 2011; Parker, Alstyne, & Jiang, 2017; Tiwana, 2014).

So far, platform literature has largely focused on provider strategies (Cennamo & Santalo, 2013; McIntyre & Srinivasan, 2017; Zhu & Iansiti, 2012) and less on how complementors appropriate value from centrally provided resources such as data. The discourse on the hidden value of 'big data' in recent years (e.g., Abbasi, Sarker, & Chiang, 2016; Chen, Chiang, & Storey, 2012; McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012) revealed the importance of data as a key resource for business and research, particularly for value creation in platform ecosystems (Tiwana, 2015). Despite the importance of data as a core resource, our knowledge about appropriation strategies applied by complementors from data infrastructures, respectively digital ecosystems, is still limited. We refer to appropriability as the complementor's effectiveness in exploiting data by processing and translating it into a solution or service to a customer (Teece, 1986). Assuming that complementors join and stay in a platform ecosystem when they are able to appropriate financial value from it, we focus on their ability to *protect their own resources* and operate on the *data infrastructure* (Pisano & Teece, 2007) – in this case the ecology of platform ecosystems providing open data.

Complementors apply data to the specific use cases of multiple target groups and finally grow the user base of a platform (Cennamo & Santalo, 2013; Iansiti & Zhu, 2007; Rietveld & Eggers, 2018). They solve concrete problems as they apply this data with IT artefacts in specific contexts (Abbasi et al., 2016; Adomavicius & Tuzhilin, 2005). As such, they behave as digital entrepreneurs who use platform data and functions in novel ways. They create new opportunities in the market and help "moving the economy closer to the technological frontier" (Sussan & Acs, 2017). Digital entrepreneurs utilize data to engage in future-shaping practices (Kelestyn & Henfridsson, 2014; OECD, 2015). At the same time, we learn from digital entrepreneurship, that complementors need to find a sustainable appropriation strategy that allows them to scale (Kelestyn & Henfridsson, 2014; Nambisan, 2016; Nambisan, Lyytinen, Majchrzak, & Song, 2017). This may inflict a conflict for complementors – appropriate value from the data infrastructure within a specific customer context in forms of a steady revenue stream while at the same time being able to scale. Hence, we raise the question: *How do complementors appropriate value from open data in platform ecosystems?*

One of the largest areas for using open data lies in bio data for healthcare and bio technology. Particularly genomic data has gained massive attention following completion of the human genome project in 2003 (e.g., Lander, 2011). This is mainly due to cost reductions for creating data: "The first human genome cost $1.3 billion, now it costs less than $700. You can do it. 17 years ago, it took almost a decade, now you can have it in 48 hours if you have the machine." (interview #11) Today, the numbers of ventures that make use of human, animal, microbiome, or plant bio data is ever more growing. At the same time, data in these domains is growing with exponential rates (Birney, Vamathevan, & Goodhand, 2017; Jagadish et al., 2014). Most data in genomics is offered on open data platforms on an anonymized basis; raw data is mostly accessible through bio banks for research purposes. Data is offered as a common-pool resource to overcome the legal challenges and technical control associated with it (Koutroumpis & Leiponen, 2013). While more and more ventures are created, the commercial use of bio data has been largely overlooked in the domain of information systems (for exceptions, see e.g., Jarvenpaa & Markus, 2018; Vassilakopoulou, Skorve, & Aanestad, 2016). At the same time, the focus on platform providers and their governance models has also prevailed in terms of biodata (Jarvenpaa & Markus, 2018; Vassilakopoulou et al., 2016; West, 2017). For researchers and practitioners, it remains unclear how these complementors appropriate financial value from bio data published on open platforms. In most cases, even the source of this data is not revealed; neither in reports nor publications. Hence, entrepreneurs "might use UniProd [a bio data platform], but not gonna cite it" (interview #2).

We shed light on how complementors appropriate value from open data in platform ecosystem by outlining how they access, process, and transfer this data along the value chain (see Table 1). For this purpose, we introduce a qualitative exploratory study in the domain of bio data, namely genomics. Based on 26 interviews across thirteen complementors, investors and platform providers, as well as complimentary secondary data we propose six value appropriation strategies for complementors in open data platform ecosystems. Finally, we infer three mechanisms of how financial value is appropriated from the data infrastructure: decontextualizating, recontextualizing, and contextualizing open data.

Appropriation Strategies	**Value Chain for Data Analysis**				
	Data Selection	**Data Integration & Preprocessing**	**Data Transformation**	**Data Mining**	**Data Interpretation**
Custom Full Service (A)					
High Quality Data Service (B)					
Vocabulary Generation (C)					
Configurable Model (D)					
Care Service (E)					
Workflow as a Service (F)					

Table 1 Appropriation strategies on the value chain. core activity (dark), infrastructure (light)

References

Abbasi, A., Sarker, S., & Chiang, R. H. L. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(2).

Adner, R., & Kapoor, R. (2010). Value creation in innovation ecosystems: How the structure of technological interdependence affects firm performance in new technology generations. *Strategic Management Journal*, *31*(3), 306–333.

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions On*, *17*(6), 734–749.

Birney, E., Vamathevan, J., & Goodhand, P. (2017). Genomics in healthcare: GA4GH looks to 2022. *BioRxiv*, 203554.

Ceccagnoli, M., Forman, C., Huang, P., & Wu, D. J. (2011). Co-creation of value in a platform ecosystem: The case of enterprise software. *Management Information Systems Quarterly*, *36*(1),

263–290.
Cennamo, C., & Santalo, J. (2013). Platform competition: Strategic trade-offs in platform markets. *Strategic Management Journal*, *34*(11), 1331–1350. https://doi.org/10.1002/smj.2066
Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, *36*(4).
Constantinides, P., Henfridsson, O., & Parker, G. G. (2018). Introduction—Platforms and Infrastructures in the Digital Age. *Information Systems Research*, *29*(2), 381–400.
Cusumano, M. A., & Gawer, A. (2002). The Elements of Platform Leaderhsip. *MIT Sloan Management Review*, (Spring), 51–58. https://doi.org/10.1371/journal.pone.0015090
de Reuver, M., Sørensen, C., & Basole, R. C. (2017, April 11). The digital platform: a research agenda. *Journal of Information Technology*, pp. 1–12.
Eaton, B., Elaluf-Calderwood, S., Sorensen, C., & Yoo, Y. (2015). Distributed tuning of boundary resources: the case of Apple's iOS service system. *Mis Quarterly*, *39*(1), 217–243.
Evans, P. C., & Basole, R. C. (2016). Revealing the API Ecosystem and Enterprise Strategy via Visual Analytics. *Communications of the ACM*, *59*(2), 26–28.
Evans, P. C., & Gawer, A. (2016). The rise of the platform enterprise: a global survey. *The Center for Global Enterprise*, (January), 1–30.
Gawer, A., & Cusumano, M. A. (2008). How Companies Become Platform Leaders. *MIT Sloan Management Review*, *49*(2), 28.
Gawer, A., & Cusumano, M. A. (2014). Industry platforms and ecosystem innovation. *Journal of Product Innovation Management*, *31*(3), 417–433.
Ghazawneh, A., & Henfridsson, O. (2013). Balancing platform control and external contribution in third-party development: the boundary resources model. *Information Systems Journal*, *23*(2), 173–192.
Iansiti, M., & Zhu, F. (2007). Dynamics of platform competition: Exploring the role of installed base, platform quality and consumer expectations. In *ICIS 2007 Proceedings* (p. 38).
Jagadish, H. V, Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., & Shahabi, C. (2014). Big data and its technical challenges. *Communications of the ACM*, *57*(7), 86–94.
Jarvenpaa, S., & Markus, M. L. (2018). Genetic Platforms and Their Commercialization: Three Tales of Digital Entrepreneurship. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.
Kelestyn, B., & Henfridsson, O. (2014). Everyday digital entrepreneurship: The inception, shifts, and scaling of future shaping practices. In *Proceedings of the Thirty Fifth International Conference on Information Systems* (p. 8). Auckland, New Zealand.
Koutroumpis, P., & Leiponen, A. (2013). Understanding the value of (big) data. In *Proceedings of 2013 IEEE international conference on big data* (pp. 38–42).
Lander, E. S. (2011). Initial impact of the sequencing of the human genome. *Nature*, *470*(7333), 187.
McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big Data. *The Management Revolution. Harvard Bus Rev*, *90*(10), 61–67.
McIntyre, D. P., & Srinivasan, A. (2017). Networks, platforms, and strategy: Emerging views and next steps. *Strategic Management Journal*, *38*(1), 141–160.
Nambisan, S. (2016). Digital Entrepreneurship: Toward a Digital Technology Perspective of Entrepreneurship. *Entrepreneurship Theory and Practice*. https://doi.org/10.1111/etap.12254
Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. (2017). Digital Innovation Management: Reinventing innovation management research in a digital world. *Mis Quarterly*, *41*(1).
OECD. (2015). *Digital Transformation of European Industry and Enterprises: : A Report of the Strategic Policy Forum on Digital Entrepreneurship*.
Parker, G., Alstyne, M. Van, & Jiang, X. (2017). Platform Ecosystems: How Developers Invert the Firm. *Management Information Systems Quarterly*, *41*(1), 255–266.
Piezunka, H. (2011). Technological platforms. *Journal Für Betriebswirtschaft*, *61*(2), 179–226.
Pisano, G. P., & Teece, D. J. (2007). How to capture value from innovation: Shaping intellectual property and industry architecture. *California Management Review*, *50*(1), 278–296.
Rietveld, J., & Eggers, J. P. (2018). Demand Heterogeneity in Platform Markets: Implications for Complementors. *Organization Science*, *29*(2), 304–322.
Sanchez, R., & Mahoney, J. T. (1996). Modularity, flexibility, and knowledge management in product

and organization design. *Strategic Management Journal*, *17*(S2), 63–76.
Sussan, F., & Acs, Z. J. (2017). The digital entrepreneurial ecosystem. *Small Business Economics*, 1–19.
Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, *15*(6), 285–305.
Tiwana, A. (2008). Does technological modularity substitute for control? A study of alliance performance in software outsourcing. *Strategic Management Journal*, *29*(7), 769–780.
Tiwana, A. (2014). The Rise of Platform Ecosystems. In *Platform Ecosystems* (pp. 3–21). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-408066-9.00001-1 M4 - Citavi
Tiwana, A. (2015). Evolutionary Competition in Platform Ecosystems. *Information Systems Research*, *26*(2), 266–281.
Vassilakopoulou, P., Skorve, E., & Aanestad, M. (2016). A Commons Perspective on Genetic Data Governance: the Case of Brca Data. In *Proceedings of European Conference on Information Systems* (p. 136).
West, J. (2017). Open Source Platforms Beyond Software: From ICT to Biotechnology. In *Entrepreneurship, Innovation, and Platforms* (pp. 337–370). Emerald Publishing Limited.
Zhu, F., & Iansiti, M. (2012). Entry into platform-based markets. *Strategic Management Journal*, *33*(1), 88–106. https://doi.org/10.1002/smj.941