The Resource Dependency of Data: A Prospective on Data Sharing in Supply Chains

Abstract: In addition to traditional production factors, such as capital and labor, data can also be considered a similar resource. Furthermore, data cannot be created and handled without the involvement of other resources. In contemporary supply chain structures, companies rarely produce all the resources required by the final product internally. Instead, they partially rely on external inputs from other supply chain participants. This proportion is reflected by the added value captured by each supply chain participant. This study examines the extent of sharing of data by companies to their supply chain participants. The study explores how these data sharing percentages would look if data was shared in the same proportion as the value added is captured by the supply chain participants. While the theoretical figures on data sharing based on the assumption of similar proportions of data sharing and value added are high, in practice no correlation can be seen between the value added contribution and data sharing within the supply chain.

Timo Seppälä

the Research Institute of the Finnish Economy and Aalto University, timo.seppala@etla.fi

Esko Hakanen,

Aalto University, esko.hakanen@aalto.fi

Ilkka Lähteenmäki

Aalto University and Oulu University, ilkka.lahteenmaki@aalto.fi

Juri Mattila

the Research Institute of the Finnish Economy and Aalto University, juri.mattila@etla.fi

Rasmus Niemi

Aalto University, rasmus.niemi@aalto.fi

Acknowledgements: This report has been drafted as a part of the Industry and Data research project by The Finnish Innovation Fund Sitra and by Digital Disruption of the Industry (DDI) research projects funded by Academy of Finland. Sitra's IHAN project builds the foundation for a fair, human-driven data economy. In the new model, data is shared through the individual's or data source's consent, leading to transparency and trust in the use of data, as well as perceived value among all network members from individuals to companies and society at large. Special thanks to Mika Pajarinen for his data prosessing work.

Keywords: Data sharing, Supply chain, Value Chain System of systems, Data as a Resource, Boundary resources, Value added

Suggested Citation: Seppala, Timo and Hakanen, Esko and Lahteenmaki, Ilkka and Mattila, Juri and Niemi, Rasmus, The Resource Dependency of Data: A Prospective on Data Sharing in Supply Chains (November 27, 2019). Available at SSRN: https://ssrn.com/abstract=3493650 or http://dx.doi.org/10.2139/ssrn.3493650

Understanding data as a resource

Thousands of terabytes of data is being controlled by different companies in their industrial supply chains. For decades, they have systematically gathered, shared and stored various kinds of data regarding their processes and business environments. Furthermore, they have been systematically integrating more connectivity features into their components, products and services, as well as the associated information and communication technology systems (Meyer & Schwager, 2007; Tao, Qi, Liu & Kusiak, 2018).

Companies gather and acquire data with various techniques and from various sources (Yin, Li, Gao & Kaynak, 2015). Internally and externally sourced data are typically developed and analyzed by companies to primarily satisfy their own needs (Fitzgerald, 2013). When companies share specific use case data, they tend to create their own company-specific partial copies of the data. (Huttunen, Seppälä, Mattila & Lähteenmäki, 2019). That said, the data can be multiplied, shared among companies using different information technologies, and reproduced in many ways (Levitin & Redman, 1998; Nikander, Mattila & Seppälä, 2018).

Data can be considered a production resource, just like human labor, machinery, capital and other such constituents (Barney, 1991; Levitin & Redman, 1998; Wernerfelt, 1984). However, data as a resource comprises different characteristics from other resources (see Levitin & Redman, 1998; Nikander et al., 2018). These characterizations include features such as *intangibility, consumability, shareability, copiability, transportability, nonfungibility, fragility, versatility, valuation, depreciability, source, renewability and storageability* (see Levitin & Redman, 1998). Additionally, perceiving data as a resource has various managerial and policy implications (Levitin & Redman, 1998).

The concept of data as a resource refers to the data controlled by an organization. Data as a capability, in turn, refers to what the companies can do with this data.[1] Perceiving data as a resource as well as a capability can empower companies to migrate towards platform-like business models (Huttunen et al., 2019; Nikander et al., 2018). The data resources of a company can be divided into two categories: proprietary data and shared data. (Rajala et al., 2018; Technology Industries of Finland, 2019).

The make-or-buy decision for data: proprietary or shared?

This study examines the extent of sharing of data by companies to other supply chain participants. The study explores how these data sharing percentages would look if data was shared in the same proportion as the value added is captured by the supply chain participants. With value added analysis we can understand the share of value added of each supplier–producer–customer relationship within supply chains (Seppälä, Ali-Yrkkö & Kenney, 2014). This breakdown of value added also helps to understand how much of companies' revenue is associated with internal and external resources (Ali-Yrkkö, Rouvinen, Seppälä & Ylä-Anttila, 2011; Seppälä et al., 2014). One clear benefit of the value added methodology for research is that it can easily be calculated for companies from publicly available annual financial reports and for industries from national statistics.

Much in the same way as companies are faced with internal-or-external (make-or-buy) decisions for conventional resources, in a similar fashion, they are also faced with *proprietary-or-shared decisions* for data. The proprietary-or-share decision is the act of making a strategic choice on whether data should be kept

proprietary or whether it can be shared. These strategic choices take place at the policy, strategic and operational levels. (Peteraf, 1993; Wernerfelt, 1984).

Methodology and data

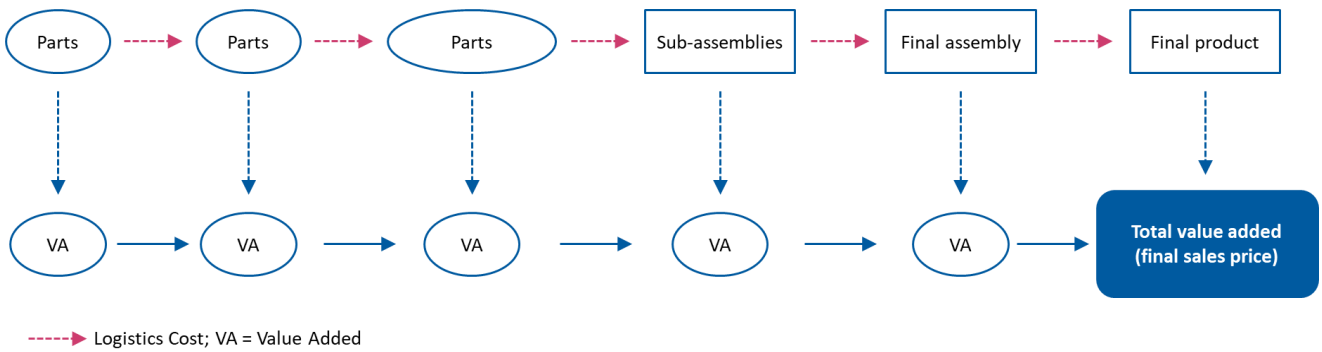
While maintaining an explorative and conceptual focus, this study aims to contribute to the theoretical discussions on resource-based management of the firm (Barney, 1991; Wernerfelt, 1984), core competences of the firm (Prahalad & Hamel, 1990; Prahalad, 1993; Javidan, 1998) and global supply chains discursions by focusing on the specifics of value-added analytics (Ali-Yrkkö et al., 2011; Seppälä et al., 2014) from the perspective of value added and data sharing. Supply chains or value chains refer to a set of activities in numerous companies that are involved in providing a product or a service (Figure 1).

The paper explores how the data sharing percentages of companies would look if data was shared in the same proportion as the value added is captured by the supply chain participants. At the company level, the value added margin created by the company is calculated as follows: value added margin=value added/output. The inverse of this margin describes what share of company's revenue is created by external supply chain participants. To this inverse share, we refer as the *supply chain contribution*.

To calculate these shares, we use the data by Eurostat (Structural Business Statistics Database were extracted on February 28th, 2019 and updated on September 27th, 2019). In the database, value added is defined as follows: Value added at factor cost is the gross income from operating activities after adjusting for operating subsidies and indirect taxes. It is an indicator in the domain of structural business statistics. (Ali-Yrkkö et al., 2011). It can be calculated as the total sum of items to be added (+) or subtracted (-): turnover (+); capitalized production (+); other operating income (+); increases (+) or decreases (-) of stocks; purchases of goods and services (-); other taxes on products which are linked to turnover but not deductible (-); duties and taxes linked to production (-)[1]. At the company level, the data were extracted from Orbis – Bureau van Dijk database (1.6.–6.8.2019).

In the future studies, we will further explore this conceptual avenue with wider scope of data and elaborate the link between the value added method and data sharing. Various methodologies exist which could also be of further interest in this regard (Glaser, Strauss & Strutzel, 1968; Levin-Rozalis, 2004; Timmermans & Tavory, 2012; Levin-Rozalis, 2010; Appendix 2).

Figure 1: Supply chain versus value chain



[1] Alternatively, it can be calculated from the gross operating surplus by adding personnel costs. (Seppälä et al., 2014).

Analysis

The analysis for the European business economy

Table 1 shows the shares of internal and external resourcing of company revenues in EU-28 are.

Table 1: The division between internal and external resources (companies in EU-28 countries), %, Data source: Eurostat (2019).

	Internal resources	External resources
All companies	48	52
Retail companies	13	87
Healthcare companies and public sector	45	55

On average, from the perspective of a single company in the EU-28 countries, in 2017 as much as 52% of company revenue came from value added created by other supply chain participants (Table 2). If we assume that data would be shared in the same proportion as the value added is captured by the supply chain participants, companies would share 52% of their data.

Table 2: Supply chain contribution in EU-28 area (averages), % of company revenue, Data source: Eurostat (2019).

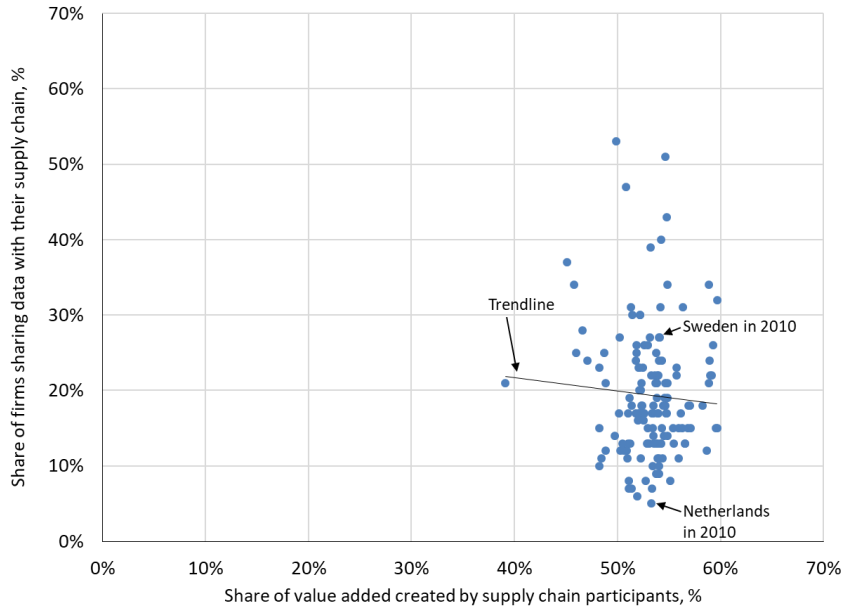
	2012	2013	2014	2015	2016	2017[a]
EU-28	54	54	53	53	53	52

[a] In Belgium, the Czech Republic and Estonia the contribution of supply chain is higher than EU-28 average.

As shown in Table A.1 (in Appendix 1), in value added terms industrial companies contribute 26% of their revenues by themselves. Assuming data was shared in the same proportion as the value added is captured by the supply chain participants, industrial companies would share the inverse proportion—that is, 74%—of their data.

Next, we consider the relationship between the contribution of supply chain and data sharing within supply chain (Figure 2). To our knowledge, the only comprehensive source concerning data sharing is ICT survey by Eurostat. This survey included the following question: Enterprises whose business processes are automatically linked to those of their suppliers and/or customers

Figure 2: The role of supply chain participants and data sharing within supply chain



Data sources: Eurostat (value added and output), Eurostat: ICT usage in Enterprises.

Note: Sharing information electronically on Supply Chain Management means exchanging all types of information with enterprises either suppliers or customers about the availability, production, development and distribution of goods or services. This information may be exchanged via websites, networks or other means of electronic data transfer, excluding e-mails not suitable for automated processing or manually typed. Data sources: Eurostat (value added and output), Eurostat: ICT usage in Enterprises.

Surprisingly, no correlation can be seen between the supply chain contribution (X-axis in Figure 2) and data sharing (Y-axis in Figure 2) within the supply chain. For example, in Netherlands in 2010, the supply chain contribution was 53% but only 5% of companies shared data with their supply chain. As another example, in Sweden in 2010, the supply chain contribution was almost identical—namely 54%— but as many as 27% of companies shared data with their supply chain. While statistics on the data sharing percentages are not available on a more detailed level of granularity, for the benefit of future prospective studies, we examine the differences in the value added percentages by capital intensity and at the company level.

The analysis by capital intensity

Based on their capital intensity, companies can be divided into three categories: asset heavy, asset medium or asset light. An asset heavy business model is one where significant capital expenditure is required to start the business. An asset light business model is a model for an organization where the company has relatively few capital assets compared to its operations. An asset medium business is somewhere in between asset light and asset heavy business models (see table 3).

For illustration, we use following industrial subsectors:

1. Asset heavy: Manufacturing of paper and paper products[2]
2. Asset medium: Wholesale and retail trade (excluding repair of motor vehicles and motorcycles)[3]
3. Asset light: Information and communication (computer programming, consultancy and related activities)[4]

Table 3: Value added of example companies of asset heavy, asset medium and asset light industries, % of company revenue

Country	Asset heavy[a]	Asset medium[b]	Asset light[c]
The Netherlands	21	11	48
Belgium	20	10	45
Bosnia and Herzegovina	20	12	52
Bulgaria	23	8	58
Spain	25	12	45
Ireland	34	12	18
Iceland	30	10	66
Great Britain	30	15	57
Italy	24	12	48
Austria	31	12	48
Greece	23	12	43
Croatia	26	10	49
Cyprus	34	13	26
Latvia	27	14	58
Lithuania	28	7	54
Malta	28	11	36
Norway	18	11	48
Portugal	22	11	51
Poland	25	12	45
France	24	10	50
Romania	22	11	56
Sweden	27	12	37
Germany	26	14	48
Slovakia	25	12	38
Slovenia	22	9	45
Finland	**15**	**10**	**51**
Switzerland	32	13	51
Denmark	31	11	49
Czech	24	6	44
Hungary	25	15	44
Estonia	25	10	53
EU-28	**24**	**11**	**47**

[a] Manufacture of paper and paper products.

[b] Wholesale; excluding repair of motor vehicles and motorcycles.

[c] Computer programming, consultancy and related activities.

The asset heavy companies contribute 24% of their revenues by themselves. The respective share for asset medium companies is 89%, and for asset light companies 53%. Assuming data was shared in the same proportion as the value added is captured by the supply chain participants, companies would share of their data according to the inverse proportion of these shares—for asset heavy companies equaling 76%, for asset medium companies 11%, and for asset light companies, 47%.

The analysis at the company level

Next, we examine differences between companies in four different industries: 1) manufacture of paper and paper products; 2) wholesale and retail trade; excluding repair of motor vehicles and motorcycles; 3) human health activities; and 4) computer programming, consultancy and related activities.

Unsurprisingly, the comparison suggests that the supply chain contribution differs between industries. However, the results reveal that also within industries, companies determine their make-or-buy decisions very differently. As a consequence, the contributions of the supply chain also differ (Tables 4, 5, 6 and 7).

Table 4: Value added of paper and paper products companies, % of company revenue

Company	Value Added
Stora Enso Oyj	28
Binderholz Gmbh	28
UPM-Kymmene Oyj	28
Essity AB	35
Ahlstrom-Munksjö Oyj	29
Sappi Limited	33
Kotkamills Oy	19
Rettenmeier Holding Ag	18
Powerflute Oy	24
Ranheim Paper & Board As	23
Average of key figure of Finnish companies	**26**
Average of key figure of international comparables	**27**

Table 5: Value added of human health activity companies, % of company revenue

Company	Value Added
Pihlajalinna Oyj	49
Sykehuset Telemark HF	66
Mehiläinen Oy	53
Nordlandssykehuset Hf	66
Suomen Terveystalo Oy	37
Helse Fonna HF	71
PlusTerveys Oy	65
Aleris Helse AS	42
Fimlab Laboratoriot Oy	46
Unilabs Norge AS	69
Average of key figure of Finnish companies	**50**
Average of key figure of international comparables	**63**

Table 6: Value added of a single wholesale and retail trade company, % of company revenue

Company	Value Added
Alko Oy	13
Systembolaget AB	14
Stockmann Oyj Abp	25
Dansk Suermarked A/S	18
Tokmanni Group Oyj	17
Norsk Butikkdrift AS	14
Lidl Suomi Kommandiittiyhtiö	16
Bestseller A/S	30
Gigantti Oy Ab	13
Telenor A/S	25
Average of key figure of Finnish companies	**17**
Average of key figure of international comparables	**20**

Table 7: Value added of a single computer programming, consultancy and related activity companies, % of company revenue

Company	Value Added
Rovio Entertainment Oyj	20
Zynga Inc.	9
Tieto Oyj	68
The Sage Group Plc.	67
F-Secure Oyj	70
McAree Ireland Ltd	13
3Step IT Oy	10
SDC A/S	44
Efecte Oyj	64
Easyvista	65
Average of key figure of Finnish companies	**47**
Average of key figure of international comparables	**40**

For prospective on data sharing / industry level value added analysis see appendix 2 for our very preliminary prospective results.

Managerial considerations

Different types of process flows

Next, we examine the supply chains and production processes in greater detail and consider what kinds of potential implications different types of things and information flows could have for data sharing.

The flow of things and information in manufacturing and service processes can be divided into five categories: 1) primary flow; 2) secondary flow (own re-use); 3) secondary flow (sold); 4) waste; 5) hazardous waste/emissions (see Table 10).

Primary flow includes inputs that are converted to desired, sellable products or services that relate to the core business of the company. In manufacturing, a typical measure of resource efficiency relates to contrasting the inflow (sourcing) and outflow (sales)–for example, the amount of ore versus the produced amount of copper.

Secondary flow (own re-use) refers to different secondary streams that feed back into the primary process. In manufacturing, a typical example is the scrap from different molding or shaping processes. Since this material may never leave the production facilities, or at least is recycled back to the original process flow, there is a good understanding of the composition and properties of the materials.

In turn, *secondary flow (sold)* contains the things and information that are not recycled back to the company's own primary process but are still considered as valuable resources. This category includes various side-streams that originate from production processes. In manufacturing, one example is the sulphuric acid that is a side-product of copper refining—a process that has a little monetary value but is a convenient way to tie up the harmful sulphur emissions.

Waste includes all things and information that have no value to the focal company. This category brings no revenue, but rather creates expenses due to waste-handling and disposal, which are often priced by the amount of waste collected. The less waste the company is left to deal with, the better. So, companies are usually happy if other parties can help to reduce or re-use their waste.

Last, *hazardous waste/emissions* include all the waste that cannot be disposed of easily. Different regulations dictate the types of waste and emissions that the firms need to control and document closely. Here, companies have to be more cautious with giving permission for third parties to re-use their waste, due to the hazardous nature of the materials.

Table 8: Summary of different characteristics of operational and process data

Data categorization	Primary flow	Secondary flow (own re-use)	Secondary flow (shared)	Waste	Hazardous waste/ emissions	Model terms of data sharing
Proprietary data	X	(X)			X	Proprietary information
Confidential data		X	X		X	Confidential information
Distributed data			X			Decentralized information
Open data				X	(X)	Open information
	Value added of internal resources			Value added of external resources		

Manufacturing process flows and data sharing

As discussed earlier in this report, the data resources of a company can be divided into two broad categories: proprietary data and shared data. To elaborate further, shared data can be divided into three subcategories: 1) data shared with trusted partners—or confidential data; 2) data shared with other stakeholders—or distributed data; and 3) data shared with anyone—or open data. (Rajala et al., 2018; Technology Industries of Finland, 2019).

Each material flow category can be associated with different types of data sharing motivations. The primary motivation for sharing the material-related data is to move materials from inferior streams toward the primary flow. Put differently, resource efficiency improves when higher fraction of the incoming materials ends up into the primary production. By sharing the information about the secondary or waste material streams, firms can identify new use cases for their inferior material streams.

The data regarding the primary material flow can related to *proprietary data*. This is especially the case when the primary material flow relates to the core business of the firm and is directly connected to the incoming revenues. In these cases, information related to this material flow can be business-sensitive and should be handled carefully.

The secondary material flow—which the firm re-uses in its own processes—can be related to *confidential data*. In many cases, these re-used secondary flows relate to testing and adjusting of machinery or to certain limitations of the production process. The increasing role of networked business processes, value-adding service offerings, and co-creation of business value mean that more and more operational processes include collaboration among different organizations. Furthermore, many of the value-adding services relate to improving material efficiency. For instance, more detailed knowledge of the incoming material properties helps to adjust the machinery and reduces the need for test batches. Thus, sharing of such confidential information with the right and relevant participants helps to improve the operational practices and to move materials from secondary to primary flow.

Conversely, when the secondary materials are not re-used inside the organization, but sold for that purpose instead, *distributed data* can help in finding the most potential customer. In addition, the buyer can find materials with specific information about their properties more valuable. However, the primary reason for opening the information about the sellable secondary flows is to gain information about the usability of those materials. By understanding the use cases for the secondary material flows, the focal firm can expose itself to external innovations and pinpoint interesting aspects in the material-related data. These benefits can lead to identifying new synergies between different firms.

To reduce the amount of waste, firms can investigate the benefits of *open data*. Such approach can help to find new viable use-cases internally or externally. As a rule of thumb, waste fraction includes items that the focal firm considers uninteresting from the economic perspective. On the contrary, since waste management and collection incurs expenses, firms may be eager to give away these materials free of charge. From this perspective, open approach to data sharing is advisable—the openness helps to find new stakeholders and potential uses for the materials, thus reducing the amount of disposable waste. The new use cases may help to identify new types of sellable secondary material flows and, thus, initiate new business opportunities.

However, various regulations stipulate that hazardous waste/emissions need to be closely controlled. The data related to this material flow can contain business-sensitive information, and/or other discreet information regarding the operational activities of the focal firm. In many cases, firms have outsourced the handling and management of these material streams and must share the related data with their partner firms. The

regulative bodies enforce policies for disclosing the data, but these requirements may relate to averages (not on specific units or streams) or may not be publicly available. Thus, it is reasonable to mainly categorize the data on hazardous waste/emissions as proprietary, of which the applicable elements are shared as confidential data, and some parts as completely open data.

Service process flows and data sharing

In service business, processes can be categorized in the same way as in manufacturing: primary flow, secondary flow (own re-use), secondary flow (sold), and waste. In service business, data can be related to different process flows in the same way as in manufacturing processes. These relations are described from the standpoint of lending by banks as follows:

Primary process flow relates to things and information that are needed in the focal service, which is part of the service provider's core business. In the case of bank lending, identification, customer's basic information, loan specification, collaterals, know-your-customer, and anti-money laundering are examples of information that are needed for the granting of a loan.

Secondary process flow (own re-use) is about the use of some other stream of things and information that may be utilized in the primary flow. For example, analyses about customers' payment information may be used to optimize the loan process. In this category, data that is formed in a certain service process may be re-used in some other service area, e.g. loaning process originated data in the sales of insurance services.

Secondary process flow (sold) includes streams of things and information that originate from the primary processes. For example, banks sell identification services to online stores and other organizations for verifying the customer's identity. Also know-your-customer (KYC) and anti-money laundering (AML) are sellable data. The European Commission's Second Payment Services Directive (PSD2) opens several secondary flow potentials for the banks to sell or trade with third-parties on customer payment and account information (with the consent of the customer). Many of the service components are side-products (e.g. KYC), which have no monetary value in the primary flow, but they have data trading value outside of that.

Waste has no positive value to the focal firm. In loans, credit loss is a typical situation where the value of the loan is negative for the originator. However, the loan may be sold to a collection agency, which specializes in debt collection. The data about customer behavior may be used afterwards in other services and also used by third parties e.g. in credit information.

Finally, hazardous waste exists in financial services as well. Data about anti-money laundering is highly classified by regulators, and therefore the data needs to be controlled and documented closely.

Discussion

While the theoretical figures on data sharing based on the assumption of similar proportions of data sharing and value added are high, in practice data sharing is restricted by different types of factors, e.g. contractual confidentiality. Typically, contracts between companies and their first tier suppliers and directs customers are not public. Due to these restricting factors, companies are not able to share their data freely, irrespective of whether they would like to or not.

The unused and unshared data in both proprietary and shared categories could potentially be very valuable to other companies. Many companies monetize data, *e.g.* by processing and selling it (LaValle, Lesser, Shockley, Hopkins & Kruschwitz, 2011; Liozu & Ulaga, 2018). Furthermore, many companies share the revenues of sold innovations, *e.g.* software which is developed from someone else's data (Parker & Van Alstyne, 2018; Rajala, Rossi & Tuunainen, 2003).

Digital-platform-related data sharing practices and may also be viable business models for companies, as shown by successful technology platforms such as Amazon, Apple and Google. These platform providers capitalize on the products and services developed by external suppliers to gain competitive advantages over the other companies (Benzell, LaGarda & Van Alstyne, 2017; Gawer & Cusumano, 2016).

In supply chains, more open data sharing practices could lead to new types of information, knowledge, business constellations and business models (Parker & Van Alstyne, 2018). The implementation of more open data sharing practices could also enable better interoperability for system-of-systems-level integration. Furthermore, this development could lead to improvements in autonomous machine-to-machine interaction, potentially enabling completely new algorithmic value creation and capturing logics, and more autonomous ways of executing tasks and running infrastructures (Huttunen et al., 2019; Mattila & Seppälä, 2015; Rajala et al., 2018). In part, this research aims to spark new discussion amongst companies regarding their data policies, strategies and data sharing, but also on the importance on developing methodological tools for understanding the value of the data sharing.

Endnotes

1 Data may belong to various actors, but it cannot be owned in the legislative sense. Data can, however, be controlled and managed. The most natural view of data management is that the actor is the one who owns the device and the service where the data are. The ownership of a device or service is the default situation of data management when no contractual arrangements or the like have been made. In this case, the owner of the device and service usually have a natural ability to prevent others from accessing the data by preventing access to the device or service. Within the freedom of contract, it can be specified who data belongs to, what kinds of access rights there are to the data, whether they are exclusive, parallel, etc. It is aimed at agreements between parties on the ownership of data and use restrictions even when no one owns the data and only restrictions on any contractual partner. The restriction of contract comes, however, from the fact that the contract cannot be binding on a third party. In the end, the contractual policies between the actors will define the relative strengths of data ownership between parties, for example how the ownership of data will be established in the autonomous smart device and service entities of the future. (Ailisto et al., 2015; Rajala et al., 2018; Seppälä et al., 2018).

2 Manufacture of paper and paper products statistics – https://ec.europa.eu/eurostat/statistics-explained/index.php/Archive:Manufacture_of_paper_and_paper_products_statistics_-_NACE_Rev._2 (information retrieved 14.10.2019)

3 Retail trade statistics – https://ec.europa.eu/eurostat/statistics-explained/index.php/Archive:Retail_trade_statistics_-_NACE_Rev._2 (information retrieved 14.10.2019)

4 Computer programming and consultancy statistics – https://ec.europa.eu/eurostat/statistics-explained/index.php/Archive:Computer_programming_and_consultancy_statistics_-_NACE_Rev._2 (information retrieved 14.10.2019)

Bibliography

Ailisto, H. (Editor), Mäntylä, M. (Editor), Seppälä, T. (Editor), Collin, J., Halen, M., Juhanko, J., Jurvansuu, M., Koivisto, R., Kortelainen, H., Simons, M., Tuominen, A. & Uusitalo, T. (2015). Finland – The Silicon Valley of Industrial Internet. *Publication of the Government's Analysis, Assessment and Research Activities, 10*(August).

Ali-Yrkkö, J., Rouvinen, P., Seppälä, T. & Ylä-Anttila, P. (2011). Who Captures Value in Global Supply Chains? Case Nokia N95 Smartphone. *Journal of Industry, Competition and Trade*, (11), 263–278.

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management, 17*(1), 99–120.

Benzell, S., LaGarda, G. & Van Alstyne, M. W. (2017). The Impact of APIs in Firm Performance. *Ssrn*, 1–38. https://doi.org/10.2139/ssrn.2843326

Fann, K. T. (1970). *Peirce's theory of abduction* Springer Science & Business Media.

Fitzgerald, M. (2013). An Internet for Manufacturing. *MIT Technology Review*.

Gawer, A. (2009). *Platforms, markets and innovation*. Edward Elgar Publishing.

Gawer, A. & Cusumano, M. (2016). Industry Platforms and Ecosystem Innovation. *Journal of Innovation Management, 31*(3), 417–433.

Ghazawneh, A. (2012). Towards a Boundary Resources Theory of Software Platforms. In *JIBS Disseration Series*. https://doi.org/10.1016/S0009-2614(95)01235-4

Ghazawneh, A. & Henfridsson, O. (2010). Governing Third-Party Development through Platform Boundary Resources. *ICIS 2010 Proceedings*, 1–18.

Glaser, B. G., Strauss, A. L. & Strutzel, E. (1968). The discovery of grounded theory; strategies for qualitative research. *Nursing Research, 17*(4), 364.

Hagiu, A. (2014). Strategic decisions for multisided platforms. *MIT Sloan Management Review, 55*(2), 71–80.

Huttunen, H. (2019). *The role of data in firm performance: A techno-economic view*. Aalto University.

Huttunen, H., Seppälä, T., Mattila, J. & Lähteenmäki, I. (2019). What Are the Benefits of Data Sharing? Uniting Supply Chain and Platform Economy Perspectives. *Etla Report,* 93.

Javidan, M. (1998). Core competence: what does it mean in practice?. *Long range planning, 31*(1), 60–71.

Katz, M. L. & Shapiro, C. (1994). Systems Competition and Network Effects. *Journal of Economic Perspectives, 8*(2), 93–115. https://doi.org/10.1257/jep.8.2.93

Kotiranta, A., Kullvik, M., Maijanen, S. & Seppälä, T. (2016). *Behind the Scenes of Operation Theatres*. Taloustieto Oy.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S. & Kruschwitz, N. (2011). Big Data, Analytics and the Path From Insights to Value. *MIT Sloan Management Review, 52*(2), 21–32. https://doi.org/10.0000/PMID57750728

Levin-Rozalis, M. (2004). Searching for the unknowable: A process of detection – Abductive research generated by projective techniques. *International Journal of Qualitative Methods, 3*(2), 1–18.

Levin-Rozalis, M. (2010). Using abductive research logic to construct a rigorous explanation of amorphous evaluation findings. *Journal of Multidisciplinary Evaluation, 6*(13), 1–14.

Levitin, A. V. & Redman, T. C. (1998). Data as a Resource: Properties, Implications, and Prescriptions. *MIT Sloan Management ReviewF, Fall*(40), 89–101.

Liozu, S. M. & Ulaga, W. (2018). *Monetizing Data*. Arizona: Value Innoruption Advisors Publishing.

Mattila, J. & Seppälä, T. (2015). Blockchains as a Path to a Network of Systems – An Emerging New Trend of the Digital Platforms in Industry and Society. *Etla Report, 45*.

Meyer, C. & Schwager, A. (2007). Understanding Customer Experience. *Harvard Business Review, 85*(2), 116–126.

Nikander, P., Mattila, J. & Seppälä, T. (2018). Kehittymättömät datamarkkinat johtavat Internetin keskittymiseen. *Digibarometer 2018*, 31–37.

Parker, G. G. & Van Alstyne, M. W. (2005). Two-Sided Network Effects: A Theory of Information Product Design. *Management Science, 51*(10), 1494–1504. https://doi.org/10.1287/mnsc.1050.0400

Parker, G. & Van Alstyne, M. W. (2018). Innovation, Openness, and Platform Control. *Management Science, 64*(7), 3015–3032. https://doi.org/10.2139/ssrn.1079712

Peteraf, M. (1993). The Cornerstones of Competitive Advantage: A Resource-Based View. *Strategic Management Journal, 14*(3), 179–191.

Prahalad, C. K. (1993). The role of core competencies in the corporation. *Research-Technology Management, 36*(6), 40–47.

Prahalad, C. K. & Hamel, G. (1990). The core competence of the corporation. *Harvard Business Review, 68*, 79–91.

Rajala, R., Hakanen, E., Mattila, J., Seppälä, T. & Westerlund, M. (2018). How Do Intelligent Goods Shape Closed-Loop Systems? *California Management Review, 60*(3), 20–44.

Rajala, R., Rossi, M. & Tuunainen, V. K. (2003). A Framework for Analyzing Software Business Models Research Questions and Research Method. *11th European Conference on Information Systems*, 1614–1627.

Seppälä, T., Juhanko, J. & Mattila, J. (29.8.2018**).** Data Ownership and Governance – Finnish Law Perspective. ETLA Brief No 71. https://pub.etla.fi/ETLA-Muistio-Brief-71.pdf

Seppälä, T., Ali-Yrkkö, J. & Kenney, M. (2014). Global Supply Chains and Transfer Pricing: Insights from a Case Study. Supply Chain Managemment: An International Journal, 19(4), 445–454.

Seppälä, T. & Kalm, M. (2013). Profiting from Product Innovation: The Product Life Analysis of Economic Geography of Value Added. *Industry Studies Woking Paper Series, 2013*(1), 32.

Star, S. L. & Griesemer, J. R. (1989). Institutional Ecology, "Translations" and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science, 19*(3), 387–420.

Tao, F., Qi, Q., Liu, A. & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems, 48*, 157–169. https://doi.org/10.1016/j.jmsy.2018.01.006

Technology Industries of Finland (2019). *Model terms of the Technology industries for data sharing*. 4. Retrieved from https://teknologiainfo.net/en/content/model-terms-technology-industries-data-sharing

Timmermans, S. & Tavory, I. (2012). Theory construction in qualitative research from grounded theory to abductive analysis. *Sociological Theory, 30*(3), 167–186.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal, 5*(2), 171–180.

Yin, S., Li, X., Gao, H. & Kaynak, O. (2015). Data-based techniques focused on modern industry: An overview. *IEEE Transactions on Industrial Electronics, 62*(1), 657–667. https://doi.org/10.1109/TIE.2014.2308133

Appendix 1

Table A. 1: Value added of manufacturing, retail and healthcare (public and private sector) industries, % of company revenue

Country	Manu-facturing value added	Retails value added	Healthcare value added[a]
The Netherlands	24	20	46
Belgium	21	15	42
Bosnia and Herzegovina	24	14	–
Bulgaria	21	13	41
Spain	26	18	48
Ireland	29	19	44
Iceland	35	21	41
Great Britain	33	21	57
Italy	25	21	47
Austria	28	17	48
Greece	20	21	57
Croatia	27	10	47
Cyprus	28	17	46
Latvia	22	16	47
Lithuania	22	14	51
Malta	27	15	32
Norway	34	22	55
Portugal	24	17	48
Poland	21	15	45
France	26	14	51
Romania	22	18	43
Sweden	28	13	49
Germany	27	18	49
Slovakia	19	19	37
Slovenia	24	10	45
Finland	**26**	**15**	**46**
Switzerland	20	13	50
Denmark	29	16	49
Czech	21	24	38
Hungary	21	18	38
Estonia	22	15	42
EU-28	**26**	**17**	**45**

[a] Human health activities including both private and public sector.

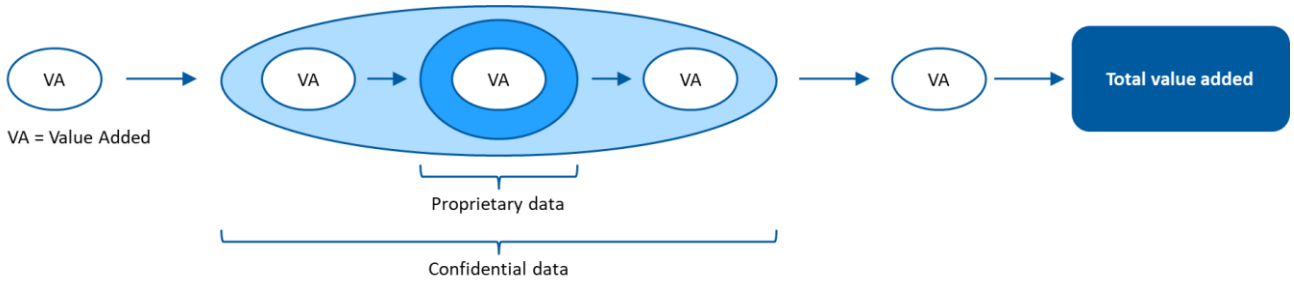
[b] Sharing information electronically on Supply Chain Management means exchanging all types of information with enterprises either suppliers or customers about the availability, production, development and distribution of goods or services. This information may be exchanged via websites, networks or other means of electronic data transfer, excluding e-mails not suitable for automated processing or manually typed.

Appendix 2

The proposition – What is the share of proprietary and confidential data? What is the share of industry and open data? Value added perspective in industrial supply chains

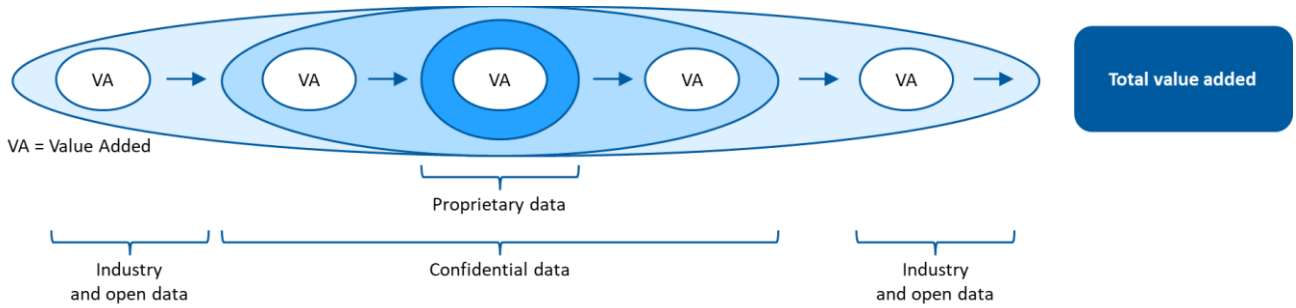
Most of the total revenues associated with the benefits of data sharing fall under the category of trusted partners – confidential information and data (company–customer & company–supplier relationships) (see figure 2). These existing relationships could easily be extended to cover 3rd party confidential relationships (company–customer and known third party; company–customer and unknown third party). However, the results might vary from industry to industry (see table 3 and 4).

Figure A. 1: Value added related to the proprietary data vs confidential data



Case example 1: The case company operates in a horizontal industry in which actors strive to provide integrated solutions for customers. The company operates in an industry which manufactures industrial and commercial machinery equipment and it produces machinery products and provides services. In the case of a company's supply chain, the value added is divided as follows, depending on the examined product or service: company internal 12–38%; supplier–company relationship 24–39%; company–customer 21–38%. The total value added associated with internal and trusted partners equals 83–89%. Respectively the revenue associated with other value chain participants outside the supplier–company–customer -relationship is 11–17%. (Seppälä & Kalm, 2013) (See figure 3; Table 9).

Figure A. 2: Value added related to the confidential data vs. other data



Case example 2: The case company operates in a vertical industry in which actors strive to provide a material service solution for customers. The company operates in an industry, which provides health care services. In the case of a company's supply chain, the value added is divided as follows, depending on the examined product or service: company internal 6–8%; supplier–company 6–12%; company–customer 71–82%. The total value added associated with internal and trusted partners equals 89–96%. Respectively the revenue associated with other value chain participants is 4–11%. (Kotiranta et al., 2016) (See figure 3; Table 9).

Assuming data was shared in the same proportion as how the value added is captured by the supply chain participants, the proportion of distributed and open data shared by the company in the first case example would be 11–17%. With a similar assumption, the respective proportion for the second case company would be 4–11%. However, due to the lack of any significant correlation between the data sharing propensity and the value added captured by companies (see Figure 2), these case figures are not generalizable to wider contexts, and do not form the basis for estimates concerning data-sharing practices at the macro level.