EGOV-CeDEM-ePart

Collective Data Analytics Capability Building Processes: a Governance Model

Boriana Rukanova*, Anneke Zuiderwijk-van Eijk**, Moorchana Das***, Yao Hua Tan****, Toni Männistö*****

*Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands, b.d.rukanova@tudelft.nl
**Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands, A.M.G.Zuiderwijk-vanEijk@tudelft.nl
*** Delft University of Technology, M.Das@student.tudelft.nl
****Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands, Y.Tan@tudelft.nl
*****Cross-Border Research Association, toni@cross-border.org

Abstract: Collective data analytics capability building offers opportunities for government organizations to develop capabilities that would be difficult to develop on their own. However, research on that topic is scarce and there is still a limited understanding of how collective data analytics capability building processes contribute to the value realization of the individual participating organizations. In this paper, drawing from the governance literature and by analyzing a case study from the customs domain we develop a governance model that allows to analyze collective data analytics capability building processes. Our governance model is a contribution to the literature on the use of data analytics in government, with the specific focus on understanding the collective data analytics capability building processes. For practitioners, the model can be used for identifying scenarios for engaging in collective data analytics initiatives in a multi-level context.

Keywords: Governance, collective, data analytics, capabilities, value, customs

Acknowledgement: This research was partially funded by the PROFILE Project (nr. 786748), which is funded by the European Union's Horizon 2020 research and innovation program. Ideas and opinions expressed by the authors do not necessarily represent those of all partners.

1. Introduction

Governments today are facing big challenges in the domain of international trade. They face increase in international trade due to developments such as Brexit and eCommerce, and at the same time they need to ensure safety and security while at the same time facilitating trade (Tan et al., 2011). To address such challenges governments are starting to explore the possibilities that big data and data analytics can offer. Big data refers to data that is high in volume, velocity and variety and that requires specific technology and analytical methods for transforming it into value (De Mauro et al.,

2016). Despite the promises of big data and analytics and successful examples from businesses, earlier research reports that many organizations have failed to reach their strategic goals after significant investments (Gunther et al., 2017) and there is a limited understanding of how social and economic value can be created (Grover et al., 2018). This is especially problematic for government organizations as they traditionally to do not have advanced data analytics capabilities in-house and the risks of failure pose a big barrier. Government agencies from different countries are now starting to explore whether they can collaborate to collectively develop data analytics capabilities. From a practical point of view such collective capability building offers opportunity to share efforts and resources to develop capabilities that would be difficult to develop on their own.

Previous research on value of data analytics has discussed value by looking at the data itself (e.g. Kim, 2015; Sammon & Nagle, 2017), or focusing more on the organizational perspective (e.g. Gunther et al., 2017; Grover et al., 2018; Rukanova et al., 2019). Of particular interest is the study of Grover et al. (2018) which examines the strategic processes that lead to data analytics value creation in an organization. Two key strategic processes can be distinguished (Grover et al., 2018), namely: (1) big data analytics capability building processes, and (2) big data capability realization processes. The latter is followed by a learning loop which can initiate new big data capability building processes. The first process (i.e. data analytics capability building process) relates to developing data analytics infrastructure (including assets such as data sources, platforms, analytics portfolio and human talent) and data analytics capabilities. The second process relates to the capability realization processes include value creation mechanisms, value targets and impacts where the impact can be seen as functional value for organization (e.g. improved performance) or in symbolic value (e.g. reputation).

Nevertheless, previous research that examines value from an organizational perspective is focused on understanding value by looking at a single organization as a unit of analysis. While in some papers it is acknowledged that managing relationship with external stakeholders is important to create value (e.g. Gunther et al., 2017), this relationship management is still seen from an internal perspective of an individual organization and its ability to manage such relationships. Thus, there is lack of research that focusses on identifying value from collective data analytics capability building processes, where multiple organizations join forces to jointly develop data analytics capabilities which they can then exploit individually in their own organizations. To address this gap, the objective of this research is to develop a governance model to support the analysis of collective data analytics capability building processes and how these link to value realization processes in individual government organizations.

2. Research Background: Multi-level and Multi-actor Governance

In this study we take a broad perspective on governance and define governance as all processes of governing, whether undertaken by a government, market, or network, whether over an entire system, formal or informal organizations, or individuals part of such a system, and whether through laws, power, contracts, norms, language (adjusted from Bevir, 2012, p. 1). Thus, governance is also initiated by other parties than governments, such as citizens, non-profit organizations, companies,

lobby organizations and associations. In fact, to realize a state of governance it is essential that multiple actors combine their efforts and apply combinations of governance arrangements.

Previous research emphasizes different aspects of governance. Given the nature of our domain and our research objective (i.e. we are interested in understanding collective data analytics capability building processes) we were particularly interested in understanding governance in a multi-actor context. One focus of the governance literature that is relevant in our study concerns multi-level governance. This is particularly interesting as in the international trade there are complex interactions among businesses, as well as national and supranational government agencies such as the EU. In multi-level governance, governance and decision-making encompass multiple levels, such as local, national and international levels of public administration (Marks et al., 1996). Another governance study that is particularly relevant in the context of this paper is networked governance. Networked governance focuses on the use of organizations and structures of authority and collaboration to assign resources to network participants, and to control collective action across the network as a whole (Provan & Kenis, 2008). In contrast to hierarchies and markets, in networks there is decentralization of power and decision-making and a blurring of roles and responsibilities (Stoker, 2018). Provan & Kenis (2008) developed three basic models of network governance, namely participant-governed networks, lead organization-governed networks and network administrative organization.

Table 1: Governance Framework for Case Study Analysis; Adapted from Provan & Kenis (2008), Emerson et al. (2012), and Ostrom (1990)

Networked governance – Three basic models of network governance (Provan & Kenis, 2008) •Participant-governed networks; •Lead organization-governed networks; •Network administrative organization)
Collaborative governance – Dimensions and components of collaborative governance Emerson et al. (2012) •System context; •Drivers; •Principled engagement; •Shared motivation; •Capacity for joint action; •Outputs / collaborative actions; •Impacts; •Adaptations
Collective governance – Eight design principles for sustainably and effectively managing common resources (Ostrom, 1990) • Clearly defined boundaries; •Proportional equivalence between benefits and costs; •Collective-choice arrangements; •Collective-choice arrangements; •Monitoring; •Graduated sanctions; •Conflict resolution mechanisms; •Minimal recognition of rights to organize; •For groups that are part of larger social systems, there must be appropriate coordination among relevant groups

Another stream of governance literature focuses on collaborative governance, which refers to "the processes and structures of public policy decision making and management that engage people constructively across the boundaries of public agencies, levels of government, and/or the public, private and civic spheres in order to carry out a public purpose that could not otherwise be accomplished" (Emerson et al., 2012). Emerson et al. (2012) developed an integrative framework for collaborative governance, which consists of dimensions such as system context, shared motivation and capacity for joint action. Each dimension contains a number of underlying components, such as mutual trust, knowledge and resources. Finally, we draw from literature concerning collective

governance and the management of commons. Ostrom (1990) investigates how communities co-operate to share resources in common pool problems and states that such problems are sometimes solved by voluntary organizations rather than by a coercive state. Ensuring collective action, however, is not straightforward as e.g. parties may have conflicting interests and pursue other goals. Ostrom (1990) shows that, under certain conditions, groups of people are capable of sustainably and effectively managing their common resources. These conditions are presented as design principles. Using the insights derived from the above-mentioned literature, we developed our governance framework (Table 1).

3. Research Approach

For this study we followed an interpretative case study approach (Walsham, 1993). In our study we are interested in data analytics and the broader organizational context where data analytics capabilities are developed. We conducted our case study in the context of the H2020 PROFILE[1] research project funded by the European Commission. The project aims to develop and demonstrate the use of data analytics for customs risk analysis. The work in the project is carried out in demonstration projects called Living Labs which provide real-life setting in which data analytics solutions are developed and piloted. A brief description of the Living Labs that we used as an empirical ground is provided in Table 2.

Table 2: Living Labs Overview

Living Labs (LL)	**Short description**
(1) Dutch LL	Focus on use of data from eCommerce platforms to cross-validate declared price of goods on customs declarations
(2) Belgian LL	Focus on behavior of traders by using data analytics and machine learning on historic data sets and external data sources
(3) MS- N MS LL	Comparing results of analytics performed on customs declaration data of two neighboring customs administrations (one in the EU and one outside the EU)
(4) EU LL	Providing an infrastructure for sharing data among customs administrations in the EU

Data was collected in the period 2018-2020 through interviews, participation in meetings and project workshops, participation in bi-weekly calls, review of project deliverables and policy documents. In our case analysis we analyzed the four Living Labs as well as the PROFILE project as a whole in order to understand the complexity of the domain and identify examples of collective capability building efforts which enabled us to build our governance model which we present in Section 4. The Living Labs are still in pilot stages and results have not yet been implemented in practice. Nevertheless, each of the Living Labs sheds light on complexities of setting up collective data analytics initiatives. The data collection and data analyses evolved through a number of iterations. The initial understanding of the empirical context guided us in our search for suitable theories. In this process we arrived at our initial governance framework (Table 1) which we further

[1] https://www.profile-project.eu/

applied as a conceptual lens to structure our empirical observations. As a result we developed our model for governance of collective data analytics capability building processes (Figure 1) discussed in the next section.

4. Results: A Governance Model of Collective Data Analytics Capability Building Processes

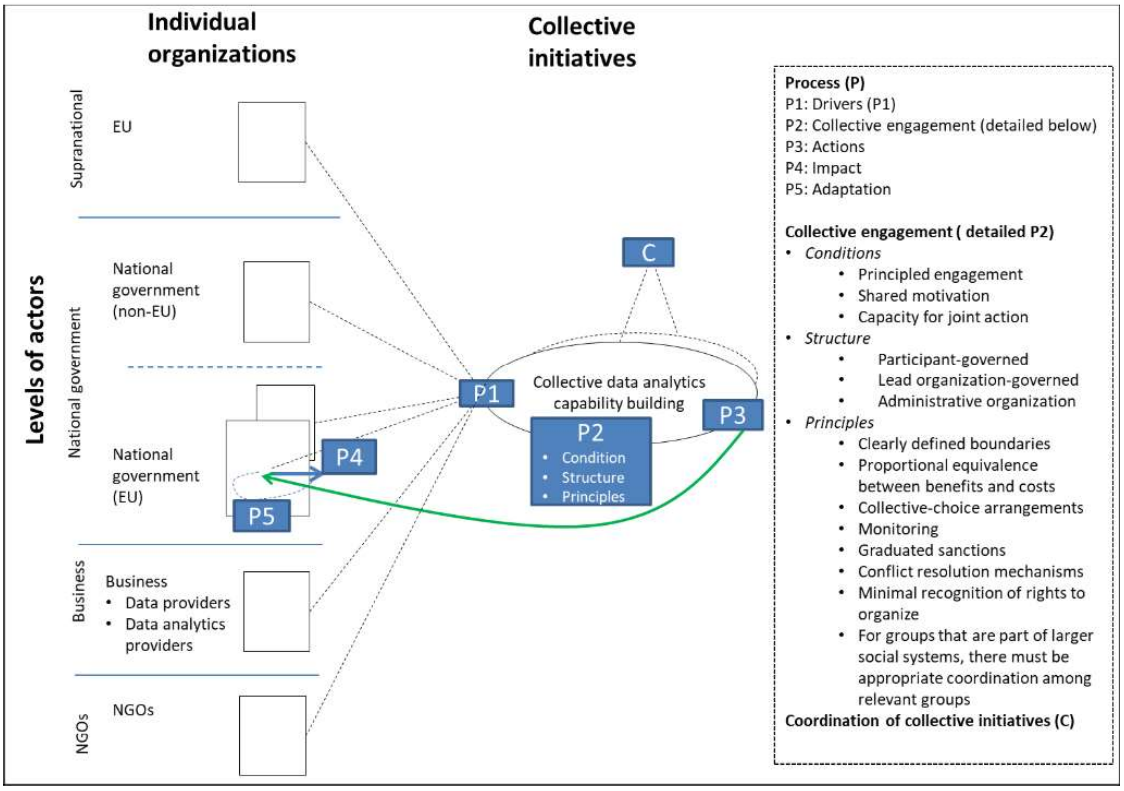
Based on the insights from literature and insights from the case domain we derived our model for governance of collective data analytics capability building processes (Figure 1). This model is intended to serve as a conceptual foundation to: (a) identify governance scenarios for collective data analytics capability building initiatives; (b) analyze specific collective data analytics capability building initiatives; (c) allow to reason how these collective capabilities developed jointly feed back into the individual organizations; (d) provide an oversite of the different collective initiatives to allow parties to reason about synergies among them. Our point of departure for developing our governance model was the model of Grover et al. (2018) where strategic data analytics processes are viewed as (1) capability building processes and (2) value realization processes, where the impact of analytics is visible in real life, followed by learning loops. In our model, however, we took part of the process related to data analytics capability building outside of the organization. The capability building takes place now as part of a collective initiative.

With this idea in mind and building on the rich empirical material from the PROFILE project, as well as the conceptual framework (Table 1), we arrived at the model as presented in Figure 1. Our model captures explicitly on the one hand individual actors (at multiple levels), and on the other hand the collective data analytics capability building initiatives. The individual actors are further divided into business and government actors, where the government actors are positioned at multiple levels, namely national and supranational. Our model distinguishes further the national actors as national governments that form part of the European Union (EU) and government actors that are outside of the EU. On the supranational level we position the EU as a supranational government. This identification of levels of actors is consistent with earlier multi-level analysis research in the area of international trade (Rukanova et al., 2015). On the collective side, our model captures the collective data analytics capability building (illustrated with an oval in our framework). The dotted oval indicates that multiple collective data analytics capability building initiatives can be started. For simplicity we will focus on explaining only one. The dotted arrows from the actors to the collective initiative suggest the diversity of actors that potentially may join an initiative. In practice we foresee that different scenarios of collective initiatives may evolve having different actor compositions.

In order to analyze a collective data analytics capability building initiative we make use of and adapt the high-level categories of Emerson et al. (2012) (see Table 1), namely: (1) drivers; (2) collective engagement; (3) actions, in terms of outputs as a result of the collective engagement; (4) impact; and (5) adaptation. These are numbered P1-5 and in the list of concepts that are listed at the right-hand side of Figure 1. The second concept, (2) collective engagement in the list above, is aimed to better understand how the collective initiative functions internally. In our model we list explicitly the relevant concepts from our framework (Table 1) that are relevant for understanding the internal

collective action processes under the concept collective engagement. More specifically under the concept collective engagement we distinguish among: (a) conditions; (b) structures; and (c) principles. Under (a) conditions we adapt several of the categories of Emerson et al. (2012). Furthermore in our model we further enrich the concept of collective engagement by adding also the three structures proposed by (Provan & Kenis, 2008) and the 8 governance principles proposed by Ostrom (1990). The full list of concepts that we use to understand collective engagements can be found in Figure 1. Finally in our model we also include the concept of coordination among collective initiatives (marked with C in Figure 1). By adding this concept to the model it becomes possible to reason about interdependencies among different collective initiatives.

Figure 18: Model for Governance of Collective Data Analytics Capability Building Processes



Some elements of our model deserve further attention. In the model of Emerson et al. (2012), the concepts impact and adaptation are related to the collective initiative. This is often the case when parties collaborate to jointly bring some desired change. In our case however the outcome of the collective data analytics capability building process (be it new analytics methods or cheaper access to new data sources) is fed back to the individual participating organizations (in our case e.g. the participating customs organizations). This is indicated with the arrow in our model pointing from the collective to the individual organizations. The individual (in our case customs) organizations are those that will deploy these outcomes in their own organizations, as part of their capability realization processes as described by Grover et al. (2018). They will combine the data analytics capabilities that they have acquired via the collective initiative together with their internal data analytics capabilities. They will then employ these combined capabilities in their processes (in case of customs in their customs risk assessment processes). By doing this they can observe the impact

(see symbol 4 in Figure 1). In terms of Grover et al. (2018), this impact can be functional or symbolic. As such the impact of the outcome of the collective process to the real world is not visible as a result of the collective process itself but becomes visible only when this output is used in by the individual organizations, which with their individual actions contribute to societal goals (e.g. better revenue collection and safety and security). By using the collective capabilities in their own processes and observing the achieved impact these individual organization then accumulate learnings and can initiate adaptations (see symbol 5 in Figure 1). These adaptations can be seen also as the learning loops in the model of Grover et al. (2018). These adaptations can then can be fed back to either the same collective initiative or can serve as a basis for initiating new collective initiatives if needed. For simplicity in our model we illustrated the feedback loop from the collective initiative to one organization only. In practice this loop is also directly relevant for all the organizations. In our case these loops would be directly relevant to national customs administrations. These loops can be also relevant for other organizations such as businesses or the EU as supranational government but these parties may not directly use the outcome of the collective process in their operational processes, but they may use them in other strategic processes such as new service delivery or drafting new policies. The impact and adaptation for these organization may be of different nature.

5. Discussion and Conclusions

In this paper, building on insights from the governnace literature and by using a case study from the customs domain we developed a governance model to support the analysis of collective data analytics capability building processes and identify how these processes relate to value realization processes for individual government organizations. Our governance model contributes to science by providing rich ground for analyzing collective capability building in a wider context and by giving insight into the complex dependencies. The societal contributions of our study are in the provision of a model that can be applied to identify scenarios for collective data analytics initiatives for government in a multi-level and multi-actor context and to aid in their governance processes. Future research can investigate the applicability of model in other domains. Future research could focus on how collective data analytics capability building processes evolve, how they can be implemented and funded.

References

Bevir, M. (2012). Governance: A Very Short Introduction. Oxford: Oxford University Press.

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. Library Review, 65(3), 122-135.

Emerson, K., & Nabatchi, T., & Balogh, S. (2012). An integrative framework for collaborative governance. Journal of public administration research and theory, 22(1), 1-29.

Grover, V., & Chiang, R. H., & Liang, T. P., & Zhang, D. (2018). Creating strategic business value from big data analytics: A research framework. Journal of Management Information Systems, 35(2), 388-423.

Günther, W.A., & Mehrizi, M.H.R., & Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. The Journal of Strategic Information Systems, 26 (3), 191-209, https://doi.org/10.1016/j.jsis.2017.07.003.

Kim, H. (2015). Big data: The structure and value of big data analytics. In Proceedings of the Twenty-First Americas Conference on Information Systems, Puerto Rico, August 13–15.

Marks, G., & Hooghe, L., & Blank, K. (1996). European integration from the 1980s: State-centric v. multi-level governance. JCMS: Journal of Common Market Studies, 34(3), 341-378.

Ostrom, E. (1990). Governing the commons: The evolution of institutions for collective action. Cambridge: Cambridge University Press.

Provan, K.G., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. Journal of public administration research and theory, 18(2), 229-252.

Rukanova, B., Tan, Y.H., Slegt, M., Molenhuis, M., van Rijnsoever, B., Plecko, K., Caglayan, B., Shorten, G. (2019). Value of Big Data Analytics for Customs Supervision in e-Commerce. In: Electronic Government, 18th IFIP WG 8.5 International Conference, EGOV 2019, 288-300.

Rukanova, B., & Wigand, R.T., & van Stijn, E., & Tan, Y.H. (2015). Understanding transnational information systems with supranational governance: A multi-level conflict management perspective. Government Information Quarterly, 32, 182–197.

Sammon, D. & Nagle, T. (2017). The data value map: a framework for developing shared understanding on data initiatives. European Conference on Information Systems. AIS Electronic Library: Guimarães, Portugal.

Stoker, G. (2018). Governance as theory: five propositions. International Social Science Journal, 68 (227-228), 15-24.

Tan, Y.-H., & Bjørn-Andersen, N., & Klein, S., & Rukanova, B. (2011). Accelerating global supply chains with IT-innovation: ITAIDE tools and methods. Springer Science & Business Media.

Walsham, G. (1993). Interpreting information systems in organizations: John Wiley & Sons, Inc. New York, NY, USA.

About the Authors

Boriana Rukanova

Dr. Boriana Rukanova is a researcher at Delft University of Technology focusing on digital trade infrastructures.

Anneke Zuiderwijk-van Eijk

Dr. Anneke Zuiderwijk-van Eijk is an assistant professor of Open Data at Delft University of Technology.

Moorchana Das

Moorchana Das is a student in the CoSEM Master program at Delft University of Technology.

Yao Hua Tan

Prof. dr. Yao-Hua Tan is a professor of Information and Communication Technology at the Delft University of Technology.

Toni Männistö

Dr. Toni Männistö is a research director at Cross-Border Research Association.