

**Identifying the value of data analytics in the context of government supervision  
Insights from the customs domain**

Rukanova, Boriana; Tan, Yao Hua; Slegt, Micha; Molenhuis, Marcel; van Rijnsoever, Ben; Migeotte, Jonathan; Labare, Mathieu L.M.; Plecko, Krunoslav; Post, Suzanne; More Authors

**DOI**

[10.1016/j.giq.2020.101496](https://doi.org/10.1016/j.giq.2020.101496)

**Publication date**

2020

**Document Version**

Final published version

**Published in**

Government Information Quarterly

**Citation (APA)**

Rukanova, B., Tan, Y. H., Slegt, M., Molenhuis, M., van Rijnsoever, B., Migeotte, J., Labare, M. L. M., Plecko, K., Post, S., & More Authors (2020). Identifying the value of data analytics in the context of government supervision: Insights from the customs domain. *Government Information Quarterly*, 38(1), Article 101496. <https://doi.org/10.1016/j.giq.2020.101496>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.



Contents lists available at ScienceDirect

## Government Information Quarterly

journal homepage: [www.elsevier.com/locate/govinf](http://www.elsevier.com/locate/govinf)

## Identifying the value of data analytics in the context of government supervision: *Insights from the customs domain*

Boriana Rukanova<sup>a,\*</sup>, Yao-Hua Tan<sup>a</sup>, Micha Slegt<sup>b</sup>, Marcel Molenhuis<sup>b</sup>, Ben van Rijnsoever<sup>c</sup>, Jonathan Migeotte<sup>d</sup>, Mathieu L.M. Labare<sup>d</sup>, Krunoslav Plecko<sup>e</sup>, Bora Caglayan<sup>e</sup>, Gavin Shorten<sup>e</sup>, Otis van der Meij<sup>a</sup>, Suzanne Post<sup>a</sup>

<sup>a</sup> Delft University of Technology, Jaffalaan 5, 2628 BX Delft, the Netherlands

<sup>b</sup> Customs administration of the Netherlands, the Netherlands

<sup>c</sup> IBM Nederland B.V., the Netherlands

<sup>d</sup> Belgian Administration of Customs & Excises, Belgium

<sup>e</sup> IBM Ireland, Ireland

## ARTICLE INFO

## Keywords:

Framework  
Data analytics  
Value  
Government  
Supervision  
eCommerce  
Capabilities  
Collective capability

## ABSTRACT

eCommerce, Brexit, new safety and security concerns are only a few examples of the challenges that government organisations, in particular customs administrations, face today when controlling goods crossing borders. To deal with the enormous volumes of trade customs administrations rely more and more on information technology (IT) and risk assessment, and are starting to explore the possibilities that data analytics (DA) can offer to support their supervision tasks. Driven by customs as our empirical domain, we explore the use of DA to support the supervision role of government. Although data analytics is considered to be a technological breakthrough, there is so far only a limited understanding of how governments can translate this potential into actual value and what are barriers and trade-offs that need to be overcome to lead to value realisation. The main question that we explore in this paper is: *How to identify the value of DA in a government supervision context, and what are barriers and trade-offs to be considered and overcome in order to realise this value?* Building on leading models from the information system (IS) literature, and by using case studies from the customs domain, we developed the *Value of Data Analytics in Government Supervision (VDAGS)* framework. The framework can help managers and policy-makers to gain a better understanding of the benefits and trade-offs of using DA when developing DA strategies or when embarking on new DA projects. Future research can examine the applicability of the VDAGS framework in other domains of government supervision.

### 1. Introduction

Customs administrations nowadays have to find a delicate balance. On the one hand they need to control the cross-border flow of goods to ensure revenue collection and safeguard safety and security. On the other hand they need to facilitate the legitimate trade and stimulate economic growth (Tan, Bjørn-Andersen, Klein, & Rukanova, 2011). In recent years, developments such as eCommerce and Brexit put additional challenges on customs administration due to the steep increase of

customs declarations that need to be controlled. For example looking at eCommerce, as a report indicates, “when the growth rates of eCommerce are taken into consideration (CAGR<sup>1</sup> of 12%, up to 18% in the high growth scenario), it becomes clear that problems such as administrative burden, non-compliance, and consequent VAT loss, distortion of competition, will only become more pressing in the near future” (EC, 2016, p.3). Furthermore, customs administrations are preparing for possible effects and an increase in customs declarations due to Brexit<sup>2</sup>. Having in mind that customs is dealing with large trade volumes, which

\* Corresponding author.

E-mail addresses: [b.d.rukanova@tudelft.nl](mailto:b.d.rukanova@tudelft.nl) (B. Rukanova), [y.tan@tudelft.nl](mailto:y.tan@tudelft.nl) (Y.-H. Tan).

<sup>1</sup> Cumulative Average Growth Rate

<sup>2</sup> See e.g. <https://customsnl-insight.nl/article/319660563>

<https://doi.org/10.1016/j.giq.2020.101496>

Received 30 January 2020; Received in revised form 28 May 2020; Accepted 15 June 2020

0740-624X/ © 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

are likely to increase even further in the future, the use of IT and modern customs risk management methods are seen as essential components of the solution<sup>3</sup>. In recent years, customs administrations are also starting to explore what possibilities data analytics (DA) may offer in this context.<sup>4</sup>

Although the business sector has been leading in the use of big data, governments are also actively exploring the opportunity to use big data to address public sector challenges (e.g. Chatfield & Reddick, 2018; Chen, Chiang, & Storey, 2012; Hagen, Keller, Yerden, & Luna-Reyes, 2019; Kim, Trimi, & Chung, 2014; Vydra & Klievink, 2019). Maciejewski (2017) identifies three government roles, where analytics can produce significant improvements. Each of these calls for different use of big data and data analytics (DA), namely: (1) public supervision, which deals with detecting and penalising non-compliance with government laws; (2) public regulation, which focuses on regulating social activities; and (3) public service delivery, which focuses on providing services or products. In this paper we focus specifically on the use of DA to support the supervision role of government, and more particularly, in the domain of customs.

Research shows (Kim et al., 2014) that although businesses and governments can potentially derive value from the massive amounts of data that they collect, they need also to overcome various challenges that pose barriers for value realisation. Kim et al. (2014) further discussed that some challenges (e.g., choosing and implementing technology to extract value from big data, and finding skilled personnel with data analytics skills) are challenges that both businesses and governments face. But at the same time, “the challenges for governments are more acute, as they must look to break down departmental silos for data integration, implement regulations for security and compliance, and establish sufficient control towers” (Kim et al., 2014, p. 81).

Looking at the eGovernment literature, a growing body of research on big data focuses on big, open, and linked government data (Bertot, Gorham, Jaeger, Sarin, & Choi, 2014; Janssen, Konopnicki, Jane, Snowdon, & Ojo, 2017; Janssen & van der Hoven, 2015; Lnenicka & Komarkova, 2018). Reflecting on these studies, the focus is often on governments opening up their data to allow further analytics to be applied. In this context, government tends to play the role of a data provider, opening up its data for broader use. In other cases, the government can be a user of data from other government agencies (nationally and internationally), businesses and non-government organisations (Gil-Garcia, 2012; Rukanova, Huiden, & Tan, 2017; Susha, Janssen, & Verhulst, 2017; Susha, Rukanova, Gil-Garcia, Tan, & Gasco, 2019). In the context of customs, the government is more a user of data to perform better control, and relies on a broad set of data sources that go beyond open government data. In the eGovernment literature, we also identified that studies have focused on themes such as big data and artificial intelligence (Pencheva, Esteve, & Mikhaylov, 2018). Research has also looked at big data and linking cities and sensors (Fraefel, Haller, & Gschwend, 2017), or big data enabled through smart-phones for public services (Anshari & Lim, 2017). There are also several studies that investigate big data analytics to create value by looking at specific domains such as smart cities (Cronemberger & Gil-Garcia, 2019), customer agility and responsiveness (Chatfield & Reddick, 2018), the value of social media data (Panagiotopoulos, Bowen, & Brooker, 2017), and creating value through open government (Attard, Orlandi, & Auer,

2017). An interesting perspective in the eGovernment literature is provided recently by McBride, Aavik, Toots, Kalvet, and Krimmer (2019), looking at open government data-driven co-creation. Looking at food inspection, this study examines how open government data can contribute to the co-creation of new public services and identifies factors that play a role when co-creation occurs where non-trivial actors are involved in the co-creation process. These factors include motivated stakeholders, innovative leaders, proper communication, an existing OGD portal, external funding, and agile development.

Nevertheless, in the current eGovernment literature, we did not find general models<sup>5</sup> on the value of big data analytics, which we could further adapt to the government supervision context. From the specific models and studies that we identified, they either focussed on specialised topics (e.g. smart cities, social media usage, customer agility) that are quite different and not directly applicable for the government supervision context, or they had a specific focus on open government data. In the case of government supervision, we are interested in a broader setting than open government data. Therefore, the main question that we set to explore in this paper is:

How to identify the value of DA in a government supervision context, and what are barriers and trade-offs to be considered and overcome in order to realise this value?

By trade-offs, we mean here, that while certain decisions to use DA may help to realise benefits in one area, they may not necessarily bring benefits – and may even hinder improvements – to other areas in the organisation.

Before proceeding further, we will elaborate on the concept of value. For the purpose of our paper, we searched for a broad definition and turned to the dictionary definitions of value. The Merriam Webster dictionary defines the word value<sup>6</sup> with a number of meanings, including: (1) the monetary worth of something; (2) a fair return or equivalent in goods, services, or money for something exchanged; (3) relative worth, utility, or importance; (4) something (such as a principle or quality) intrinsically valuable or desirable (sought material values instead of human values); (5) a numerical quantity that is assigned or is determined by calculation or measurement. For the purpose of this paper, we adopt the high-level definition of value as discussed in (3), namely relative worth, utility, or importance. The view on value as something relative is key, as there are different concerns and considerations and perspectives when analysing value of DA, and it is revealing these complexities that we are interested in, rather than value in specific monetary terms, or numerical quantities.

To address our research question, in this paper, we develop a *framework for identifying the Value of Data Analytics in Government Supervision*. We will refer to it as the VDAGS framework in short. We developed this framework by following an iterative approach (Eisenhardt & Graebner, 2007). The theoretical underpinning of the framework builds upon two recently published and leading research models on the value of big data analytics from the information systems (IS) literature (Grover, Chiang, Liang, & Zhang, 2018; Günther, Mehrizi, Huysman, & Feldberg, 2017). For the further empirical extension, we build predominantly upon two case studies from the customs domain. More specifically, we examined innovation projects, called Living Labs, for developing data analytics in two leading customs administrations in Europe, namely Dutch customs and Belgian customs.

The remainder of this paper is structured as follows. In Section 2, we provide a brief overview of research on big data and analytics. In Section 3, we present the theoretical models that we use as a basis for

<sup>3</sup> See e.g. The SAFE Framework of Standards of the World Customs Organisation (WCO), <http://www.wcoomd.org/-/media/wco/public/global/pdf/topics/facilitation/instruments-and-tools/tools/safe-package/safe-framework-of-standards.PDF?la=en>, and Electronic customs-  
[https://ec.europa.eu/taxation\\_customs/general-information-customs/electronic-customs\\_en#heading\\_1](https://ec.europa.eu/taxation_customs/general-information-customs/electronic-customs_en#heading_1)

<sup>4</sup> e.g. see World Customs Organisation News, <https://mag.wcoomd.org/magazine/wco-news-91-february-2020/bacuda/>, last visited 10-5-2020.

<sup>5</sup> By general, here, we mean not defined for a specific domain, and also not limited to open government data only.

<sup>6</sup> See. [https://www.merriam-webster.com/dictionary/value?utm\\_campaign=sd&utm\\_medium=serp&utm\\_source=jsonld](https://www.merriam-webster.com/dictionary/value?utm_campaign=sd&utm_medium=serp&utm_source=jsonld), last visited 23/01/2020).

the development of our VDAGS framework. In Section 4, we discuss our method, as well as the steps that we followed for the VDAGS framework development. We present our framework in Section 5, and subsequently we demonstrate its use in Section 6. We end the paper with a discussion and conclusions.

## 2. A brief overview of research on big data and data analytics

While we do not aim to be exhaustive here, the literature discussed in this section provides insights into the domain of big data and analytics. Big data and data analytics have received a lot of attention over the last decade, and businesses, government, and research organisations are examining the transformative power of this technology (Kim et al., 2014). Big data can be seen as “the information asset characterised by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value” (De Mauro, Greco, & Grimaldi, 2016, p.133). Big data analytics is “the application of advanced analytic techniques to very big data sets” (Russom, 2011, p.4). Big data has caused a shift in the traditional way of analysing data: it has been considered to be a breakthrough in technology which brings many new opportunities (e.g. Chen et al., 2012; Fichman, Dos Santos, & Zheng, 2014; Günther et al., 2017; Sivarajah, Kamal, Irani, & Weerakkody, 2017).

The availability of data has boosted the development of data analytics methods and techniques such as big data analytics, text analytics, web and net analytics (Chen et al., 2012). There are different types of analytics describing how the results from the analytics are used (Sivarajah et al., 2017), such as: (a) descriptive analytics, which helps us to understand what has happened; (b) predictive analytics, which aims to identify what is likely to occur in the future; (c) prescriptive analytics, or analytics that helps with responding to ‘now what?’ and ‘so what?’ questions; (d) inquisitive analytics, which aims at helping to comprehend why something is happening; and (e) pre-emptive analytics, which examines the question of what needs to be done. The analytics carefully examines the data in order to find patterns or exceptions (Krittika, Vishvakarma, Sharma, & Lai, 2017; Wang, Gunasekaran, Ngai, & Papadopoulos, 2016).

However, big data and analytics also bring a lot of challenges (see e.g. Chen et al., 2012; Fichman et al., 2014; Sivarajah et al., 2017). Sivarajah et al. (2017) distinguish three types of challenges: (1) data challenges, i.e. related to the characteristics of the data itself such as volume, velocity, veracity, variability etc.; (2) process challenges or challenges encountered when processing the data such as data acquisition and warehousing; data mining and cleansing, data aggregation and integration etc.; (3) management challenges related to such topics as privacy, data ownership, security, data governance etc. Next to that Kim et al. (2014) point to a number of other challenges, namely, challenges related to: (1) choosing and implementing technology to extract value; and (2) finding skilled personnel with the right data analytics skills. In addition to that, governments face further challenges of having to deal with data that not only comes from multiple channels (such as social networks, the Web), but also from different sources, from different countries, institutions, and departments. In particular, sharing data and information between countries is a special challenge, since there are various legal systems and procedures. Furthermore, as Kim et al. (2014) further argue, compared to businesses, governments face more acute issues to deal with. Issues include: breaking down silos among different departments for data integration; implementing regulations for security and compliance; and having to establish control towers. A major regularity challenge was recently posed by the General Data Protection Regulation (GDPR), which entered into force in 2018 in Europe and aimed at protecting personal data. This regulation places significant restrictions on the kinds of data that can be shared and used for analytics purposes (Zarsky, 2016).

## 3. Conceptual foundations for the development of the VDAGS framework

In this section, we describe how we identified and selected the models that served as a conceptual foundation for the development of our VDAGS framework, and we briefly introduce these models. In the method section, we will explain the iterative process of how we used these models, in combination with the cases in the different stages of the framework development, to arrive at the final VDAGS framework.

As discussed in the introduction, in the eGovernment literature, we did not identify general modes on the value of DA that we can customise for the government supervision, and the domain-specific models that we identified were limited for our purposes. We, therefore, turned to the broader information systems (IS) literature, which examines the use of IS in an organisational context, in search of suitable models<sup>7</sup>. For the identification of relevant studies, we were driven by a number of criteria as follows. (1) The models or frameworks had to be published in leading IS journals. For the selection of journals, we used the basket of 8 leading IS journals defined in the Association of Information Systems Senior Scholars' basket of journals.<sup>8</sup> (2) We searched for models that have received recognition in the IS domain. We, therefore, were interested in models that have been highly-cited. (3) We searched for general models rather than models that focus on very specific aspects or specific domains. (4) We were interested in models that allow room for adaptation to be able to tailor the models for our purpose. We conducted a SCOPUS search where we searched on title. We searched for journal articles in the period 2010–2020. We conducted searches with the keywords: (1) “value” and “analytics”, or (2) “big data” and “value”. Search (1) resulted in 131 results; search (2) resulted in 92 results. For each search, we sorted the results by the number of citations. Subsequently, for all the results, we screened the papers with more than 10 citations to check in which journal they appeared. We subsequently identified those that appeared in the basket of 8 IS journals. The results of these two lists were partially overlapping. When combining these lists, we arrived at a combined list consisting of 5 papers that appeared in the basket of IS journals. Even though we searched in the period 2010–2020, the papers that we identified in our short list were all published in the period 2015–2018, indicating that the topic on the value of DA is starting to gain attention also in the general IS domain only recently. The citation range of papers with above 10 citations was 119 the highest, and 13 the lowest, with four of the papers having above 40 citations. In order not to miss some new models that have been recently published in the basket of 8 IS journals and had a lower number of citations, we repeated the search for a shorter period 2016–2020<sup>9</sup>. We screened the results again to identify publications from the basket of IS journals, but this time we looked for articles with less than 10 citations. As a result, we identified 2 more articles which we added to the list. This resulted in 7 selected articles (see Annex B for an overview), which we then reviewed in detail. We further analysed the paper in terms of the other two criteria, namely criteria (3), i.e., that the paper offers general models; and criteria (4), i.e., that the models are suitable for further adaptation. Regarding (3), four of the models were more specific (i.e., they focussed on specific topics such as data usage, customer analytics, social media data, and relationships between business analytics systems and customer relationship management system) and we, therefore, did not consider them further. The other three models were more general in nature, and

<sup>7</sup> We chose the IS literature, as it examines the use of IS in an organisational context. At this stage, therefore, we did not look at purely management or purely technical computer science literature.

<sup>8</sup> <https://aisnet.org/page/SeniorScholarBasket>

<sup>9</sup> From 2015 on, we already had identified the papers with more than 10 citation. In the second search, we took a shorter period to be able to screen the papers more efficiently.

we further focused on examining those. We found that two of the three models were particularly interesting as they were research frameworks that served for drafting further research agenda on the topic. Therefore these models were also particularly suitable with respect of critical (4) for further adaptation. They also covered complementary aspects (one was taking an inter-relationship perspective, the other process perspective), which meant that the two models already covered a broad range of complementary issues. We, therefore, decided to proceed with these two models as main theoretical models that we use for the development of our VDAGS framework. The third model was also relevant. It also covered the process perspective (in that sense, there was overlap with one of the other two models discussed above). It also contained some additional elements which we considered as useful addition. Therefore, while we based our VDAGS framework on the first two models, we also included some elements from the third model in our adaptation.

The first model on which we build upon is the model of big data value realisation (Günther et al., 2017). Günther et al. argued that there is still a limited understanding of how organisations translate the potential of big data and analytics into value. They propose that to realise value from big data, it is imperative for organisations to continuously realign work practices, organisational models, and consider external stakeholders interest. In the model of big data value realisation that they propose, Günther et al. position the concept of value in the middle, and they propose that to address the value of big data and analytics, organisations need to look at the interrelationships at three levels, as follows: (a) work practice level, i.e., working with big data analytics in practice; (b) organisational level, i.e., developing organisational models; (c) supra-organisational level, i.e., dealing with stakeholders' interests. In addition, the authors argue that portability and interconnectivity are prerequisites for establishing the foundation for big data. This model is interesting, as it puts value at its centre, and value is understood in a broader sense by examining the interdependencies between the work practices where DA is actually used, the organisational context where these work practices take place, and the interactions with and influences from external stakeholders. This positioning of value as interdependency between the different levels also fits very well with our definition of value as relative and had a very good fit with our empirical context. Günther et al. support the argument that the perception of the value of big data for organisations depends on their strategic goals in using big data<sup>10</sup>.

The second model on which we build upon is the model of Grover et al. (2018) which provides a strategic process perspective on understanding value of DA. The Grover et al. model takes a process view, examining how big data analytics capabilities and realisation processes evolve over time and add value. Grover et al. realise that big data analytics plays a strategic role in an organisation. They consider that organisations can pursue two types of values, namely functional value, and symbolic value. Functional value is seen as performance improvements that result directly from adopting big data analytics. Symbolic value is derived from identifying the effects of investing in big data analytics. The central concepts in the model of Grover et al. are: (1) capability building processes; (2) capability realisation processes; (3) learning loops, also referred to as 'learning by doing' or 'co-evolutionary adaptation'. *Capability building processes* cover: (a) the big data analytics infrastructure (which includes big data assets such as data assets and platforms; analytics portfolio and human talent), and (b) big data analytics capabilities (i.e., ability to integrate, disseminate, explore and analyse big data). *Capability realisation processes* include (a) value creation mechanisms (e.g., transparency and access; discovery and experimentation; prediction and optimisation); (b) value targets (organisational performance; business process improvement; product and

service innovation; consumer experience & market enhancement); and (c) impact in terms of functional value and symbolic value. Besides, the model includes moderating factors that influence the capability realisation processes. These moderating factors include factors like strategy, leadership, trust, technology and industry context, governance support, data-driven culture and competitive dynamics.

The third model is the model of Seddon, Constantinidis, Tamm, and Dod (2017). This model consists of two parts, a process model and a variance model (including factors). The process model also includes learning. The variance model includes factors driving benefits from each project and focusing on the short-term, as well as long-terms organisational benefits. The multiplicity of individual projects is linked to the organisational benefits from data analytics improvement.

In our framework, we used the models of Günther et al. (2017) and Grover et al. (2018) as the basis for our framework, and we extended the process perspective with elements from the model of Seddon et al. (2017)<sup>11</sup>. These models provided a rich conceptual basis for understanding value, and they covered complementary perspectives. To deal with the complexity, during the framework development, we gradually introduced the models and applied them to the empirical context starting with one model and one Living Lab and subsequently increasing the complexity (see Section 4.2).

#### 4. Method

For the development of our framework, in this study, we adopted an interpretative and contextualist case study approach (Klein & Myers, 1999; Orlikowski & Baroudi, 1991; Walsham, 1993). Interpretive studies are "aimed at producing an understanding of the context of the information system, and the process whereby the information system influences and is influenced by the context" (Walsham, 1993, pp 4-5). In our study we were interested in big data analytics and the context and the processes through which data analytics influences this context. In particular, we were interested in understanding what value DA brings to a government organisation in the context of government supervision, and what are benefits and trade-offs for achieving that. Reflecting on the types of theory, Gregor (2006) identifies five theory types, namely theories for: (1) analysis; (2) explanation; (3) prediction; (4) explanation and prediction; and (5) design and action. As the topic that we are investigating is still not well researched, the VDAGS framework that we develop in this study is predominantly intended as a framework for analysis and explanation, therefore reflecting the first two theory types. The development of our VDAGS framework progressed in several stages and was developed in an iterative manner, where the empirical context guided us in the search of relevant theories, which in turn enabled us to structure our observations. At the same time, our empirical context allowed us to adapt the theoretical models that we identified. Such an iterative approach allows for the development of theories that are deeply informed by the empirical context (Eisenhardt and Graebner (2007)). In the next section we present: (1) our empirical context; (2) the iterative process between the theories and the empirics that we followed, and the different stages through which the development of the VDAGS framework progressed.

<sup>11</sup> While the model of Seddon et al. (2017) is different compared to Grover et al. (2018), they do have some similarities: both include a process perspective, and both cover key factors. For our framework we selected the Grover et al. (2018) as it is a research framework and is more applicable for further adaptation. However, in our adaptation, to the process perspective of Grover et al. (2018), we also added insights from the model of Seddon et al. (2017) specifically to capture the short-terms and long-term perspective and the cumulative effect of multiple data analytics projects.

<sup>10</sup> This definition is in line with the broader definition that we adopt in this paper, namely that value is seen as a relative concept.

#### 4.1. Empirical context: the PROFILE project and the four living labs

The empirical context for this study was provided by the PROFILE research project funded by the European Commission's Horizon 2020 Research programme. The project brings together customs administrations from five EU countries (the Netherlands, Belgium, Sweden, Norway and Estonia), leading technology and data analytics providers (IBM, TNO, FFI, FOI, Inlecom), the European Commission's Joint Research Centre, associations and academic partners. The aim is to develop and test, in real life, innovative DA solutions to improve customs risk assessment processes. All customs partners involved in the project have an interest in data analytics innovation. The PROFILE project is structured around demonstration projects (referred to as Living Labs (LLs)), conducted in various EU countries as real-life settings for developing and piloting innovative DA solutions. The Living Labs research approach "takes a development view of innovation and studies novel technologies in a complex real-world setting" (Higgins & Klein, 2011, p.32).<sup>12</sup> Four Living Labs have been set up in the context of the PROFILE Project. Table 1 provides an overview of the four Living Labs and the use of data analytics.

This study is predominantly based on an in-depth analysis of the first two Living Labs of Dutch and Belgian Customs respectively. The other two Living Labs were not analysed in detail. However, through the PROFILE project we also had access to these Living Labs and we used insights gained as additional inputs for this study. We selected the cases based on convenience and theoretical sampling. We had access to the Living Labs via our involvement in the PROFILE project. Our participation in the project allowed us to closely monitor concerns and trade-offs related to the use of DA. This would have been very difficult to investigate in cases where we did not have such rich access. The Living Labs of the PROFILE project were not set up at random but were part of a deliberate project design. The Dutch and the Belgian cases were focusing on the use of DA internally in customs administrations in Europe, but these two administrations were also interested in exploring the benefits of collective data analytics capability building. Therefore, the first two living labs initially focussed on DA development in their individual organisations and subsequently they concentrated on the collective aspect. The other two Living Labs were focussed on the collective process from the start. Therefore the PROFILE project design and the way the Living Labs were set up provided complementary cases that were suitable for theoretical sampling to extend theory. We started the initial framework development based on a single case, namely the Dutch Living Lab. This case was selected based on theoretical sampling as it provided opportunity to an unusual research access (Eisenhardt & Graebner, 2007; Yin, 1994) where we could get very rich access to the empirical context due to established long-term relationships and physical proximity of the researchers with the empirical context provided by Dutch Customs. As will be explained in the next section, this case allowed us to construct the first version of our framework. However, from a theory-building perspective, multiple-case studies allow for a more substantial base for theory building (Yin, 1994). We, therefore, subsequently used theoretical sampling to include subsequent cases. We added the Belgian Living Lab, as it allowed us to examine the use of data analytics in another leading customs administration in Europe, but which made different choices on piloting with analytics, namely using machine learning to analyse trader behaviour. This different context in terms of country and analytics approach provided the opportunity for validating and further extending the initial framework. The major part of the framework development was done using these two cases and in iteration loops. Subsequently, we examined the other two cases, also based on theoretical sampling as they provided opportunities to use contrasting cases to the Dutch and the Belgian Living Lab to further

extend theory. The new cases were contrasting in a sense that their primary focus was on the collective aspect of developing data analytics to realise value. In contrast, the first two cases were initially focussed on exploring how they can derive value from DA by concentrating on their own administration and only later in the process they focussed on the collective. Therefore, while the richness of the analysis of last two Living Labs that we added was more limited compared to the Dutch and Belgian Living Labs, from a theoretical perspective, it was worth included them in this study. They reconfirmed the significance of the collective view for examining the value of DA empirically.

The data collection related to the Dutch and the Belgian Living Labs was done as follows: participation in project meetings; review of key project documents and deliverables; bi-weekly conference calls related to each Living Lab; face-to-face meetings with experts from the risk analysis and data analytics teams; as well as participation in workshops with targeting officers. Besides, a series of dedicated interviews were conducted with customs experts from both Belgian and Dutch customs to understand the current processes and data analytics developments. Demonstration sessions by the external DA providers were attended in the Living Labs to follow the DA developments and progress of the LL. Several dedicated sessions were organised with experts on performance measurement to understand the priorities, concerns and considerations involved in measuring performance and DA results. The data collection related to the other two Living Labs was performed via general project meetings where the results of these Living Labs were discussed, via individual interviews and review of key project deliverables. Annex A provides an overview of key experts with whom we interacted in the context of this study.

The data collection and data analysis were performed in an iterative manner. During the various stages of data collection, detailed meeting notes were taken, the findings were analysed via the lens of our theoretical models as the model development progressed. The results of the analysis were sent back to the customs experts for comments and further clarification. This allowed us to identify inaccuracies in the interpretations and correct those. Multiple sources of evidence were used for triangulation of findings. In this process, we identified themes which we used subsequently to make the general theoretical models that we used as a conceptual basis for our framework more specific for our context. The different versions of the framework were subsequently presented and discussed with experts from the Living Labs, and the feedback was used to revise the framework. In the next section, we provide further details on the iterative process that we followed to arrive at the final framework.

#### 4.2. Iterative process and stages for the VDAGS framework development

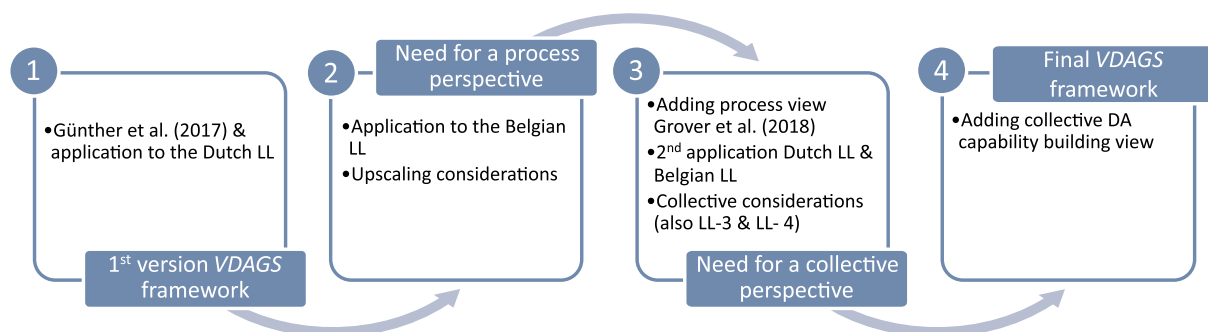
The process that we followed for the VDAGS framework development can be roughly divided into four stages, as presented in Fig. 1. In Stage 1, we started with the Günther et al. big data value realisation model, and we applied and extended it in the context of the Dutch LL. We chose to start with the Günther et al. model instead of the Grover et al. model as it was particularly suitable to capture the context for government supervision. Especially the work practice level was very suitable to be extended to capture the customs risk assessment process. The supra-organisational level was suitable to capture interactions with external parties. We considered that it is essential to capture the specific context first before we focus on the processes.

We used the levels (i.e. work practice, organisational, and supra-organisational) to structure our empirical observations from the case. At the same time, we used themes that emerged from the Dutch LL to make the model more concrete. We included a specific customs risk assessment process at a work practice level, and based on themes that emerged from the case, we further added categories to each of the levels. This resulted in an initial version of our VDAGS framework (see Stage 1 in Fig. 1 from the framework development process). This version was presented at the eGov'2019 conference (Rukanova et al., 2019).

<sup>12</sup> See also <http://tiny.cc/8s564y> for other examples of applications of Living Lab approach.

**Table 1**  
Overview of the living labs.

Living Lab	Overview	Focus (individual organisation/collective)
LL-1: Dutch Living Lab (Dutch LL)	Focus on customs risk assessment of eCommerce flows. Data analytics used is web data retrieval of price information from eCommerce platforms for cross-validation of price information in eCommerce declarations. Use of DA in the customs risk assessment process as a support tool for a human decision-maker.	Initial stage: focus on the individual organisation Later stage: individual and collective
LL-2: Belgian Living Lab (Belgian LL)	Focus on the use of DA for analysing the behaviour of operators. Focus on machine learning on historic data sets and using external data sources for analysing the behaviour of traders. Use of DA to enhance risk assessment software.	Initial stage: focus on the individual organisation Later stage: individual and collective
LL-3: Sweden- Norway Living Lab (S&N LL)	Comparing aggregated results of the data analytics performed on customs declaration data of two neighbouring customs administrations (one in the EU and one outside the EU).	From the beginning focused on collective (2 customs administrations)
LL-4: EU Living Lab (EU LL)	Focus on developing and piloting of an infrastructure for sharing data among customs administrations in the EU. Only governments can access this infrastructure. Initially intended for exchanging bulk data among member states which can be used by the participating member states for their own data. analytics.	From the beginning focused on collective (multiple customs administrations)



**Fig. 1.** Iterative stages for the VDAGS framework development.

Subsequently, in Stage 2, we applied our VDAGS framework to the Belgian LL. The Belgian LL allowed us to examine the applicability of our VDAGS framework in another context and allowed us to extend the framework with the new findings.

As the Living Labs proceeded, upscaling from a Living Lab to the organisational context became more relevant. At this stage, it became clear that the framework that we had so far was useful to elicit the value of data analytics in a specific customs context. It also allowed for the identification of barriers and trade-offs. Nevertheless, it was limited and did not allow to understand the Living Lab upscaling processes. The framework was too static and lacked the process perspective to capture the dynamic processes of how the capabilities developed in the Living Labs extend and build upon earlier DA capabilities developed in the customs organisations over time. At this point, it was suitable to extend the initial framework with a process perspective. Subsequently, in the next phase (Stage 3), we extended the initial framework to better capture the processes perspective and we adapted the Grover et al. strategic process value model for that purpose. A question that arose at this time of the framework development was how to integrate the new perspective into our initial framework developed so far. We decided to use views. The initial part of the framework was focussed mostly on interdependencies between levels. We, therefore, labelled it *the interdependency view*. We added the process perspective which focusses on DA capabilities processes over time as *the process view*. These two views allowed us to look at the same phenomenon but from different perspectives. Furthermore, the use of views allowed us to keep the underlying logic behind the conceptual models that form the basis of each of the views to a large extent independent. The adding of the process view also resulted in the second cycle of data collection and data

analysis. In the second iteration, we first applied the extended model to the Dutch LL, now taking into account the process view. This allowed us to gain a better understanding of the data analytics capability building processes preceding the Dutch LL and to better understand the Dutch LL as a new cycle in that process. Subsequently, we conducted a second round of data collection and data analysis of the Belgian Living Lab. The extended framework that resulted based on the second iteration round with the Living Labs turned to provide for a richer understanding than the initial version. However, based on the interaction with the Living Labs another limitation of the framework became visible. The Belgian and the Dutch LLs were entering a stage where they started to discuss collaboration to exchange algorithms and data sets. This collective aspect was not possible to capture explicitly with the current framework. The framework that we had developed was mainly developed from the point of view of a single organisation. While our framework included the aspect of supra-organisational level, which was adopted from Günther et al., this aspect focussed predominantly on how one organisation manages its relationships with external stakeholders. This aspect did not explicitly cover the collaborative processes for collective data analytics capability building. At this stage of the framework development we examined the other two Living Labs (LL-3 and LL-4) as contrasting cases to extend the theory. In these two cases, the collective process was central to the Living Labs. This gave us the further assurance that it is essential to include a new view, namely the *collective capability building view* in our framework. While further development of this view is out of the scope of this paper, our study showed empirically that this is a missing aspect for elaborating value of DA that needed to be explicitly added. As a result of this iterative process, we arrived at our final framework, which is presented in the next section.

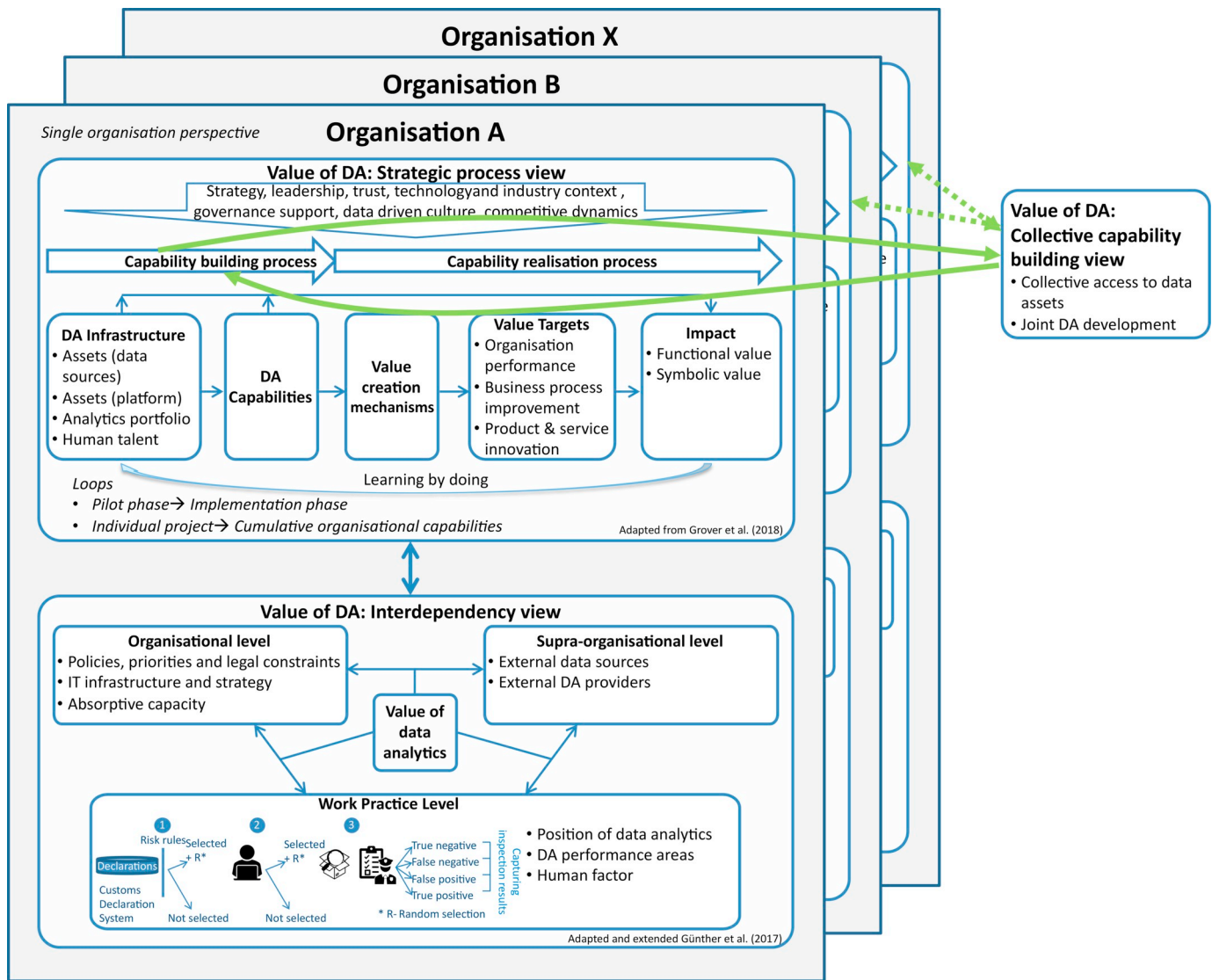


Fig. 2. VDAGS framework.

## 5. Result: a framework for identifying the value of DA in government supervision

In this section, we present the *VDAGS* framework (see Fig. 2) that we developed based on the process discussed in Section 4. The framework incorporates three views to analyse the value of DA in the context of government supervision, and related barriers and trade-offs, namely: (1) the *interdependency view*; (2) the *strategic process view*; and (3) the *collective capability building view*.

### 5.1. Interdependency view

The first view in our *VDAGS* framework is the *interdependency view* (see bottom part of Fig. 2, interdependency view).

As discussed earlier, as a basis for the *interdependency view* we build upon Günther et al., which postulates that the value of DA can be understood by examining the interdependencies between work practice level, organisational, and supra-organisational levels. In our conceptualisation, we made the following adaptation to the levels<sup>13</sup>. First

<sup>13</sup> Günther et al. (2017) include also portability and interconnectivity in their model. In our model in order to simplify the visualization we absorb this aspect under IT strategy.

of all, the *work practice level* is further developed to explicitly include the customs risk assessment process<sup>14</sup>. This explicit inclusion of the customs process was instrumental in capturing the context where DA is used. Furthermore, three additional concepts were added under work practice level to capture aspects that emerged from the case, namely: (1) the *position of DA in the customs clearance process*; (2) areas of desired *performance* improvements in the customs process with the aid of DA; (3) the *human factor* in the socio-technical customs clearance process.

Second, the *organisational level* was further developed to include specifically the following concepts: (1) *IT infrastructure and strategy*; (2) *absorptive capacity*; (3) *policies, priorities and legal concerns*. Finally, the *supra-organisational level* was further operationalised to explicitly include concepts related to (1) *external data providers* and (2) *external DA providers*.

### 5.2. Strategic process view

The second view in our *VDAGS* framework is the *strategic process view* (see the top part of Fig. 2). It is based on the strategic process

<sup>14</sup> For presentation purposes the detail explanation of the customs risk assessment process that is part of the work practice level in our framework will be explained in Section 6.1.



perspective of Grover et al. (2018). It captures the key concepts from the Grover et al. model: the two broad arrows in the process view of Fig. 2 are those representing the *Capability building processes* and the *Capability realisation processes*. Under each of the processes, we also captured the concepts that Grover et al. included in their model, as we discussed in Section 3. The vertical arrow pointing down to the capability realisation processes includes the *moderating factors* as identified in the model of Grover et al. The arrow in Fig. 2 from *impact* to *DA infrastructure* labelled *learning by doing* was also adopted from Grover et al., and illustrates the various learning loops in the process. In our adaptation we excluded details that populated some of the boxes under the processes in order to simplify the model.

Our extension to the Grover et al. model is that we added the concept of *loops*, where we added the following loops: (1) *pilot phase* → *implementation phase*; (2) *individual project* → *cumulative organisational capabilities*. We added the first loop based on our empirical insights and the DA considerations moving from a pilot phase of the Living Labs towards implementation. As we will demonstrate with the examples later, there are different considerations in the different phases. Inspired by Seddon et al. (2017), we added the second loop, *individual project* → *cumulative organisational capabilities*, which enables us to trace how different individual project loops contribute to cumulative capability building in the organisation.

### 5.3. Collective DA capability building view

The third view in our VDAGS framework is the *collective DA capability building view*. Fig. 2 shows how the *collective DA capability building view* extends the strategic process view. The arrow pointing from *capability building processes* in the *strategic process view* to the *collective capability building view* in Fig. 2 represents that, as part of a specific cycle, an organisation may decide to engage in collective capability building with other organisations to develop DA capabilities in collaboration. The collective process happens outside of the specific organisation, and other organisations can also join in (indicated by the dashed arrows in Fig. 2). This collective process can be focussed on (but not limited to): (1) collective access to data assets (e.g., sharing data among customs organisations or jointly securing access to business data source); or (2) joint DA development. Once the collective capability building process develops new capabilities, the next step is to bring these capabilities back into each of the individual organisations (denoted by the arrow, pointing from the *collective capability building view* to the *capability building process*, as part of the *strategic process view* of the specific organisation). These new capabilities then become part of the capabilities of the specific organisation, which can utilise them in its own capability realisation processes. This logic applies to each of the organisations participating in the collective process. For the sake of simplicity, in Fig. 2, we make the interactions explicit via the solid arrows from the *strategic process view* to the *collective view* and back for only one of the organisations. For the other organisations, we used double-sided dashed arrows in order not to complicate the figure; these dashed arrows, however, are intended to capture the same types of interactions as represented by the solid arrows for Organisation A. Together, these three views of our VDAGS framework provide richer insights, compared to each of the views alone. In the next section, we provide a demonstration of the VDAGS framework based on case examples.

## 6. Demonstration of the framework

In this section, we demonstrate the use of the framework that we developed. We start by providing an introduction to the customs risk assessment process.

### 6.1. Introduction to the customs risk assessment process

For a better understanding of the context of our study, it is important to provide high-level insights into the customs domain and specifically the customs risk assessment process (see also Fig. 3). This high-level process is a rather standard process for conducting risk assessment in customs when goods enter the EU. When goods enter Europe, a customs declaration for these goods needs to be submitted to the customs declaration system. Once the declaration has been submitted, an automated risk analysis is performed, based on risk rules predefined in risk assessment software (Step 1). These risk rules have normally been developed by human experts, and many of them reflect mandatory rules issued by various EU Directorate Generals, such as customs (DG-TAXUD) and food and product safety (DG-SANTE).

After the customs declarations have been assessed by the software, a list is generated, which marks declarations that are considered risky based on the risk rules. This initial risk assessment produces a long list of declarations. In the next step (Step 2), a human targeting officer further analyses the list of declarations and makes a final selection for inspection. After the targeting officer has made the final choice, the list of declarations selected for inspection is sent to the inspection team, which is responsible for carrying out the actual inspections. Subsequently, the inspection team inspects the packages (Step 3) and enters the inspection results into the inspection reporting system. The goods inspected would either have been identified as suspicious based on the risk rules by the targeting officer, or would have been randomly selected from the deselected population that arises from Steps 1 and/or Step 2. In Fig. 3 we use the symbol R to indicate random selections. In line with detection theory statistics, in the case of suspicious goods, the outcome of the inspection could be that something wrong was indeed found (a hit): this is referred to as a *true positive* (TP) selection. If nothing suspicious was found in goods that were considered suspicious, it is referred to as a *false positive* (FP). The result of the inspections performed on the randomly selected goods could be a *true negative* (TN), i.e., the goods were not selected as suspicious and were indeed not suspicious. The result could also be a *false negative* (FN), i.e., the goods were considered as not suspicious in the selection process but in reality something wrong was found during the random inspection. The existence of false negatives is very important, as it means that goods which were suspicious, were not identified in the selection process: this indicates that the risk rules should be improved so that next time this type of goods will be identified as suspicious.

Having explained the current process, it is also important to highlight that customs currently face a huge disproportion between the number of declarations submitted and the number selected for inspection. Only a very small percentage of declarations are selected after Step 1, and the number is further reduced in Step 2 by the targeting officer. Of the declarations for which the goods finally undergo physical inspection, only a small percentage result in true positives (actual hits). Reducing false positives would be an improvement, but that would be limited to the declarations selected after Step 2 and would not affect the other declarations released during the earlier steps. Among the declarations released in Step 1 and Step 2 that are not considered for further inspection, customs administrations currently know very little about how many fraudulent declarations they miss. The purpose of the random selections is to provide some information on that, but random inspections are few. This poses major challenges for customs, and there is a huge disproportion between declared goods, expected fraudulent declarations, goods selected for controls, goods that are eventually controlled, and fraudulent goods found. The current process relies on risk rules software for automatic selection and the expertise of the targeting officer. When engaging in data analytics projects, customs aim to add insights from DA algorithms to the risk assessment process, in addition to the risk rules. In the next section, we demonstrate the use of the VDAGS framework, starting with the interdependency view.

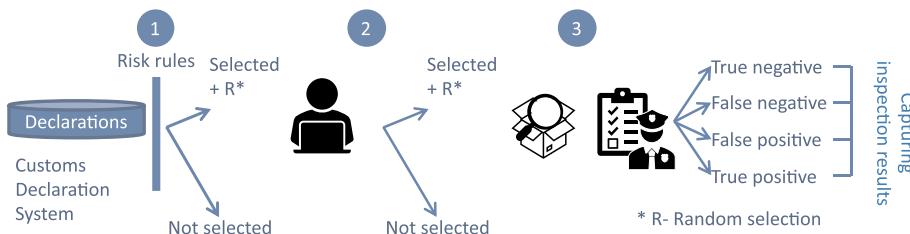


Fig. 3. Customs risk assessment process.

## 6.2. Interdependency view

In this section, we structure the analysis along the three levels from the interdependency view, namely: work practice, organisational, and supra-organisational levels.

### 6.2.1. Work practice level

At a work practice level, one fundamental decision when embarking on DA projects to improve the customs risk assessment process is where to place the data analytics. In the VDAGS framework this is referred to as the *Position of DA*. Based on the cases we identified different positions in the customs risk assessment process where customs can deploy DA (see also Fig. 3), namely:

- (1) Step 1 of the risk assessment process, where DA can be used to improve the risk assessment software to provide more accurate automated selections; this case is piloted in the Belgian LL;
- (2) Step 2 of the risk assessment process, where DA can be used as a support tool to help the human targeting officers to make a final decision on which declarations to select for further inspection and which to release; this case is piloted in the Dutch LL;
- (3) Step 3 of the risk assessment process, where DA can be used to provide support during the inspection process (e.g., DA of scanned images).

Thus customs administrations have various options regarding where to deploy DA in the customs risk assessment process, which step in the process to support, and whether to use analytics to enhance the software or to support the human experts.

At work practice level, there are also various performance areas where customs may want to achieve improvements. This is captured with the concept *performance areas* in our VDAGS framework. In the Dutch LL, the goal is to achieve a reduction in *false positive* cases. In the Belgian LL, the focus of using DA is on finding more cases where something is wrong (from the full population). Furthermore, a clear distinction is made in the Belgian LL when looking at performance areas: it is possible to make different choices and by using data analytics to increase one area does not necessarily mean improving another. Such choices regarding performance include aspects such as: whether to focus on quantitative measures (e.g., catching more fraudulent cases), or qualitative measures (e.g., catching cases of high-level fraud). Another option is to aim that the algorithms aid at handling a steep increase in volumes while keeping the other performance levels the same.

A third consideration when looking at *work practice level* is related to the *human factor*. This factor was deemed important in both Living Labs. The customs risk assessment process is a social-technical process where some parts of the process and outputs are supported by IT, and others are performed by humans. Explainability is a big issue when using DA for customs, and whether the outcome of the DA is taken in the next step of the customs clearance process where humans are involved depends on whether the human experts trust the result of the analytics. This is a key issues, as even if the algorithms perform very well, if the results are not taken up in the next step in the social-technical process to influence the final selection, this will constrain the effects of DA and the value that can be realised in performance improvement. Fig. 4

visualises an example of dependency between the position of DA and performance areas in the Dutch LL.

In the Dutch LL, the choice of positioning DA at Step 2 (see Fig. 4, indicated with a circle around step 2) of the customs risk assessment impacted the areas where performance improvement can be achieved. DA when deployed at this position of the customs process can help to reduce the *false positives* (see the oval shape around false positives in Fig. 4). It will be of limited use, however, for detecting more illegal trade that is part of the population that was already excluded in Step 1.

This is only one example. In the Belgian case, the DA is positioned at the beginning of the customs process, and this offers more opportunities for looking at different performance areas. Still also in this case, there are different choices to be made (e.g., to focus on high-value hits rather than many hits of low-value consignments). These different choices will result in different outcomes from deploying DA in the process, and the realisation of impact in the real world would also depend on whether or not the human experts involved in the follow-up steps in the process trusts the analytics and uses the result.

### 6.2.2. Organisational level

While developing DA at work practice level may identify a high number of fraudulent cases, issues at organisational level may influence the effect of DA. Based on the case analysis we identified factors which we grouped into three categories, namely (1) priorities, policies and legal constraints; (2) absorptive capacity; and (3) IT infrastructure and strategy, which may affect the value that DA brings at work practice level.

Fig. 5 below provides one illustration of how organisational policies can influence the effect of DA at work practice level. Due to limited customs resources, customs administrations define organisational policies regarding a maximum number of inspections that can be performed per day. This maximum thus acts as an upper boundary to the improvements that DA can achieve at work practice level. In Fig. 5, this dependency is shown with the arrow between organisational and work practice level. The oval at Step 3 at work practice level marks the effect of the organisational policy on the inspection capacity at organisational level on the third step of the customs process at work practice level. Subsequently, the arrow from Step 3 to Step 2 of the customs process at work practice level shows the constraining effect of the restricted capacity on the possibilities of the DA deployed at Step 2. This means that even if the analytics deployed at Step 2 could identify many more fraudulent packages, in practice, many of them would not be inspected due to limited inspection resources.

As a result, the effects of the analytics may not be utilised to its fullest potential. Going a step further, based on legislation or EU requirements, national customs administrations need to inspect shipments if they are hit by some of the risk rules defined based on policies or legal grounds. This further limits the inspection capacity for other shipments that are identified as fraudulent by e.g., using analytics, therefore further influencing the value of DA that can be realised in practice. Furthermore, customs organisations can put other priorities on whether to put more resources for detecting fiscal fraud or security threats or other threats. These priorities can change over time, and depending on these changes, DA that is able to identify fiscal fraud (like the one piloted in the Dutch LL) could be more valuable at one point in time than another.

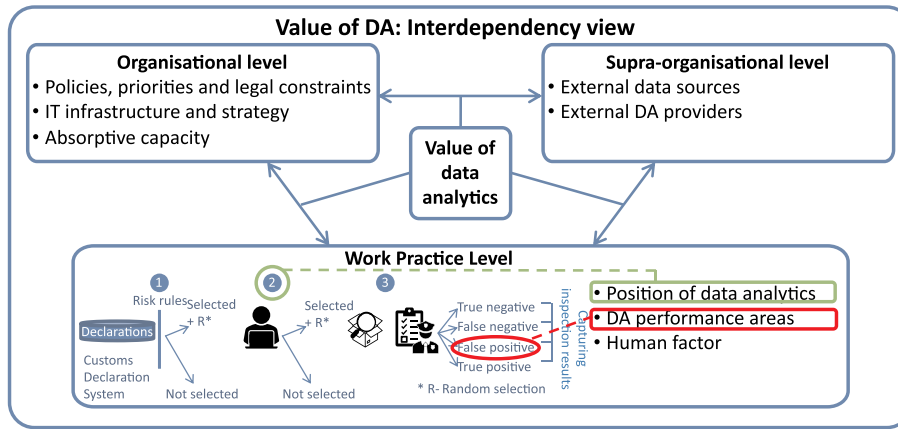


Fig. 4. An example of interdependency between the position of DA and DA performance areas.

Other factors at organisational level that can influence what is possible to achieve at work practice level include the IT Infrastructure and strategy and absorptive capacity. Data analytics relies heavily on IT. In the Dutch LL case, an explicit decision was made to develop the DA solution without taking any constraints of the existing IT infrastructure into account. The benefit of this approach is that it enables the exploration of possible solutions to be unconstrained by the existing IT infrastructure and capabilities. The Belgian LL adopted a different approach. The results of the DA models can, to some extent, be directly fed into the operational risk assessment software. This allows for faster loops in the operational process, where it is possible to evaluate the performance of the algorithms and use the feedback. This is very valuable as it allows for much faster learning loops. The possibility of feeding the results of the DA models directly into the operational risk assessment software system, however, critically depended on the availability of IT to support that. The DA methods that are deployed now are relatively simple. If more complex analytics methods are to be developed in the future, this would also require further investment in the IT to accommodate the more sophisticated analytics. The development of new IT is guided by the organisation's IT strategy. This relationship between DA and DA strategy, and IT and IT strategy is very important, needs to be made explicit and carefully considered.

Another organisational aspect that can affect on what can be realised on a work practice level is what we refer here to as absorptive capacity or as the ability of organisations to be prepared to absorb DA innovations. What we see in the Living Labs is that both the Dutch and Belgian customs have invested in hiring DA experts and forming DA teams. The teams are relatively small, but by collaborating closely with the experts involved in the customs risk assessment process, who have domain knowledge, they are able to experiment with new data analytics

approaches and innovate. Another important aspect that became clear during the study, however, was that there are two groups of experts that provide input to the automated risk rule software used at Step 1 of the customs risk assessment process. First, there are the risk rule experts who define the risk rules. Second, there are the DA experts that develop additional models based on the results of DA. These two groups of experts have differences in paradigms and ways of thinking, but work towards improving the same customs risk assessment process. These differences need to be acknowledged, and organisational measures can help to achieve a more aligned way of working between the two expert groups to achieve better results and benefit from the potential offered by DA.

6.2.3. Supra-organisational level

Next to the organisational level, the supra-organisational level can also affect what is achieved at work practice level with DA. Under supra-organisational level we identified two important aspects, namely (a) the external data sources, and (b) the external DA providers.

When developing DA, a key question is what kind of data will be used for the analytics. In the Dutch LL case, in order to perform the intended data analytics, it is necessary to obtain price data from eCommerce platforms and websites to cross-validate the prices on customs declarations. The choice of this type of data and the potential value that it can bring at work practice level, however, triggered issues and dependencies at the supra-organisational level (see the arrow between work practice and supra-organisational level in Fig. 6) related to accessing external data sources.

In many cases, eCommerce platforms would not allow robots to crawl their websites. An alternative way of accessing price information from eCommerce platforms was identified, namely via an Application

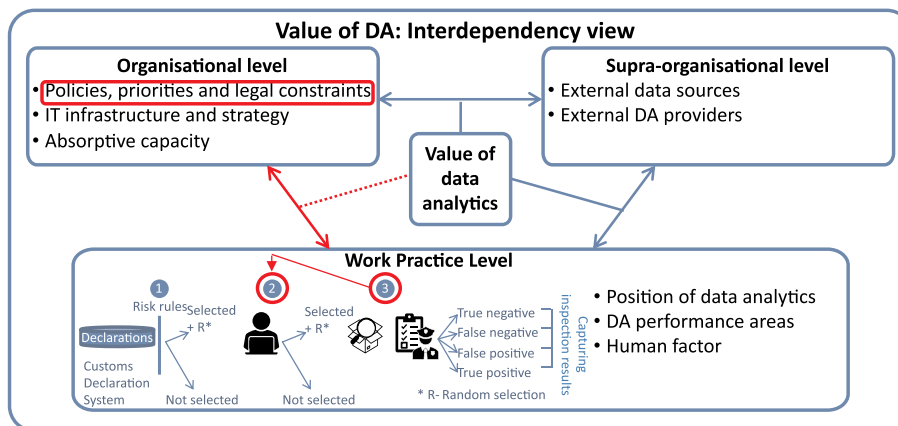


Fig. 5. Organisational policies on maximum inspection capacity act as upper bound for what can be achieved with DA at work practice level.

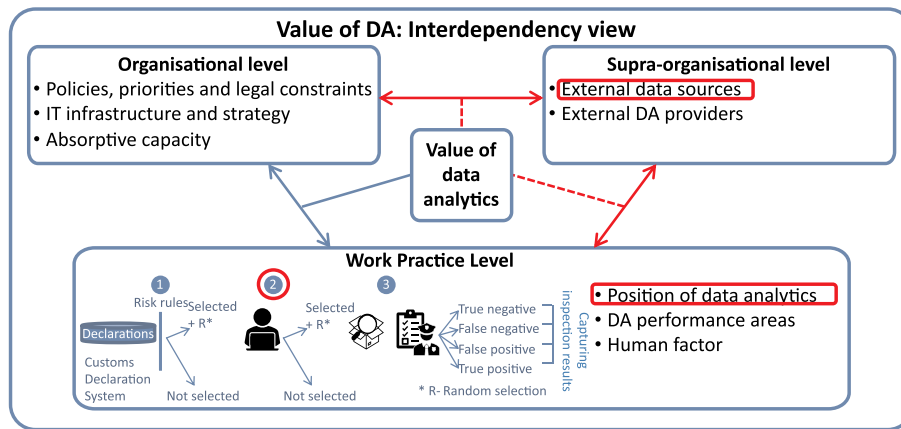


Fig. 6. An example of interdependency with the supra-organisational and the organisational levels.

Programming Interface (API). These eCommerce platforms have defined terms and conditions, setting out how organisations can access this information via APIs. There are specific conditions for piloting which may be quite different when it comes to operational use. Therefore, negotiating access to the external data sources also triggers dependency between the supra-organisational and organisational level (see the arrow between supra-organisational and organisational levels in Fig. 6) for arranging terms for data access. Therefore, no matter how valuable the external data is for realising improvements at work practice level, these improvements would not be possible to realise unless the necessary access at organisational level is secured.

Another decision that customs organisations face is whether to develop DA capabilities internally or outsource them to external DA providers. In the cases several concerns were raised when it comes to external DA provider. First, developing DA for customs requires very specialised domain knowledge to know which data fields it makes sense to combine, and it may take a lot of time to transfer the domain knowledge to external companies. Second, the quality of the data in customs systems is not very high and as long as data quality is not improved it makes little sense to invest in complex analytics by very experienced DA experts. Third, if external DA providers develop more sophisticated analytics in the future, this may set new requirements for the existing IT infrastructure, which in turn may set new requirements for updating IT systems and the related investments. Finally, GDPR concerns are a serious barrier to engaging with external DA providers and constrain what data can be exchanged. Legal efforts would also be required on both sides to agree on how the data can be exchanged. Engaging with an external DA provider brings benefits, as such companies have developed extensive DA capabilities that can be beneficial for bring improvements at work practice level. There are nevertheless concerns as discussed above, and customs organisations need to be aware of the concerns and the efforts needed to manage their relationships with external DA providers.

### 6.3. Strategic process view

We will now expand our analysis by taking the strategic process view. The strategic process view allows us to take a process perspective and view the Living Lab developments in the historical context of DA developments in a specific customs organisation. Looking at Dutch Customs in retrospect, a very strategic DA project on which Dutch customs embarked in the past was ACXIS<sup>15</sup>. ACXIS was an EU-funded research project aiming among other things at developing DA solutions for the automated image interpretation of X-ray images of containers. In the ACXIS project, Dutch customs was part of a consortium, in which

other organisations were responsible for developing the algorithms. Let us look at the process view, and focus on the *loop pilot* → *implementation* at *time t1* (see Fig. 7a)<sup>16</sup>.

At *t1*, we will examine only the pilot stage from this loop (see Fig. 7a, under the loop in the process view, pilot is marked with a green shape). At *t1*, Dutch Customs was mostly interested in the *capability realisation processes* (right green arrow of the strategic process view in Fig. 7a) and not so much in the *capability building processes* (left dashed arrow in Fig. 7a). The *capability building* was done by another organisation that acted as an external DA provider (marked with dashed shape at supra-organisational level, Fig. 7a). During the pilot, positive results from the use of analytics were realised (marked with green in the process view under *Impact* in Fig. 7a). At *t2* (see Fig. 7b), when the R&D project ended and the pilot moved towards implementation (see the red shape under implementation in the loop of the strategic process view), the relationship with the external DA provider was dissolved (marked with red at supra-organisational level in Fig. 7b). The necessary capabilities were not covered any more (the capability building arrow in Fig. 7b is also marked with red as Dutch Customs did not have this capability covered internally). As a result, the capability realisation processes were also influenced negatively (also marked with red in the strategic process view in Fig. 7b). The impact from the algorithms started to disappear (marked with red under *Impact* in Fig. 7b). At *t3*, Dutch Customs did invest and internally developed the missing capabilities (the capability building processes are marked now with green in Fig. 7c), which allowed to regain control over the process.

This specific DA expertise has now been expanded, but there are remaining steps to be overcome<sup>17</sup>. Some important lessons have been learned from the ACXIS project. First of all, algorithms are not finished products that are developed once by an external provider and then serve their purpose. They need to be maintained and continuously re-trained with new data sets to remain relevant. Organisations either need to maintain a relationship with an organisation that has the capabilities to maintain the algorithms or to engage in capability building internally. This would enable them to maintain the algorithms and continue realising value. Second, image interpretation requires very different DA capabilities compared to what is now developed in the Dutch LL and the web data retrieval. Seen in this broader historical

<sup>16</sup> Showing the figures next to each other allows to trace high-levels effects over time. These effects are explained in the text.

<sup>17</sup> Getting image data onto the customs system, where declaration and inspection reports are available, remains a step that needs to be overcome. In addition, the image data, required for training and maintenance, needs to be annotated with data from the inspection reporting system, and/or data on the expected types of goods from the declaration system. This process needs to be done automatically, not manually, i.e., the scanning process needs to be related automatically to the declaration and inspection processes.

<sup>15</sup> <https://www.acxis.eu/>

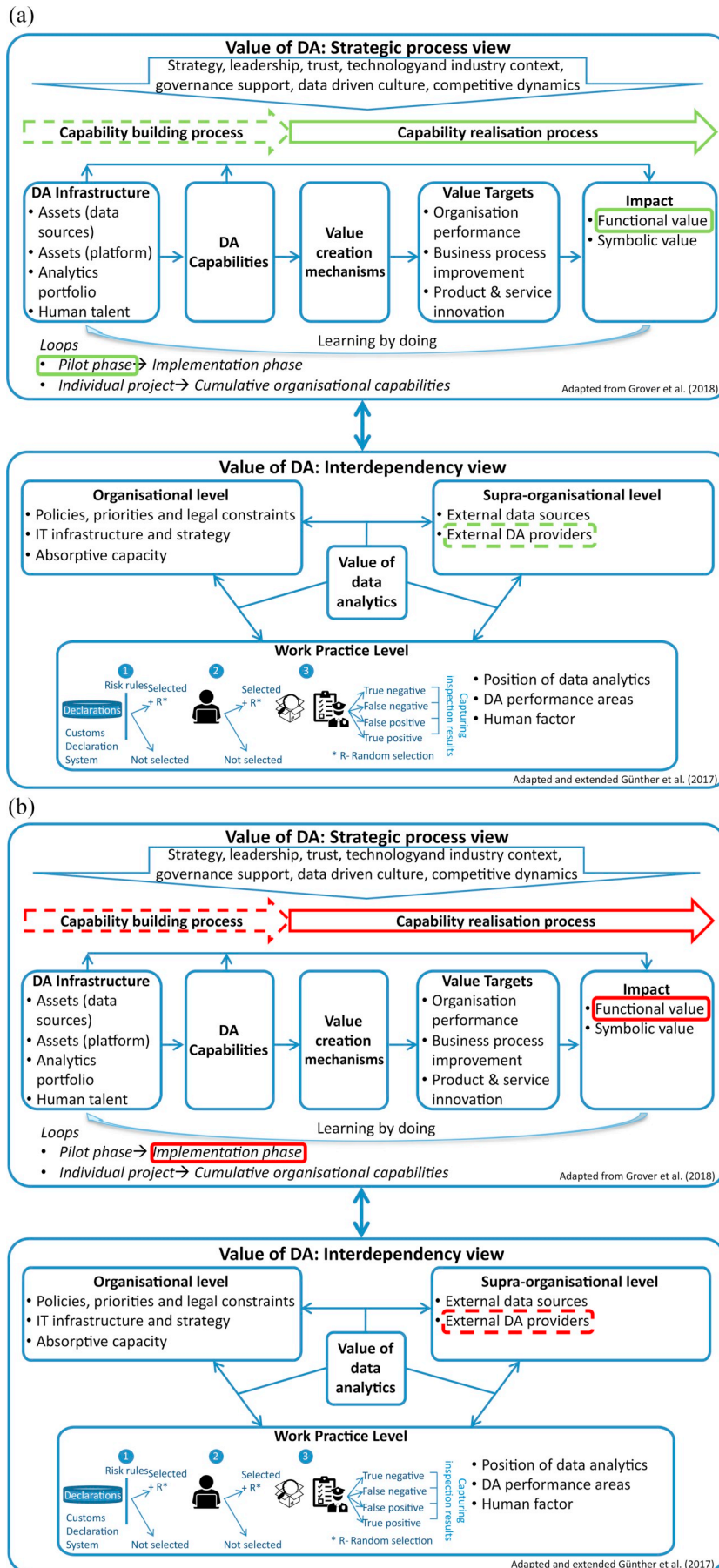


Fig. 7. a. Pilot stage at t1. b. Implementation stage at t2. c. Implementation stage at t3.

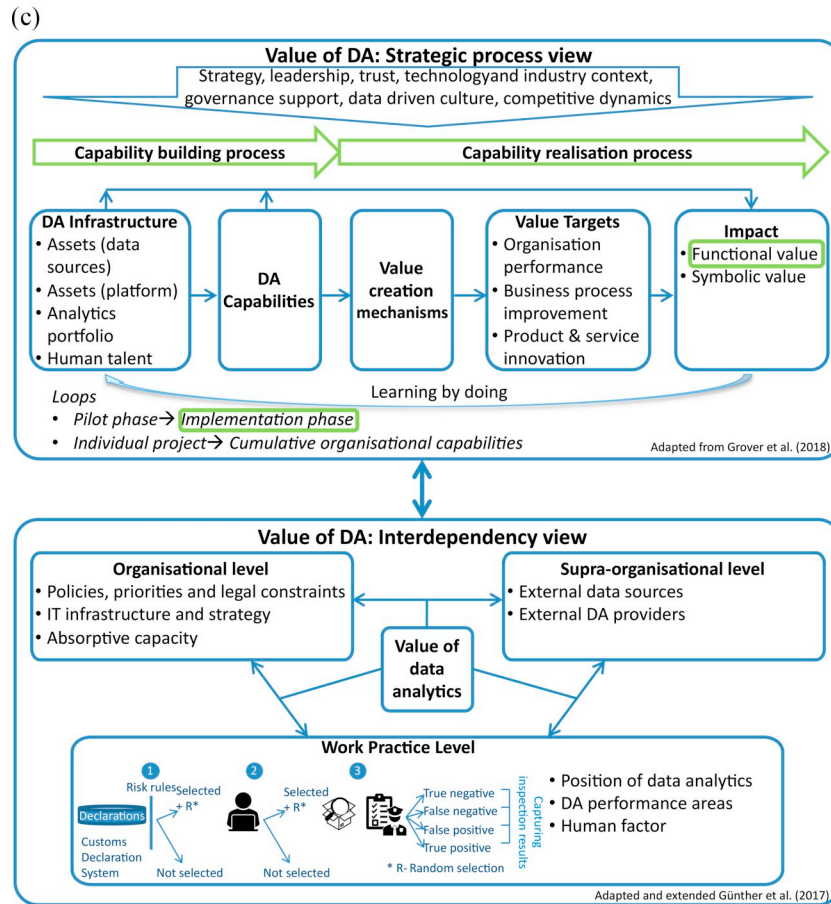


Fig. 7. (continued)

context, and taking the second loop from the process view into account (i.e., the loop from the individual projects → cumulative organisational capabilities) the Dutch LL can be seen as the next step in expanding the DA capabilities at Dutch customs with new capabilities related to web data retrieval. This web data retrieval DA also requires specific knowledge and skills and specific data sets that, in turn require specific skills and capabilities for accessing and processing data from the web. At the moment Dutch customs are doing that by using the external capabilities of external DA providers to explore the potential. In the future, they will need to make strategic decisions on whether to deploy the web data analytics developed in the Dutch LL and how to incorporate these capabilities in the broader portfolio of the DA capabilities of Dutch customs.

Looking at the Belgian case, a process perspective also helps us to understand the DA developments in the Belgian LL in the broader context of DA capability building and capability realisation processes over time. Reflecting on DA capability building processes, in a short period of time, Belgian customs invested in DA infrastructure by hiring human talent, investing in IT infrastructure by adjusting their risk assessment system to enable input from DA to be fed in the system in near operational setting – something which the old IT system was unable to accommodate, as it was not flexible enough. The Belgian LL can be seen as the next loop in the DA capability building process. In collaboration with external DA providers, the aim of the Belgian LL was first of all, to develop more sophisticated DA models to analyse traders' behaviour based on machine learning using historical data sets. Second, to develop analytics based on external data sources. These processes lead to accumulating new DA capabilities, and the internal DA experts are working together closely with the external DA providers. This pilot also enables access to, and experimentation with, new data assets (i.e., the external data sources secured through the project). Analysing the cumulative organisational capabilities over time allows

customs to see how capabilities are added to allow for further value realisation. This could allow customs to identify missing capabilities that can be further developed in the future.

#### 6.4. Collective capability building view

Looking at the strategic process view and at the cumulative organisational capabilities of Dutch Customs (marked with green in Fig. 8, Organisation B), Dutch Customs has developed capabilities related to image interpretation that can be used at Step 3 of the customs process. Via the Dutch LL, Dutch customs is developing new DA capabilities related to web data retrieval for Step 2 of the customs process (see the green circles around Step 2 and Step 3 in Fig. 8, Organisation B). At the same time Dutch Customs is also interested in the DA that the Belgian LL is developing for Step 1 of the customs risk assessment process (marked with green circles at Step 1 in the work practice level of Fig. 8, Organisation A). This analytics uses historical data and external data sources to analyse trader behaviour and improve the automatic risk assessment software.

Through the agreement to collaborate and collectively join forces in the DA developments Dutch Customs has the opportunity (marked with dashed circle around Step 1, Fig. 8 Organisation B) to gain access to new DA capabilities, i.e., the DA algorithms developed in the Belgian LL (see Fig. 8, the arrows from Organisation A to Organisation B via the collective view). In this way, Dutch Customs can add these to its analytics portfolio instead of having to develop the algorithms from scratch. This opens new opportunities to share efforts and costs in developing DA solutions and is an important scenario to consider when assessing costs and benefits for obtaining access to new DA capabilities. While undoubtedly running the algorithms in the Dutch environment may require adaptation, sharing the initial algorithms and experiences already holds the potential for saving development efforts. The area of

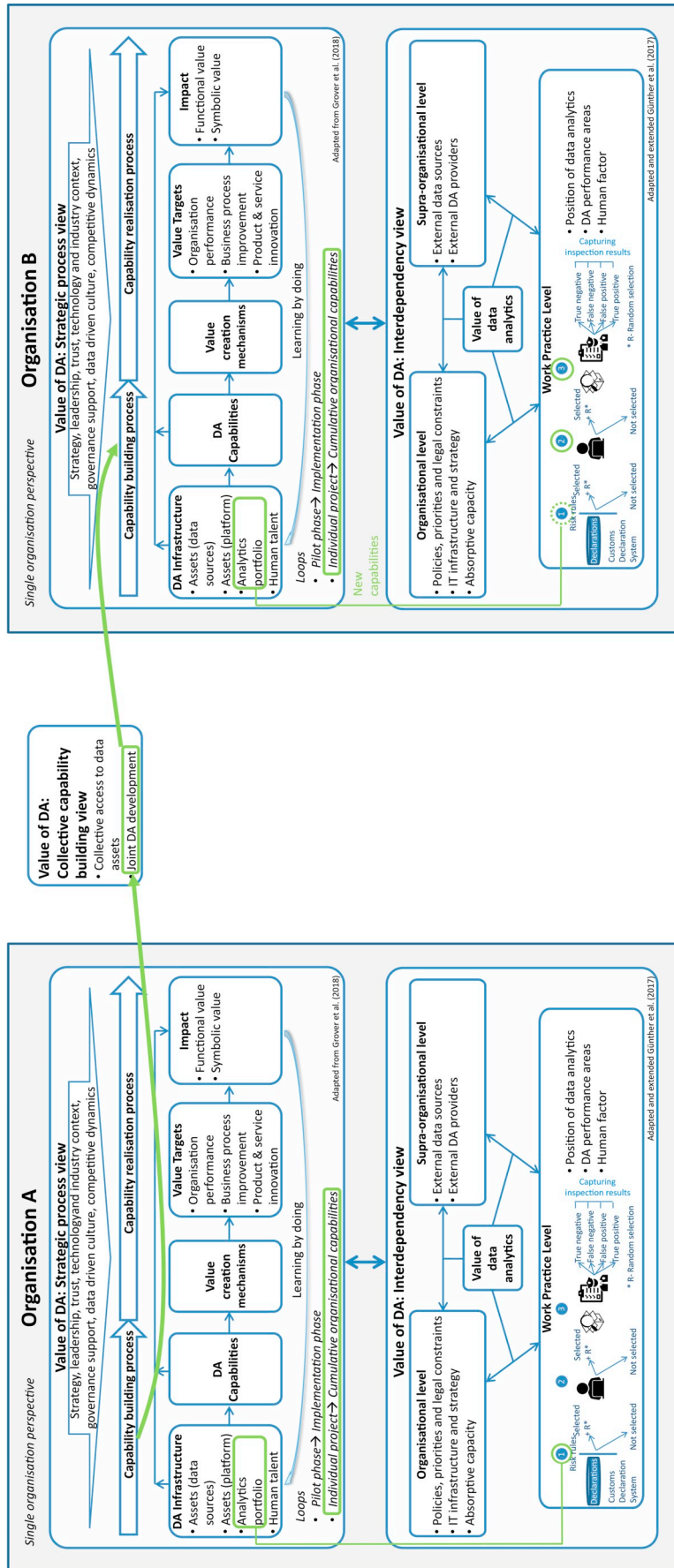


Fig. 8. Potential for Dutch Customs to obtain new capabilities via the Belgian I.L.

sharing algorithms is only one area where customs administrations can benefit from collective efforts. Other areas that we identified from the Living Labs include access to external data sources, as there are overlapping data sources that are interesting for customs administrations in Europe. Instead of every customs organisation individually approaching each of the external providers and negotiating access and possibly also prices, this access to data assets is also an area to approach collectively. It also includes data preparation, including data cleansing and making the data sets suitable for analytics use for customs. This is a resource-consuming task that can be also done collectively. Last but not least, there is a need for a technical infrastructure to allow for sharing data and algorithms, and this infrastructure can also be developed collectively. LL-4 that we examined is focussing on the collective aspect for developing such an infrastructure. Such collective infrastructures are also being developed at EU level. As our analysis demonstrates, by using the VDAGS framework, we can reveal and make explicit a large number of considerations and dependencies that impact how value form DA can be viewed and perceived. At the same time, the strategic process view and the collective view allow customs organisations to identify missing capabilities and explore new strategies for collective engagement to advance the value realisation from DA in their own organisations.

## 7. Discussion and conclusions

The main question that we set out to explore in this paper was: *How to identify the value of DA in a government supervision context, and what are barriers and trade-offs to be considered and overcome in order to realise this value?* To address the research question, building upon leading models from the IS domain and case studies from the customs domain, we developed our VDAGS framework for identifying the value of DA in the context of government supervision. This framework comprises three different views for understanding the value of DA, namely: (1) the *interdependency view*, which enables us to examine the value of DA by looking at the interdependencies between work practice levels and interactions with the organisational and supra-organisational levels; (2) the *strategic process view*, which enables us to examine value by taking a longitudinal perspective and tracing DA capability building and DA capability realisation processes over time and examining the learning loops; (3) the *collective capability building view*, which enables us to examine value by looking at collaboration with other organisations which may provide access to capabilities which would be difficult for a single organisation to realise alone. These views taken together allow reasoning about complex interdependencies and provide more profound insights compared to considering each of the views in isolation.

At this point, it is worth reflection on our view on the concept of the *value* of DA as a relative concept. By applying the VDAGS framework, we see that the value of DA is not static and something that can be clearly defined. There are many dependencies that influence the value of DA and how it is perceived. As we have illustrated with examples, what can be ultimately realised as improvement at work practice level is heavily dependent on developments at the other levels. These include issues such as organisational policies and priorities (influenced also by EU legislation), as well as limitations on inspection capacity. Next to that supra-organisational factors such as availability of external data sources and the organisational arrangements, or lack thereof to secure access to that data, also play a role. Our analysis also shows that value is time-dependent. Change in priorities such as a shift from fiscal priorities to a greater focus on safety and security can increase or decrease the immediate value of DA solutions that are tailored more towards one or the other. Furthermore, DA solutions that show high performance in the R&D phase do not necessarily retain their value in the upscaling phase, unless there is a clear plan of how the capabilities needed would be sustained. At the same time, in the long run, accumulating different capabilities internally, as well as capitalising on capabilities developed collectively, provides a broader range of options and choices for

customs to act in a dynamic context and have more possibilities to deploy analytics and realise benefits.

As Kim et al. (2014) point out, some challenges that businesses and governments face relate to choosing the right technology and having data analytics personnel. These are challenges that we also saw in our cases. But as Kim et al. identified the more acute challenge that governments face is that they must look to break down silos for data integration, implement regulations for security and compliance, and establish sufficient control towers" (Kim et al., 2014, p. 81). These are precisely the issues that customs also faces in our cases, where data that is useful for analytics is available often outside the organisation either with other customs organisations or with external data providers.

Furthermore, EU regulation and organisational policies put strong demands on the risk rules and the physical checks associated with these rules, which, in combination with the limited inspection capacity that customs has, influences the improvements that may be possible to achieve with analytics. Finally, for customs organisations, it is essential to link all these various data sources. This would allow customs to obtain an overarching view on the traders, on their transactions, on the types of goods they trade, to be able to better target the fraudulent flows, and facilitate the compliant flows. In this context, the VDAGS framework helps to structure and capture lessons learned and implications for practice (see also Annex C) and provides an analytical lens, which can be a useful tool also for other customs administrations that plan to embark on DA projects.

The VDAGS framework provides a rich conceptual basis for analysis and understanding. It is clear that Data Analytics experts are essential for developing analytics solutions. At the same time, many other parties need to understand the effects of data analytics and what is possible to achieve. These parties include the management, the data analytics experts, the IT department, the experts translating EU and national policies and legislation into requirements for risk rules, and the experts implementing risk rules into the system. Next to that other parties include the legislators and other governing bodies at EU level, which set the legal basis which could enable, or constrain, what is possible to achieve with DA in customs. The VDAGS framework can serve to elicit these different perspectives and promote a shared understanding of the opportunities and trade-offs, which is essential for choosing strategies for action.

From the point of view of research, by further developing and integrating two leading research models from the IS domain (Grover et al., 2018; Günther et al., 2017) and extending them with findings from the cases, we have developed the VDAGS framework, a novel framework for identifying the value of DA that is specific to the government supervision context. This study can be seen as an addition to the eGovernment literature which has explored value and data analytics in specific eGovernment contexts such as smart cities (Cronemberger & Gil-Garcia, 2019), social media data (Panagiotopoulos et al., 2017), and open government (Attard et al., 2017). Our specific contribution on the value of DA is to the domain of government supervision and, more specifically, the domain of customs.

A limitation from the point of view of the empirical context is that our study is focussed on the domain of customs. Some of the elements in the VDAGS framework, especially the operationalisation of the work practice level, are therefore very specific to the customs domain. Nevertheless, we found out that such a level of detail was useful for reasoning about the position of DA and further analysis. Therefore it may be useful for other domains to make the process where analytics is deployed explicit. Future research can examine the applicability of the VDAGS framework to other domains of government supervision, which may allow for enhancing the framework and making it more applicable in a broader range of government supervision contexts. Interesting research questions would include whether aspects of customs supervision can be generalised to other types of government supervision. For example, in most countries, customs supervision is carried out in close collaboration with food and product safety inspection agencies, and DA



can play a role there as well. In further research, therefore, the *VDAGS framework* can be further applied and extended to other domains of government supervision.

In our study, we also made specific choices of theoretical foundations. More specifically, we focussed on the IS domain, and highly-cited general research models that appeared in the basket of 8 IS journals. Posing these limitations helped us to search in a targeted way. By using this strategy, we were able to find suitable models that we further adapted for our purpose. Further research can look broader into other domains beyond the IS, as well as into a broad range of sources. This may reveal new aspects that may be relevant, which would allow us to further develop and enhance the *VDAGS framework*. In addition, as discussed in Section 4, in terms of the theory types of Gregor (2006), the *VDAGS framework* was developed to serve as a framework for analysis and understanding. Further research can focus on advancing the framework towards a theory for prediction or design and action. In our study, we identified that the collective DA capability building

processes and their link to the DA capability processes in a single organisation are very interesting but not yet well understood. Further research can focus on addressing this gap as well.

### 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. The authors would like to thank the PROFILE consortium partners for the fruitful collaboration and discussions on the topic in the broader project context, as well as the GIQ editor and the anonymous reviewers for the constructive comments and guidance during the review process. This is an extended paper of an earlier paper that was presented at the eGov'2019 conference (Rukanova et al., 2019).

### Annex A. Overview of key experts consulted

Table A1  
Overview of experts consulted from the Dutch LL (LL-1).

No	Experts LL-1	Role
1	Dutch Customs Expert	Secretary of the Innovation Coordination Group and Senior advisor Data & Analytics Project leader of the Dutch Living Lab
2	Dutch Customs Expert	Senior scientific staff member, Dutch Customs Laboratory
3	Dutch Customs Expert	Data scientist
4	Dutch Customs Expert	Data scientist
5	Dutch Customs Expert	Chair of the Coordination Group Innovation
6	Dutch Customs Expert	Expert risk targeting eCommerce
7	Dutch Customs Expert	Head of Trade Relations
8	Dutch Customs Expert	Manager eCommerce developments
9	Dutch Customs Expert	Data analytics expert
10	Dutch Customs Expert	Expert risk targeting eCommerce
11	IBM expert	Executive IT architect
12	IBM Expert	Web data retrieval expert
13	IBM Expert	Natural Language Processing expert
14	IBM Expert	Manager
15	IBM Expert	Data analytics expert

Table A2  
Overview of experts consulted from the Belgian LL (LL-2).

	Expert LL-2	Role
1	Belgian Customs Expert	Project leader in PROFILE of the Belgian Living Lab Attaché, risk management department
2	Belgian Customs Expert	Attaché/Data miner, Department for risk analysis and data mining
3	Belgian Customs Expert	Attaché/Data miner, Department for risk analysis and data mining
4	Belgian Customs Expert	Director, Department for risk analysis and data mining
5	Belgian Customs Expert	Legal expert, Belgian Customs Administration
6	IBM Expert	Executive IT architect
7	IBM Expert	Machine learning expert
8	TNO expert	Expert on data linking and data semantics
9	Various TNO experts	Various TNO experts which were involved on a on-need bases for specific analytics.

Table A3  
Overview of experts consulted from the Sweden- Norway LL and the EU LL (LL-3 & LL-4).

	Expert LL-3 & LL-4	Role
1	Sweden-Norway LL	Key contact via the Swedish DA research partner with expertise on data analytics Interactions with different representatives of this LL during general project meetings.
2	EU LL	Key contact with the project manager of the EU Living Lab Interactions with Data Scientists involved in the EU LL during general project meetings.

Annex B. Overview of the short list of papers that resulted from the search of the basket of 8 IS journals (last updated 1-5-2020)

No	Title	Authors	Year	Journal	Pages	Citations	Observations	Single organisation/collaborative focus
1.	How the use of big data analytics affects value creation in supply chain management	Chen, D.Q., Preston, D.S., Swink, M.	2015	Journal of Management Information Systems	32(4), pp. 4–39	119	Specific, focus on big data usage	Single organisation
2.	Debating big data: A literature review on realising value from big data	Günther, W.A., Rezazade Mehrizi, M.H., Huysman, M., Feldberg, F.	2017	Journal of Strategic Information Systems	26(3), pp. 191–209	103	General; Research framework	Single organisation but acknowledges supra-organisational
3.	Creating Strategic Business Value from Big Data Analytics: A Research Framework	Grover, V., Chiang, R.H.L., Liang, T.-P., Zhang, D.	2018	Journal of Management Information Systems	35(2), pp. 388–423	64	General; Research framework	Single organisation
4.	How does business analytics contribute to business value?	Seddon, P.B., Constantinidis, D., Tamm, T., Dod, H.	2017	Information Systems Journal	27(3), pp. 237–269	41	General	Single organisation
5.	Advanced Customer Analytics: Strategic Value Through Integration of Relationship-Oriented Big Data	Kitchens, B., Dobolyi, D., Li, J., Abbasi, A.	2018	Journal of Management Information Systems	35(2), pp. 540–574	13	Specific, focusses on customer analytics	Single organisation
6.	Assessing value creation in digital innovation ecosystems: A Social Media Analytics approach	Suseno, Y., Laurell, C., Sick, N.	2018	Journal of Strategic Information Systems	27(4), pp. 335–349	9	Specific, analysis focusses on social media data	Acknowledges relationship to consumer and professional domain; focus on interactions but not on the collective/collaborative capability building process
7.	Reconceptualising synergy to explain the value of business analytics systems	Someh, I., Shanks, G., Davern, M.	2019	Journal of Information Technology	34(4), pp. 371–391	2	Specific, focus on relationship between business analytics systems and customer relationship management system	Acknowledges synergetic relationship business analytics system and customer relationship management system, no explicit focus on collaborative capability building processes

**Annex C. Summary of key lessons learned and implications for practice**

Views from the V-DAGS framework		Key concepts form the VDAGS framework		Elaboration on lessons learned and implications for practice	
Interdependency view	Work practice level	Position DA, DA performance areas and human factor	DA can be deployed in different places of the customs process; and different analytics methods can be used. Decisions, where to place the DA in the process, may have implications on the performance areas and in order for the effect of the analytics to be visible in practice, human experts form an essential part, as they need to trust the analytics and use these results in the next step of the socio-technical process.	Organisational policies (e.g., defining limits on a maximum number of inspection), priorities (focus on more fiscal or more security threats), and legal requirements imposed by other bodies such as the EU (e.g. obligatory checks that are required by law) may pose upper boundaries of what can be achieved with DA in practice.	Use of DA in customs can benefit from quick learning loops and from integrating and linking external government and business data sources, however, this requires investment in IT. It is, therefore, important to align the DA strategy with the organisational IT strategy. Organisationally customs administrations have different expert groups (risk rule experts and DA experts) that aim to bring improvement to the same risk assessment process. Close collaboration between these groups can allow for better synergies to benefit from the potential of DA.
	Organisational level	Policies, priorities and legal concerns			External data sources offer opportunities for customs to link data and use it to enhance their risk assessment process. Organisational efforts, however, are needed for engaging with external data providers for negotiating terms and conditions for accessing the data. Efforts are also required for linking the data, and learning where it provides value. Furthermore resources for developing the IT environment to be able to handle inputs from these data sources will also be necessary.
	Supra-organisational level	Implications for aligning DA strategy and IT strategy IT infrastructure and strategy	External data providers		Engaging with external DA providers may bring benefits for customs. Developing DA for customs, however, requires in-depth customs knowledge and knowledge of the customs processes. It is, therefore, critical that customs experts are actively involved. Furthermore, the issue of data quality in the customs systems needs to be further addressed in order to be able to perform more advanced analytics.
Strategic process view	Strategic process view	Pilot project → implementation phase	External DA providers		Algorithms are not finished products. While during the pilot stage algorithms that show improved performance may be developed, these algorithms need to be maintained to continue to perform well. If external DA capabilities are used in the R&D phase, it is essential that these capabilities are also secured during the implementation stage (either via external parties) or covered internally to continue to contribute to the value realisation processes.
Collective capability building view	Collective capability building view	Individual project → cumulative organisational capabilities Collective access to data assets Joint DA development			Mapping the capabilities developed in the individual projects to the customs risk assessment process and following the cumulative effects over time would allow customs to oversee these cumulative effects and identify missing capabilities and areas for further developments. By being aware of their own cumulative capabilities and areas for potential further development, customs organisations can engage in a targeted way in collective capability building activities with other customs organisations for pulling efforts and resources to develop these capabilities collectively. These collective efforts can be in various areas such as jointly securing access to external data sources, jointly developing a technical infrastructure to allow exchange of external data, jointly developing the necessary cleansing and transformation of the data to make it readily used for analytics, jointly developing algorithms. In the future these collective efforts can also be directed towards the legislative bodies like the EU to create more room for analytics.

## References

- Anshari, M., & Lim, S. A. (2017). E-government with big data enabled through smartphone for public services: Possibilities and challenges. *International Journal of Public Administration*, 40(13), 1143–1158.
- Attard, J., Orlandi, F., & Auer, S. (2017). Exploiting the value of data through data value networks. *Proceedings of the 10th international conference on theory and practice of electronic governance* (pp. 475–484). ACM.
- Bertot, J. C., Gorham, U., Jaeger, P. T., Sarin, L. C., & Choi, H. (2014). Big data, open government and e-government: Issues, policies and recommendations. *Information Policy*, 19, 5–16.
- Chatfield, A. T., & Reddick, C. G. (2018). Customer agility and responsiveness through big data analytics for public value creation: A case study of Houston 311 on-demand services. *Government Information Quarterly*, 35(2), 336–347.
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165–1188.
- Cronemberger, F., & Gil-Garcia, J. R. (2019). *Big data and analytics as strategies to generate public value in smart cities: Proposing an integrative framework Public Administration and Information Technology*. Vol. 35, 247–267.
- 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.
- EC (2016). VAT aspects of cross-border e-commerce - options for modernisation final report – Lot 1 economic analysis of VAT aspects of e-commerce. October 2015. [https://ec.europa.eu/taxation\\_customs/sites/taxation/files/vat\\_aspects\\_cross-border\\_e-commerce\\_final\\_report\\_lot1.pdf](https://ec.europa.eu/taxation_customs/sites/taxation/files/vat_aspects_cross-border_e-commerce_final_report_lot1.pdf).
- Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, 50(1), 25–32.
- Fichman, R. G., Dos Santos, B. L., & Zheng, Z. (2014). Digital innovation as a fundamental and powerful concept in the information systems curriculum. *Management Information Systems Quarterly*, 38(2), 329–353.
- Fraefel, M., Haller, S., & Gschwend, A. (2017). Big data in the public sector. Linking cities to sensors lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics). *Lecture Notes in Computer Science*, 10428, 276–286.
- Gil-Garcia, J. R. (2012). Towards a smart state? Inter-agency collaboration, information integration, and beyond. *Information Policy*, 17(3, 4), 269–280.
- Gregor, S. (2006). The nature of theory in information systems. *Management Information Systems Quarterly*, 30(3), 611–642.
- 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., Mehri, 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>.
- Hagen, L., Keller, T. E., Yerden, X., & Luna-Reyes, L. F. (2019). Open data visualizations and analytics as tools for policy-making. *Government Information Quarterly*, 36(4), Article 101387.
- Higgins, A., & Klein, S. (2011). Introduction to the living lab approach. In Y. H. Tan, N. Björn-Andersen, S. Klein, & B. Rukanova (Eds.). *Accelerating global supply chains with IT-innovation* (pp. 37–54). Berlin, Heidelberg: Springer.
- Janssen, M., Konopnicki, D., Jane, L., Snowdon, J. L., & Ojo, A. (2017). Driving public sector innovation using big and open linked data (BOLD). *Information Systems Frontiers*, 19, 189–195.
- Janssen, M., & van der Hoven, J. (2015). Big and open linked data (BOLD) in government: A challenge to transparency and privacy. *Government Information Quarterly*, 32, 363–368.
- Kim, G., Trimi, S., & Chung, J. (2014). Big-data applications in the government sector. *Communications of the ACM*, 57(3), 78–85.
- Klein, H. K., & Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Quarterly*, 23(1), 67–93.
- Krittika, Vishvakarma, N. K., Sharma, R. R. K., & Lai, K. K. (2017). Linking big data analytics to a few industrial applications: A conceptual review. *Journal of Information and Optimization Sciences*, 38(6), 803–812.
- Lnenicka, M., & Komarkova, J. (2018). Developing a government enterprise architecture framework to support the requirements of big and open linked data with the use of cloud computing. *International Journal of Information Management*, 46(2019), 124–141.
- Maciejewski, M. (2017). To do more, better, faster and more cheaply: Using big data in public administration. *International Review of Administrative Sciences*, 83(1S), 120–135.
- McBride, K., Aavik, G., Toots, M., Kalvet, T., & Krimmer, R. (2019). How does open government data driven co-creation occur? Six factors and a ‘perfect storm’; insights from Chicago’s food inspection forecasting model. *Government Information Quarterly*, 36(1), 88–97.
- Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, 2(1), 1–28.
- Panagiotopoulos, P., Bowen, F., & Brooker, P. (2017). The value of social media data: Integrating crowd capabilities in evidence-based policy. *Government Information Quarterly*, 34(4), 601–612.
- Pencheva, I., Esteve, M., & Mikhaylov, S. J. (2018). Big data and AI – A transformational shift for government: So, what next for research? *Public Policy and Administration*. <https://doi.org/10.1177/0952076718780537>.
- Rukanova, B., Huiden, R., & Tan, Y. H. (2017). Coordinated border management through digital trade infrastructures and trans-national government cooperation: The FloraHolland case. *Proceedings of eGov conference, lecture notes of computer science* (pp. 240–252). Springer.
- Rukanova, B., Tan, Y. H., Slegt, M., Molenhuis, M., van Rijnsoever, B., Plecko, K., Caglayan, B., & Shorten, G. (2019). In I. Lindgren, & et al. (Vol. Eds.), *Value of Big Data Analytics for Customs Supervision in e-Commerce*. 11685. *Value of Big Data Analytics for Customs Supervision in e-Commerce* (pp. 288–300). Springer, Cham: Electronic Government EGOV 2019. Lecture Notes in Computer Science.
- Russon, P. *Big Data Analytics.TDWI Best Practices Report, fourth quarter, 2011. (2011). Available on-line at: https://tdwi.org/research/2011/09/~media/TDWI/TDWI/Research/BPR/2011/TDWI\_BPRReport\_Q411\_Big\_Data\_Analytics\_Web/TDWI\_BPRReport\_Q411\_Big%20Data\_ExecSummary.ashx; last visited 11-5-2020.*
- Seddou, P. B., Constantinidis, D., Tamm, T., & Dod, H. (2017). How does business analytics contribute to business value? *Information Systems Journal*, 27(3), 237–269.
- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. *Journal of Business Research*, 17, 263–286.
- Susha, I., Janssen, M., & Verhulst, S. (2017). Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development. *Proceedings of the 50th Hawaii international conference on system science, IEEE Computer Society, Big Island, HI, January 4th-7th, 2017a* (pp. 2691–2700).
- Susha, I., Rukanova, B., Gil-Garcia, R. J., Tan, Y. H., & Gasco, M. (2019). Identifying mechanisms for achieving voluntary data sharing in cross-sector partnerships for public good. *Proceedings of the 20th Annual International Conference on Digital Government Research* (pp. 227–236).
- Tan, Y.-H., Björn-Andersen, N., Klein, S., & Rukanova, B. (2011). *Accelerating global supply chains with IT-innovation: ITAIDE tools and methods*. Springer Science & Business Media.
- Vydra, S., & Klievink, B. (2019). Techno-optimism and policy-pessimism in the public sector big data debate. *Government Information Quarterly*, 36(4), 101383.
- Walsham, G. (1993). *Interpreting information systems in organizations*. New York, NY: John Wiley & Sons, Inc.
- Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, 176, 98–110.
- Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Newbury Park, CA: Sage.
- Zarsky, T. Z. (2016). Incompatible: The GDPR in the age of big data. *Seton Hall Law Review*, 47, 995.
- Boriana Rukanova** is a researcher at Delft University of Technology. She has been working on a series of EU-funded innovation projects (ITAIDE, CORE, PROFILE, PEN-CP) in the area of international trade and eCustoms. Her research interests include value of data analytics for government supervision, business-government information sharing to create public value, initiation and upscaling of digital trade infrastructure innovations, public-private collective innovations. Her research work appears in leading international journals, including *Government Information Quarterly*, *European Journal on Information Systems*, *Electronic Markets*, *Transforming Government: People, Processes and Policies*, *World Customs Journal*, *Supply Chain Forum*, *Complex Systems Informatics and Modeling Quarterly*, book chapters and proceedings of international conferences in the area of eGovernment and Information Systems.
- Yao-Hua Tan** is full professor of Information and Communication Technology at the Delft University of Technology. He is program director of the Master programme in Customs and Supply Chain Compliance at the Rotterdam School of Management. He was also Reynolds visiting professor at Wharton Business School. He has published five books and over 220 conference papers and journal articles. He was coordinator and scientific director for the EU funded ITAIDE (2006–2010), CASSANDRA (2010–2014) and CORE (2014–2018) projects. He was also the vice-chair (2012–2014) of the Committee on Trade of the Trade Division at the United Nations Economic Commission for Europe. He also regularly acts as an expert for various Dutch government committees and the European Commission.
- Micha Slegt** completed his bachelor in Organic Chemistry in 1996. At DSM Research and AKZONOBEL he then contributed to research in the field of polymer chemistry and pharmaceutical production. His PhD study in organic photochemistry started in 2001 and was completed in 2006 at Leiden University. In 2005 Micha joined the laboratory unit of the Customs administration of the Netherlands. First as the responsible chemists for the polymer, paper, ink and paint products, later, in 2009, to study detection and selection technology for Customs applications. In this capacity he has contributed to the FP7 projects CRIM-TRACK, ACXIS, CORE and C-BORD is contributing to the Horizon 2020 project COSMIC, PEN-CP and BORDERSENS. In 2013 he was awarded the World Customs Organization Diplôme d’Honneur for his innovative work. Representing his administration, Micha is a member of the EU Customs Detection Technology Project Group of EC/DG TAXUD.
- Marcel Molenhuis** is a Senior Advisor Data & Analytics at Dutch Customs and Secretary of the (national) Coordination Group for Innovation. Marcel graduated in 2015 at TU-Delft as Master Customs and supply chain compliance. He is currently the Work Package Leader for the Dutch Living Lab of the PROFILE EU Project.
- Ben van Rijnsoever** is member of IBM’s Global Centre of Competence for Public Sector as technical subject matter expert for Customs, Immigration and Border Management. IBM, being a major and global player in this industry, has organized its resources in sectors, where the Public Sector is an important sector. Customs, Immigration and Border Management is an important part of this industry, with projects and references all over

US, AP, and EMEA. Customs is at the heart of the IBM Footprint. In his role as technical subject matter expert, he works worldwide with local IBM teams to bridge knowledge and expertise across those teams. He manages the portfolio of eCustoms Suite applications, including a Customs Declaration Management System and Advanced Cargo Analytics for risk management. He was the lead architect of IBM's Supply Chain Visibility Platform, which is the predecessor of the Global Trade Digitisation offering and is involved in EU Customs research activities.

**Jonathan Migeotte** is a national and international Coordinator for the Risk Management department of Belgian customs since 2017. As such he is the project leader for Belgian customs in a number of European innovation and research projects such as PROFILE, PEN-CP and SSTL. At the national level, he actively participated in the development of the new selection engine of Belgian customs, SEDA 2.0. Before that, he 2.0 worked as a targeting officer at Liège airport from 2013 to 2016, allowing him to build operational and strategic knowledge in the field of targeting, E-commerce, and express couriers.

**Mathieu L.M. Labare** is a data-miner at the Risk management & datamining department of the Belgian Customs. Before joining the SPF Finances in 2018, he worked for more than 12 years as a researcher in astro-particle physics for two international neutrino physics experiments. He got his PhD in experimental physics at the Université Libre de Bruxelles (2006-2010), developing original statistical methods to optimize the detection of extreme low signals, whilst controlling simultaneously the false discovery rate of an observation and naturally accounting for the trial factor inherent in multiple testing. As a postdoctoral researcher at the Vrije Universiteit Brussel (2010–2013) and at the Ghent University (2013–2018), he consolidates his expertise in Bayesian statistics, as well as in programming and database management. He was also lecturing "Statistics and Data Analysis" at the bachelor students in Physics and Astronomy of the Ghent University.

**Krunoslav Plecko** is a Solution Architect in IBM with Global Cloud Acceleration team. He has 12 years of experience as a System and Software engineer, DevOps tech lead, Applied Researcher in large multinational environments. His main interest is automation, cloud computing and development. Krunoslav received a Masters degree in Computer science with major in Data Science from University College Dublin.

**Bora Caglayan** is a data scientist and applied researcher at IBM Ireland Innovation Exchange. He has 10 years of experience in machine learning and software engineering.

Bora received his B.Sc., M.Sc., PhD degrees in Mechanical Engineering, Software Engineering and Computer Science respectively from Bogazici University. His research interests include empirical software engineering, recommender systems and applications of AI.

**Gavin Shorten** has a PhD in Biomedical Engineering from Trinity College Dublin and a Bachelors Degree in Electronic Engineering from Trinity College Dublin and the Technical University of Dublin respectively. During his academic endeavours Gavin published extensively on the design of embedded systems, analogue & digital signal processing and the application of analytics technologies to biomedical use cases. He published extensively in well-reputed international conference and journals. After graduating from academia in 2011 Gavin joined a start-up company as a Configuration and Release Engineer in Intune Networks. Gavin joined IBM Ireland in April 2013 to work in the Ireland Lab Central Business Services team and was a founding partner in the creation of an Ireland Lab R&D mission named the Innovation Exchange (IX). As part of the IX Gavin has fulfilled several roles including as a Program, Business Development and People manager with overall responsibility for direction and management of the IX. During this time IBM Ireland became the most successful Irish based industry participant in the ICT and Secure Societies strands of the Horizon 2020 framework. Most recently he has become the Cloud & AI Foundry manager which is a new initiative to grow commercial and monetary value from the IBM Ireland Labs R&D activities. Gavin continues to have a keen interest in the technical aspects of R&D publishing and patenting occasionally.

**Otis van der Meij** completed his BSc degree in Systems Engineering, Policy and Management at Delft University of Technology in 2019 after which he moved to Barcelona for his MSc degree in Business Analytics at Universitat Pompeu Fabra. In 2019 Otis worked also as a research assistant for the PROFILE EU-funded research project.

**Suzanne Post**, a fourth year student, received her bachelors with honors in Technische Bestuurskunde in the summer of 2019 at the faculty of Technology, Policy and Management. Currently she is working as a research assistant for the PROFILE EU-funded research project. Next semester she will be president of the student Introduction Board whom organize the annual introduction programme for new bachelor students of TU delft.