

## Value of Big Data Analytics for Customs Supervision in e-Commerce

Rukanova, Boriana; Tan, Yao-hua; Slegt, Micha; Molenhuis, Marcel; Van Rijnsoever, Ben; Plecko, Krunoslav; Caglayan, Bora; Shorten, Gavin

**DOI**

[10.1007/978-3-030-27325-5\\_22](https://doi.org/10.1007/978-3-030-27325-5_22)

**Publication date**

2019

**Document Version**

Final published version

**Published in**

Electronic Government - 18th IFIP WG 8.5 International Conference, EGOV 2019, Proceedings

**Citation (APA)**

Rukanova, B., Tan, Y., Slegt, M., Molenhuis, M., Van Rijnsoever, B., Plecko, K., Caglayan, B., & Shorten, G. (2019). Value of Big Data Analytics for Customs Supervision in e-Commerce. In I. Lindgren, M. Janssen, H. Lee, A. Polini, M. P. R. Bolívar, H. J. Scholl, & E. Tambouris (Eds.), *Electronic Government - 18th IFIP WG 8.5 International Conference, EGOV 2019, Proceedings* (pp. 288-300). (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 11685 LNCS). Springer. [https://doi.org/10.1007/978-3-030-27325-5\\_22](https://doi.org/10.1007/978-3-030-27325-5_22)

**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.

***Green Open Access added to TU Delft Institutional Repository***

***'You share, we take care!' - Taverne project***

**<https://www.openaccess.nl/en/you-share-we-take-care>**

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.



# Value of Big Data Analytics for Customs Supervision in e-Commerce

Boriana Rukanova<sup>1</sup>(✉), Yao-Hua Tan<sup>1</sup>, Micha Slegt<sup>2</sup>,  
Marcel Molenhuis<sup>3</sup>, Ben van Rijnsoever<sup>4</sup>, Krunoslav Plecko<sup>5</sup>,  
Bora Caglayan<sup>5</sup>, and Gavin Shorten<sup>5</sup>

<sup>1</sup> Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands  
{b. d. rukanova, y. tan}@tudelft.nl

<sup>2</sup> Customs Administration of the Netherlands, Amsterdam, The Netherlands

<sup>3</sup> Customs Administration of the Netherlands, Rotterdam, The Netherlands

<sup>4</sup> IBM Nederland B.V., Amsterdam, The Netherlands

<sup>5</sup> IBM Ireland, Dublin, Ireland

**Abstract.** Big data and analytics have received a lot of attention in e-government research over the last decade and practitioners and researchers are looking into the transformative power of this technology to create for example competitive advantage, and increase transparency. Recent research points out that while parties are aware of the transformative power of this technology, understanding the value that this technology can bring for their specific organizations still remains a challenge. Data analytics is in particular interesting to support supervision tasks of governments. Here we take the customs supervision as a typical example where data analytics is used to support government in its supervision role. The main question addressed in this paper is: *How to understand the value of big data analytics for government supervision?* To address this question this research builds upon a case study where big data analytics solutions are developed and piloted as part of the PROFILE EU-funded research project. We adapt and utilize a recently published integrated model of big data value realization of Günther et al. [5] as a conceptual lens to structure the case findings. As a result we develop a more detailed model for analyzing value of big data specifically in the context of customs as an example of a specific domain of government supervision. This research contributes to the eGovernment literature on articulating value from big data analytics, particularly focusing on the role of government supervision.

**Keywords:** Data analytics · Big data · Value · Customs · Government supervision · Web data retrieval · e-Commerce

## 1 Introduction

Big data and data analytics have received a lot of attention over the last decade and practitioners and researchers are looking into the transformative power of this technology to create for example competitive advantage or increase transparency. Big data has been considered to be a breakthrough technological development which brings opportunities, as well as challenges (e.g. [2, 3, 5, 12]). Big data is seen as “massive

amount of digital data being collected from all sorts of sources, is too large, raw, or unstructured for analysis through conventional database techniques” [8, p. 78] and analytics can help to generate new insights.

Although the business sector has been leading in using big data governments are also actively exploring the opportunity to use big data to address public sector challenges [2, 8]. A growing body of eGovernment research on big data focuses on Big Open and Linked Government Data [1, 6, 7, 9]. Often in these studies the focus is on government providing public access to their data to allow for applying further analytics; for example to increase transparency, or to extract further value from government data. In this context the government is more in a role of a data provider, opening up their data for wider use. In other cases, government can be seen as a user of information from business and non-government organization in order to create public value [4, 13]. Prior research [10] distinguishes three roles that a government can play which call for different use of big data and analytics in government. These roles are (a) *public supervision*, which deals with identification of irregularities (e.g. legal incompliance) and the respective responsive action; (b) *public regulation*, which focuses on regulating social activities and relations by using e.g. permits, prohibitions or orders; (c) *public service delivery* which focusses on providing services or products. In this study we focus on examining the use of big data and analytics by government in its supervision role where government is a user of big data to improve the supervision process.

As argued by some researchers [8] parties recognize that being able to create value from big data “represents a new form of competitive advantage” [8, p. 85]. However in a recent study [5] it is concluded that although big data has been considered to be a breakthrough technological development there is still limited understanding of how organizations translate its potential into actual social and economic value. The main question that we explore in this paper is related to understanding value of big data analytics for government in its supervision role or in other words: *How to understand the value of big data analytics for government supervision?*

To address this question we use the domain of customs, which is one example of a domain where government acts in its supervision role<sup>1</sup>. This research builds upon a case study in the customs domain where big data analytics solutions are developed and piloted as part of the PROFILE<sup>2</sup> EU-funded research project. The case studied in this paper concerns e-Commerce goods imported to The Netherlands. One of the issues with these streams is undervaluation, i.e. the declared value of the goods is lower than the actual value for which the goods are bought. Undervaluation is one of the examples of fiscal fraud. The effect of this undervaluation is loss of revenue from duties and taxes levied when goods are imported in the EU, translating in loss of revenue for the national and EU budgets. The pilot aims to investigate how contextual information could be collected, for example from e-Commerce websites in China and the US, and how this information can be used in the customs process to support targeting officers in their risk assessment on undervaluation. In our analysis we adapt the model of big data

<sup>1</sup> Other domains where governments act in their supervision role include areas such as tax, and food safety.

<sup>2</sup> <https://www.profile-project.eu/>.

value realization of Günther et al. [5] and use it as a conceptual lens to structure the case findings. Although the actual case in this paper is undervaluation, the societal value of this data analytics innovation is much broader. The same customs data analytics solutions that work for identifying undervaluation can also be used to identify unsafe and dangerous goods, such as counterfeited medicines that do not cure. Hence, the acute need for customs to develop data analytics solutions are a significant contribution to making our society a safer and more secure world.

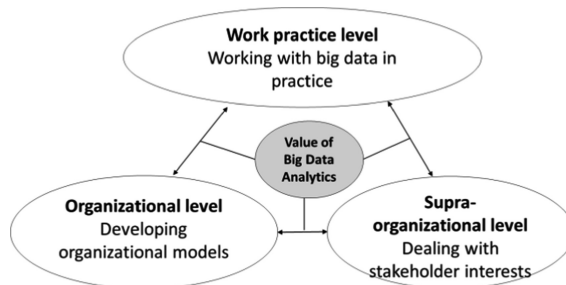
The remaining part of this paper is structured as follows. In Sect. 2 we present our conceptual framework, followed by our interpretative case methodology in Sect. 3. Section 4 presents our case findings and we end the paper with discussion and conclusions.

## 2 Conceptual Framework

As argued in literature [8] parties recognize that being able to create value from big data “represents a new form of competitive advantage” [8, p. 85]. However in a recent study, based on a thorough literature review investigating the issue of value realization from big data, Günther et al. [5] conclude that although big data has been considered to be a breakthrough technological development there is still limited understanding of how organizations translate its potential into actual social and economic value. To address this issue, Günther et al. [5, p. 202] propose that it is imperative for organizations to “continuously realign work practices, organizational models, and external stakeholders interest to realize value from big data”. The authors [5] formulated a number of propositions and propose an *Integrated model of big data value realization*. In their model Günther et al. [5] position *social and economic value of big data* in the middle and they propose that in order to address the value of big data, parties need to look at the interrelationships among three levels as follows: (a) *work practice level*; i.e. working with big data analytics in practice; (b) *organizational level*; i.e. developing organizational models; (c) *supra-organizational level*; i.e. dealing with stakeholders’ interests. As directions for further research Günther et al. [5] call for further empirical research to examine the cross-level interactions and alignments, and that is what we do in this paper.

Before continuing further it is worth elaborating on the concept of value. Günther et al. [5] support the argument that the perception of value of big data for organizations depends on the strategic goals of the organizations for using big data. Günther et al. [5] give various examples of areas of social and economic value that organizations may pursue. Examples of social value include enhanced transparency, prevention of fraud, improved security, improved wellbeing through better healthcare and education. Examples of economic value include increase in profit, business growth, competitive advantage. In the case that we analyze later in this paper the focus is on improvements in fiscal fraud detection due to under valuation of the value of the imported goods from e-Commerce transactions. In the wider societal sense reduced fiscal fraud also means that more duties are collected to finance the EU and the national budgets which can then be distributed for services to citizens.

We consider the model of Günther et al. [5] as a useful conceptual model that could be instrumental also for understanding of value of data analytics in the context of government supervision as it positions value at the center and it brings a broader perspective on examining value by looking at the interactions among *work practice*, *organizational*, and *supra-organizational levels*. For the purpose of our analysis we adapted a simplified version of the model of Günther et al. [5] in order to keep the analysis manageable when applying it to the empirical case<sup>3</sup>.



**Fig. 1.** Adapted model form Günther et al. [5]

In our adaptation (see Fig. 1) we chose to stick to the main elements of the Günther et al. [5] model, namely positioning the concept of *value* at the center and looking at the inter-relationships among *work practice*, *organizational* and *supra-organizational levels*. In our case we use the term *value* to refer both to *social* and *economic value* and we chose to use the terminology *Value of Big Data Analytics* as big data on its own is of little value and becomes valuable when analyzed or combined and analyzed to produce some new insights.

### 3 Method

To address our main question we used an interpretative case study method [14]. Data collection took place in the period June 2018–January 2019. The data collection took part in the context of the PROFILE EU-funded research project, where the goal is through the use of four real-life demonstration projects (called Living Labs) conducted in different EU countries to explore the potential of data analytics for customs. The Living Labs research approach “takes a development view of innovation and studies novel technologies in complex real-world setting” [15, p.32]<sup>4</sup>. This study focusses on

<sup>3</sup> In order to keep the model manageable we also did not include explicitly portability and interconnectivity which are part of the original Günther et al. [5] model but they are implicitly taken into considerations.

<sup>4</sup> For earlier application of the Living Lab approach in the context of IT innovation in international trade and customs see e.g. [11]. See also <http://tiny.cc/8s564y> for other examples of applications of Living Lab approach.

the Dutch Living Lab, where Dutch Customs, IBM and the university partners are working together to develop and test data analytic innovations for the import of e-Commerce goods to The Netherlands. The aim of the Dutch Living Lab is to explore whether data analytics solutions can help customs to detect fiscal risks related to undervaluation. The data analytics solution that is used in the Dutch Living Lab is developed by IBM and is a web retrieval tool using a contextualization engine, a piece of software that can search on e-commerce websites in China and the US, and analyze for example price information of a product. The purpose of this web retrieval tool is to provide Customs officers with additional online publicly available information to cross validate the price that is declared on the import declaration.

While our empirical insights were predominantly gained from our interactions with the Dutch Living Lab they were further informed by our broader involvement in the PROFILE project, where more data analytics pilots are conducted. As a result of these iterations we identified an initial list of issues and considerations related to the value of data analytics in the customs process. As a next step we used the adapted model (see Fig. 1) based on Günther et al. [5] as a lens to structure our findings. In this process we arrived at a number of observations as follows:

1. We needed a further detailing of the *work practice level* (the customs process), as the value of data analytics can vary depending on where in the process data analytics is used and in what part of the process performance improvement is desired.
2. We needed further detailing on the organizational level to better capture issues related to value of data analytics that we identified. In this further detailing we added (a) outsourcing to other data analytics (DA) providers, (b) existing IT systems, (c) priorities, policies, capacity and other legal constraints.
3. We needed further detailing of the supra-organizational level by looking at (a) external data providers, (b) other customs; and supra-national (such as the EU).

Based on these findings we extended the adapted model (Fig. 1) and we arrived at our more detailed model for analyzing value of big data analytics in the customs process (see Fig. 2 below).

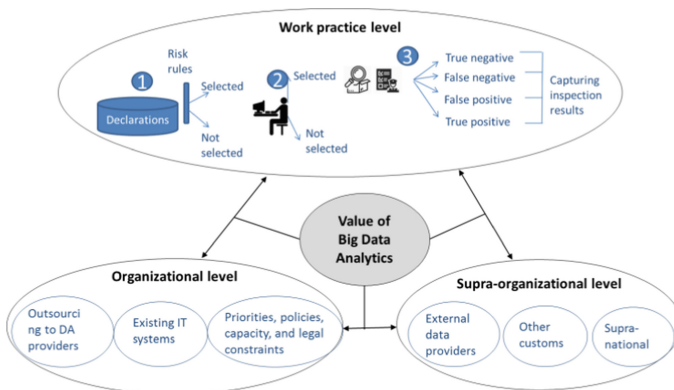


Fig. 2. A model for analyzing value of big data analytics in the customs supervision process

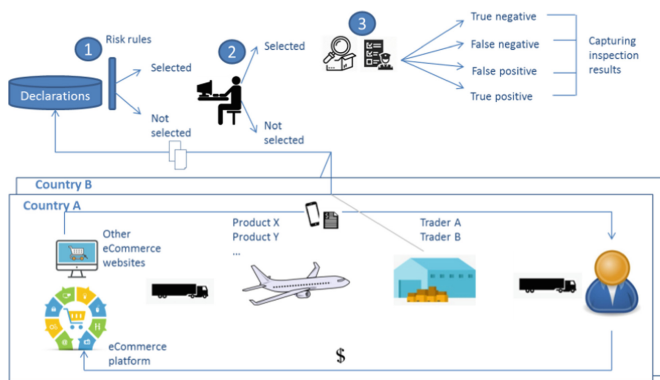
## 4 Case Findings

In this section we start by discussing the *Work practice level* by describing the e-Commerce customs process in The Netherlands and by explaining the big data analytics use in the Dutch Living Lab. We then demonstrate how our models (i.e. Fig. 2) is applied by using examples from the case.

### 4.1 Work Practice Level: The e-Commerce Customs Process in the Netherlands

Figure 3 below captures on a high level the e-Commerce customs process in The Netherlands. The bottom part of the figure illustrates an e-Commerce process, where a customer buys a product on the Internet. This could be via an e-Commerce platform (e.g. Alibaba) or directly via a e-Commerce website. The buyer pays for the goods via the e-Commerce platform or directly to the e-Commerce website. The seller on its turns sends the products to the customer.

Before the goods are sent to the final customer an e-Commerce declaration needs to be submitted to the declaration system of Dutch Customs. The submission of the e-Customs declarations is normally not done by the seller or the buyer but by a logistics service provider or a trader that handles the declaration procedures when the goods enter The Netherlands.



**Fig. 3.** High-level overview of the e-Commerce context and the customs inspection process for e-Commerce declarations in The Netherlands

Once the declarations are available in the declaration system an automated risk analysis is performed based on pre-defined risk rules software (marked as step 1 in Fig. 3). As a result a list is generated which marks declarations that are suggested to be risky, and hence selected for inspection. As a next step (marked as step 2 in Fig. 3) a targeting officer further analyses the declarations and the list of declarations that were automatically selected for inspection from the system and makes a final list of packages to be inspected. One important aspect in this process is that in making the final



selection decision the targeting officer is limited in the number of declarations that can be selected for inspection. This maximum number is defined by customs policy. This maximum number is determined by the capacity limitations of the inspection team to carry out inspections. After the targeting officer makes the final choice the refined list of declarations that are selected is provided to the inspection team. Subsequently the inspection team inspects (marked as step 3 in Fig. 3) the packages and documents the inspection results. The goods that are inspected were either selected as they were suspicious or they were randomly selected. For the suspicious goods the outcome of the inspection can be either that something wrong was indeed found (hit), resulting in a true positive (TP) selection. In case nothing suspicious was found in goods that were considered suspicious we have a case of false positive (FP). One of the goals that would indicate improved targeting with use of data analytics is to find ways to reduce the false positive inspections or the cases where customs invests resources to inspect a package but at the end nothing was found. Reduction of false positive is important as it leads to unnecessary use of customs resources and delays the trade flows.

Inspections also take place based on random selection in order to check whether in the flow of goods that were not identified as suspicious by the risk rules there are also cases where something is wrong with the goods. The results of the inspections performed on the goods that are randomly selected can either be true negative (TN), meaning that the goods were not selected as suspicious and indeed they were not suspicious. The result however could be also false negative (FN), meaning that the goods were considered as not suspicious in the selection process but in reality something wrong was found during the inspection. The availability of false negatives means that goods which were suspicious were not identified in the selection process. Apart from reducing false positives selections, data analytics can be used to better reduce false negative cases which are not detected by the current system. In fact, an improved balance between false positive and false negative cases generated by the system may be achieved.

## 4.2 Data Analytics in the Customs Process

IBM developed a prototype that allows the user to input a description of any item from a customs declaration and the user is then presented with other potentially useful contextual information about this particular item. For example, in the first instance this could be information sourced from e-Commerce websites via API's but it is envisaged that this can be extended to include insights gained from previous declarations and inspections.

After inputting a search string, the user is presented with an analysis of other similar product descriptions related to the search query and a statistical analysis of attributes associated with that item (e.g. price) is performed on the data set that is retrieved from the e-commerce website via the API. The high-level architecture is presented in Fig. 4. The user interface allows the customs officer to be informed via a Customs Portal. The main components of the system highlighted in Fig. 4 are:

- Server-side (back-end) component which is retrieving item data (e.g., price, weight) from e-Commerce websites, performing analytical part and providing API (application programming interface).
- Client-side (front-end) component designed for user interaction and data visualization.

The proof-of-concept tool has a page with search bar functionality. Results are generated based on search term, visualized as box plot to display range of values together with some statistical metrics like min, max and median. Using the initial box plot results display, a user can choose to change the view to the tabular data format. Table view contains product description, price in currency retrieved from web data extraction component and additional column with price converted to Euro currency to show both values. In the pilot initial functionality was added where every row with product description is a hyperlink to the sources of information. The targeting officer can choose this option to get complete information about the product.

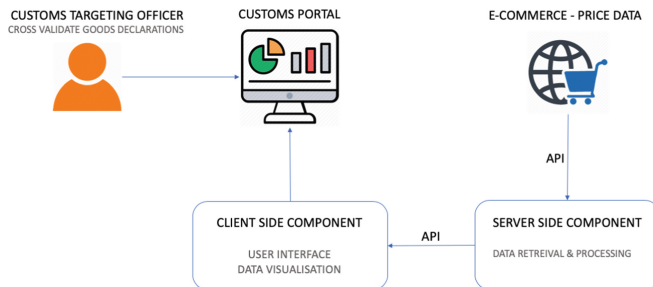
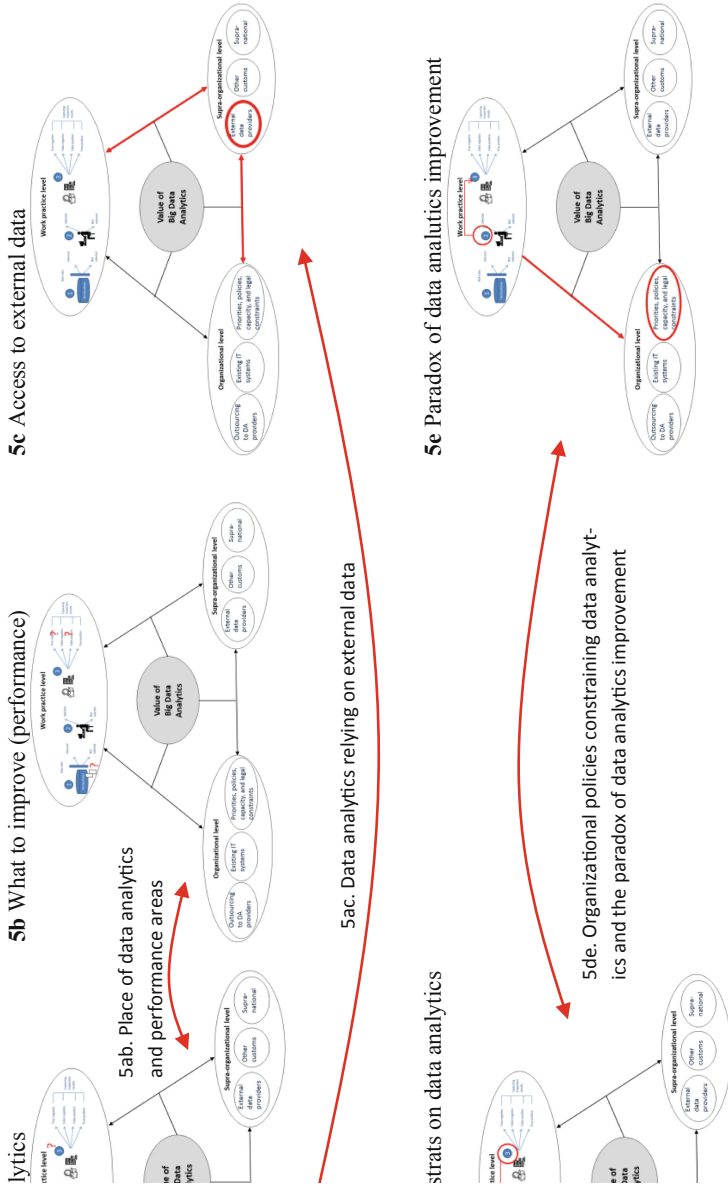


Fig. 4. High-level architecture of Customs Portal (dashboard)

### 4.3 Illustration of How the Model for Analyzing Value of Big Data Analytics in the Customs Supervision Process Can Be Applied

Figure 5 below uses a number of examples identified in our case to demonstrate how the model for analyzing value of big data analytics in the customs supervision process (Fig. 2) can be applied.

One of the decisions with which Dutch Customs was confronted was where in the customs process to place data analytics (see Fig. 5a). There are different possibilities where to include data analytics in the customs process, namely: (1) at the beginning of the customs process on the full set of customs declarations; (2) in the middle of the customs process, where pre-selection of declarations has already been made via running automated risk rules software and a sub-set of the declarations is presented to a targeting officer for further risk analysis. In this case data analytics can be used as a support tool for the targeting officer; (3) at the end of the customs process where data analytics (e.g. on scans) can be used to support the inspection process. In the Dutch Living Lab a decision was made to deploy data analytics in the second step, i.e. to provide decision support to the targeting officer. As such the immediate value of the data analytics solutions would be (on a work practice level) for the targeting officer.



**Fig. 5.** Examples of application of the model and dependencies between decision aspects

To that extent the tool can be considered as a decision support tool which has big data analytics as an important element. In this case big data analytics is required because large number of web pages on The Internet with unstructured data need to be searched and analyzed on price information.

Taking a process improvement perspective however Dutch Customs also needed to decide how to measure the performance improvement brought by the data analytics on the customs process (See Fig. 5b). In the discussions with Dutch Customs three areas for measuring performance improvements were identified, i.e.: (a) to reduce the number of *false positive* inspections (i.e. reduce the number of boxes which are selected and there is nothing wrong); (b) to reduce the number of *false negatives*, meaning catching more illicit trade; (c) to handle *large increase in volumes* of declarations. In the Dutch Living Lab we found out that there is an inter-dependence between the decision on where to place data analytics in customs process and the decision on what aspects to focus when measuring performance improvement (see arrow 5ab in Fig. 5). Namely the choice that was made in the Dutch Case (i.e. to use data analytics in the second step of the customs process as a decision support tool for the targeting officer) allows to measure performance effects on the reduction on false positives and possibly also effects on handling large volumes. In discussions with performance measurements experts from Dutch Customs however it became clear that placing the data analytics later in the process (in our case at step 2, as a support tool for the targeting officer) makes it difficult to realize data analytics improvements related to the reduction of false negatives (i.e. catching more illicit trade), as a large number of the declarations are already pre-filtered earlier in the process, before data analytics is applied.

When deciding where in the process to deploy data analytics (in our case in step 2 of the process) a next question was: What kind of data to use for data analytics? In the Dutch case it was decided to use price data obtained from e-Commerce websites to cross-validate the price on customs declarations. The choice of this type of data and possible value that it can bring on a work practice level however triggered issues related to data access (see Fig. 5c, and arrow 5ac) at the *supra-organizational* and *organizational level*. On a supra-organizational level in the case we found out that in many cases e-Commerce platforms would not allow robots to crawl their websites. An alternative way to access the price information from e-Commerce platforms was identified, namely via APIs and there are defined terms and conditions of these e-Commerce platforms of how organizations can access this information via APIs. This triggers issues at *organizational level*, as Customs needs to decide whether to accept the terms and conditions and under what circumstances. This shows that even if using data analytics and external data could bring value at a work practice level, for making this possible issues at *supra-organizational* and *organizational level* need to be solved. If these are not solved, even if potentially the data analytic solution could bring value to the work practice level, due to the dependencies with the other two levels this value may not be possible to realize in this specific customs situation.

Another issue that emerged from the discussions was that data analytics improvements are constrained by the existing organizational models and policies (see Fig. 5d). In the Dutch case the available resources in customs put an upper limit to the maximum packages that can be selected for inspection per day, as there are limited resources of customs officers who could actually inspect the packages.

This maximum number of inspections is therefore an upper bound and even if data analytics can make the targeting officers very effective in selecting only packages that lead to true positive and they may be able to identify a large number of packages where things are wrong, the effects that data analytics will have on the whole customs process are limited by the inspection capacity that is set as policy from the customs organization. This shows how policies on an *organizational level* put an upper bound of what is possible to achieve (and the value that this data analytics solution can bring) at *work practice level* in this specific organization.

Another interesting issue that we identified in the case (see also Fig. 5e) is related to what we call here the paradox of data analytics improvement. To take a simple example. Customs may be able to inspect 100 packages per day. With the current methods some of these packages will be false positive, others will be true positive. For the false positive packages the inspection officer spends limited time, as it is mostly opening the package and identifying that nothing is wrong. In case of true positives there is much more processing time per package, as the customs officer needs to start procedures and follow-up activities. If with data analytics the customs inspection process is improved, and the false positives are reduced, with the available resources customs will be able to carry out less inspections (i.e. less than 100), as it takes more time to process packages where there is something wrong. Thus increased efficiency due to data analytics can overload inspection capacities. The arrow 5de in Fig. 5 shows this complex interdependency, where the value of the data analytics can be bounded by the available organizational capacity but also vice versa, that improvements with data analytics could influence the organizational model and available capacity.

## 5 Discussion and Conclusions

The main question that this paper set to explore was: *How to understand the value of big data analytics for government supervision?* As Günther et al. [5] argue it is “imperative for organizations to continuously realign work practices, organizational models, and external stakeholders interest to realize value from big data” [5, p. 202]. Building on Günther et al. [5] and on a case study of using data analytics in the customs domain we developed a model for analyzing value of big data analytics in the customs supervision process and demonstrate how it can be applied. As our analysis demonstrates there is no simple answer on what is the value that data analytics can bring in a specific customs supervision process and benefits at one level have effects on other levels. It is the understanding of these complex interdependencies from multiple level perspective that allows us to reveal the multiple considerations and effects which give a more complete picture for decision makers and policy makers about what value of data analytics is for their specific organization. This study is limited to the customs domain and one case study. Further research looking at the use of data analytics in other cases from the customs domain as well as from other domains where government acts in its supervision role could help to determine the applicability of the model in other contexts and refine the model. This will allow for a more complete view on how to analyse value of data analytics in the context of government supervision. Further research can specifically focus also on elaborating the *organizational* and *supra-organizational level*

in more detail. Follow-up research can focus on elaborating elements from our model which we only highlighted here but which we were not able to go in detail (e.g. *outsourcing to other data analytics providers*). Another issue is that while many businesses companies do have highly skilled personnel to perform data analytics, many customs organizations do not have such in-house capabilities. Further research can examine what are successful organizational models where customs have successfully been able to incorporate such capabilities in their organizations. Next to that the supra-organizational level also deserves further attention, especially in the customs domain the links to other customs administration, as well as the link to the EU as a supra-national body would bring additional insights.

**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 also the PROFILE consortium partners for the fruitful collaboration and discussions on the topic in the broader project context.

## References

1. Bertot, J.C., Gorham, U., Jaeger, P.T., Sarin, L.C., Choi, H.: Big data, open government and e-government: issues, policies and recommendations. *Inf. Policy* **19**, 5–16 (2014)
2. Chen, H., Chiang, R.H.L., Storey, V.C.: Business intelligence and analytics: from big data to big impact. *MIS Q.* **36**(4), 1165–1188 (2012)
3. Fichman, R.G., Dos Santos, B.L., Zheng, Z.: Digital innovation as a fundamental and powerful concept in the information systems curriculum. *MIS Q.* **38**(2), 329–353 (2014)
4. Gil-Garcia, J.R.: Towards a smart State? Inter-agency collaboration, information integration, and beyond. *Inf. Policy* **17**(3, 4), 269–280 (2012)
5. Günther, W.A., Mehrizi, M.H.R., Huysman, M., Feldberg, F.: Debating big data: a literature review on realizing value from big data. *J. Strat. Inf. Syst.* **26**(3), 191–209 (2017). <https://doi.org/10.1016/j.jsis.2017.07.003>
6. Janssen, M., Konopnicki, D., Snowdon, J.L., Ojo, A.: Driving public sector innovation using big and open linked data (BOLD). *Inf. Syst. Front.* **19**, 189–195 (2017)
7. Janssen, M., van der Hoven, J.: Big and open linked data (BOLD) in government: a challenge to transparency and privacy. *Gov. Inf. Q.* **32**, 363–368 (2015)
8. Kim, G., Trimi, S., Chung, J.: Big-data applications in the government sector. *Commun. ACM* **57**(3), 78–85 (2014)
9. Lnenicka, M., Komarkova, J.: Developing a government enterprise architecture framework to support the requirements of big and open linked data with the use of cloud computing. *Int. J. Inf. Manag.* **46**, 124–141 (2019)
10. Maciejewski, M.: To do more, better, faster and more cheaply: using big data in public administration. *Int. Rev. Adm. Sci.* **83**(1S), 120–135 (2017)
11. Rukanova, B., et al.: Beer living lab – intelligent data sharing. In: Tan, Y.H., Björn-Andersen, N., Klein, S., Rukanova, B. (eds.) *Accelerating Global Supply Chains with IT-Innovation*, pp. 37–54. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-15669-4\\_3](https://doi.org/10.1007/978-3-642-15669-4_3)
12. Sivarajah, U., Kamal, M.M., Irani, Z., Weerakkody, V.: Critical analysis of Big Data challenges and analytical methods. *J. Bus. Res.* **17**, 263–286 (2017)

13. Susha, I., Jannsen, M., Verhulst, S.: Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: taxonomy development. In: Proceedings of the 50th Hawaii International Conference on System Science, Big Island, HI, January 4th–7th, 2017a, pp. 2691–2700. IEEE Computer Society (2017)
14. Walsham, G.: Interpreting Information Systems in Organizations. Wiley, New York (1993)
15. Higgins, A., Klein, S.: Introduction to the Living Lab Approach. In: Tan, Y.H., Björn-Andersen, N., Klein, S., Rukanova, B. (eds.) Accelerating Global Supply Chains with IT-Innovation, pp. 31–36. Springer, Heidelberg (2011). [https://doi.org/10.1007/978-3-642-15669-4\\_2](https://doi.org/10.1007/978-3-642-15669-4_2)