



Brussels, 13 March 2017.

**DATA ANALYSIS – WCO THEME OF 2017**

**Results of Secretariat's research on Big Data  
and how it can support risk management capabilities**

**(Item IV. (a) on the Agenda)**

**I. Introduction**

1. The joint PTC/Enforcement Committee meeting held in March 2016 discussed Big Data and its relevance to Customs administrations. The WCO Secretariat subsequently encourage WCO Members and other stakeholders to share their experiences with Information and Communication Technology (ICT) including Big Data.
2. In addition, the WCO Secretariat conducted a study that examines how Big Data can support Customs administrations. This work is closely related to the deliverables specified in the WCO's 'Digital Customs Work Plan' which links Big Data with a phased implementation and gradual consolidation of ICT systems in Customs.

**II. Big Data concept**

3. Unlike conventional data, Big Data encompasses a variety of data, ranging from transactional data for commercial use to everyday communications over social media. Big Data can be structured (prepared for a certain use) or unstructured (yet to be adjusted for use). Certain capabilities are thus required to incorporate such diverse data into certain Customs tasks.
4. When it comes to analysis, key enablers are advanced analytical techniques ('analytics') that may help to make sense of the latent risk associated with the subjects of control. With Big Data analytics, Customs administrations can in advance potentially carry out targeting exercises from a wider perspective. Such efforts correspond to the WCO's annual theme for 2017: "Data Analysis for Effective Border Management".

**III. Research paper**

5. To raise awareness to the potential of Big Data, the WCO Research Unit has prepared the research paper 'Implications of Big Data for Customs: How It Can Support Risk Management Capabilities' (appended as Annex to this document). The paper includes information on WCO Members' practices in this area, which is based on submissions by Canada, Hong Kong China, New Zealand, the United Kingdom and the United States.

**IV. Action expected from the PTC**

6. The PTC is invited to:
- take note of the outcome of WCO Secretariat's research on the applications of Big Data, including Member initiatives in this domain; and
  - encourage WCO Members to continue sharing their Big Data experiences.

\*  
\*       \*  
\*

## **Implications of Big Data for Customs How It Can Support Risk Management Capabilities**

### **I. Introduction**

The term 'Big Data' embraces a broad category of data or datasets that, in order to be fully exploited, require advanced technologies to be used in parallel. Many big data applications have the potential to optimize organizations' performance, including the optimal allocation of human or financial resources in a manner that maximizes outputs.

Nowadays businesses deal with a large amount of data collected through interactions with current or potential customers, in an effort to raise their operational efficiency and expand the market frontier. With respect to e-commerce, online retailers have continually updated and analysed the data concerning consumers' behaviour, trying to uncover latent consumption propensity – the information used to figure out what they should merchandise and how. Armed with these data, they have been able to assess risks in terms of consumer credit to minimize their potential loss in revenue. Widespread use of communication devices has significantly boosted the volume of transactions involving online sale of consumer goods, while simultaneously boosting the frequency and volume of 'B to C' data transmissions.

With the prevalence of electronic devices connected to the Internet (e.g. smartphones, tablets, smart TVs, wearables, and in-vehicle infotainment devices), big data exists almost everywhere today. Social media generates huge data every second, the only way to arrest this traffic would be if the countless users stopped delivering, sharing or posting everything from text messages to recorded videos over the Internet at once. The Internet of Things technology has incorporated many kinds of physical goods (e.g. home appliances, security cameras and garbage containers) into the big data applications. Where trade in goods is concerned, stakeholders such as manufactures, shippers and logistics operators have focused on ensuring that the vast array of data ranging from personal transaction history to the location of containerized goods can be put to practical use, with a view to providing quality service and enhancing the connectivity to be reflected in the supply chain.

The purpose of this paper is to discuss the implications of the aforementioned big data for Customs, particularly in terms of risk management. To ensure that better informed and smarter decisions are taken, some Customs administrations have already embarked on big data initiatives, leveraging the power of analytics, ensuring the quality of data (regarding cargos, shipments and conveyances), and widening the scope of data they could use for analytical purposes. This paper illustrates these initiatives based on the information shared by five Customs administrations: Canada Border Services Agency (CBSA); Customs and Excise Department, Hong Kong, China ('Hong Kong China Customs'); New Zealand Customs Service ('New Zealand Customs'); Her Majesty's Revenue and Customs (HMRC), the United Kingdom; and U.S. Customs and Border Protection (USCBP).

### **II. Understanding Big Data**

Big data entails huge datasets which are considered 'too big' to complete the necessary work within an acceptable waiting time by relying on traditional data management and processing models. Although such a description is often associated with the concept, there is no single agreed definition of 'Big Data'<sup>1</sup>. It is, however, interesting to note that the term was characterized by Gartner<sup>2</sup> as "high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation."<sup>3</sup> This explains in a concise manner several data-specific features, often referred to as the "3-Vs" of big data: volume, velocity, and variety.

Volume usually denotes the size and scale of individual datasets, while often referring to the aggregate amount of data on earth. One study demonstrated that in 2012, approximately 2.8 zettabytes (ZB) of data were created and replicated, projecting the volume to reach 44 ZB by 2020<sup>4</sup>. Velocity is considered to include both speed and frequency with which data can be created, updated and processed. Variety is somewhat synonymous with diversity in that data can be diverse in format, semantics, origin and medium. Indeed, big data can be obtained from a wide variety of sources, and it is said that at least eighty percent of corporate (or business-relevant) data are unstructured<sup>5</sup>.

### Box 1. The "3-Vs" of Big Data

#### Volume

- From 2005 to 2020, the digital universe will grow by a factor of 300, from 130 exabytes (EB) to 40,000 EB, or 40 trillion gigabytes (more than 5,200 gigabytes for every man, woman and child in 2020). Until 2020, the digital universe will about double every two years<sup>6</sup>.

International System of Units (SI)			International Electrotechnical Commission (IEC)-approved prefixes for binary multiples (for data processing and data transmission)			
kilo	k	10 <sup>3</sup>	kibi	Ki	2 <sup>10</sup>	= 1 024
mega	M	10 <sup>6</sup>	mebi	Mi	2 <sup>20</sup>	= 1 048 576
giga	G	10 <sup>9</sup>	gibi	Gi	2 <sup>30</sup>	= 1 073 741 824
tera	T	10 <sup>12</sup>	tebi	Ti	2 <sup>40</sup>	= 1 099 511 627 776
peta	P	10 <sup>15</sup>	peti	Pi	2 <sup>50</sup>	= 1 125 899 906 842 624
exa	E	10 <sup>18</sup>	exbi	Ei	2 <sup>60</sup>	= 1 152 921 504 606 846 976
zetta	Z	10 <sup>21</sup>	zebi	Zi	2 <sup>70</sup>	= 1 180 591 620 717 411 303 424
yotta	Y	10 <sup>24</sup>	yobi	Yi	2 <sup>80</sup>	= 1 208 925 819 614 629 174 706 176

- A major factor behind the expansion of the digital universe is the growth of machine-generated data, increasing from 11% of the digital universe in 2005 to over 40% in 2020<sup>7</sup>.
- A very large contributor to the ever-expanding digital universe is the Internet of Things with sensors all over the world in all devices creating data every second<sup>8</sup>.

#### Velocity

<sup>1</sup> Cleary (2017).

<sup>2</sup> An IT-related marketing company headquartered in Stamford, Connecticut, the United States.

<sup>3</sup> Gartner – 'IT Glossary', <http://www.gartner.com/it-glossary/big-data> (last visited 7 March, 2017).

<sup>4</sup> EMC Press Release, <https://www.emc.com/about/news/press/2012/20121211-01.htm> (last visited 7 March, 2017).

<sup>5</sup> Id.

<sup>6</sup> EMC (2012).

<sup>7</sup> Id.

<sup>8</sup> Rijmenam (2013).

- Every minute 100 hours of video is uploaded on YouTube. Every minute over 200 million emails are sent, around 20 million photos are viewed and 30,000 uploaded on Flickr, almost 300,000 tweets are sent and almost 2.5 million queries on Google are performed<sup>9</sup>.
- The essence is an increasing speed of data production and of performance that data-driven businesses need to benefit from their data. The challenge is not just to store streams of data but to transform fast-flowing data into a resource that fosters innovation and improves decision-making processes<sup>10</sup>.

### **Variety**

- In the past, all data were structured, neatly fitting in columns and rows. Nowadays, 90% of the data being generated by organisations are unstructured data<sup>11</sup> either not having any pre-defined data model or being organised in a pre-defined manner<sup>12</sup>.
- Being able to manage and extract insights from unstructured data is essential to effective big data deployment<sup>13</sup>.

In addition to the above, 'veracity' (as the forth "V") is often referenced as well. It consists of conventional data that can normally be obtained from certain familiar sources and prepared to meet the needs of users. Unlike these data, big data is generally incomplete or imperfect by itself, hence uncertain or error-prone<sup>14</sup>, and users are unlikely to suspect that their data at hand have been obtained from unreliable sources or already touched by unknown parties.

Big data can be subsumed to some extent under the Internet of Things – the networked connection of physical objects that are not computers in the classic sense. A variety of items, including household products and community facilities, have been equipped with information terminals and thus become a provider of raw or initial data, focused on their regular use, helping to facilitate certain business tasks, public or private. If such goods for everyday life were to represent themselves digitally, they would certainly add more value to the benefit and utility of their own, while transforming the physical world into one big information system<sup>15</sup>.

Big data raises a question of whether 'small data' exists and what it is. According to one legal scholar, "[a]lmost invariably, big data expressly or implicitly precludes human storage and processing capacity – if a human can comprehend the data without computing and algorithmic assistance, it is not big data." He adds, "[a]s a result, a small data world involves things that humans can create and grasp using human judgement alone"<sup>16</sup>. Big data, therefore, does fall within the scope of data from which no one can gain any insight without certain data management capabilities, including the instruments of analytics.

To sum up, big data is anything but voluminous data, not a solution by itself. Furthermore, big data is intrinsically worthless without the applications that surround it. It is not until certain tasks requiring informed decision are identified that big data applications are shaped and put into practical use.

---

<sup>9</sup> Id.

<sup>10</sup> Adamov (2015).

<sup>11</sup> Rijmenam (2013).

<sup>12</sup> Gerasimou (2016).

<sup>13</sup> Id.

<sup>14</sup> See AAAS-FBI-UNICRI (2014), at 8.

<sup>15</sup> See Techopedia, 'Internet of Things (IoT)', <https://www.techopedia.com/definition/28247/internet-of-things-iot> (last visited 7 March, 2017).

<sup>16</sup> Hu (2015), at 798.

**Box 2. Canada Border Services Agency (CBSA), part 1<sup>17</sup>**

**Big Data for CBSA**

- CBSA collects data from both internal and external sources. Internally collected data include closed-circuit television (CCTV) video streams. Externally, the CBSA takes advantage of open source data and external situational awareness data.
- CBSA's work on big data has mainly focused on analysing high-volume structured data, while future-facing efforts aimed at addressing complex problems are being made to align the Administration's capabilities in analytics with the broader definition of big data, including its velocity and variety.

**III. Basics of Big Data Analysis**

Governments and private sector entities have applied big data to their respective organizational goals, expecting it to provide new knowledge, to drive value creation and to foster new business processes, well-informed decision making and potentially future developments. However, big data alone cannot generate any additional value. A noteworthy description can be seen in a recent report delivered to the US President: "Big data is big in two different senses. It is big in the quantity and variety of data that are available to be processed. And, it is big in the scale of analysis (termed "analytics") that can be applied to those data, ultimately to make inferences and draw conclusions."<sup>18</sup> Davenport (2014) states that a sea change is required in IT attitudes and activities within large organizations. He finds it "typical to hear that 75 to 80 percent of an analyst's time is spent sourcing, cleaning, and preparing data for analysis," and adds, "[t]his is true in both companies and government organizations." Additionally, he quotes from a leader of US Defense Department's "Data to Decisions" strategic initiative: "without dramatic improvements, data volumes from next-generation sensors and the complexity of integrated systems will far outpace the ability of analysts to consume it."<sup>19</sup> It is, therefore, necessary for organizations to build in certain techniques for data analysis (e.g. text analytics, data mining, statistics, and natural language processing), as well as related capabilities in order to make available data ready for use.

Noting that big data cannot be handled and apprehended with human capacity alone, a key approach is to employ machine learning, a method of data analysis that automates analytical model building. Using algorithms<sup>20</sup> that iteratively learn from data, machine learning allows computers to find hidden insights without being explicitly programmed where to look<sup>21</sup>. Mayer-Schönberger and Cukier (2013) consider that big data is not about trying to let computers to 'think' as humans do, but rather about 'applying math' to huge quantities of data in order to 'infer probabilities' that, for instance, a received email is spam, certain typed words are misspelled, or a person jay-walking will make it across the street in time, affecting how an automobile's self-driving function works. Programs with such functions "perform well because they are fed with lots of data on which to base their predictions" and "are built to improve

<sup>17</sup> Written information provided by Canada.

<sup>18</sup> PCAST Report (2014), at ix.

<sup>19</sup> Davenport (2014), at 19.

<sup>20</sup> Algorithms have already been in practical use, permeating our daily life (e.g. webpage recommendation engine, speech recognition system and self-driving car).

<sup>21</sup> SAS, 'Machine Learning: What it is & why it matters', [http://www.sas.com/en\\_us/insights/analytics/machine-learning.html](http://www.sas.com/en_us/insights/analytics/machine-learning.html) (last visited 7 March, 2017).

themselves over time, by keeping a tab on what are the best signals and patterns to look for as more data is fed in”, the authors explain<sup>22</sup>.

‘Machine learning’ techniques are often incorporated into ‘data mining’ exercises; however, these terms differ in many ways. Data mining, as often considered a discovery process, revolves around analysis of a cluster of data from different perspectives, summarizing them into useful information. Machine learning is more centered around making predictions, based on certain trends and patterns that have been already identified through data analytics<sup>23</sup>. Apart from its predictive functions, machine learning is also of use to data mining when it performs the tasks of undirected association and classification, helping to extract knowledge and meaning from the flood of data. Forms of big data vary from text message to motion picture, whether raw or organized. In view of challenges associated with mining such complex aggregations of data, the strength of machine learning algorithms does matter<sup>24</sup>.

#### **IV. Policy Considerations for Big Data**

Possible applications of big data in the realm of public policy vary from tax fraud detection and prevention to threat assessment for heightened security. It is, however, necessary to inform policy makers and administrators that there are certain cross-cutting issues in this domain – for one thing, protection of privacy; for another, data security.

Mayer-Schönberger and Cukier (2013) state, “big data allows for more surveillance of our lives while it makes some of the legal means for protecting privacy largely obsolete”; and “it also renders ineffective the core technical method of preserving anonymity”<sup>25</sup>. A person would certainly feel uncomfortable or even scared if his or her information could be used for unknown purposes. Therefore, governments requiring data on personal identification and activities should put practical measures in place to make sure these data are used only for certain legitimate purposes and not shared by irrelevant parties. Legal arrangements (including the guidelines and/or disciplinary codes concerning the treatment of data, the obligation of confidentiality, public-private partnership, inter-governmental cooperation etc.) should be of direct relevance, and technical solutions in this regard should be explored as well. However, such privacy concerns could make it difficult for government agencies to facilitate the exchange of necessary data<sup>26</sup>.

Enhanced connectivity with other entities for data-gathering purposes will allow an organization to benefit from informed and evidence-based decision making. However, establishing permanent external channels could lead to less robust security parameters of computer and network systems, making them less protected from potential hackers or crackers who seek confidential information.

---

<sup>22</sup> Mayer-Schönberger and Cukier (2013), at 12.

<sup>23</sup> McNulty (2014).

<sup>24</sup> Oberlin (2012), 10<sup>th</sup> para.

<sup>25</sup> Mayer-Schönberger and Cukier (2013), at 170.

<sup>26</sup> CBSA is of the view that, while accessing data in a timely manner has been identified as one of the Agency’s biggest issues with respect to big data, one of their key challenges linked to data acquisition arises from privacy limitations across jurisdictions, which could make it difficult for them to carry out horizontal reporting. (Written information provided by Canada)

Another more pressing concern is the potential of cybercrime (or cyberattack) to disrupt the organization's information processing facilities, which could suspend its whole business operation. This is typically represented by a 'distributed denial-of-service' (DDoS) attack – to swamp a computer server with trivial (but thousands or millions of) requests aiming to make it impossible to serve legitimate parties. Such a sizable digital assault exploits a 'botnet', an aggregate of computer terminals (located throughout countries) that have been infected with malware (or a computer virus, spyware etc.) and thus awaiting the remote perpetrators' command. With the advancement of Internet of Things (IoT) technology, it has become easier to create bigger botnets as larger numbers of related devices, such as routers, CCTV cameras, thermostats, networked TVs and smart enabled vehicles, can be 'enslaved'<sup>27</sup>.

According to TechTarget<sup>28</sup> (2016), IoT devices can operate as data-collecting sensors (sending information to a central service and receiving instructions back from it to adjust usage), as well as interacting with other devices to employ localized, swarm intelligence<sup>29</sup> algorithms to respond to local conditions without having to interact with a centralized service. They strongly recommend that corporate entities complete an authentication process for each of the known IoT devices as well as the source of software running on the devices, ensure access controls over those devices and design IoT software analytics with an eye on anomaly detection. These efforts ought to mitigate the risk of a malicious attacker spoofing an IoT device to implement a DDoS attack on the IoT network<sup>30</sup>.

## **V. Big Data from Customs' Perspective**

Several Customs administrations have adopted approaches to the application of big data. An initiative brought into practice by U.S. Customs and Border Protection (USCBP) is applying the concept of big data technologies to its own data in a centralized location. At the present stage, simply assembling the data and preserving the appropriate privacy and security controls is a major undertaking. USCBP recognizes that the tools for big data are still a maturing set of technologies, and thus the greatest practical return is expected with adequate investment. Also, USCBP is of the view that the greatest advantage of the effort to put big data into practical use is the ability to marshal its diverse data holdings and make them available for analysis in ways that would have been too expensive using conventional technologies from storage and computational standpoints<sup>31</sup>.

When it comes to cargo information, for instance, the format of data needs to be fitted for Customs use. According to the provisions of the SAFE Framework, "[w]hile EDI using the international standard UN/EDIFACT is still one of the preferred interchange options, Customs shall also look at other options such as XML", and "even the use of e-mail and telefax could provide a suitable solution"<sup>32</sup>. Data in the form of XML or e-mail are generally perceived as semi-

---

<sup>27</sup> Naughton (2016).

<sup>28</sup> A technology media company headquartered in Newton, Massachusetts, the United States.

<sup>29</sup> 'Swarm intelligence' (SI) is an artificial intelligence term referring to the collective behavior of decentralized, self-organized systems.

<sup>30</sup> TechTarget (2016).

<sup>31</sup> Written information provided by the United States.

<sup>32</sup> WCO Safe Framework of Standards, June 2015, III. 2.6.4.

structured data. Within the scope of Big Data are not only semi-structured but almost any types of data, structured or not. In this regard, the term 'multimedia data' could be applied since it refers to a certain category of data that consist of text, audio, video, animation and the like. Customs administrations thus need to build capabilities of receiving, storing and processing such heterogeneous forms of data, which could subsequently be amalgamated for analytical purposes.

In order for Customs administrations to implement a big data strategy, knowledgeable experts, whether in-house or outsourced, are necessary. Furthermore, a strategy on how to engage specialists in the relevant fields, including cyber security and incident readiness, should be elaborated. The government of New Zealand is currently undertaking a major IT project that will use analytics for risk assessment. Although its capability in this regard has yet to be fully developed, New Zealand Customs has found it extremely important to build a strong connection between the IT professionals that built the capability (still 'offline', sand-box environment to test models) and the analytics team. Now the Administration is of the view that 'co-location' of its analytics team (manager/member) with the IT work group has been instrumental in prioritizing analytics data requirements, considering that there will be a barrier for analytics teams where it can be a long wait for data requirements to be met<sup>33</sup>.

**Box 3. New Zealand Customs Service<sup>34</sup>**

**Joint Border Management System (JBMS)<sup>35</sup>**

- Once built, the JBMS will enable New Zealand Customs to use analytics in order to carry out smarter and more selective targeting of trade in real time. The Administration is building the system together with the Ministry for Primary Industries.
- So far, New Zealand Customs has built a system for traders to lodge entries through a single portal (Trade Single Window)
- Using the legacy system, New Zealand Customs has developed an analytics environment in which new profiles and models for risk assessment can be built (offline capability). The Administration is also building a system to deploy the profiles and models against the traders' lodgements for risk assessment in real time (online capability). The development of the offline analytics capability has proceeded at a faster pace.

**Outlook for Transformation**

- New Zealand Customs' vision for 2020 is to provide intelligence-led service. With analytics which is critical to that vision, the insight can be used to transform the way of doing work.
- Analytics has the potential to disrupt (in a positive way) how the Administration carries out its core business. If that potential is realized, predictive analysis will shape how the Administration makes both strategic and operational decisions and how it allocates its resources (particularly for intelligence, investigations, revenue assurance and the frontline).
- Analytics can also be used to improve customer service and design. It has the potential to transform overall Customs service.

Different divisions/sections of a government agency could maintain and use their respective data in an exclusive manner, neglecting to share such data across business units and

<sup>33</sup> Written information provided by New Zealand.

<sup>34</sup> Id.

<sup>35</sup> Updated information is available at <http://www.customs.govt.nz/features/jointbordermanagement/pages/default.aspx> (last visited 7 March, 2017).

geographies. By contrast, Hong Kong China Customs' effort is remarkable in that it took on an initiative to 'centralize' and 'consolidate' all the relevant data that would have otherwise remained disparate, thus being rendered less useful. Since 2012, the Administration has put in place a centralized information repository, providing a centralized data warehouse to store various operational data coming through nine information systems which are respectively used for certain regulatory purposes<sup>36</sup>.

**Box 4. Customs and Excise Department, Hong Kong, China<sup>37</sup>**

**Central Information Repository System**

- Hong Kong China Customs' Central Information Repository features a Data Warehouse which stores and deals with the data bearing the characteristics of high volume, high variety, and high velocity. The volume of data processed by the warehouse in 2012 was 3 terabytes, which has increased to 12 terabytes in 2015. The more data it processes, the more like a big data management tool it may look.
- The warehouse offers a unified platform in order for users to uncover insights and trends in multiple fields with cross-system data analysis. The data can be stored up to seven years, facilitating users to formulate long-term and operational planning.
- The operational master of the system provides an infrastructure for timely and efficient sharing of data among various systems within the Administration. It standardizes the interface mechanism among systems on a harmonized data model, thus reducing the cost for system maintenance and improving the efficiency of system support.

**Application of Big Data**

- Besides the data captured through Customs regulatory activities, the Data Warehouse also stores the data that are transferred from other government departments (e.g. Business Registration particulars from the Inland Revenue Department, Closed Road Permit information from the Transport Department) for verification purposes.
- By so doing, the warehouse enhances the data quality and serves as a unified platform for data analysis, rendering to all users at different levels more accessibility to statistics and more efficiency in research.
- The warehouse routinely generates reports on-line to give users specific and latest information for their immediate reference. The reports can be generated in different formats so as to facilitate further analysis. It also allows customization of interface in accordance with users' preference.
- With the information made available, the Administration can gain insights for not only immediate decisions in enforcement (such as risk management, operational planning, and resource allocation) but organizational management and strategy planning in the long run.
- For future development, the Administration is considering an intensive automation of data mining, which will be put on the list of study for enhancement and optimization of operational efficiency.

Applying big data in a systematic manner could highlight some sorts of chronic problems that have been faced by Customs for years. Her Majesty's Revenue and Customs, the United Kingdom (HMRC), for instance, have reflected on themselves to notice that they might have relied on inaccurate, incomplete and misleading data (including those coming from the wrong source and/or at the wrong time). The Administration has therefore taken steps to obtain the

<sup>36</sup> Written information provided by Hong Kong, China.

<sup>37</sup> Id.

proper assurance systems and the right sources to ensure the accuracy and integrity of their data at hand, which could otherwise create a false sense of security<sup>38</sup>.

**Box 5. Her Majesty's Revenue and Customs (HMRC)<sup>39</sup>**

**Community System Providers (CSPs)**

- With the creation and development of the CSPs – trade systems at ports and airports, the UK has been tackling the challenge of handling huge volumes of data associated with commercial freight movements and the related physical movement of the freight prior to Customs clearance.
- The UK model is uncommon, even on an EU level, as HMRC has developed a symbiotic relationship with commercial operators and thus has its anti-smuggling and targeting technology embedded in these commercial systems to allow us to take the broadest view of the relevant large data flow.
- The concept has even been extended to apply in-out controls directly within economic operators' systems<sup>40</sup>.

**CORE Project**

- David Hesketh, head of Customs research and development at HMRC and leader of the CORE Project, explains the background: according to experience garnered from previous projects aiming at tackling the problem of 'data inaccuracy', "commercial supply chain visibility solutions" are "[typically] web-based, electronic data 'pipelines' that can link the seller, buyer and interested parties; as well as the controlling authorities"<sup>41</sup>.
- The project is delivering seamless, integrated data pipelines that provide the right data from the right source at the right time. The ongoing R&D work with the project is aimed at gaining access to commercial data flows much earlier in the supply chain to enhance the targeting capability, as well as providing increased assurance to compliant business.
- The project is based on their perception that, if benefits can be fed-back from Government to the providers of such data, then the volume of correct data will increase and so will the ability to de-risk the participating companies and create compliant models – an example of what 'good' data could contribute.
- In HMRC's view, it has built data sets to international standards; and it cannot adopt 'Big Data' principles if the data are inaccurate, late and from the wrong sources.

According to Baida (2016), "data is a means, not an end"; 'information', from which people gain insights that can form the basis of confident decisions, "is data that has been processed, organized, structured or presented in a meaningful context"; and 'digitization' represented by the abundance of sensors and 'smart' technologies has resulted in an explosion of available data<sup>42</sup>. In general, Customs administrations have not yet been fully prepared to use such data to extract the information which could result in improved decision-making. This could be explained by the fact that they normally look to their 'internal' data collected through goods declarations or obtained by way of regular interventions, while trying to distill these data to focus on key information points.

<sup>38</sup> Written information provided by the United Kingdom.

<sup>39</sup> Id.

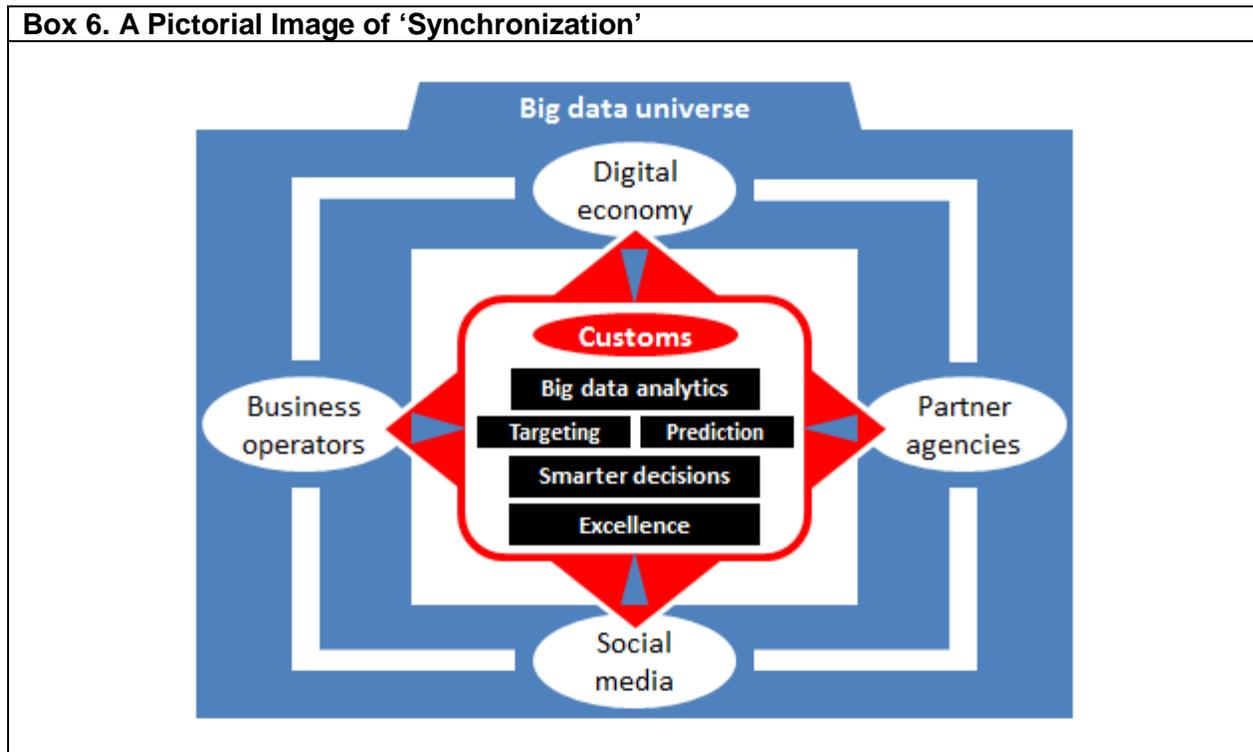
<sup>40</sup> Economic operators are represented by Fast Parcels Operators – commercial operators approved by HMRC so as to deal with Customs procedures (declarations) on behalf of importers/exporters.

<sup>41</sup> Core 2nd Newsletter, <http://www.coreproject.eu/newsletters/core-2nd-newsletter-august-2015/seamless-integrated-data-pipelines-by-david-hesketh-hmrc.aspx> (last visited 7 March, 2017).

<sup>42</sup> Baida (2016), at 30.

When it comes to big data, the primary issue is how, and to what extent, to diversify the types and sources of data to create a data-rich repository within an organization. The more data available for Customs administrations, the more chances to combine and correlate different categories of data, allowing them to identify unknown trends and patterns concerning the subjects of control (i.e. cargoes, conveyances and persons), as well as to verify and ascertain the accuracy of data in question. The faster they update data, the more likely they are to notice and respond to possible frauds and irregularities in a timely manner. Essentially, ensuring that Customs are equipped for the practical application of big data would simply require that they take advantage of the characteristics of big data.

Acquisition or 'internalization' of big data for Customs purposes can be achieved by 'synchronizing' the activities of Customs with the dynamism of the big data universe, which should be ever-evolving externally but may have been assimilated to some extent into another authority (either Customs or its partner agency) wherever it is located. It is a 'symbiotic' relationship (as specified by HMRC) between Customs and the private sector (ranging from common carrier to social networking service) that should unlock a variety of data accessible via multiple avenues. Such an information infrastructure cannot be put in place without a solid legal foundation, particularly in view of privacy/confidentiality concerns and cybersecurity questions. Human resources are also required to address these issues and thus should constitute an important element of all the technical capabilities that Customs administrations need to develop.



## **VI. Challenges and Opportunities with Big Data**

Customs administrations will face multiple challenges as they try to keep pace with developments in data analysis. Canada Border Services Agency (CBSA) has been reviewing its

internal environment to maximize the benefit of the use of available data. The Administration considers that leveraging big data will require addressing challenges associated with workforce development, data management and system infrastructure. As for workforce development, it contends that unlocking potential for big data usage will require strategically developing their workforce by identifying high potential data analysts that can be hired early in their career and grow professionally inside the organization. It is not enough for CBSA to have such in-house experts only; that is, there is the necessity to enhance the skill and capacity of current and overall staff through further training and routine access to analytical tools. The Administration also recognizes the importance of enhancing efficiency in data management, while considering developing a centralized and specialized function within the organization to coordinate and facilitate the acquisition of data and systems. In order to enhance its analytical capacity, the Administration also finds it necessary to carry out further integration of datasets across the different sections within the organization<sup>43</sup>.

**Box 7. Canada Border Services Agency (CBSA), part 2<sup>44</sup>**

**Way forward**

- In order to store and process big data which it regards as a core enabler, CBSA currently uses an Enterprise Data Warehouse (EDW). The Administration finds it important to focus on minimizing barriers and broadening access to the EDW.
- Development is continuing on the existing EDW to integrate additional sources and types of information. Very recent additions to the EDW include open-source datasets that will improve their ability to assess the risks regarding both goods and people. Additional projects in the exploratory phase are biometric screening, facial recognition, automated lie detection and predictive modelling which will be brought together to the traveller screening process and fed into their big data footprint.
- By integrating large datasets into the EDW and expanding access to it, the Administration will be able to gain greater insight into big data. This will lead to actionable predictions on high-risk goods and people, as well as on their previously unidentifiable relationships.

***International Standards***

Big data contains clusters of raw data that are disordered and untidy. More technically, it relies on a broad mix of both structured and unstructured data formats. This represents a stark change for Customs administrations usually accustomed to structured data; that is, data from electronic declarations for example are well defined by the WCO Data Model, and the EDIFACT family of standards. The 'standards' defining big data would be much more varied, and largely outside of government's control.

It is clear that any intentions to exploit the potential of big data will require Customs, on the one hand, to have a strong mastery of its own structured data management and, on the other hand, to advance into other areas to exploit new data as well. The standards are still being written for the areas of commerce and industry; therefore, Customs administrations seeking data from the business community will be encouraged to focus their attention on new emerging standards, with a view to benefiting from the 'sense-making' potential of big data.

***Risk Management***

<sup>43</sup> Written information provided by Canada.

<sup>44</sup> Id.

In recent years, an increasing number of private companies have been keeping track of their customers' historical records to make better and quicker purchase-and-credit decisions. Government agencies could also apply such techniques to their services. Using big data analytics, they can identify the behavioral patterns of persons as well as the implications<sup>45</sup>. With passenger data (that might not necessary fall within the scope of big data), for example, Customs and other border agencies can consolidate and update travelers' histories in flight over time to identify and flag certain travelers who should be considered high-risk.

A similar approach can be taken to control the cross-border movement of cargoes, including those being transported by land and those by sea. Unlike passenger controls, however, cargo risk management could involve big data in a different manner.

It is more likely that marine cargoes, whether in bulk or containerized, are exploited for concealing contraband goods and other illicit trade articles on a larger scale. Risk rating has already been widely incorporated into Customs' regular enforcement operations, and are thus typically reflected in the outcomes of a cargo selectivity exercise. Customs nonetheless can hardly pinpoint the specific consignment that should be seized among other cargoes awaiting Customs clearance at seaport. Whole quantity inspection is almost impossible (or infeasible/impracticable) over these large-sized cargoes, even though it could be employed for screening passengers' check-in luggage.

With this in mind, it may be suggested that Customs administrations should seek to enrich their existing data that have accumulated over time, thereby obtaining a bigger picture of the current risks they face. With advanced statistical analytics, for instance, they would be able to correlate the internal data on the parties involved in trade (collected through their regular interaction with traders, brokers, shippers, consignees etc.) with other categories of data on the in-depth histories of private entities in great numbers (as seen in Passenger Name Records or PNRs, whose personal data can be easily merged with other metadata in the financial or communications domain). This could enable them to uncover the schemes for certain crimes involving trade (from smuggling to false declaration), while helping them to make sense of possibly related crimes (committed to avoid tax or to disguise illicit monetary transaction) that require intensive examination or audit.

**Box 8. Canada Border Services Agency (CBSA), part 3<sup>46</sup>**

**How Big Data being used at CBSA**

- CBSA is taking advantage of internal and external data assets to gain better insight to support its facilitation and security mandate. The National Targeting Centre is currently using big data to identify high risk travellers and goods.
- The advanced analytics team<sup>47</sup> is leading multiple big data pilot projects to enhance predictive analytics, visualize information and improve the use of the Administration's biometric data. The team's members benefit from access to the Administration's Experimental Computing Environment, which allows them to tailor software and hardware to any given piece of work.

<sup>45</sup> Prow (2014).

<sup>46</sup> Written information provided by Canada.

<sup>47</sup> The team resides in the Information, Science and Technology Branch (comprising CBSA with other five brunches and one group).

- The EDW – earlier mentioned – brings together information from various data sources of mainly internal data for data analytics. The EDW is currently being used to support CBSA’s targeting operations in the interdiction of travellers and goods prior to the arrival of them at a port of entry within the territory.
- In the commercial realm, project funding has served as a catalyst to advance the analytics within CBSA. Advanced analytics is used in the commercial program to identify and intercept high risk shipments and improve targeting. (Customs-to-Business Partnership)

Data analytics that enables Customs to identify and trace suspicious persons and goods would have a greater impact on the management and accomplishment of day-to-day operations in view of the deployment of personnel and related equipment. With the trends and patterns concerning the flow of the subjects of control, Customs would hopefully learn to anticipate incoming threats, allowing optimal or prioritized allocation of certain limited organizational resources.

Considering the features of big data, however, data analytics should not be understood as an absolute all-purpose tool for Customs. Big data, for instance, provides another basis for verification (thus enriching the pool of data available and potentially usable for Customs in its decision-making); nonetheless, alternative sources of data that might have been generated by unrelated parties for certain non-regulatory purposes are not always equipped with the same level of verifiability as Customs declarations. This requires Customs officials, including those wishing to use big data to its fullest potential, to have a more nuanced view of data in general.

More specifically, big data may not provide the analysts with absolute “yes-no” answers, but it could otherwise provide a snapshot of trends, patterns, gaps and potential indicators. Customs staff thus would have to tackle certain crucial questions about data accuracy, availability and admissibility in decision making, before being able to use big data. Here, it is worth recalling that HMRC puts great emphasis on the quality, or the accuracy and integrity, of available data, taking account of the risk of creating a false sense of security. Passive and mechanical application of whatever data without caution and diligence could rather jeopardize Customs’ continuous and consistent effort to accumulate knowledge and take steps in a strategic manner.

### ***Internet of Things***

All goods moving across borders fall under the control of Customs. The advantage of applying big data can be well demonstrated by focusing on the Internet of Things (IoT) that may be considered a key enabler for Customs to capture the location, condition and status of traded goods in real time. This is illustrated with advanced sensor technology that allows logistics service providers to detect any irregularities occurring in or around the cargoes in transit, thus helping to enhance supply chain security. As such, a containerized cargo being once regarded as ‘low-risk’ or ‘risk-free’ can maintain the same condition until it is delivered at the destination unless any suspicious intervention is detected during the time of transport.

Customs administrations are not necessarily supposed to ‘watch’ the current movements of cargoes/shipments on their part (to make sure there is no accident) but rather expected to strengthen the cooperative relationship with certain stakeholders (shippers, carriers, forwarders etc.) that have employed the IoT applications, with a view to promptly obtaining any information

that corresponds to certain risk factors. In other words, Customs administrations would be able to concentrate on their analytical tasks, while having commercial operators function as 'caretakers'. It seems a possible way forward that certain roles and responsibilities are shared between Customs administrations and private sector entities.

Keeping track of the state of consignments over time will positively affect Customs' 'selectivity' criteria, resulting in expedited clearance of imports, as well as benefiting all stakeholders. At the same time, Customs will be able to react to the incidences of any irregularity in a timely manner, while preventing them from leading to serious offences being committed within the territory.

### ***E-commerce***

In the e-commerce environment, enormous data concerning the sale of goods and services are generated day by day, and regularly stored, processed, managed and utilized by online marketplaces and related intermediaries (e.g. payment gateways, express cargo shippers/distributors) that have been directly serving consumers. This is one of the areas where businesses leverage big data to maximize their profits. Applying such data as being associated with social and economic factors concerning online trade participants (involved in 'B to C' or possibly 'C to C' transactions in particular) could make Customs more informed in evaluating the 'compliance' level of private/individual purchasers – the parties who have been less focused under the 'brick and mortar' trading system. By the same token, Customs would be more informed about the traded goods and thus better prepared, for one thing, to tackle emerging risks to human health and safety.

E-commerce is a very data-rich environment, in which Customs could potentially of embrace voluminous data concerning trade in goods. Unlike conventional forms of trade, certain centralized marketplaces do manage and update the records of the ongoing and past transactions in a systematic manner. If their data were created and effectively shared on a transactional basis, it could allow Customs to make even more informed and quicker decisions, facilitating the clearance of shipments (traded online) as well as ensuring revenue collection (with tax amount to be properly assessed). Online retail (or online shopping) is thriving in certain countries (e.g. China, Germany, India, Brazil and the United Kingdom)<sup>48</sup>; therefore, e-commerce data recorded from individual transactions might have a potential for Customs to confidently expedite the clearance process to the fullest extent possible.

### ***Security Concerns***

Big data has implications for security and intelligence. Certain police forces analyze some aspects of communications through social media (also referred to as SNS) trying to anticipate security threats (that have been latent but could be imminent). Such efforts have had an impact on how they perform patrolling duties in terms of specificity in time and location. As such, their approaches to big data commonly feature 'predictive analytics' – an area of data mining as well as an application of statistical techniques, especially from machine learning with 'predictive policing' algorithms. Predictive analytics, which is aimed at predicting unknown events

---

<sup>48</sup> Statista, <https://www.statista.com/topics/871/online-shopping/> (last visited 7 March, 2017).

of interest that might be critical for community protection, has exploited the updates of numerous people's 'words and actions' being considered 'variables'. It has been running on ever-evolving big data (or 'stream data') deriving from day-to-day human communications; therefore, the question of how to prevent abuse of privacy/secretcy needs to be addressed.<sup>49</sup> Another public concern is whether relevant data could be used for criminal proceedings, rather than securing society against imminent threats.

In the United States, a screening system named Automated Virtual Agent for Truth Assessments in Real-time (AVATAR)<sup>50</sup> has been in place within the framework of the Trusted Traveler programs which is aimed at enhanced convenience of preapproved or low-risk passengers/travelers entering or reentering the States with certain dedicated lanes and kiosks. A virtual officer (called 'avatar'), who resides in the AVATAR-administered kiosks that USCBP has installed at various ports of entry, interviews all the persons applying for Trusted Traveler status and then provides human officers with automated feedback via their tablets and smartphones, allowing them to conduct subsequent interviews for follow-up in depth<sup>51</sup>.

Marr (2016) explains the mechanism of the AVATAR system as follows.

In order to make probabilistic judgements about whether a subject person is telling the truth, the system uses three different sensors that have been built into the machine's cabinet: an infrared camera that records data on eye movements and pupil dilation at 250 frames per second; a video camera that monitors the person's body to detect its suspicious twitches or habitual movements matching what people tend to show when they are hiding something; and a microphone that records voice data, listening for subtle changes in the pitch and tone of voice. The virtual agent's further interactions with the subjects of immigration inspection generate new personal data for the system to compare against its ever-growing and constantly updating "Big Database".

In sum, AVATAR undergoes a process that can be seen in many other big data projects using predictive modelling techniques. As it examines more people, it learns more about the facial, vocal and contextual indicators that tend to be present when a person is lying or acting deceptively. The machines' level of accuracy in finding suspicious travelers who needs to be flagged will steadily increase if they are given the right data and algorithms<sup>52</sup>.

### ***Emerging Threats***

In recent years, Customs has been urged to tackle emerging and complicated areas of threats as represented by money laundering and terrorist financing. Criminals exploit legitimate trade (including banking systems related to trade finance) to disguise their illicit proceeds. Such crimes, often referred to a trade-based money laundering (TBLM), involve a number of schemes

---

<sup>49</sup> See Engelen (2015).

<sup>50</sup> AVATAR was developed by researchers at the University of Arizona's National Center for Border Security and Immigration. According to the Center, AVATAR is designed to flag suspicious or anomalous behavior that warrants further investigation by the Department of Homeland Security (DHS)'s human agent in the field. The University of Arizona – 'AVATAR', <http://borders.arizona.edu/cms/projects/avatar-automated-virtual-agent-truth-assessments-real-time> (last visited 7 March, 2017).

<sup>51</sup> Greenemeier (2012).

<sup>52</sup> Marr (2016), at 111-115.

that have been contrived to complicate the documentation of a series of trade transactions. It is estimated that annual TBML exceeds billions of dollars and is growing each year<sup>53</sup>.

Red flag indicators for potential TBML, which may allow Customs officers to detect the fraud in real time, include false reporting such as over valuation or undervaluation of the goods concerned. The work required to identify and mitigate TBML risks is extremely laborious and, due to its manual nature, prone to human error. It is, therefore, advantageous to pursue automation in performing this sort of task<sup>54</sup>; that is, big data analytics may be of great help to keeping track of transboundary flows of illicit funds that wrongdoers tend to hide by exploiting international trade.

TBML schemes often involve trade finance for which banking institutions handle lots of documents, structured or unstructured. In order to extract data from unstructured sources, text analytics tools need to be developed. It is also necessary to transform these data into certain usable information and ensure accuracy and relevance of that information. To that end, algorithmic and statistical transaction monitoring techniques would help identify patterns and linkages potentially indicative of TBML activities (as often seen in improper price manipulation or transfer pricing)<sup>55</sup>.

Applications of big data for enforcement purposes could be harnessed to encourage Customs to collaborate with other relevant authorities. Regarding anti-TBML best practices, a stand-alone report issued by the Financial Action Task Force (FATF) specifically calls on countries to allow the investigative authorities (prosecutors, police forces, public safety agents etc.) access, directly or indirectly, and on a timely basis, to trade data and relevant financial information<sup>56</sup>. In this context, Customs and Tax administrations could explore the possibility of shaping the relevant database, inter alia, on taxpayers' trade practices and related activities for banking transactions, while employing data analytics to seek attributes commonly found in the respective Administrations' datasets. The same techniques could be utilized for comparing export information with tax declarations to detect discrepancies – another FATF-recommended practice<sup>57</sup>.

## **VII. Conclusion**

Big data requires certain capabilities including tools for analytics, in addition to an infrastructure with relevant technologies. In this regard, Hong Kong China Customs has developed the capability to store and manage a large volume of datasets, while enabling timely and efficient data sharing among the separate systems within the Administration. USCBP marshalled its diverse data holdings to make them available for the personnel conducting analysis. CBSA has built a data warehouse to consolidate data (mostly internal) from various sources (which the Administration has regularly used) for analytical purposes, recognizing the

---

<sup>53</sup> USICE, 'ICE leads trade-based money laundering investigations', <https://www.ice.gov/trade-transparency> (last visited 7 March, 2017).

<sup>54</sup> PwC US (2015), at 13.

<sup>55</sup> Id, at 13-15.

<sup>56</sup> FATF/OECD (2008), para. 21.

<sup>57</sup> Id. para. 13.

necessity of furthering data integration to strengthen the capacity for analysis. These efforts can be referred to as 'centralization' of data, which is aimed at extracting more clues from a greater pool of data.

There are some different priorities in promoting big data strategies. For HMRC, data quality has been at the top of the agenda. It is therefore considered that tangible benefits fed-back to the trade community could become a possible incentive for them to provide the Administration with accurate data<sup>58</sup>. New Zealand Customs considers that analytics is critical to realizing the Administration's vision for the next few years, which is aimed at intelligence-led decision-making and operations, possibly changing how their tasks are performed<sup>59</sup>. USCBP estimates that its big data platform could have the potential to provide practical solutions (like predictive analytics)<sup>60</sup>. CBSA finds it necessary to explore additional sources and types of information that could form the basis of actionable predictions<sup>61</sup>.

HMRC practices illustrate a certain progressive partnership with the private sector. It has been collaborating with the supply chain intermediaries which generate, store and transmit a variety of data for their business use. The Administration has synchronized its Customs intelligence platform with the business operators' systems to make sense of risks from a wider perspective<sup>62</sup>. This would probably enhance its risk identification/assessment capacity while meeting the necessity of cross-checking data for accuracy. From another viewpoint, it could be said that businesses also have borne some responsibility in pursuit of effective border control – the common goal for both public and private sectors<sup>63</sup>.

Big data is not a static concept as it has ever-evolving applications and a potential to produce a sea change in how Customs field officers perform their duties, affecting how organizational resources are allocated to support their tasks to be accomplished. There is nonetheless a concern about what could happen if a Customs administration blindly pursues the incorporation of the data being collected externally into its day-to-day operation without examining the veracity of those data. A false sense of security and threats caused by improper/imprecise assessment of risks will almost invariably result in decreased Customs capacity in intelligence-led enforcement. As such, applications of big data alone will not necessarily guarantee that Customs enforcement becomes smarter.

Last but not least, Customs could compromise the security of its information management system that should affect every facet of the service, as the network is built up with outside entities. Internet of Things technology, among other things, is of direct relevance to this problem since it allows the Customs intelligence platform to interact with multiple devices (that could be large in number and located across the globe), while making it susceptible to unwanted actions or destructive forces – typically represented by cyberattacks. Bearing this in mind, it is worth considering how Customs administrations should prepare to cope with such future

---

<sup>58</sup> Written information provided by the United Kingdom.

<sup>59</sup> Written information provided by New Zealand.

<sup>60</sup> Written information provided by the United State.

<sup>61</sup> Written information provided by Canada.

<sup>62</sup> Written information provided by the United Kingdom.

<sup>63</sup> Besides, HMRC's current challenge includes further collaboration with stakeholders toward earlier submission and sharing of cargo-related information. For instance, e-commerce could be discussed in this context due to the nature of transactional data (that always 'precede' the physical distribution).

contingencies. For one thing, enhanced resilience of the Customs data repository should certainly strengthen the Administration's business continuity processes, which is crucial to the sustainability of international trade in times of emergency.

\*  
\*       \*  
\*

## **References**

1. Cleary, Duncan (2017), 'Big Data Analytics for Better Revenue Administration', IMF Public Financial Management Blog, <http://blog-pfm.imf.org/pfmblog/2017/02/big-data-analytics-for-better-revenue-administration.html> (last visited 7 March, 2017).
2. EMC (2012), Press Release, 'New Digital Universe Study Reveals Big Data Gap: Less Than 1% of World's Data is Analyzed; Less Than 20% is Protected', <http://www.emc.com/about/news/press/2012/20121211-01.htm> (last visited 7 March, 2017).
3. Rijmenam, Mark van (2013), 'Why The 3V's Are Not Sufficient To Describe Big Data', Dataflog, <https://dataflog.com/read/3vs-sufficient-describe-big-data/166> (last visited 7 March, 2017).
4. Adamov, Abzetdin (2015), 'Understanding Big Data: 3Vs of Big Data', posted on his blog, <https://aadamov.wordpress.com/2015/04/24/understanding-big-data-3vs-of-big-data/> (last visited 7 March, 2017).
5. Gerasimou, Vicky (2016), 'Big Data and the 3Vs: What is the fourth 'V' and what are the implications for not embracing it?', Think Big, <https://www.thinkbiganalytics.com/2016/03/29/big-data-3vs-fourth-v-implications-not-embracing/> (last visited 7 March, 2017).
6. AAAS-FBI-UNICRI (2014), *National and Transnational Security Implications of Big Data in the Life Sciences* (Joint project), American Association for the Advancement of Science, available at <https://www.aaas.org/report/national-and-transnational-security-implications-big-data-life-sciences>.
7. Hu, Margaret (2015), 'Small Data Surveillance v. Big Data Cybersurveillance', Washington & Lee Legal Studies Paper No. 2016-6, PEPPERDINE LAW REVIEW, Vol. 42: 773, available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2731344](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2731344).
8. PCAST Report (2014), *Big Data and Privacy: A Technological Perspective*, Executive Office of the President, US.
9. Davenport, Thomas H (2014), *Big Data at Work: Dispelling the Myths, Uncovering the Opportunities*, Harvard Business Review Press, US (ISBN: 978-1-42216-816-5).
10. Mayer-Schönberger, Viktor and Cukier, Kenneth (2013), *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, John Murray, UK (ISBN: 978-1-84854-792-2).
11. McNulty, Eileen (2014), 'Understanding Big Data: Machine Learning', Dataconomy, <http://dataconomy.com/understanding-big-data-machine-learning/> (last visited 7 March, 2017).

12. Oberlin, Steve (2012), 'Machine Learning, Cognition, and Big Data', CA (unavailable at the corporate website).
  13. Naughton, John (2016), 'Why the internet of things is the new magic ingredient for cyber criminals', The Guardian, <https://www.theguardian.com/commentisfree/2016/oct/02/brian-krebs-ddos-attack-google-protection-cybercrime> (last visited 7 March, 2017).
  14. TechTarget (2016), *Time to Toughen Up for IoT*, SearchSecurity.com e-publication, available at <http://searchsecurity.techtarget.com/> (last visited 7 March, 2017).
  15. Baiba, Ziv (2016), 'Start small, think big: big data!', WCO NEWS N. 79, World Customs Organization.
  16. Prow, Charles (2014), 'Putting Analytics at the Border Would Increase Security, Convenience (Industry Perspective)', Government Technology, <http://www.govtech.com/federal/Putting-Analytics-at-the-Border-Would-Increase-Security-Convenience-Industry-Perspective.html> (last visited 7 March, 2017).
  17. Engelen, Errol van (2015), 'Big Data Will Effectively Fight Terrorism In The World', Dataflog, <https://dataflog.com/read/big-data-will-effectively-fight-terrorism/785> (last visited 7 March, 2017).
  18. Greenemeier, Lary (2012), 'Avatar Officer Installed at Arizona-Mexico border Station', Scientific American, <http://www.scientificamerican.com/article/avatar-officer-installed-mexico-border/> (last visited 7 March, 2017).
  19. Marr, Bernard (2016), *Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results*, John Wiley & Sons, UK (ISBN: 978-1-119-23138-7).
  20. PwC US (2015), *Goods Gone Bad: Addressing Money-laundering Risk in the Trade Finance System*, available at <http://www.pwc.com/us/en/risk-assurance-services/publications/trade-finance-money-laundering.html> (last visited 7 March, 2017).
  21. FATF/OECD (2008), *Best Practices Paper: Best Practices on Trade Based Money Laundering*, FATF Guidance Document.
-