Data analysis techniques for enhancing the performance of Customs

Danilo Desiderio

Abstract

One of the most powerful tools available to Customs to reconcile the functions of controlling the international movement of goods with the needs of trade facilitation is represented by data collection and analysis techniques. These techniques are supported by the use of statistics, algorithms and other mathematical tools, as well as by adequate IT systems for their treatment. If properly used, they can allow Customs to act in a targeted way to achieve their institutional objectives more efficiently. Customs authorities can improve the effectiveness of controls and their overall performances not only by analysing the traders' historical activity and the number of past frauds detected, but also by using additional sources of information, both internal and external to the administration. The reality, however, is that today most customs administrations use data analysis almost exclusively for conducting risk management and risk scoring activities. Instead, a holistic approach suggests that modern Customs should use such techniques also for facilitating trade, not only by minimising obstacles for operators in terms of fluidity of their operations, but by observing and analysing their behavioural patterns to introduce simplifications in customs procedures aimed to make them more user-friendly.

1. Data analysis, data warehousing and data mining

Data analysis can be defined as the process of transforming raw data into usable information, then into knowledge, in order to add value to the statistical output (Dabbicco & Di Meglio, 2011). It consists in systematically applying statistical and logical techniques to describe, summarise and compare information so that it can be used to efficiently drive priority-setting, decision-making, performance measurement, budget planning and forecasting operations.

The WCO (2018) *Draft guidance on data analytics* describes data analytics as the process of analysing datasets in order to discover or uncover patterns, associations, and anomalies from sets of structured or unstructured data, and to draw practical conclusions. The document also recommends the adoption of data analytics strategies within Customs to improve the use of available data and information, in view of expediting their decision-making process and increase their facilitation results. An example is the strategy for data analytics developed by the Canada Border Services Agency (CBSA, 2018), aimed to maximise its capacity to extract, process and evaluate all data and information needed for the efficient accomplishment of its tasks.

Data warehousing, on the other hand, is a data management technology for collecting data from multiple sources, aggregating, organising and storing them in a single repository so that they can be effectively mined. Conversely, data mining refers to the automated exploration of the data using artificial intelligence

paradigms such as machine learning or agent-based network modelling. Based on the data maturity lifecycle, data mining may start with internal available data and local providers. This could be enriched on case-by-case basis with auxiliary data available from other sources, followed by coupling external datasets.

2. Data analysis for effective border management

The World Customs Organization (WCO) dedicated the 2017 International Customs Day to data analysis for effective border management. Every year, the WCO selects a specific theme to stimulate the reflection of the global customs community on the use of new techniques and work methods that allow Customs to increase their performance, in particular through the dissemination and comparison of best practices. In 2016, the International Customs Day was opened by the WCO with the theme 'Digital Customs: progressive engagement', where customs agencies were urged to exploit the latest technologies to effectively carry out their mandate. On that occasion, WCO members were invited to use all the latest techniques and technologies, such as cloud computing or blockchain, to increase their efficiency and to improve coordination and data exchange mechanisms with other stakeholders, such as other government agencies involved in cross-border movements of goods and the trade community.

3. Internal and external data available to Customs to improve overall operations

In carrying out their routine tasks, Customs collect and generate on a day-to-day basis a huge amount of data that in most cases is underutilised by such administrations. The efficient exploitation of information in the possession of Customs is the result of a complex process, where data is interpreted, evaluated, compared and amalgamated with other information held by the administration; received from other government agencies, foreign customs administrations, regional and international organisations and private operators; or extracted by other external sources. In developing countries, additional information can be obtained by the private inspection companies conducting pre-shipment inspections (Laporte, 2011).

The need to complement internal sources with the external ones, is inversely proportional to the capacity of Customs to generate data internally. For instance, many customs administrations in developing countries, not disposing of internal databases (e.g. customs valuation databases) or having reduced internal sources of information (e.g. records on the offences and frauds detected in the past), must necessarily use additional sources to support their operations and decision-making processes (Cariolle et al., 2017). Further steps consist of integrating such data and information into the customs risk management systems and use them to develop indicators for measuring performance, both in terms of efficiency in revenue collection and trade facilitation. These actions are known (given their highly strategic and operational value) as 'customs intelligence' activities.

Data analysis, data warehousing and data mining techniques, as well as information exchanges, can greatly increase Customs' tasks of revenue collection and protection of collective interests, enhancing their ability to detect irregular declarations and illegal or suspicious movements of goods, persons and financial flows. Advanced data analytics, such as predictive analytics[1], can enable Customs to risk/rank import and export transactions and create risk scores in real time, thus facilitating compliant traders while intercepting fraudulent shipments. These techniques, when combined with information exchanges (e.g. with foreign customs administrations and other cross-border regulatory agencies) can maximise performances of Customs in the fight against frauds and other illicit activities as a result of more targeted action with minimal impediment to trade.[2]

There is, however, another aspect that must be considered. Data analytics techniques are usually used by Customs for risk management and risk scoring activities, for revenue collection, and to protect collective interests. A holistic approach suggests, however, that these techniques can also be used for facilitating trade. For example, the data generated during pilot tests of new procedures, where the behavioural patterns of private operators can be observed and analysed in order to introduce simplified procedures designed to make them more user-friendly. Another example is the time release studies (TRSs), strategic studies carried out by Customs, realised with the participation of private sector stakeholders and following a precise methodology[3], measuring the time from arrival of cargo until its physical release, for each point of entry or exit in the customs territory, so that bottlenecks in the clearance process can be easily identified and corrective actions implemented, specifically through the redesign of procedures in view of further simplifying or facilitating trade operations.

4. Major constraints in exploiting data analysis, warehousing and mining techniques

In order to efficiently exploit data analysis and data mining techniques, Customs needs to address two main constraints:

1. availability of staff with adequate knowledge and skills in data mining and who are familiar with algorithms, predictive analysis, probability models and other statistical targeting techniques

2. availability of adequate IT systems for data treatment and management.

The recruitment or identification within the customs administration of qualified staff with data mining and intelligence analysis skills can be a daunting task but is a solution that can significantly contribute to enhancing an administration's performances by improving its risk mapping and risk targeting capability. To this end, modern customs administrations should aim to create an environment that is conducive to advanced data analytics through the establishment of multidisciplinary teams made up of staff who have computer, statistics, mathematics and social science skills. This kind of approach is being pursued today by some countries, like India, where the National Academy of Customs, Indirect Taxes and Narcotics (NACEN) recently proposed a plan to recruit data mining experts and establish, within the customs administration, a Knowledge and Analysis Centre for data mining and analysis purposes.

5. Sources of data and information available to Customs

The potential sources of data and information available to Customs are today particularly numerous because of the recent advances in information and communication technologies that have drastically changed the learning behaviour of such administrations.

A list of such sources is provided in Table 1.

Table 1: Sources of information

Source	Description	Notes
Cargo and goods declarations	Cargo and goods declarations, with supporting documents, are transmitted to Customs by traders and/or their customs representatives.	Data contained in cargo and goods declarations allow Customs to learn from the traders' historical activity, so that they can conduct predictive analytics with statistical rigor aimed at forecasting their behaviour in future operations.
Results of inspections carried out by customs offices	Results of inspections can be fed into data warehouses and risk management systems and may give rise to follow-up action.	In developing countries, additional information can be extracted by the certificates issued by private inspection companies conducting pre-shipment inspections.
Data and information received from other cross-border regulatory agencies and/ or regional organisations	An example of data and information received from regional organisations is given by the 'INF-AM' notifications, which are communications periodically sent by the European Anti-Fraud Office (OLAF) to the customs administrations of the European Union (EU) member states reporting on cases of fraud related to the evasion of customs duties.	Once received an INF-AM notification, the concerned customs administration is entitled to re-route the result of its risk analysis from the green (no control) or yellow (documentary control) to the red channel (physical inspection of the consignment).
Databases owned by Customs and other data held by Customs	These are, for instance, data generated during pilot tests of new procedures, and databases on customs value, on counterfeit goods or goods suspected of violating other intellectual property rights, on customs decisions regarding the classification and origin of goods, lists of registered or trusted traders (e.g. AEO), etc.	Customs valuation databases are typical risk assessment tools used by customs administration to assess potential risk regarding the truth or accuracy of the import declarations submitted by operators, by comparing the declared value to previously accepted customs values. In the EU a central AEO database is available to the customs authorities of the EU member states that allows the adjustment of the results of risk analysis for AEO-traders through a reduction of risk scoring for such operators at the moment they lodge a customs declaration.
Internal analysis or studies	An example is given by the time release studies (TRSs), which are used by Customs for re-engineering their procedures to further facilitate trade. Another example is the mirror analysis (i.e. studies where the imports reported by the importer country are compared to the exports declared by the exporter country in order to detect discrepancies in quantity, weight or declared value leading to the detection of possible irregularities).	Import/export statistics generated by Customs can also be compared with statistics produced by other organisations, such as the Commodity Trade Statistics Database (COMTRADE) of the United Nations or the WTO Data Portal, so to identify further discrepancies (undervaluation or tariff slippage).

| External databases and other open source platforms | An example is the Customs Enforcement Network (CEN), a database developed by the WCO containing information on customs seizures and offences, as well as pictures required for the analysis of illicit trafficking in the various areas of Customs' competence; as well as the WIPO (World Intellectual Property Organization) Global Brand Database, a freely accessible online database with information on trademarks registered under the Madrid System for the international registration of trademarks, on appellations of origin registered under the 1958 Convention of Lisbon on the protection of appellation of origin; and on emblems protected under the 1883 Paris Convention of on the protection of intellectual property. | In the EU, the European Anti-Fraud Office (OLAF) has developed a Container Status Messages (CS) database recording the movements of containers transported on maritime vessels and an Import, Export and Transit (IET) directory containing data on goods entering, transiting and leaving the EU. These tools are aimed to strengthen the analytical capabilities of national customs authorities and OLAF in detecting fraudulent operations.

Another example is the Center for Advanced Defense Studies (C4ADS), a non-profit organisation carrying out analysis on illicit trade and security risks at a global level (available online), which collaborates with the WCO in mapping illegal international trade flows in some key sectors such as: cultural heritage; drugs; environment; intellectual property rights, revenue, safety and security. |

6. Conclusions

Customs administrations deal with an increasing amount of data and information, most of it made available by the recent advances in ICT. In order to efficiently handle such flows, so that they can be strategically used for informing their decisions, two preliminary activities need to be conducted:

1. diagnose and select, on the basis of qualitative criteria, those data and information that are more suitable to guide their activities in an intelligent way

2. integrate and amalgamate such information and data with those already in their possession and with those received from other government agencies, regional and international organisations and private operators.

The second activity, in turn, implies the need to harmonise and standardise data and information collected, possibly by developing appropriate IT infrastructure for their interchange/interoperability (e.g. single window). To this end, it must be remembered that the WCO developed the WCO Data Model, a set of standardised and harmonised data and electronic messages that enables the sharing of information between Customs and other cross-border regulatory agencies and can be used by them to accomplish formalities for the arrival, departure, transit and release of goods in international cross border trade.

Finally, another factor that needs to be considered is the presence of adequate legislation or regulation giving legal force and ensuring the authenticity of the documents electronically transmitted by private operators to Customs and the other cross-border regulatory agencies (e.g. electronic signature laws), as well as legislation or regulation aimed to protect the confidentiality of the data handled by such authorities.

References

Cariolle J., Chalendard C., Geourjon A.-M., & Laporte B. (2017). Going beyond analysis of internal data to support customs modernization: a case study in Gabon. *Études et Documents*, no. 23, CERDI. Retrieved from https://halshs.archives-ouvertes.fr/halshs-01677266/document

Canada Border Services Agency (CBSA). (2018). *2018–19 departmental plan*. Quebec: Public Service Commission of Canada.

Dabbicco, G., & Di Meglio, E. (2011). *Data analysis and national accounts*. Retrieved from https://ec.europa.eu/eurostat/cros/content/ps2-poster-8-ntts-2011-s2_en

Han, C-R., & McGauran, R. (2014). Tracing trails: implications of tax information exchange programs for customs administrations. *World Customs Journal, 8*(2), 3–14.

Jamaica Customs Agency. (2018). Data analysis for effective border management. *Trade Beat, 10*(8). Retrieved from https://www.jacustoms.gov.jm/sites/default/files/docs/Tradebeat/TradeBeat%202018/TradeBeat%20-%20Data%20Analysis%20Updated%20Reissue%20Nov%202018.pdf

Laporte, B. (2011). Risk management systems: using data mining in developing countries' customs administrations. *World Customs Journal*, *5*(1), 17–27.

Mikuriya, K. (2017). Data analysis: seizing opportunities for effective border management. *WCO News*, *82*, 7–9.

National Academy of Customs, Indirect Taxes and Narcotics (NACEN), India. *Data warehousing and data mining for improvement of custom administration in India – lessons learnt overseas for implementation in India.* NACEN. Retrieved from https://nacen.gov.in/resources/file/downloads/569778c84ce95.pdf

Victorien Gnogoue, R-C. (2017). Mirror analysis, a risk analysis support tool for customs administrations. *WCO News, 82*, 16–18.

World Customs Organization (WCO). (2018). *Draft guidance on data analytics*. Brussels: Permanent Technical Committee, WCO.

World Customs Organization (WCO). (2018). *Guide to measure the time required for the release of goods*, (version 3). Brussels: WCO. Retrieved from http://www.wcoomd.org/-/media/wco/public/global/pdf/topics/facilitation/instruments-and-tools/tools/time-release-study/time_release-_study.pdf?db=web

World Customs Organization (WCO). (n.d.) *Sharing connecting & bridging*. Brussels: Customs Enforcement Network (CEN) brochure. Retrieved from http://www.wcoomd.org/-/media/wco/public/global/pdf/topics/enforcement-and-compliance/tools-and-instruments/cen/cen-brochure.pdf?db=web

Notes

1 Predictive analytics is defined as analysis that uses data and algorithms to answer the question 'Given past behaviour, what is likely to happen in the future?'

2 An example of how the exchange of information between Customs can simplify trade is the development a web platform, or of interlinked national databases through which two or more customs administrations can verify the authenticity of the documents submitted by traders (e.g. certificates of preferential origin). In this context, it is also worth mentioning the initiative developed by the ICC World Chambers Federation of a website that offers a list of accredited Chambers of Commerce and Customs Authorities the possibility of verifying the authenticity of certificates of (non-preferential) origin online (https://certificates.iccwbo.org).

3 WCO, 'Guide to measure the time required for the release of goods', Version 3, 2018. With particular regard to the Central Asian Countries, WCO also developed, under the Central Asia Regional Economic Cooperation (CAREC) Program a 'Guide to Time Release Studies for Central Asian Countries'.

Danilo Desiderio

Danilo Desiderio is an international expert specialising in customs regulation and trade facilitation. He is head researcher of the Centre for Study and Services of the Italian Council of Customs Brokers, serves as consultant to various institutions, organisations, governmental agencies and trade associations in the EU and abroad, especially in Africa.