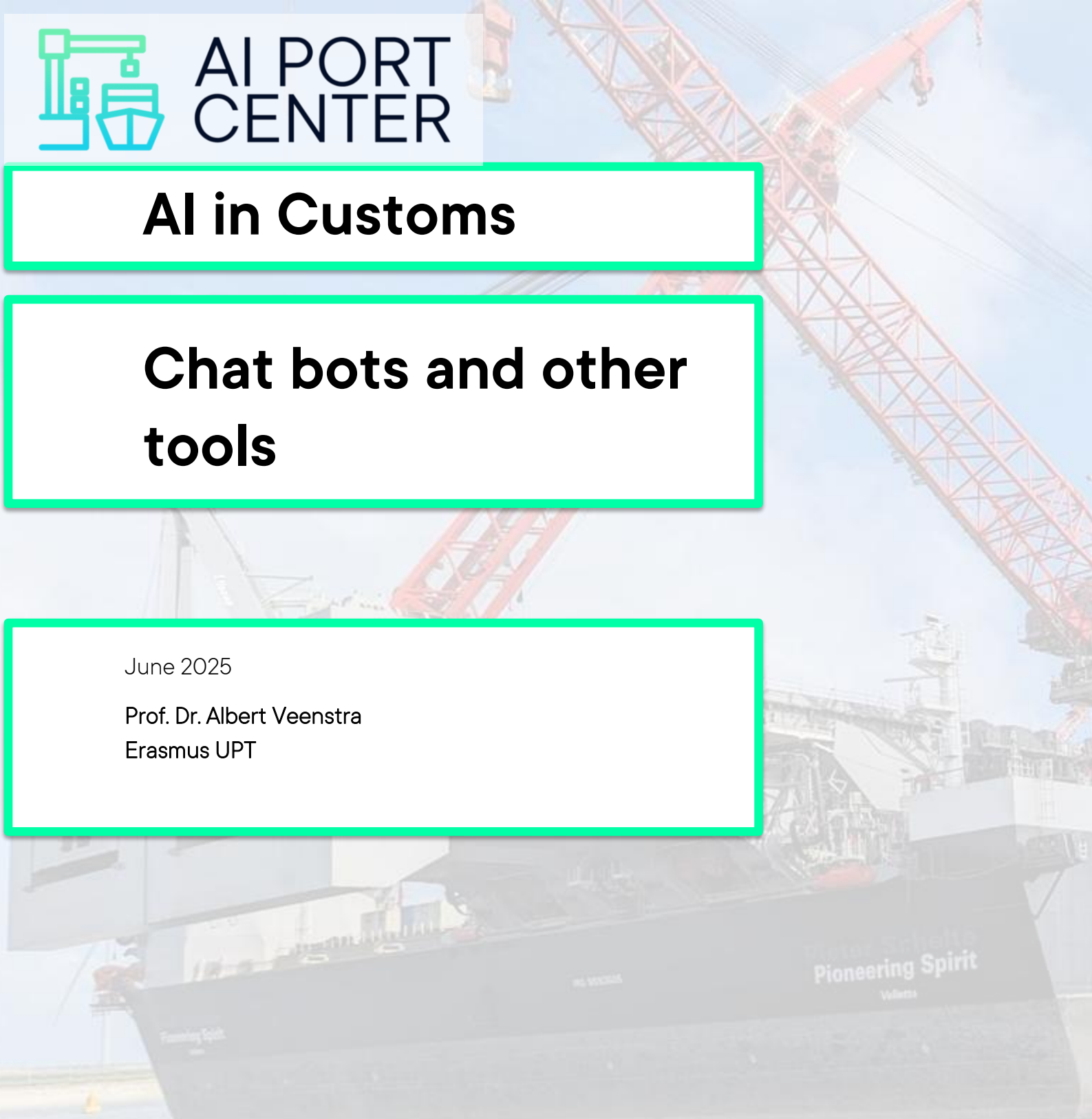
AI PORT CENTER

AI in Customs

Chat bots and other tools

June 2025

Prof. Dr. Albert Veenstra
Erasmus UPT



TU Delft

Erasmus University Rotterdam

Contents

1. Assignment

The original proposal for this project was submitted in Spring 2024 with the aim to look into current development of chat bot technology in the customs domain. The work in this proposed project could only partially be carried out. The development of the chat bot technology was carried out by a partner in this project, but the case study on applications in Dutch Customs did not get much traction. As a result, the direction of the project was re-oriented towards a more general look at AI applications in the customs domain. Below we further elaborate on this. This work was sponsored by the AI Port Center, which is part of the Convergence program from TU Delft and Erasmus University.

In this report, we will provide a description of the customs chat bot technology that is being developed by the development team of Lithuanian organisation Customs Clear, under the direction of Enrika Naujoke. This organisation has been in direct contact with Dutch Customs to demonstrate the approach and discuss possible collaborations. Parallel to this, the plan was to discuss the need for chat bot technology with representatives of both customs and trade in the Netherlands.

In the course of 2024, some interaction with various customs professionals revealed that Dutch customs is working on its own chat bot application, and that this application is part of the disentangling of Dutch Customs and the Dutch Tax Office, more particularly the disentangling of the customer service telephone service. This is an ongoing process, however, that has its own speed and dynamic, which made this process unsuitable to study in an academic setting.

To redirect the project, two decisions were made: the budget was reduced somewhat, and the focus was shifted to include the study of the potential of other AI applications. The latter was primarily oriented towards advanced data analytics in the customs domain. Part of the work in this new section of the project was carried out through an Honours project in the MSc Supply Chain Management at the Rotterdam School of Management for customs software company Fiton BV. In addition, discussions were held at the Trade Compliance Conference on 17 June 2025 in Veenendaal, co-organized by evofenedex and Customs Knowledge, on the use of data in the customs domain.

This report provides an overview of these various project activities, as well as a brief outlook on AI applications in the customs and trade domain in the future (with thanks to insightful comments of Ramunas Sablinskas of CustomsClear).

2. Developments in Chat Bots

In this section, we will provide an insight in current developments regarding chat bots in the customs domain. We will not present a deep dive into the chat bot domain in general. However, we will start with some early ideas on chat bot applications in the customs domain and describe how this technology is currently made available for customs practice.

Origin

For the last fifteen years, a large number of research projects have explored new ways of applying physical technology, data exchange and analysis tools and business concepts to the customs domain. Early project were EU funded projects such as IT-Aide, INTEGRITY, Cassandra. In these projects, much of the attention was directed towards data exchange concepts, such as the now famous 'data pipeline' model[1].

In parallel to these projects, one of the more visionary people of Dutch Customs, Frank Heijmann, also developed ideas about the broader enforcement strategy for customs, as well as all kinds of new ideas about technology and tools customs would need to efficiently carry out their role. Somewhere around 2014/2015, this led to the idea to create a automated customs assistant called 'Robin'. At that time, this was no more than a look and feel design in a customs strategy presentation. However, in the core, this idea was based on the insight that much of the declaration based interaction between business and customs is 'repeat-business' where the same processes are carried out over and over again. This also means the same problems occur, and the same questions are asked over and over again. This is why, currently, Dutch Customs has a call centre in place, where these questions can be addressed, and possibly answered satisfactorily.

Automating this question and answer process makes a lot of sense, especially when this is already happening in many day to day situations. In 2014/2015 this was relatively new, but nowadays, chat bots are virtually everywhere.

CustomsClear

CustomsClear is the brand name of a Lithuanian company, CC Learning, that offers educational and informational services and solutions in the customs and trade domain. CustomsClear has created a knowledge sharing platform for customs professionals: www.customsclear.net. Part of its activities is a software development company for advanced customs applications.

As a basis for its applications, CustomsClear has been working on developing suitable applications using AI techniques for the customs domain. The unfiltered application of chat bots in the customs domain will not produce acceptable results, since the answer generating engine in these tools is essentially a probabilistic model (the so-called large language models) generating readable text. This means that similar or exactly the same questions will result in different answers, depending on how they were prompted, when they were prompted, and in what sequence of questions they were asked. This is not acceptable in the legalistic domain of customs and trade practice.

CC Learning is project lead of a EU funded project called "Development of personalised adaptive systems for the continuous learning of customs professionals". This project runs from 2023-2026, and has led, among others, to a proof of concept for Lithuanian Customs of a working AI assistant. The core technology for AI assistants developed in the context of this project is called Picoco.

Picoco[2]

[1] Klievink, B., Van Stijn, E., Hesketh, D., Aldewereld, H., Overbeek, S., Heijmann, F., & Tan, Y. H. (2012). Enhancing visibility in international supply chains: the data pipeline concept. *International Journal of Electronic Government Research (IJEGR)*, *8*(4), 14-33.

[2] See also www.picoco.ai

This report is too short to do full justice to the technology that is driving the Picoco AI Assistant platform. We limited ourselves to some remarks on the core structure of the tool, and some applications.

The customs domain in the EU has a number of challenges. One of these is to obtain answers to questions that are grounded in the formal legal texts of relevant customs and trade regulation. This requires questions to result in answers that are legally exact and unchanging. A second challenge is that the regulation in the EU is published in 24 official languages. This means that, even though these documents are considered to be equal from a legal point of view, translation from one language to another can produce slightly different meanings in the answers of questions as well.

In the Picoco platform this was addressed by building the platform on the basis of two pillars. The first pillar is a so-called 'knowledge box' that contains the relevant legal regulations. Questions and answers on the basis of this information is not based on large language models, but on database search technology. This means that in this part of the platform, the information to answer questions is exact, and can be linked back to source documentation.

The second pillar is a large language model that will be fed with the questions ànd answers, but that can be directed to structure the answer text according to a variety of templates and forms. The answers can be supplied in any of the 24 EU languages, regardless in which language the questions were asked.

On top of this, the tool is configured in such a way that it only provides answers to questions related to the customs domain. This prevents the tool to create answers for which it has no knowledge base.

Lithuanian Customs has started to use this tool from 31 December 2024 onwards. They have named their application Matas.

Demonstrations

The Picoco tool was demonstrated to both Dutch Customs and to Dutch customs software provided Fiton.

Current state of affairs on chat bot and other AI development in Dutch Customs

The development of AI based tools and applications is not necessarily a new development in (Dutch) Customs. For more than two decades, advanced scanning equipment has been introduced in the customs domain that generates pictures – with rather large file sizes - which require analysis. This development got a strong impulse after 9/11.

More or less in parallel, automated image recognition solutions were being developed already. As part of this development process, customs agencies recognized that more data results in better learning. Therefore, projects intended to combine output of multiple scanning operations in various customs agencies and training tools together were organized as well.

Currently, the state of this technology, including image recognition, has advanced sufficiently to proceed with the development of new supervision and enforcement concepts. One important reason for this is that any customs agency, and therefore also Dutch Customs, is facing personnel shortages. This means that relying on trade to carry out strategically chosen verification and control activities is part of the current customs operational strategy.

From a supervision strategy and risk management perspective, this means that modern supervision entails that tasks such as scanning and possibly first line inspection may be carried out by businesses, while second or third line inspections are still carried out directly by customs professionals. These kinds of concepts are currently the subject of innovation projects. The main insight that derives from this is that the introduction of AI tools will, inevitably, result in organizational adaptations. Introducing AI tools potentially requires changes in team structures, managerial controls, and the distribution of responsibilities.

Further innovative developments in the area of AI at Dutch Customs are related to chat bot applications for the 'belastingtelefoon douane'. Here, Dutch Customs followed the same principles as Customs Clear. Part of the innovation trajectory also included some interaction with Lithuanian Customs, who are using the technology developed by Customs Clear.

There are two insights that merit mention here. First of all, questions asked by businesses to Customs are not always of a legal or technical nature. Rather, they are procedural: when is an inspection planned, when can a decision be expected, when will the customs professional come to our premises? This means that the basic regulatory documentation is not always a sufficient basis to answer these questions.

A second insight resulted from some experimentation. Hallucination of chat bots is a well-known problem[3]. Dutch customs has also experienced this in the early applications of chat bots, where even with careful 'knowledge boxing', answers would still be generated based on generic input from the internet. This means that the development of these types of chat bots needs to proceed carefully, and a good first application step is probably to keep the use within the professional customs organization, and only in a second phase, open up the tool for the trade public. In this way, customs professionals can use, but also weigh in on, the results generated by the chat bot platform, and field potential erroneous outcomes.

Finally, as any other organization, Dutch customs sees some merits in the use of large language model applications (Copilot, ChatGPT, Grok and the like). Policy advisors, for instance, frequently need to summarize various regulatory documents for policy or decision making processes. It is already known that, with some care, this can be done with the widely available AI tools. Until recently, however, these were not available in the Customs IT toolbox. Some sanctioned experimenting with these kinds of tools is currently ongoing.

Two final remarks on challenges related to the development of AI tools in the customs domain. The first of these is that the current state of affairs has shown the continuing need to have experienced humans (previously *het geoefend oog*) involved during the application of AI-based selection and detection models. These models do learn, but they also make mistakes, especially in early stages of application. These mistakes need to be 'trained away' in order to manage or reduce the workload in follow-up processes. For every new tactical target or risk, this human-machine based reinforcement learning cycle needs to be present in order to obtain optimal outcomes. Another reason why training and retraining will remain relevant, even in the long run, is that the composition of trade flows is not stable. Product groups appear and disappear, or the value, design or origin of goods changes. This means that adjustment of the 'knowledge base' of AI applications in the trade compliance domain will always be necessary.

A second remark is that Dutch Customs already sees the productivity benefits of the current use of AI tools. This type of gains is an important driver for innovation activities in general. This means that 'potential productivity gains' may be a dominant criterion to look for new, currently unexplored, applications. Any sort of repetitive, semi-standardized process, may qualify for at least an exploration of AI tools. Processes where spoken information needs to be recorded in databased may benefit from a wider use of speech-to-text tools. Second- and third-line inspection processes (collecting additional documentary information, decision-making on deconsolidation of pallets and boxes, preparation of appeals procedures may all at some point benefit from AI applications.

[3] See, for instance, Metz, C. (2023). Chatbots May 'Hallucinate' More Often Than Many Realize. *International New York Times*, NA-NA, 14-11-2023.

3. The road towards AI applications in the Customs Domain

Chat bot applications are not the only option for AI tooling in the customs and trade domain. The use of AI tools in reading X-Ray scanning images is currently being developed and filters into practice. See for instance, reports on this by Dutch Customs[4], Lithuania Customs[5], Taiwan Customs[6]. AI tools can also provide support in reading paper documents, automate data cleaning and consistency checks, and generate parts of entire declarations according to all kinds of different formats. This is not yet common practice, however.

Transport and logistics software company Fiton, based in Zwijndrecht provides a suite of software tools for forwarding, warehousing and customs. Fiton is exploring the possibilities to enhance their software product for the customs domain with AI based features. They have introduced this challenge to the RSM Honours-program, where a group of students will investigate this challenge and come up with some solutions.

Their report is briefly summarized here. The full report is attached to this report as an annex.

Introduction
The generic process of reporting to customs in international trade is based on declarations. It is well known that this is a domain in which errors are frequent and may occur for a wide variety of reasons. If this industry is moving forward with self-learning mechanisms and other applications of artificial intelligence, a better understanding of the occurrence of errors is required.

This study is carried out in collaboration with customs software provider Fiton. Fiton has a vision to embrace AI applications in the development of its declaration platform. As a first step, this study will take a deep dive into the classification and structure of errors in customs declarations as they are processed by clients of Fiton in their declaration platform. This error classification model will be the outcome of this study.

Literature review
The literature view provides an overview of various methods to help identify errors in customs declarations. The idea behind this is that the AI tools developed in the customs domain will primarily contribute to the identification of (human) errors.

Part of the input for this is some literature on classification and error codification. The AI literature indicates that tools such as a confusion matrix are helpful to evaluate the performance of algorithms. Customs itself employs a mix of rule based systems, intelligence based risk profiling and expert judgement to identify errors in declarations and shipments. The current customs declaration system DMS has a large number of built-in controls on the input of businesses, and this results in some 260 standardised error types that this system can identify.

Some further study reveals that both industry and governments have taken a wide variety of measures to reduce or eliminate errors. A final step in this process is the voluntary disclosure of errors that Dutch Customs allows for companies who report previously unidentified errors in their declarations in good faith.

Error classification
This project has developed an approach that allows Fiton to shift from a error correction process to an error prevention processes through four steps: detection, classification, prioritization and reduction of human errors. The first three of these are carried out as part of this project.

Detection in this project is done on data obtained from the Fiton system, but augmented with error codes based on errors identified by customs. Classification was done based on human error theory to obtain a sound and robust error classification. The approach to establish prioritization of errors was grounded in failure model and effect analysis

[4] Smarter Supervision with the Help of AI Models and Autodetection | Articles | About Customs Administration of the Netherlands
[5] Lithuanian Customs is developing an AI-based X-ray image analysis system · Novian
[6] Applying AI assistance in X-ray Inspection-Customs administration

combined with multi-criteria decision making models. As a weighting method, the M-CRITIC method was used in combination with ABAC (see Sharkasi & Rezakhah, 2022[7]).

The application of this approach was done on a data set of 350.000 declarations between February 2022 and June 2025. In total, this sample contained declarations flagged with some 98 different error codes, with the top 10 errors representing about 80% of the sample. The declaration data allowed for some calculation of the relative importance of errors, by looking at the total value of the declaration.

The top 10 errors included errors such as 'document missing', 'verification deadline unknown', 'document not permitted', 'missing invoice', and various codes for value errors, as well as the code 'more than 10 validation errors'. This seemingly structured communication on errors turned out to be only partially structured. In many cases, however, the error code was not the final answer to the rejection of the declaration. Some further work therefore has to be done on creating a proper error classification overview.

Eventually, the following four error clusters were identified:
1. Missing or invalid data
2. Missing of invalid type of document
3. Logical inconsistency
4. Structural issues

Logical inconsistencies follow from requirements that certain declaration fields need to be filled in together, with related data. Structural issues follow from requirements on timing, deadlines, or overall submission problems where the number of errors exceeds a threshold. Together these four clusters hold the 30 most frequently occurring error types.

The application of the M-CRITIC method in combination with ABAC then results in a ranking of the most significant errors. The error that comes out as the most important is FM2 'verification deadline unknown/preceding document expired'. The other two errors in the top-3 are also related to documents. The error ranked lowest in importance (in the top 10) is an error related to weight inconsistencies (net weight > gross weight).

The analysis also identified from challenges and shortcomings. For instance, the detection complexity was difficult to measure without detailed access to the system data. Also, data completeness was a limitation. Much of the timing information, for instance, was not available: time between declaration submitted and response of customs, for instance. More accurate measures for the magnitude of the impact than just customs value should also be developed.

Discussion

The study shows that formal concepts for error classification are useful in the customs domain. Many of the errors conform with predictions from human error theory. Most of the important errors are related to documentation and required information.

This information on the translation of error codes reported by customs to the platform of Fiton should contribute to a better error detection service on top of the declaration submission engine of Fiton. Part of the immediate future work is to make system information, such as time stamps, more easily available for error detection and analysis.

[7] Sharkasi, N., & Rezakhah, S. (2022). A modified CRITIC with a reference point based on fuzzy logic and hamming distance. *Knowledge-Based Systems*, *255*, 109768.

4. A note on the role of data quality

This section is based on the keynote address at the Trade Compliance conference 17 June 2025 Veenendaal, and the discussion on 'data quality as a service' in a break-out session.

The domain of customs and trade is perceived as a data rich domain (see, for instance, the premise of the EU project Profile[8], but also the discussion in the previous chapter). It is therefore not surprising that complex new methods and tools are contemplated for both government agencies and business practice to deal with this wealth of data. One area where this is particularly conspicuous is the field of border crossing e-commerce.

This chapter therefore examines the origin of some of these advanced data management ideas, and the way these ideas are introduced in the customs and trade regulation domain. In this chapter, we will examine the possible consequences for companies, and the possibilities for no regret decisions.

Data in global supply chains

This is not the place to present an extensive review of all literature on data use in supply chains. However, as a *pars pro toto* we refer to Kamble & Gunasekaran (2020)[9], who present an overview of literature on the use of data in supply chains. They direct the attention almost exclusively to the application of advanced data tools in the development of performance measures and metrics with the aim to enhance the performance of the supply chain. They find that the current body of literature suggests that the impact of advanced data analytics is positive.

If the search for literature is extended from keywords Data Analytics + Supply Chain to include Compliance, a few relevant papers are found. For instance, Ramachandran et al. (2021)[10] suggest that in compliance management in supply chains, these kinds of advanced analytics tools also provide new ways of managing adherence to regulation. This holds both for businesses and for government agencies that have to enforce the rules. Specifically, they argue: "… in data-driven digital supply chains, the regulators may be able to leverage digital technologies, such as artificial intelligence and machine learning to automate the compliance verification process"[11].

Apart from this forward thinking research, the interaction between data, and data quality, one the one hand, and business operations on the other has also been studied on a more pragmatic level. Bozic et al (2024)[12], for instance, argue that there is a definite link between the quality of product master data and logistics performance. In the customs domain, this is perhaps stating the obvious. Every customs and trade practitioner knows that problems with data (i.e. declarations or the underlying documentation) can potentially have significant operational consequences when goods get stuck at a border.

To complete this really brief overview of data and supply chains, we point out that there has been some fifteen years of research on solutions, particularly in the customs domain to resolve the acquisition and transfer of data with the purpose to avoid experiencing operational consequences. The so-called data pipeline (see Klievink et al, 2012 footnote 1), is one of these ideas. In later years, businesses such as Maersk and IBM tried, and failed, to develop a practical, global, version of this, in the Tradelens project[13]. Currently, regulatory developments originating from the circular economy program of the EU, called Ecodesign for Sustainable Products Regulation (ESPR)[14] puts forward a new idea for data exchange in

[8] www.profile-project.eu

[9] Kamble, S. S., & Gunasekaran, A. (2020). Big data-driven supply chain performance measurement system: a review and framework for implementation. *International journal of production research*, *58*(1), 65-86.

[10] Ramachandran, G. S., Deane, F., Malik, S., Dorri, A., & Jurdak, R. (2021, December). Towards assisted autonomy for supply chain compliance management. In *2021 Third IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA)* (pp. 321-330). IEEE.

[11] Ibid.

[12] Božić, D., Živičnjak, M., Stanković, R., & Ignjatić, A. (2024). Impact of the Product Master Data Quality on the Logistics Process Performance. *Logistics*, *8*(2), 43.

[13] A.P. Moller - Maersk and IBM to discontinue TradeLens, a blockchain-enabled global trade platform | Maersk

[14] Ecodesign for Sustainable Products Regulation - European Commission

supply chains in the form of a so-called Digital Product Passport. This device might also bring new ways for supervision agencies at the border to receive or gain access to detailed product data in international supply chains.

Data for compliance

The insights from academia and the multitude of projects involving customs agencies have filtered into the collective conscious in a recent initiative to review customs supervision for the European Union. This so-called Wise Persons Group delivered its vision for customs in the EU in 2022 with the report "Putting more Union in the European Customs; Ten proposals to make the EU Customs Union fit for a Geopolitical Europe".[15]

In this document, 'data' plays a prominent role: in the 50-page document, the word *data* is mentioned 190 times. The general vision towards data is that the customs and trade domain represents risks for society that need to be mitigated, and collecting and processing data is an effective way to do that. There are challenges, however:

> "The poor availability and quality of the data submitted to Customs, the lack of a common data warehouse, and the low level of data sharing across Customs administrations lead to fragmentation and makes it extremely difficult to properly manage risks through data analytics at both national and EU level."[16]

This statement introduces the idea that data in itself is not enough: the quality of the data has to be of sufficient level, in addition to properly collecting and storing it, and sharing it for different enforcement responsibilities. While this sounds rather obvious, in the customs domain this is a relatively new idea. That is because from a legal perspective, customs law requires the declarant to submit a correct declaration. In other word, data in declarations has always been considered to be 'faultless'. In practice, of course, this is not the case. Putting this insight forward in relation to a review on customs practice in the EU is therefore a bit of a watershed moment.

It may not come as a surprise, therefore, that one of the ten recommendations of the Wise Persons Group is to develop a new approach to data, which includes ideas such as cross-validation, but also a centralised EU facility to collect and store data.

In an appendix to the report, some further ideas are discussed related to data quality:

> " … Data quality can be improved by developing algorithms for checking the accuracy of data submitted or by creating incentives for filling in accurate data. Customs procedures also need to be simplified and data required from trade must be assessed and rationalised by deleting unnecessary data requests, eliminating redundant requests and adding the necessary but missing data elements."[17]

This statement is relevant because it introduces the idea that data verification can be subject to automation, just as has already been suggested in the academic domain (see Ramachandran et al, 2021, ibid). In addition to this the statement shows the rather tedious challenge of steering the data requirement towards a better collection of relevant data, where now data is received that is unnecessary, while some missing data still needs to be collected. To some extent, *better* data is also *other* data.

These early ideas earned a central role in the practical solution for a better collection and storage of data: the EU Customs Data Hub. This data hub is a central pillar in the current EU customs reform. See, for a more detailed discussion on the idea of the data hub, the website of the EU Customs reform[18].

Explicit in the preliminary design of this data hub is the use of machine learning, artificial intelligence and human intervention to achieve – for authorities - a 360-degree overview of supply chains and the movement of goods from and to the common market. There benefits offered in return: interacting with only one portal for all of the EU, and, for those

---

[15] https://taxation-customs.ec.europa.eu/document/download/e5326383-2e8d-4d0e-9025-ddf262e9df6e_en

[16] Ibid, pg 22.

[17] Ibid, pg 44.

[18] EU Customs Reform - European Commission

businesses who can be fully transparent, significant simplifications in the obligation to submit declarations to authorities (the so-called Trust & Check status).

Guaranteeing the quality of data

The narrative until now is that the trade and customs domain is a data rich environment, and that theoretical and academic ideas of how to leverage this wealth of data are filtering into the customs domain, and are becoming very concrete elements of the new vision of customs regulation in the EU. In other words: not only data, but also advanced data analytical tools will become a reality in the future of customs and compliance operations.

This discussion requires two further pieces of the puzzle. First of all, while there is high degree of faith in the problem solving capabilities of advanced tools, especially the use of AI, it is also becoming well known that AI has drawbacks. IBM, in one of its opinion pieces, discusses so-called AI bias[19]. AI bias is the occurrence that the application of AI tools may result in biased or skewed outcomes, and in this way distort output and potentially cause damage to people, processes or businesses. IBM attributes this to 'human biases that skew the original training data or [the] AI algorithm'. In other words, our world is not neutral or unbiased, and data collected from that world will inevitably transfer this bias to the application of AI tools.

Examples of this relatively close to home are noted in the report of the Accounting Office of the city Rotterdam in 2021 called 'Gekleurde technologie' in which it was noted that some of the decision supporting tools used in the municipal administration could result in untransparent and prejudiced decision, severely disadvantaging already disadvantaged communities and groups of citizens.

The solution, as IBM notes it, is a combination of using the right model, make sure the development teams are inclusive, monitor results closely, ànd 'train with the right data'.

This forms a bridge to the second piece of the puzzle: identifying the 'right' data in the customs domain. We already know for quite some time that the quality of the data customs received through declarations is problematic. The Wise Persons Group concluded this, but authors such as Hofman et al (2021)[20], also argue – in the context of increased use of data analytics by governments in general – that quality of data, as well as value derived from data intensive methods need to be addressed. Heijmann et al (2020)[21] discuss the same point, more specifically in the context of customs. A solution, for government agencies such as customs, it was thought, was to develop and rely on additional sources beyond the data they are legally entitled to. These data sources, according to Hofman et al (ibid) may exhibit the same problems, however.

Data quality issues in the customs domain are generic and globally recognised. Nugraha et al (2024)[22] discuss a case of data quality in Indonesia. Mayega et al (2024)[23] discuss the data quality situation for the Uganda Revenue Authority.

This situation is aggravated by ongoing developments in EU regulation (and many other countries) that attempt to attain societal goals through stricter regulation: eradicating child labour, avoiding deforestation, solving the carbon leakage problem inherent in the EU-ETS system, and so on. This type of regulation introduces new data elements (and new data formats) that need to be collected, transferred from high up the supply chain to downstream, and – at the border – their presence (if not their correctness) checked by Customs. In other words, we are adding more and more data to the information stream that is the basis for government supervision of business activities.

---

[19] See What Is AI Bias? | IBM

[20] Hofman, W., Migeotte, J., Labare, M., Rukanova, B., & Tan, Y. H. (2021). Using business data in customs risk management: Data quality and data value perspective. In *International Conference on Electronic Government* (pp. 271-287). Cham: Springer International Publishing.

[21] Heijmann, F., Tan, Y.H., Rukanova, B., Veenstra, A. (2020). The changing role of Customs: Customs aligning with supply chain and information management. World Customs Journal, 14 (2)

[22] Nugraha, T. F., Wibowo, W. S., Genia, V., Fadhil, A., & Ruldeviyani, Y. (2024). A Practical Approach to Enhance Data Quality Management in Government: Case Study of Indonesian Customs and Excise Office. *Journal of Information Systems Engineering & Business Intelligence*, *10*(1).

[23] Mayega, J., Waiswa, R., & Nabuyondo, J. (2024). How Clean is Customs Data? Data Management in Uganda Revenue Authority.

The 'AI world' is thinking about solutions. Yu et al 2024)[24] discuss a survey on what they call failure analysis and fault injection in AI systems. Chan et al (2022)[25] present a survey of work on the impact and possibilities to mitigate training machine learning models on faulty data. They point out that there are techniques to deal with such problems, such as Ensemble Learning, where multiple models are training in parallel, and outcomes are summarized in some way (basically averaged). So there is some work going on to recognise that training AI tools on low quality or faulty data can be dealt with. This research is still, however, in its early stages.

Data quality as a service

Here, we aimed to discuss another solution direction: the work the business community can do to improve the quality of their data, and the way in which their service providers in the customs domain could possibly take this on as an additional business proposition.

To start with, the generic customs service provider, the so-called customs broker, is a party that performs a very specific business service on behalf of its clients: it submits customs declarations to customs organizations in its own country, or in all countries of the EU. As part of this service, follow-up interaction with customs agencies (for additional information, checks or cargo hold-ups) may also be involved.

This profession has standardised trading conditions (in the Netherlands called the Dutch Forwarding Conditions[26]), maintained by their industry association Fenex (now a sub-association under TLN). These conditions state, regarding quality of information (art 9:3):

> The Client guarantees that the information and documents that it provides are correct and complete and that all instructions and goods that are made available comply with current legislation. The Freight Forwarder shall not be obliged but shall be entitled to investigate whether the information provided is correct and complete.

In these conditions, the responsibility for quality of information is wholly put on the client (i.e. the party who needs the customs declaration to be submitted), and the broker may, but is not obliged to, check this information. This does create issues, because the broker is, to some extent, liable for any quality issue with the declaration, because they, and not their client, has a direct relationship with customs[27].

So to presume a role of a customs broker as a data manager is not immediately obvious. On the other hand, there is a number of customs brokers that specifically deal with e-commerce flows for primarily foreign (read: Chinese) clients. Their experience is that the data is often so poor that they have to engage in significant data reparation efforts to even be able to submit declarations. They therefore do this, urged and monitored by Dutch customs, but knowing – at the same time – that they do not have much legal coverage for this under the regular Fenex conditions.

On 17 June 2025, a discussion session was held with some 30 representatives of Dutch forwarding and customs broking companies. First some basic insights into data quality as a service was presented. The main elements of a data management proposition are:

1. Data profiling
2. Validation
3. Cleansing
4. Stewardship.

---

[24] Yu, G., Tan, G., Huang, H., Zhang, Z., Chen, P., Natella, R., ... & Lyu, M. R. (2024). A survey on failure analysis and fault injection in AI systems. *ACM Transactions on Software Engineering and Methodology*.

[25] Chan, A., Gujarati, A., Pattabiraman, K., & Gopalakrishnan, S. (2022, June). The fault in our data stars: studying mitigation techniques against faulty training data in machine learning applications. In *2022 52nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)* (pp. 163-171). IEEE.

[26] 20180501-Nederlandse-expeditievoorwaarden-Engels-met-logo.pdf

[27] This holds particularly for the situation where the broker indirectly represents their principal. But we will not go into the complicated issue of direct and indirect representation.

The e-commerce brokers mentioned above clearly offer 'cleansing' services already. Validation is represented in a lot of the built in controls in the current Customs declaration system DMS. Data profiling and data stewardship as possible propositions may be relatively new ideas in the customs domain.

Looking more specifically at data quality technical services, activities include formulating and enforcing data quality rules, monitoring ex-post as well as keeping data quality rule meta data in place, both on the applicability of different rules at different times, but also on cleansing activities performed.

The room was then confronted with a number of questions regarding this proposition. These were presented to them by Mentimeter. In total 24 persons present answered the questions.

Question 1: Are there presently data quality services?

- No, we don't (33%)
- Yes, we have this/we do this (33%)
- Thinking about it (25%)
- No answer (8%)

The parties answering yes here were a number of customs brokers offering their service to e-commerce parties. These brokers explained that they have to work on data quality improvement as a condition to their business. The majority, however, does not really experience with data quality related services.

Question 2: Can data quality be a service in the customs domain?

- Yes (92%)

This question clearly shows that the sector, and its service providers see the potential, even though this question does not suggest they should be the party to offer these services (see for that question 4). The next question therefore tests an import ingredient of a data quality management approach: the data quality rules.

Question 3: Does this sector have *Data Quality Rules?*

- Yes, these come from Customs (39%)
- Yes, some companies have them (30%)
- No, not really (30%)

This question reveals that, although companies see the potential, they seek the origin of data quality rules not among themselves, but 'outside': with Customs, or with other businesses.

Question 4: Can the customs broker offer data quality services?

- Yes, definitely (74%)
- No, certainly not (26%)

The sector seems to think, in majority, that the classic customs broker could engage in data quality services. Part of this group, presumably, are the brokers who already do this. There is no consensus, however. There are parties who think the customs broker should not do this.

Finally, an open question was asked about the most attractive element of the proposition. This resulted in a variety of comments, with some referring to a specific proposition ('setting data rules and cleansing would be a nice proposition'), and many others just focusing on a single part of the proposition: cleansing, monitoring, profiling. Only a single comment was made with respect to stewardship of data.

In conclusion

In conclusion, this chapter argues that data quality is a critical element in trade and customs compliance, and that this will only gain more importance going forward. The customs and trade ecosystem sees the opportunities for data quality services as a new business proposition, and in some cases, this business proposition has already developed as part of a greater service offering.

However, the ideas to develop this proposition as a general dimension of customs broking is not very strongly present yet. The industry sees the potential, but the thought process to make this more concrete is in its very early stages.

5. Closing remarks

The report takes stock of the developments of AI applications in the customs domain. Two main developments can be observed. The first is a long running program to develop Ai based tools to analyse X-ray images of scanning equipment. For this type of application, models and approaches are reaching maturity, and this will drive further and more advanced AI applications in the customs domain.

The second development is to use chat bot technology for 'frequently asked questions' in the highly legalised international trade arena. Here, the technology has to achieve two aims at the same time: first to generate answers that are correct and consistent, and second, to present these answers in specific ways suited to the party who asked the question. The latter means that templates answers may be called for, but also answers in multiple languages. Current chat bot technology developed for the custom domain, for instance by a consortium led by CustomsClear in Lithuania can do this.

The introduction of advanced technologies, such as AI applications, also brings a heavy reliance on the quality of data. This report therefore also reports on two activities that delve into this in a bit more detail. The first of these is a project by RSM students for a customs software company with the aim to identify errors in customs declarations. This study found that standardized error identification and classification is very well possible, especially if system information from the declaration submission platform is leveraged.

The second activity is an exploration on the potential development of data quality management as a service. This activity presents a desk research analysis as well as some first results of a small survey among customs brokers and other customs service providers on the merits of data quality management as a new business model for customs brokers. The general observation is that this is definitely a possibility, but much work remains to be done to formalize the general conditions for data quality management in the customs domain.

In conclusion, we can say that AI applications are gradually being introduced in the customs and trade domain, with some successful applications, such as chat bots and image recognition. There are also high expectations of further, advanced, applications on both the business and government side. The new customs regulation revision will drive this to a considerable extent. This will, however, also require a more detailed discussion on the quality of the underlying data, development of new business models for data quality management and models for formalized data improvement. Here, much of the work remains to be done.

What will the future bring?

Much of the discussion in this report reflects current developments, and part of a vision on how customs and trade works that has not essentially changed from hundred years ago. The major advance, with the Union Customs Code in 2016, is that digital submission of declaration data became the standard.

There are developments, however, that might put the movement of goods in a new perspective, much in the same way as the movement of people has become relatively simple. In the same way as people can have a 'passport', goods or shipments may, in the future, also obtain a passport and in this way prevent any other type of documentary process to move them across borders.

The basic construction of this 'passport' for goods is already underway: the European Ecodesign for Sustainable Products regulation contains the notion of a so-called digital product passport. The purpose of this regulation is to gather and provide the information required for more effective recycling. We have argued elsewhere (Veenstra et al., 2025[28]) that this way of employing digital product passports may also satisfy all kinds of other new regulatory requirements, coming our of the regulatory initiatives such as the Corporate Sustainable Reporting Directive (CSRD), and the EU Deforestation Regulation (EUDR).

[28] Veenstra, A.W., Rukanova, B. & Tan, Y.H. (2025). EU Deforestation regulation, challenges for businesses and authorities and the potential role of Digital Product Passport and Digital Infrastructures. Report for Topsector Logistics.

Currently, the EU Customs revision does contain ideas for a centralised data hub, but this does not include a vision on the facilitation of digital product passports … yet. We think rethinking the wider use of digital product passports could possibly be the real EU customs revision that will make the call from the Wise Persons Group a reality.

Appendix

Report by Ane Babibanga, Babette Hulsmann and Kristyan Lolev (June 2025). Honours project, Rotterdam School of Management.

AI PORT CENTER

TU Delft

Erasmus University Rotterdam