Revenue Administration's Strategy on Artificial Intelligence

Final Report

Technical Support Instrument

Supporting reforms in 27 Member States

Funded by
the European Union

This document was produced with the financial assistance of the European Union. Its content is the sole responsibility of the author(s). The views expressed herein can in no way be taken to reflect the official opinion of the European Union.

The project is funded by the European Union via the Technical Support Instrument, managed by the European Commission Directorate-General for Structural Reform Support.

This report has been delivered in August 2022, under the EC Contract No. SRSS/2018/01/FWC/002. It has been delivered as part of the project “Preparing an artificial intelligence (AI) strategy for the Estonian revenue authority”.

Directorate-General for Structural Reform Support
REFORM@ec.europa.eu
+32 2 299 11 11 (Commission switchboard)
European Commission
Rue de la Loi 170 / Wetstraat 170
1049 Brussels, Belgium

Table of Contents

Table of Figures

Disclaimer

1. Executive Summary

This project aimed to explore the feasibility of using human-centric AI within the Estonian Tax and Customs Board (ETCB) and to develop a strategy to do so. It was requested by Estonia through the Technical Support Instrument (TSI) provided by the European Commission's Directorate-General for Structural Reform Support (DG REFORM), in the context of the DG's mission of supporting the preparation and implementation of growth-enhancing administrative and structural reforms within the Member States.

To assist the ETCB to improve its capacity to design, develop and implement reforms concerning digitalization and AI, this project was conducted across five phases. The inception phase initiated the project and aimed to ensure that the stakeholders had a common understanding of the approach to the assistance that was to be provided. During the second phase, the feasibility study, the team gathered examples of AI use cases implemented in public institutions and created a high-level mapping of the ETCB processes (across 11 departments and 4 sub-departments). Using the use cases and the pain points highlighted in the processes as inspiration, 44 potential applications of AI were identified within the ETCB. A prioritisation exercise was performed to rank identified opportunities, with the 5 most promising topics analysed further in a business case.

This was followed by the development of strategic choices and an action plan in the third phase, forming an indicative roadmap to move from the current state to the desired future state of the ETCB. Two AI-driven solutions were selected for the fourth phase to create technical descriptions, including requirements, desired functionalities, a visual prototype, and a proposed architecture that would support the solutions. Finally, in this closing phase summarised in this deliverable, the project is assessed against the planned outcomes and recommended follow-up actions are documented.

Through this project, the ETCB has developed a clear roadmap for delivering the set of potential AI solutions explored, as well as addressing specific enablers to deliver future solutions. Additionally, it has gained technical descriptions and a visual prototype for two promising new AI solutions that could be developed. Furthermore, the activities performed have allowed the ETCB to develop an improved understanding of the feasibility of using AI and improved strategic capabilities concerning the integration of AI, providing the necessary encouragement to reach a new level of AI maturity.

2. Introduction

This chapter covers the following elements:
The introduction of the deliverable and its context
The objectives of the project

This final report is the concluding deliverable of the project *Revenue Administration's Strategy on Artificial Intelligence* (REFORM/SC2021/016), conducted to benefit the Estonian Tax and Customs Board (ETCB), with support from the European Commission's Directorate-General for Structural Reform Support (DG REFORM). This section aims to introduce the overall project context, its objectives and main outcomes.

Context

Artificial Intelligence (AI) has emerged as a general-purpose technology. The recent progress achieved in the AI field has been extraordinary, across many different sectors. As a result, the EU has put AI at the top of the agenda.

To foster the development and uptake of artificial intelligence, the Commission presented new rules and actions in its first legal framework on AI (proposal for a regulation) and a revised Coordinated Plan with Member States. It[1] outlines the necessary policy changes and investments at the Member States' level in order to strengthen the development of human-centric, sustainable, secure, inclusive and trustworthy AI.

In this context, several Member States have developed or are working towards strategies to support AI. 24 Member States, including Estonia, have already committed to joining forces on AI and entered a strategic dialogue through the European AI Alliance, a broad multi-stakeholder platform, to engage in open discussion of all aspects of AI development and its impact on the economy and society.

In May 2019, an expert group led by the Ministry of Economic Affairs and Communications and the Government Office presented proposals on advancing the take-up of Estonia's national AI strategy. This strategy is, in the European Union's coordinated AI action plan context, synchronised with and supporting relevant EU-level activities. Estonia's AI strategy relies on four pillars: boosting AI in the government, in the economy, build skills along with research and development, and develop the legal environment.

As of the end of 2020, there were at least 52 AI initiatives implemented in the Estonian public sector, with at least 18 still in progress.[2] At the ETCB, this includes an initial analysis of whether AI could be used to identify likely cases of potential "envelope wages" being paid, to identify suitable cases for inspection. While the ETCB also has prior examples of digitalisation and being data-driven, previous efforts have mostly been architectural in nature and expert models rules-based, not yet fully exploring the opportunities made available by AI.

In this light, and in line with the overall EU and Estonian AI strategy, the ETCB requested tailor-made technical support under the Technical Support Instrument (TSI) to help it in:

- exploring the feasibility of using human-centric AI; and

[1] EC, Coordinated Plan on Artificial Intelligence 2021 Review. See https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review
[2] Ministry of Economic Affairs and Communications, Use cases. See https://en.kratid.ee/kasutuslood

- developing a strategic framework that would facilitate the ETCB's transition to a higher level of digitalisation through AI.

Objectives and outcomes

The general objective of the service contract was to contribute to institutional, administrative, and growth-sustaining structural reforms in Estonia, while the specific objective was to assist the ETCB to improve its capacity to design, develop and implement reforms concerning digitalization and AI.

As an outcome of the activities performed during the project and results delivered, the ETCB has, among other benefits, gained:

- an improved understanding of the feasibility of using AI and improved strategic capabilities concerning the integration of AI;
- improved technical knowledge about AI-driven solutions;
- a clear set of actions in the form of an indicative roadmap to deliver selected AI solutions;
- the ability to showcase its vision for two new AI solutions in the form of visual prototypes; and access to better documentation of the organisation's key processes and architecture.

In order to achieve this results, the project's team worked on the activities and deliverables presented in the next sections.

3. Deliverables

This chapter will describe:
The results of each phase of work
The methodologies utilised

In this section, the deliverables created in each phase are described and put in the global project's perspective. The methodologies utilised and the activities performed are also highlighted.

The project structure followed the planned activities laid out in the first deliverable, the inception report (Phase 1) and continued with the feasibility study (Phase 2) which presented an overview of the existing AI use cases relevant in the context of tax and customs and a high-level mapping of ETCB's processes and services. This was followed by a prioritisation exercise to rank the identified opportunities from a value and a feasibility point of view. Five of the most promising areas were selected and an analysis of any bottlenecks preventing the further development of AI solutions was undertaken.

Once the feasibility study was completed, the strategic choices and action plan (Phase 3) were created, to define a path from the current to the desired future state. A sequential roadmap has also been presented, highlighting the potential next steps of AI implementation in the ETCB and how those might be enabled by other complementary elements.

Subsequently, the ETCB narrowed down to two selected AI-driven solutions, for which the various technical aspects needed to implement these were captured (Phase 4). This led to a list of requirements, a clear definition of the desired functionalities, as well as a proposed reference architecture that would support these solutions.

Finally, this consolidated report (Phase 5) was created and will consist of a summary of all previous phases.

Deliverable 1 - The Inception Report

The Inception Report ensured that all stakeholders have a common understanding about the services provided within the scope of the project and the way in which they were delivered. It presented the overall approach of the project and described for each phase what the tasks being performed were and the outputs being delivered:

	PHASE 1 Inception	PHASE 2 Feasibility study on using AI in specific tax and customs matters in the organisation	PHASE 3 Develop strategy & action plan to integrate AI in the organisation	PHASE 4 Technical description for 2 AI-driven solutions	PHASE 5 Final report
TASKS	**T1.1** Organise the project kick-off meeting and deliver a presentation **T1.2** Draft the and submit inception report	**T2.1** Collect in an overview existing AI use cases and good practices in the worldwide public sector **T2.2** Conduct a high-level mapping of ETCB's processes and services, identify and rank areas **T2.3** Develop business cases for four to five selected topics **T2.4** Identify bottlenecks or barriers and formulate recommendations **T2.5** Analyse how AI-driven solutions should be in accordance with the GDPR (high-level legal analysis) **T2.6** Draft and submit deliverable 2	**T3.1** Develop the strategy on integrating AI in ETCB **T3.2.** Develop the action plan for implementing the strategy **T3.3.** Draft and submit deliverable 3	**T4.1** Define the requirements (possibly user stories and/or non-functional visualised user experience prototype) **T4.2** Define the functionalities **T4.3** Define the architecture **T4.4** Draft and submit deliverable 4	**T5.1** Draft and submit the final report
DELIVERABLES	**D1** Inception report	**D2** Feasibility study report	**D3** Strategy and action plan report	**D4** Technical description report	**D5** Final report
METHODS & TOOLS	EVD PM (Deloitte project management methodology) Proposal Issue analysis	Desk research and data collection methods BPMN Value and feasibility criteria Now-how-wow matrix	AI expertise AI maturity model Deloitte's best practices in AI and cognitive	Agile methodology User personas Deloitte reference architecture	Project closure checklist
TIMING	Month 1-2	Month 1-7	Month 7-11	Month 8-11	Month 11-12

Figure 1 - Project Structure

In addition to this project sequence, the Inception Report presented the different methodologies leveraged, which are also highlighted in each section dedicated to the specific deliverables.

Deliverable 2 - The Feasibility Study

This feasibility study identified opportunities for the ETCB to use AI-driven solutions, in order to address certain tax and customs challenges. To do so, it detailed human-centric and trustworthy AI implementations in other governmental institutions, split by the main benefits they provided:

- **Engagement**: Tax and customs agencies receive millions of calls, mails and chat requests annually. This puts a significant burden on the institution administrative staff to maintain a high degree of customer satisfaction. Several countries created chatbots or similar automated answering systems to provide businesses and citizens with information and assistance in the easiest and most accessible way. The different fields of Natural Language Processing have been heavily leveraged to simplify and improve the accessibility of governmental resources.

- **Insights**: AI can provide tax and customs institutions with a variety of insights necessary for revenue collection and battling black market activity. Predictive analysis, for example, uses taxpayers and fraud information to determine which factors are driving tax non-compliance which helps policy makers and tax regulators to take adequate measures. Some institutions are also using analytics to improve the traceability and identification of some hidden assets (such as cryptocurrencies). In the customs field, examples of surveillance systems using video analytics systems are numerous, using technologies including behavioural analytics and counting solutions, designed to suit crowded and complex environments.

- **Automation**: In tax-related applications, Natural Language Processing can be used to automate the processing of citizens' applications and open text format information by extracting useful information out of documents. Classification and predictive algorithms can also help to perform a first screening, reducing ultimately the required human workload. Some countries are also using deep learning techniques to mimic the back-office tasks of human workers, such as extracting data, filling in forms, moving files, etc.

Based on the processes and services analysis conducted through interviews with responsible departments, the aggregated lists of use cases implemented in other similar organisations and inspirational cases considered by the ETCB, the team presented a long list of AI-driven solutions, split by type of technology and impacted department:

Use Cases Repartition		Human Resources	Audit	Legal	Tax	Communication	Intelligence	Customs	Administration	Investigations	Mngmt
Natural Language Processing	Insight	2	0	1	0	0	1	0	1	0	0
	Engagement	0	0	0	0	1	0	2	0	0	0
	Automation	1	4	1	0	1	0	2	1	1	1
Machine Learning	Insight	2	1	0	1	2	0	3	2	2	2
	Engagement	1	0	0	0	1	0	1	0	0	0
	Automation	1	0	0	0	0	0	0	2	0	0

Figure 2 - Use Cases Repartition

Shortlisted use cases

Through two workshops with the ETCB executive management, 5 use cases were selected and later refined by the team:

1. Amount at Risk

[REDACTED]

This use case therefore aims to use supervised learning and leverage the ETCB's historical data to improve the accuracy of the risk detection process. This would result in a higher proportion of risky cases being captured, as well as reducing time lost on cases that do not turn out to present a risk.

2. X-ray Inspection

The ETCB and its Customs department are tasked with making sure that legal trades function as smoothly as possible throughout Estonia's borders, that no illegal or restricted goods are passing through and that customs formalities are respected. The efficiency of these processes is even more critical due to the fact that the Estonia border is also the frontier of the European Union's external border.

In this context, artificial intelligence can be applied to image recognition and threat detection tasks, to improve the customs checks that are currently mostly manually performed tasks. This use case would aim to utilise the x-ray images generated when a container or vehicle is checked at the border, through automating the visual analysis by flagging high-risk items and assigning a risk percentage to determine whether a manual control should be performed. This would not only reduce the time spent on visual analysis, but also improve consistency in the quality of customs controls caused by inexperience or human error.

3. Network Fraud Detection

[REDACTED]

Using the available information on relationships between individuals and companies, different algorithms can be used to detect suspect connections across a whole dataset and introduce a new risk factor into the risk analysis. Network analysis models can detect abnormal behaviours, potential areas of risk based on other known risky related individuals, and ultimately help dismantle fraud rings that were not previously detected.

4. Consignment Code Text

The ETCB's Customs department verifies the information on customs declarations issued when importing goods in Estonia. Those include a consignment code, assigned depending on what item is being declared, which is known to be a complex task requiring deep customs knowledge. It defines the amount of customs duties and if restrictions are to be applied. The department highlighted that it would be beneficial to have a solution that would generate the correct consignment codes based on invoice inputs, item description, any additional declaration data and/or the results of previous customs declarations. This could be used either as a tool used by declarants or as a control tool internally by the ETCB to identify where incorrect consignment codes might have been used.

5. Tax Payers Clustering

The ETCB and its Tax department have, as one of their primary goals, to improve tax knowledge and compliance of taxpayers in Estonia. This is done through conducting trainings, instruction sessions and providing consulting to individuals and firms alike, as well as through campaigns and communication.

However, identifying the right groups of people for specific topics is a complex manual task and one of the most time-consuming activities. To assist with these processes, a machine learning model could cluster the Estonian taxpayers, to better understand similarities in needs, who needs training, how to tailor it and who to reach for, beyond simply looking at demographic information.

Identified bottlenecks

Having detailed the above shortlisted use cases, the team highlighted potential bottlenecks and barriers to their implementation, based on information collected from process mapping, stakeholder interviews and previous attempts to develop AI solutions. The bottlenecks identified can be categorised into the following areas:

- Cultural change: AI solutions are currently taken as technical niche projects that should fit into existing business processes, without the need for business change.
- Communication: The ETCB must improve the way it manages artificial intelligence related projects and should nurture technical capabilities to communicate with outsourced partners on AI and other innovation projects.
- Technical: The ETCB employs many analysts, but none of them are dedicated data scientists. This means that the ETCB is currently not capable of creating AI solutions by itself and complicates the maintenance of outsourced AI solutions. Some of the necessary tools required to implement AI are also missing from the ETCB technology stack.
- Data: At the moment, there is no central data governance in place, which makes it difficult to ensure consistency in the data. There are also no trusted processes to anonymise data.
- Siloed implementation: There is no systematic approach for fostering innovation and there is no designated team, who's assignment would be to take the innovation idea into practice. When it comes to the developed PoC's, there is no system to preserve the knowledge and competences created during their creation.
- Buy-in: a lack of a clear vision and strategy from the top management, and no clear project valuation methodology discourage mid-level managers to take up artificial intelligence projects.
- Legal Context: from data sharing restrictions to GDPR considerations, the legal context creates numerous potential bottlenecks which should be carefully scrutinized.

Deliverable 3 - The AI Strategy

This phase established where the ETCB currently stands in its capability to integrate and leverage AI, by scoring the organisation across different dimensions, shown below:

[REDACTED]

Indicative roadmap

To address the aspects presented above, a list of actions have been formed and sequenced, resulting in the summarised roadmap below:

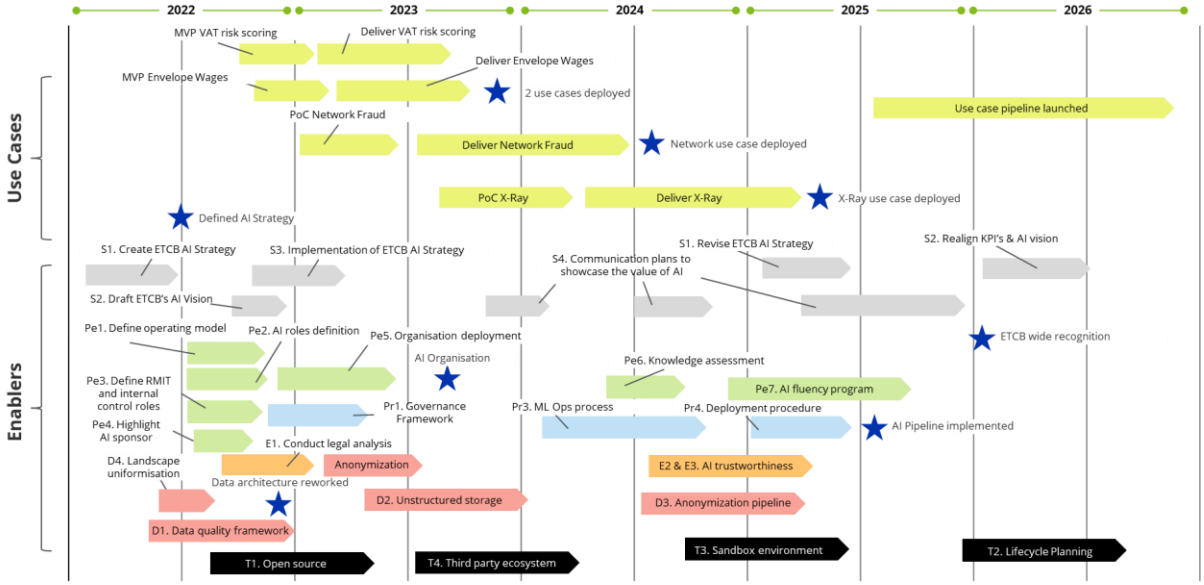


Figure 3 - High Level Roadmap

The "Use Cases" section of the roadmap refers to two strategic decisions taken by the ETCB leadership. First, the existing proof-of-concepts should be leveraged and built upon in the short term, to showcase tangible results. Secondly, out of the 5 shortlisted use cases detailed in the feasibility study, the "Network Fraud Detection" and the "Automated X-Ray Analysis" should be prioritised over the next few years.

To deploy these proof-of-concepts and two new use cases, "Enablers" have been identified. These are elements which are essential to focus on if AI is to be properly deployed within the ETCB. In addition to being critical to the development of the identified use cases, those "Enablers" will allow for long-term innovation to be fostered.

Together the actions form an indicative roadmap to be followed, in order to implement the first operational AI use cases and pave the way for future innovations to be deployed. While the exact timing will depend on other internal and external factors, such as resource and budget constraints, speed of decision-making and conducting each individual task, as a whole it outlines the order in which actions must be taken to ensure that beneficial AI solutions have the appropriate enablers in place to allow them to be properly implemented.

Deliverable 4 - The Technical Description of Two Use Cases

In this phase the two selected use cases were detailed from a process, functionalities and requirements perspective.

X-Ray Automated Analysis

The X-Ray automated analysis process is currently prone to human errors and of varying quality depending on the level of training and experience of the customs officer. Along with the physical inspection required if something is detected and filing of paperwork, each truck stopped takes the customs officers around one hour, which in the face of increasing volumes of cross-border trade limits the number of controls that can be performed. A new process has been defined, embedding AI visual image analysis:

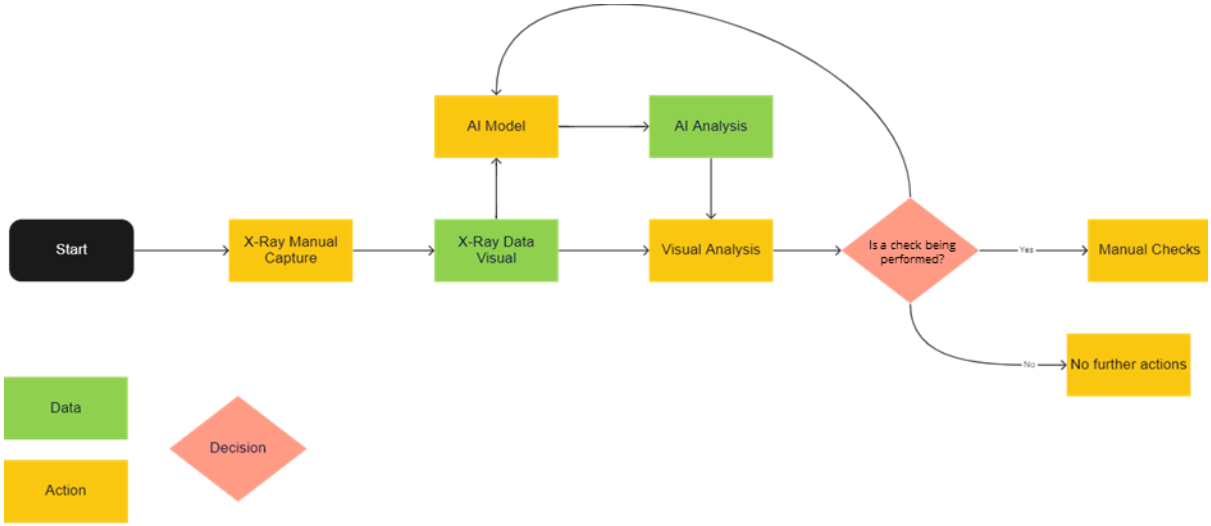


Figure 4 - Process Map - Future X-Ray

The process is conducted by a Customs officer in the field. Even though the need for manual visual analysis will diminish as the model improves, human intervention, to (at a minimum) launch the analysis and perform physical checks, will still be needed.

With regards to the functionalities of the AI solution added in the process, several functionalities were prioritised through interviews with relevant stakeholders:

- Ability to make the analysis vary depending on the type of X-Ray machine being used;
- Automated X-Ray image capture and loading;
- Automated assessment of the loaded X-Ray image from a risk perspective;
- Ability to highlight the items which are deemed risky by the model.

These have been exemplified in a clickable visual prototype developed using the prototyping software, Figma.

Network Fraud Detection

The ETCB has access to various data on taxpayers and their connections. These currently fuel rules-based models, which help analysts to highlight specific risks. However, using the same information, more sophisticated models can be developed to detect suspect connections across a whole dataset.

Network analysis models can detect abnormal behaviours and help dismantle fraud rings that were not previously detected. The revised process, involving an AI model, is outlined below:

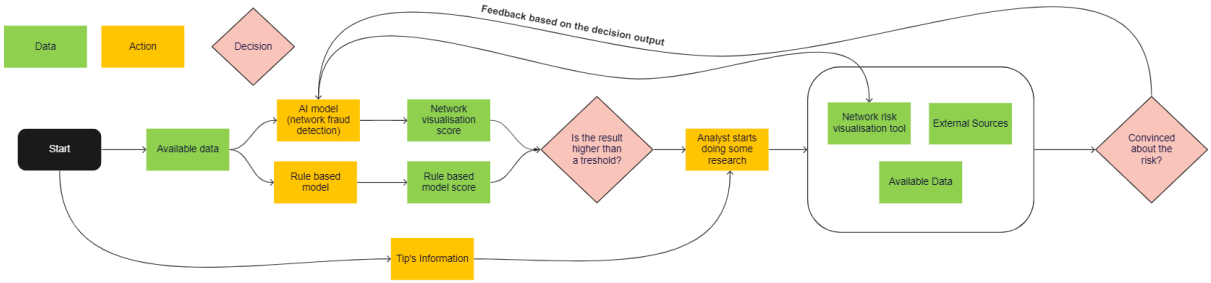


Figure 5 - Process Map - Future Network

The main modifications are the addition of an AI-based risk model to perform the scoring and the possibility for analysts to access an application providing all of the information available to the ETCB, highlighting specific risks related to the individual or company's network, as well as allowing to browse and explore those relationships.

Similar to the previous use case, key functionalities were highlighted, being:

- Generating a network-based risk score and highlighting risky connections in the data;
- Visualising the relationships between individuals and companies, particularly those that are higher risk;
- Quickly accessing information on individuals/companies from existing data sources and models.

These have been taken into consideration and visualised through a clickable prototype developed in Figma.

Requirements and high-level functions

Having detailed the two use cases selected by the ETCB leadership, the current ETCB data architecture was analysed from the perspective of the use cases' expected functionality and requirements, which are summarised below:

#	High-Level Functions of the AI-driven solution	XAA[3]	NFRD[4]
1	Image analysis	x	
2	Network analysis and network visualisation		x
3	Risk assessment, risk scoring and risk visualisation	x	x
4	Continuous evaluation and recalibration of risk models, automatic (re)defining of risks and composite models	x	x

#	High-Level Requirements of the AI-driven solution	XAA	NFRD
1	Data quality and governance	x	x

[3] XAA = X-ray Automated Analysis

[4] NFRD = Network Fraud Risk Detection

2	Centralised and uniformised database		x
3	Unstructured data storage	x	x
4	Metadata integration	x	x
5	Intuitive user interface	x	x

#	High-Level AI techniques used in the solution	XAA	NFRD
1	Unsupervised learning - Clustering	x	
2	Unsupervised learning - Network Analysis		x
3	Supervised learning - Regression	x	
4	Supervised learning - Classification	x	x

Based on this, an overview of a proposed architecture was then developed for the ETCB that would allow to fulfil the above requirements and to enable the highlighted functionality, as outlined below.

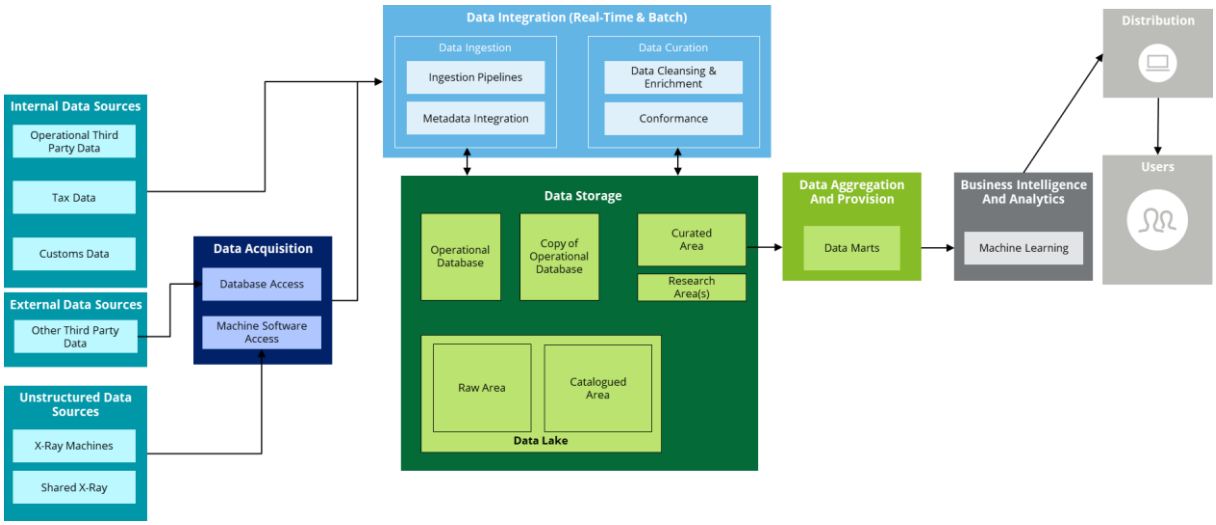


Figure 6 - ETCB Proposed Architecture

To progress from the current ETCB architecture to the proposed one, we highlight the following key changes required:

- **Enable the storage of unstructured data**: this will be required to work with x-ray images and potential future use cases that might leverage other types of data.

- **Add a data acquisition layer**: accessing third-party data should be done in a governed and safe way. To include this information in a connected way with the rest of the available data, an acquisition layer must be added as the starting point of the data flow.

- **Put in place an ETL process**: to integrate, transform, and load data into their storage solution, an ingestion pipeline is required. This will provide a means of getting data out of the source systems and into the raw data storage area and operational database.

• Create and enforce a data cleansing routine: this will be necessary to provide physical data checks and basic validations, in order to construct a curated area which will store the labelled images and the structured data in a format fit for usage by the machine learning algorithms.

• Enable users to access specific data marts: the ETCB should create data marts containing the curated and tested data used to train the AI models. That subset of data should then be shared with the end-users.

• Govern and control the created AI models: the ETCB must include in its architecture, a model repository. This will be used to store the production model and other models being tested/developed by the ETCB. A monitoring solution will also be needed to ensure the model is working as expected.

• Develop the ETCB distribution layer: as described in the prototypes, the use cases will require custom applications. The ETCB should include in its architecture the technical capabilities and technological solutions to enable the creation of such applications.

4. Lessons Learned

This chapter will describe:
The lessons learned during this project

During the course of the project, several lessons have been learned in performing the analysis and producing the deliverables. These are important findings, secondary to the formal outputs of the project, that should be considered in moving ahead with the defined initiatives:

- Firstly, regarding the positioning of AI in the ETCB, although there are areas where improvements can be made, the ambition to leverage AI within the organisation is evident. Importantly though, it should not be seen as an "all-or-nothing" type of initiative, as the ETCB has stated that its goal is to start by being use case targeted and to focus its efforts on step-by-step actions and results in key focus areas.

- Directly stemming from the first lesson and the decision to be use case focused is the need to be results-focussed and having something tangible to showcase. The prototypes have been met with positive reactions and the strategic focus to deploy the existing proof-of-concepts has been given a strong push. Such showcasing of achievements and visible results will foster great enthusiasm and motivation to move further along the AI maturity curve.

- In parallel to the willingness to display results quickly, the necessity of working on enabling capacities such as a data governance, AI prioritisation processes or coding repositories to foster long term innovation is also a key finding and should continue to be considered as the ETCB moves along in the roadmap.

- Some AI use cases and the capabilities they provide can also serve as enablers. Accordingly, they should not be analysed purely as siloed business initiatives. Such an example is the different uses of a chatbot which, alone do not bring significant value and require significant first-time investment, while if one of them is already developed, can allow for each subsequent one to be accelerated and bring even greater value.

- Developing an AI use case requires involvement from multiple different structures within an organisation. When analysing the potential bottleneck and barriers to AI implementation, it was clear that the whole organisation needed to be onboard with a project to make it a success. This means that it cannot be an initiative driven solely by IT specialists or by the Intelligence department. Business departments are critical in realising the intended benefits and in ensuring the solutions do not become stuck infinitely in PoC phase.

- To get the buy-in of different business departments, sponsorship from the top management is key. The leadership must champion AI and be able to showcase the innovations' benefits to the ETCB, while also ensuring proactivity in leading innovation is appropriately rewarded.

- Most of the considerations pertaining to AI are not ones that only the ETCB faces, which is why collaboration with other national and international institutions is highly relevant. European institutions particularly, are at the forefront when it comes to legislation, AI regulation and data

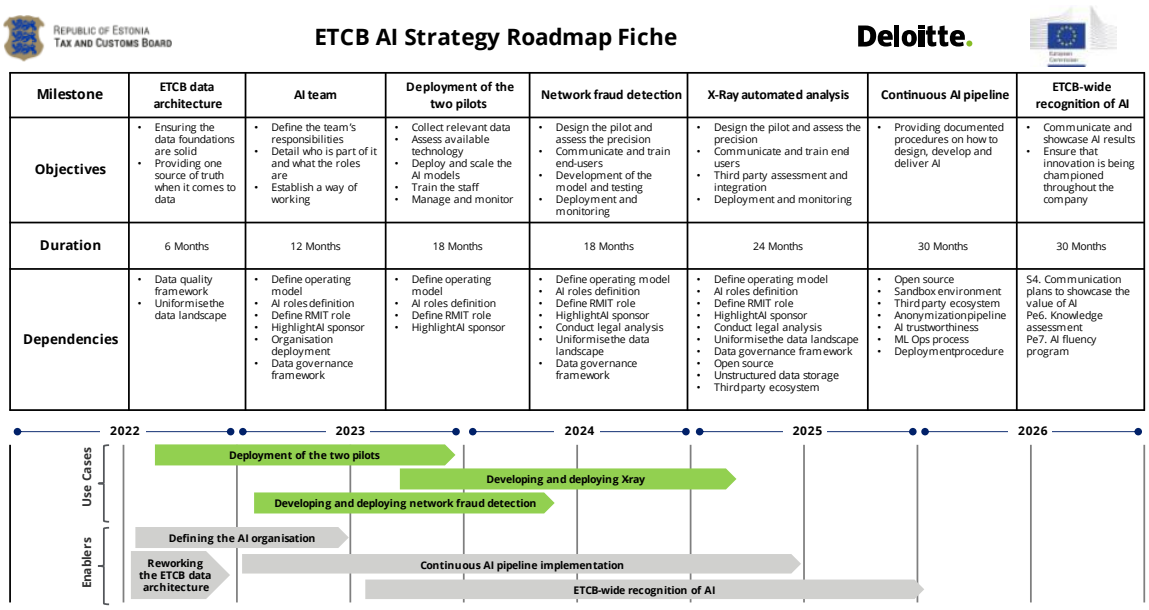
standardisation. Opportunities for shared capabilities and data-sharing should therefore be strongly considered.

- Finally, the benefits cannot be enabled without investments in architecture, both financial investment into technical solutions, but also time investment into strategic planning. These technology and data architecture components are essential in effectively integrating AI into the ETCB's solutions. Decisions such as continuing to use SAS or creating a testing environment must be analysed from an AI capability perspective as they impact the way that data scientists will work on a daily basis.

5. Final Recommendations

This chapter will describe:
The project's conclusion and last recommendations

The lessons learned, described in the previous chapter, are primarily backward-looking findings during the project that should be considered. The objective of this chapter is to shape clear action points for the future. In this context, the roadmap fiche of deliverable 3 appears essential and provides a clear view of what should be worked on:



Milestone	ETCB data architecture	AI team	Deployment of the two pilots	Network fraud detection	X-Ray automated analysis	Continuous AI pipeline	ETCB-wide recognition of AI
Objectives	• Ensuring the data foundations are solid • Providing one source of truth when it comes to data	• Define the team's responsibilities • Detail who is part of it and what the roles are • Establish a way of working	• Collect relevant data • Assess available technology • Deploy and scale the AI models • Train the staff • Manage and monitor	• Design the pilot and assess the precision • Communicate and train end-users • Development of the model and testing • Deployment and monitoring	• Design the pilot and assess the precision • Communicate and train end users • Third party assessment and integration • Deployment and monitoring	• Providing documented procedures on how to design, develop and deliver AI	• Communicate and showcase AI results • Ensure that innovation is being championed throughout the company
Duration	6 Months	12 Months	18 Months	18 Months	24 Months	30 Months	30 Months
Dependencies	• Data quality framework • Uniformisethe data landscape	• Define operating model • AI roles definition • Define RMIT role • HighlightAI sponsor • Organisation deployment • Data governance framework	• Define operating model • AI roles definition • Define RMIT role • HighlightAI sponsor	• Define operating model • AI roles definition • Define RMIT role • HighlightAI sponsor • Conduct legal analysis • Uniformisethe data landscape • Data governance framework	• Define operating model • AI roles definition • Define RMIT role • HighlightAI sponsor • Conduct legal analysis • Uniformisethe data landscape • Data governance framework • Open source • Unstructured data storage • Thirdparty ecosystem	• Open source • Sandbox environment • Thirdparty ecosystem • Anonymizationpipeline • AI trustworthiness • ML Ops process • Deploymentprocedure	S4. Communication plans to showcase the value of AI Pe6. Knowledge assessment Pe7. AI fluency program

Figure 7 - AI Strategy Roadmap Fiche

Along each of the ETCB's objectives are the dependencies that have been highlighted, which form the final recommendations of the project:

- **Define the AI strategy**: this report is the final deliverable of the project, and the ETCB needs to take action to build upon the detailed strategy, especially when it comes to drafting its vision pertaining to AI and how this vision fits into the global ETCB strategy.

- **Rework the ETCB data architecture**: to ensure that the development and deployment of AI use cases can begin the data architecture of the ETCB must be reviewed. This should encompass the definition of the data quality requirements, the start of the data governance efforts and the ETCB data landscape uniformization.

- **Set-up and organise the AI team**: the documentation of the team's organisational model, the assignment of leadership roles and the definition of how the ETCB will interact with external partners should be concretely defined.

• Start with the proof-of-concepts: First, the existing proof-of-concepts should be leveraged and built upon in the short term, to showcase tangible results. This will require collecting, cleansing and uniformising data, as well as developing, deploying, and monitoring the models. These activities will enable the established team to test its procedures and develop future-proof ways of working.

• Develop the two shortlisted use cases: as detailed in the feasibility study, the "Network Fraud Detection" and the "Automated X-Ray Analysis" use cases should be the ones to prioritise following the proof-of-concepts. As the "X-ray" use case will require a more complex technology, time to allow storage of unstructured data and collaboration with other international institutions (for example to access labelled data), it has been suggested to first launch the development of the "Network Fraud Detection" use case, for which data and technology is already more available.

• Design an AI continuous implementation pipeline: leveraging the experience built while developing the different AI use cases, the ETCB should create processes which can simplify future implementation. The ETCB should focus on providing data scientists with a data anonymisation pipeline and a safe sandbox environment while enabling contractors to work with its data in a safe and compliant way. Procedures defining how to value, deploy and maintain AI models should also be worked upon, to standardise and make the AI related processes more efficient.

• Communicate results: during all steps of this roadmap, communication will be critical. When the first results from the proof-of-concepts are generated, enablers put in place or new solutions developed, they should be showcased not only to its future users and stakeholders, but to the rest of the organisation, educating and inspiring others in the organisation to innovate.

Annex - Project Fiches

This annex outlines for communication purposes, a visual project summary fiche summarising the project's context, objectives, main activities and results.

Revenue Administration's Strategy on Artificial Intelligence

REFORM/SC2021/016

Project Title	Estonian Revenue Administration's Strategy on Artificial Intelligence
Request for Service ID	REFORM/SC2021/016
Beneficiary	Republic of Estonia Tax and Customs Board (ETCB)
Supporting Institution	Directorate General for Structural Reform Support (DG REFORM)
Support Instrument	Technical Support Instrument (TSI)
Contractor	Deloitte

Context

The mission of the European Commission's Directorate-General for Structural Reform Support (DG REFORM) is to provide support for the preparation and implementation of growth-enhancing administrative and structural reforms by mobilising European Union (EU) funds and technical expertise.

Estonia requested tailor-made technical support under the Technical Support Instrument (TSI) to help the Estonian Tax and Customs Board (ETCB) in:

- exploring the feasibility of using human-centric AI and;
- developing a strategic framework.

The technical support would facilitate ETCB's transition to a higher level of digitalisation through AI.

The project is closely linked to the EU AI agenda and other EU priorities. Furthermore, the project is in line with Estonia's AI strategy. In May 2019, an expert group led by the Estonian Ministry of Economic Affairs and Communications, and the Government Office presented proposals on advancing the take-up of Estonia's national Artificial Intelligence (AI) strategy. Estonia's AI strategy relies on four pillars: boosting AI in government, in the economy, building skills along with research and development, and developing the legal environment.

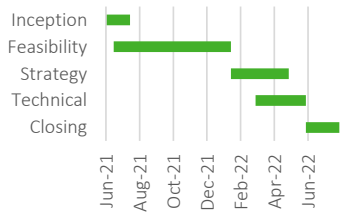
Objectives

The general objective of the service contract was to contribute to institutional, administrative, and growth-sustaining structural reforms in Estonia.

The specific objective of the service contract was to assist the ETCB to improve its capacity to design, develop and implement reforms concerning digitalization and AI.

Project overview

The project was split into five phases and spanned 14 months, with an end date of August 2022.



1. Inception Phase
The project was initiated and resources mobilised.

2. Feasibility Phase
Existing relevant AI use cases were collected and a high-level mapping of ETCB processes performed across 11 departments and 4 additional sub-departments. Bottlenecks were identified based on prior pilot projects and 44 potential applications of AI identified within the ETCB. This was followed by a prioritisation exercise to rank identified opportunities, with the most promising topics analysed further in a business case.

3. Strategy Development Phase
Once the feasibility study was completed, the strategic choices and action plan were developed, forming a clear vision and indicative roadmap from the current state to the future state.

4. Technical Description Development Phase
Two AI-driven solutions were selected to create technical descriptions, including requirements, desired functionalities, a visual prototype, and proposed architecture that would support the solutions.

5. Closing Phase
The project was closed and assessed against the planned outcomes. Recommended follow-up actions were formulated.

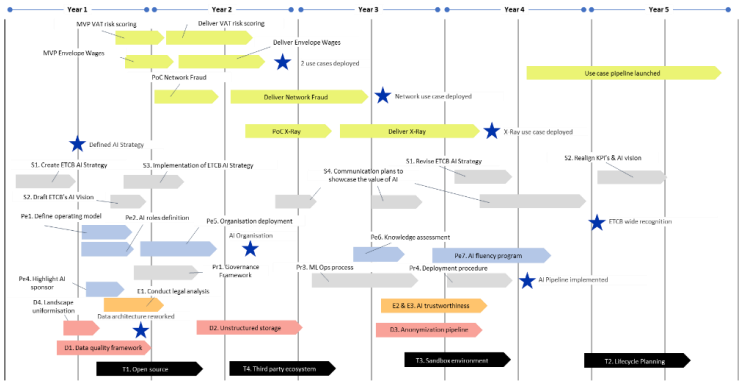
Key outcomes

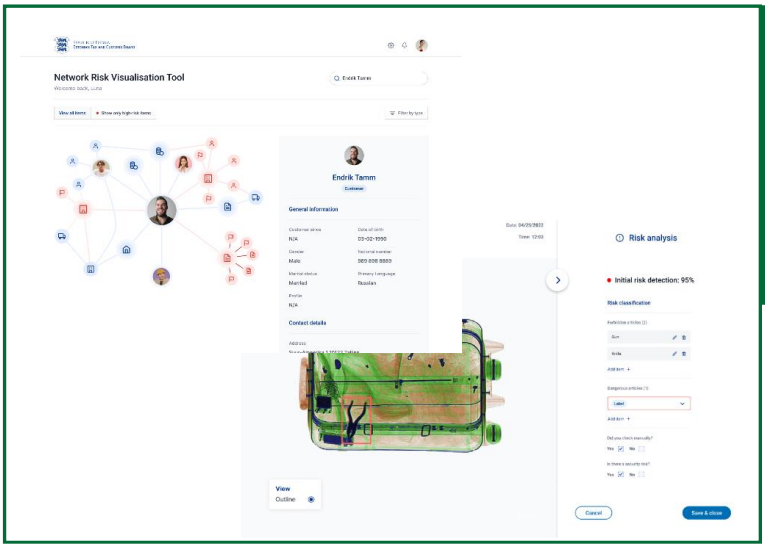
As an outcome of the activities performed during the project and results delivered, the ETCB has gained:

1. an improved understanding of the feasibility of using AI and improved strategic capabilities concerning the integration of AI.
2. improved technical knowledge about AI-driven solutions.
3. a clear set of actions in the form of an indicative roadmap to deliver selected AI solutions.
4. the ability to showcase its vision for two new AI solutions in the form of visual prototypes.
5. access to better documentation of the organisation's key processes and architecture.

Indicative roadmap

The ETCB now has a clear roadmap for delivering a specific set of AI solutions explored, as well as the enablers that will need to be addressed to deliver such future AI solutions.





Visual prototypes of X-ray Automated Analysis and of Network Fraud Risk Detection

Additionally, a technical description of two promising new AI solutions has been developed, including a visual prototype that provides the ETCB with an opportunity to showcase its ambitions.

Summary Fiche

REVENUE ADMINISTRATION'S STRATEGY ON ARTIFICIAL INTELLIGENCE
Project title
Preparing an artificial intelligence (AI) strategy for the Estonian revenue authority.
Summary
With this project, DG REFORM provided technical support to the Estonian Tax and Customs Board (ETCB) with the aim of (a) exploring the feasibility of using human-centric AI and (b) developing a strategic framework. The technical support would facilitate the authority's transition to a higher level of digitalisation through AI.
Context
AI is a promising new technology that could help revenue authorities in addressing certain persistent or emerging tax and customs challenges that might require a qualitative leap in process automation and data analytics. The development of a strategic vision is however a prerequisite for making full use of the possibilities that AI can offer whilst proactively reflecting on new risks and complexities inherent to AI. The project is linked to the Commission's AI and tax agendas.
Support delivered
At the end of the 14-month project, the authority has at its disposal: • A feasibility study on using AI in specific tax and customs matters and in the ETCB's organisation; • A strategy and action plan to integrate AI in the ETCB; • Technical descriptions for two AI-driven solutions, specifically X-ray Automated Analysis and Network Fraud Risk Detection, including requirements, desired functionalities, a visual prototype and proposed architecture that would support the solutions. The project is funded by the European Union via the Technical Support Instrument and implemented by Deloitte in cooperation with DG REFORM.
Results achieved
The technical support aided the ETCB in setting out a path for its further digitalisation through AI. The Estonian authority has, among other benefits, gained: • an improved understanding of the feasibility of using AI and improved strategic capabilities concerning the integration of AI; • improved technical knowledge about AI-driven solutions; • a clear set of actions in the form of an indicative roadmap to deliver selected AI solutions; • the ability to showcase its vision for two new AI solutions in the form of visual prototypes; and • access to better documentation of the organisation's key processes and architecture.
 This project is funded by the European Union via the Technical Support Instrument and implemented by the European Commission.

References

This annex provides the Table of Figures for all references and captions included in this document, as well as the Acronyms and Abbreviations used.

Table of Figures

Acronyms and Abbreviations

Acronym/ Abbreviation	Detail
AI	Artificial Intelligence
DG REFORM	European Commission Directorate-General for Structural Reform Support
EC	European Commission
ETCB	Estonian Tax and Customs Board
EU	European Union
GDPR	General Data Protection Regulation
KPI	Key Performance Indicator
ML	Machine Learning
NLP	Natural Language Processing
PoC	Proof-of-Concept
TSI	Technical Support Instrument

Use Case: Chatbots for tax-related queries
Country: Sweden
Technology: NLP, NLU, Sentiment analysis

Conversational AI chatbots for Tax-related queries

Business Problem - Context

- The Tax information office at the Swedish Tax agency receives about **3.5 million calls and 500,000 emails** annually. This puts a significant burden on the tax agency's administrative staff to maintain a high degree of customer satisfaction
- Swedish Tax Agency wanted to provide businesses and citizens with **information and assistance in the easiest and most accessible way**, while reducing the burden on their administrative staff
- Providing information and assistance would increase the **level of tax compliance**. Thus, the agency wanted to pilot a technology driven service

Use Case Description

- In 2019, Swedish Tax Agency (STA) developed a chatbot called 'Skatti' available in STA's website round the clock to provide **personal assistance to the taxpayers** to help answer questions around their tax returns
- The chatbot was trained by certified AI trainers who were experienced in answering complicated questions related to **population registration and personal taxation**
- In future, the agency plans to **train Skatti in several languages** to test speech as a supplement to text chat. They are also exploring the possibility of escalating to human chat while logged in

Related Technologies

- Using Natural Language Processing (NLP), AI algorithms can pre-process user input messages into tokens that are linguistically representatives. Normalization techniques are applied to remove human mistakes or slight deviations from a certain word. Those normalized tokens can then be grouped using pattern matching techniques to get an understanding of the user's request
- Natural language understanding (NLU) comes into play to extract intents and entities from the processed text inputs. Responses databases are then queried in order to generate response candidates from which another algorithm selects the most suited one
- In chatbots, sentiment analysis are also often used to analyze the user's experience and transfer the conversation to a human if deemed necessary

Benefits

- Skatti query handling rate is 15,000 queries/month
- Skatti could resolve about 78-83% of queries
- 40% of conversations with Skatti took place outside the working hours of Tax Information office, improving customer service in non-working hours

Implementation Complexity

- Skatti has gained expertise in answering questions pertaining to tax and few other subject areas The challenge is to further widen the scope to understand common business-related queries
- The chatbot is currently trained to communicate in Swedish. Its ability to communicate in other languages is limited

Main Sources

- Swedish Tax Agency Skatti chatbot: https://www.ai.se/en/news/artificial-intelligence-improves-swedish-tax-agencys-customer-service-0
- Skatti chatbot implementation and benefits: https://www.iota-tax.org/sites/default/files/documents/publications/Books/iota_book_2020_may_applying_new_technologies_and_digital_solutions_in_tax_compliance-online.pdf
- Implementation complexity: https://www.thelocal.se/20190319/its-time-to-file-your-taxes-and-heres-when-you-can-get-your-rebate/

Use Case: Sorting customer enquiries
Country: Denmark
Technology: Text mining, NLP, Machine learning

AI for sorting incoming customer enquiries

Business Problem - Context

- The Copenhagen Municipality receives multiple queries from citizens all through the day and a manual system is not capable of addressing and forwarding the concerns to the relevant officer. A huge wait time and backlog in calls **adversely affects the customer's experience**
- Forwarding the cases to the relevant case officer is subject to errors and depends on the knowledge base of the individual taking the call
- Collecting basic background data of citizens to aide in the query resolution process and an analysis of the same is a **time-consuming and tedious processes**
- The Copenhagen Municipality wanted to develop AI algorithms to efficiently handle queries from the citizens

Use Case Description

- In 2017, Copenhagen municipality developed an artificial intelligence algorithm to automatically redirect the customers who called the municipality helpline to one of the three relevant group of employees
- The AI algorithm redirects the call based on citizen's **personal information and background information** on previous tax requirements and reminders
- The citizen's information is automatically transferred to the relevant employee in a pop-up window
- Artificial intelligence was also used to **sort incoming posts and emails**, with the algorithm performing **text analysis,** to have an easy access to the information of the complaints and concerns of the callers

Related Technologies

- Using Natural Language Processing (NLP), AI algorithms can pre-process user input messages into tokens that are linguistically representatives. Normalization techniques are applied to remove human mistakes or slight deviations from a certain word. Those normalized tokens can then be grouped using pattern matching techniques to get an understanding of the user's request.
- Machine Learning or ML is a sub-set of Artificial Intelligence (AI), it utilizes data samples and algorithms to mimic the humanized way of learning. Based on these learning and feedbacks, the machine continuously improves its accuracy in application. Machine learning algorithms perform **text mining and text analysis** to classify the text contained in the E-mails

Benefits

- Reduction of manual labor and improvement in time management
- Significant reduction of wait time and call backlogs

Implementation Complexity

- The lack of sufficient data of citizens continue to pose challenges regarding the output produced by AI algorithms
- Algorithm development is a complicated, detailed and time-consuming task that demands expertise and concentration

AI Black Box

- The caller should be notified about the information being collected and retained by the Municipality
- There should be clarity regarding the rationale and method in which the AI achieves its results to reduce the asymmetric distribution of information between the caller and the Municipality

Main Sources

- AI use case in Copenhagen Municipality: https://norden.diva-portal.org/smash/get/diva2:1375500/FULLTEXT01.pdf
- AI Black Box: https://norden.diva-portal.org/smash/get/diva2:1375500/FULLTEXT01.pdf
- Machine learning in text analysis: file:///C:/Users/akshdixit/Downloads/IJCRT2002258%20(1).pdf

Use Case: Chatbots
Country: France
Technology: NLP, NLU, Sentiment analysis

Chatbots, fueled by NLP and applied in Tax Administration

Business Problem - Context

- Each year, tax declarations represent a **huge wave of incoming requests** for help, which can overload governments administrative staff. There are a lot of **recurring questions**, which are very repetitive to solve. Finding the correct document can be very cumbersome for the citizens, making the research and the filling process sometimes frustrating. The DRFIP (France's regional direction for public finance) wanted to **ease the declaration process to improve the user experience** while removing part of the burden it was causing to the administrative employees

Use Case Description

- In 2021, for the annual tax declaration campaign, the French administration launched **a virtual assistant (a chatbot)**
- It enables an **easier treatment of simple and often recuring questions** usually answered by the staff, enabling them to free time for more complex tasks. In addition, the chatbot can direct citizens towards the form that will fit their situation, **simplifying the declaration procedure**
- Internally developed, answers and questions have first been prepared by humans before getting fully automated. **Machine learning techniques** are now being applied for **the program to learn from the user's reviews** and to improve upon usage to enhance the service provided

Related Technologies

- Using Natural Language Processing (NLP), AI algorithms can pre-process user input messages into tokens that are linguistically representatives. Normalization techniques are applied to remove human mistakes or slight deviations from a certain word. Those normalized tokens can then be grouped using pattern matching techniques to get an understanding of the user's request
- Natural language understanding (NLU) comes into play to extract intents and entities from the processed text inputs. Responses databases are then queried in order to generate response candidates from which another algorithm selects the most suited one
- In chatbots, sentiment analysis are also often used to analyze the user's experience and transfer the conversation to a human if deemed necessary

Benefits

- Citizens can be answered in their native language
- In Los Angeles, after having implemented a chatbot, they noted a reduction of mails received of 50%
- Service can be maintained 24/7
- 53% of customers would rather message than call a service agent
- By 2022, EUR 8 billion will be saved from businesses using chatbots

Implementation Complexity

- Chatbots technology is mainly developed in English
- Cost of a tailored chatbot can highly vary depending on the customization and on the business requirement
- Chatbot management costs need to be considered, especially regarding continuously learning solutions

Main Sources

- French state tax website: https://www.impots.gouv.fr/portail/www2/minisite/declaration/chatbot.html?8
- French tax virtual assistant: AMI, le nouvel assistant virtuel des impôts - Chambre de Consommation d'Alsace (cca.asso.fr)
- Chatbots architectures: How do chatbots work? An overview of the architecture of a chatbot (bigdata-madesimple.com)
- Chatbots use cases & benefits: 4 ChatBot Examples of How They Improve Customer Experience (apruve.com)
- The state of chatbots: Chatbots - The Beginners Guide to Chatbot Technology | Drift
- Facebook host of data: Facebook Messenger for Business: communiceren met klanten | Facebook for Business
- Los Angeles chatbot use cases: Los Angeles, Microsoft Unveil Chip: New Chatbot Project Centered on Streamlining (govtech.com)

Use Case: Educational loan grant
Country: Norway
Technology: Machine Learning

ML model for Norwegian student loan process

Business Problem - Context

- The Norwegian Government allows students, living away from parents, to convert up to 40% of their educational loan into a grant. Subsequently, the government has been **receiving several cases** of misreported residential status from students looking to take advantage of this scheme. The government is seeking to resolve this problem through **intelligent and effective mechanism** to quickly and correctly verify the sample of applications

Use Case Description

- In 2018, the government had put **approximately 15,000 loan applications** data of students into the machine learning (ML) model to pull out a sample of students that are likely to misreport their residential status
- Based on the output, the model receives a feedback on the respective applications. This information is utilized by the ML model to identify and learn probability markers or traits to **make better selections** of loan applications for further scrutiny
- This process is looped back into the model with additional sample reports to improve sample selection in the next iteration

Related Technologies

- Machine Learning or ML is a sub-set of Artificial Intelligence (AI) and utilizes data samples and algorithms to mimic the humanized way of learning. Based on these learning and feedbacks, the machine continuously improves its accuracy in application
- ML utilizes statistical methods and algorithms to make predictions, classifications, and develop insights that drive the decision-making within the application or use case. The more sample data it utilizes, the better its decision making becomes
- These models are generally utilized in the space of big data when quick and accurate decisions or predictions are required by the user

Benefits

- Effective selection of cases for verification compared with a randomly selected sample
- In 2018, the utilization of this ML based model saved NOK 38.4 million in house inspection for the Norwegian Government

Implementation Complexity

- For the model to work effectively, good quality and volume of data is required
- Shortage of trained data scientists and specialists for developing and modeling the algorithm is a challenge

AI Black Box

- At initial stage, the set of probability marker used to train the model has not been reported

Main Sources

- AI use case for loan funds: AI revealed more loan fund cheating - Norway Today
- Norwegian Ministry website: The National Strategy for Artificial Intelligence - regjeringen.no
- AI benefits: Kunstig intelligens gir knallresultater i Lånekassen | by Johan Fu | Bransjebloggen 3min
- AI challenges: 96% of organizations run into problems with AI and machine learning projects – TechRepublic
- Machine learning and openness in digital management (Thesis): https://www.duo.uio.no/bitstream/handle/10852/69362/1/ENDELIG-UTKAST-MASTEROPPGAVE-FINF5002.pdf

Use Case: Tax Evasion
Country: USA
Technology: Data Analytics

Data analytics to help identify Crypto tax evasion

Business Problem - Context

- With **virtual currencies** like Bitcoin and nonfungible tokens taking center-stage in the global economy**, tax evasions** are becoming more difficult to track as there are various avenues available to people for storing their currency and not accounting for the same to the Internal Revenue Services of the USA (IRS)
- Hidden transactions of **cryptocurrency** has emerged as a new and evolved method for money laundering, and this has raised the requirement for digital asset auditing and new skills in asset tracing globally

Use Case Description

- In 2021, the IRS has launched an operation " Hidden Treasures" to **detect patterns** of tax evasion from a structural standpoint by utilizing AI and Data Analytics. They collect data from the citizens through Personally Identifiable Information (PIIs), dental records, etc., which is then sorted by these technologies to identify the assets owned by the individuals
- The IRS is also facing employee shortage, so AI and Data Analytics technologies are being utilized to assist and automate the department through monitoring of billions of digital transactions thereby reducing workload on IRS agents
- Additionally, the IRS Criminal Investigation has been involved with the tax authorities of four countries viz. Canada, UK, Australia and the Netherlands known as the "J5", through which they aim to address the non-compliance issues in **cryptocurrency and professional enablers**

Related Technologies

- Data Analytics refers to deriving valuable insights from a set of data (qualitative and quantitative). This involves extracting data and categorizing it in order to derive various patterns, relations, and connections
- Artificial Intelligence (AI) is a technology through which machines can simulate human intelligence and perform activities and task that require human intelligence such as learning, reasoning, problem solving etc.
- This is achieved through a certain sets of algorithms and rules, these AI systems can learn from the iterations of the task by utilizing the sample data from the computers

Benefits

- Improved tax estimation by identifying assets possessed by an individual
- Better traceability and identification of hidden currencies, risk areas and trends
- Real time analysis and traceability of digital transactions

Implementation Complexity

- Encryption, hacking and Virtual Private Networks (VPNs) might create problems to track cryptocurrencies

AI Black Box

- The American government has been working on data privacy and susceptibility to hacking and data manipulation of sensitive personal information that might lead to security threat

Main Sources

- Crypto Tax evasion: https://www.accountingtoday.com/news/irs-turns-to-data-analytics-to-track-crypto-tax-evasion
- AI at the US Department of Homeland Security: https://emerj.com/ai-sector-overviews/artificial-intelligence-homeland-security/

Use Case: Customs management
Country: Saudi Arabia
Technology: Facial and Image recognition

AI-based surveillance system applied in Customs

Business Problem - Context

- There has always been a continuous flow of goods in and out of Saudi Arabia through Al Hadithah, the transit point between Saudi Arabia and Jordan. Thus, customs department of Saudi Arabia wanted to **develop an AI-based surveillance system to ensure a smooth flow of goods through the check-point**. The objective was **to improve security at the transit point** along-with **productivity of the customs officers**

Use Case Description

- The customs department of Saudi Arabia, **collaborated with iOmniscent,** a leading company in the field of video analytics and surveillance solutions based out of Australia, **to develop an AI-based video analytics technology** to monitor the flow of goods and people through AL Hadithah.
- Implemented in 2021 across 200 cameras at the border, the technology included **behavior analytics and counting solution** which were designed to suit crowded and complex environments
- To improve security and productivity of customs officers, **multi-lingual license plate recognition system and face-recognition system** were developed **that enabled automatic detection and recognition of registration numbers of vehicles.** This system could recognize license plates of 130 countries

Related Technologies

- **Use of AI in video analytics** leverages deep learning advancements and processes signals from digital video by using a unique algorithm to perform security-related tasks such as identifying impostors, tracking people and objects, and detecting behaviors
- **Facial recognition** systems uses biometrics to map facial features from a video and compares the information with a database of faces to find a match
- **License plate recognition** systems consist of a frame-grabbing instrument that captures images with the aid of optical character reading (OCR) tools and translates the image pixels into visually readable digits

Benefits

- The behavior analytics and counting solution can work in very crowded environments. This increases the accuracy of detecting suspicious activities in crowded environments
- The face recognition system can work at long distances

Implementation Complexity

- Security component of the system is vulnerable to cyber attacks
- There may be inaccuracies in identification due to micro-changes in facial patterns and expressions, along with poor lighting and low resolution of images

AI Black Box

- The government recognizes issues with biases in AI algorithms and will formulate regulations to prevent them
- The government also plans to use AI in an innovative, responsible and ethical way to fulfil its Vision 2030 objectives

Main Sources

- Border security report - https://border-security-report.com/managing-customs-on-the-saudi-arabian-jordan-border/
- WIPO magazine - https://www.wipo.int/wipo_magazine/en/2018/05/article_0002.html

Use Case: Tax Frauds
Country: USA
Technology: ML, NLP, Graph Analytics

Software powered by ML and NLP to detect tax frauds

Business Problem - Context

- The US Internal Revenue Services (IRS) were **manually tracking and detecting criminal tax issues** such as tax fraud, hidden assets, money laundering, identity theft, and other non-compliances through revenue agents which was a **time taking process**
- Additionally, budget cuts and attrition increased burden on these agents. To solve this, the IRS decided to deploy an automated process to successfully identify instances of fraud and illegality before financial loss occurs

Use Case Description

- IRS in partnership with Palantir Technologies, an American software company that specializes in big data analytics, in 2018 developed and implemented an **AI program to mine data** from tax returns, property returns, bank reports, and even social media accounts. Post which, algorithms were employed to identify patterns and instances of taxpayer noncompliance
- The software will also use **natural language processing technology** that enables a computer to read and translate filings and contemplate their meaning, which will help predict IRS success in its appeals process

Related Technologies

- Machine Learning (ML) is a sub-set of Artificial Intelligence (AI), it utilizes data samples and algorithms to mimic the humanized way of learning. Based on these learning and feedbacks, the machine continuously improves its accuracy in application
- Using Natural Language Processing (NLP), AI algorithms can pre-process user input messages into tokens that are linguistically representatives. Normalization techniques are applied to remove human mistakes or slight deviations from a certain word. Those normalized tokens can then be grouped using pattern matching techniques to get an understanding of the user's request
- Graph Analytics is a tool used to apply algorithms to help analysts understand the relationship between graph database entries for identifying outliers, insights, value within the data

Benefits

• The technology was able to identify 296 suspected cases claiming USD 1.3 million	• Previously, 84% of these cases were not flagged during manual review. Thus, it helps uncover blind spots accurately

Implementation Complexity

- The system needs timely software and database updates to uncover new or existing blind spots that persist and facilitate fraudulent activities
- Manual intervention is required to fix bugs and errors in the system

Main Sources

- AI to help IRS efficiently identify tax crimes: https://aronsonllc.com/artificial-intelligence-to-help-resource-strapped-irs-more-efficiently-identify-tax-crimes/

Use Case: Taxation Process
Country: Australia
Technology: Robotic Process Automation and Deep Learning

RPA, deep learning support partial automation of tax filing process

Business Problem - Context

- In 2014, The Australian Taxation Office (ATO) observed that Australian tax filing process was **complex and tedious** for its citizens
- This created a lot of problems for the citizens and the government alike, as there were cases of **fraud and incorrect filings** been observed. To tackle the tax complexity problem, the government decided to implement AI-based systems to reduce time and effort

Use Case Description

- In 2016, Australian Taxation Office (ATO) developed a **highly automized tax-return process** for their official myTax website where the data provided to the ATO is used to prefill annual tax returns, so citizens won't have to locate and re-enter the information
- Predictive models and 'bots' are used to automate income tax forms processing. The predictive models are built on **AnalyticsNet** and the ATO uses 'bots' that are configurable software set up to perform certain tasks that the ATO can assign and control
- The ATO is performing **real-time analytics** and running **nearest neighbor models** over myTax data to better understand client behavior and prompt them to make certain choices in case any anomalies are detected by using a **Deep Learning technology** called the Artificial Neural Network (ANN)

Related Technologies

- Robotic Process Automation (RPA) combines APIs and UI to integrate and perform repetitive tasks between enterprise and productivity applications. This technology mimics the back-office tasks of human workers, such as extracting data, filling in forms, moving files, etc.
- Deep Learning, a subset of Machine Learning (ML), improves automation, performs analytical and physical tasks without human intervention. This technology is generally deployed to input and process unstructured data such as texts and images
- Artificial Neural Networks (ANN) are a subset of deep learning algorithms. They are comprised of node layers, input/output layers, and hidden layers that are assigned a weight or a threshold value.
- The nearest neighbor model analyses the activity of clients and compares it to groups with similar characteristics ('nearest neighbors') to determine what is likely to happen next

Benefits

- Reduced process time for form filing due to automation
- Reduced errors by means of "nearest neighbor" technique
- Reduced backlogs of pending tax claims by performing end to end business processes on a larger scale with very little human interaction

Implementation Complexity

- Human intervention is required in complex tax forms
- Slow load times due to heavy traffic during tax filing periods

AI Black Box

- ATO uses Google's open-source deep learning model framework, Tensorflow to enable real-time response requirements for the production environment
- **Robotic Process Automation:** Predictive models can identify the fix needed and then the bot can put the information in a file for ICP to process and release the form

Main Sources

- ATO kills e-tax: https://www.cio.com/article/3494354/ato-kills-e-tax-expands-mytax.html
- A New Approach to automating services: https://iorder.com.au/publication/Download.aspx?ProdID=1-F7AOIQH-P1

Use Case: Virtual assistant solution for self-service
Country: Australia
Technology: NLU, Conversational AI

Virtual assistants partially replace tax call-center services

Business Problem - Context

- Australian Tax Office (ATO) recorded **high call volumes** in the call-center during the start of the new Australian financial year (July-October). The officials realized the need to **expand its self-service offerings** for available to the clients and reduce load on call-centers

Use Case Description

- In 2016, ATO in collaboration with Nuance Communications, an American multinational computer software technology corporation, developed Alex, a **virtual assistant solution** on the ATO website
- Alex uses Natural Language Understanding, conversational dialogue and advanced resolution strategies to provide **contextual responses** within its scope of understanding. If Alex is unable to respond to certain queries, the conversation is passed to a live agent without the knowledge of the client
- The focus was to provide an **enhanced service offering** that can respond to the client based on their requirements and be printed and used for future client reference

Related Technologies

- Natural language understanding (NLU) comes into play to extract intents and entities from the processed text inputs. Responses databases are then queried to generate response candidates from which another algorithm selects the most suited one
- Conversational AI refers to the technologies that the users can talk to such as chatbots, voice assistant, etc. This technology uses ML, Natural Language processing, and big data to mimic human interactions through recognizing text or speech inputs across several languages

Benefits

- Within 18 months of its implementation, Alex engaged in **2 million conversations** with first contact **resolution rate of 88%**
- Contributed to **AUD 9.7 million** in client red tape reduction savings yearly and **8-10% reduction** in contact center calls

Implementation Complexity

- As ATO decided to scale up the utilization of this technology, it encountered more complex queries. To counter this, the ATO and Nuance identified gaps in Alex's knowledge base using hundreds of conversation transcripts as reference

Main Sources

- Nuance Communications: https://www.nuance.com/content/dam/nuance/en_au/collateral/enterprise/case-study/cs-australian-tax-office-en-au.pdf

Use Case: Cross-border tax evasion
Country: Australia
Technology: OCR and NLP

Machine learning, analytics to identify cross-border tax frauds

Business Problem - Context

- Paradise papers, a global investigation that took place into the offshore activities of certain individuals and entities, explored the leak of 13.4 million files (1.4 terabytes of data) of confidential electronic documents relating to offshore investments
- Subsequently, the Australian Tax Office (ATO) partnered with international agencies to **acquire sources of data** and identify possible Australian links. The ATO observed that a lot of time was taken to **manually review** different types of files associated with Paradise Papers on such a large scale as the **information came in various forms**, such as emails and Word documents

Use Case Description

- In 2017, the ATO used machine learning and other analytics **to understand the large amount of data** that came from the Paradise Papers
- Natural Language Processing techniques combined with scalable cluster computing infrastructure was used for:
  - Handling diverse file types
  - Converting the files into machine-encoded text using Optical Character Recognition(OCR),
  - Indexing and automatically or semi-automatically extracting names and other information of interest

Related Technologies

- Using Natural Language Processing (NLP), AI algorithms can pre-process user input messages into tokens that are linguistically representatives. Normalization techniques are applied to remove human mistakes or slight deviations from a certain word. Those normalized tokens can then be grouped using pattern matching techniques to get an understanding of the user's request
- Artificial Intelligence (AI) is a technology through which machines can simulate human intelligence and perform activities and task that require human intelligence such as learning, reasoning, problem solving, etc.
- Optical Character Recognition (OCR) is the electronic conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document

Benefits

- This has enabled the ATO to **reduce the amount of time and wastage** of resources on finding irrelevant data
- Enhanced effectiveness of **identifying higher-risk transactions** with the help of predictive and prescriptive risk models

Implementation Complexity

- The technology utilizes training data for decision-making. If the training data is biased, this might reflect in the deployed system. Thus, the system need knowledgeable humans to supervise their operation on an ongoing basis

AI Black Box

- Training sets of data were created to train ML models to detect patterns and relationships between people and entities using the Panama Papers and the ATO's internal data by searching for keywords

Main Sources

- A new approach to automating services: https://iorder.com.au/publication/Download.aspx?ProdID=1-F7AOIQH-P1
- ATO Website: https://www.ato.gov.au/Media-centre/Media-releases/ATO-statement-regarding-the--Paradise-Papers-/

Use Case: Airport security
Country: UK
Technology: Facial recognition, Deep learning

AI powered Facial recognition for airport security

Business Problem - Context

- The worldwide passenger numbers is reported to be **doubled**, estimated to be **over 8 million**, by the year 2037. This rise in the number of passengers will put a lot of pressure on airports in terms of processing passengers securely and safely
- The UK airport authority has identified the need to rely on technologies to offer more intuitive ways of processing rising numbers of passengers, but without compromising the **safety**

Use Case Description

- Aurora AI, a flexible AI development platform utilizing **Deep Learning technology**, has been implemented in UK's airport since 2012. The authorities, in 2019, have implemented a technology which utilizes both **Face recognition and AI together**
- The camera moves in a figure of '8' and captures image from multiple angles. This camera also includes superconducting detectors**, Kinetic Inductance Detectors**, which are extremely sensitive to photons in the millimeter-wave part of the spectrum and sends these images to AI for further processing
- AI scans the images and **generates a score** for each. This array of scores is then assessed against a threshold to determine if a threat has been detected or not

Related Technologies

- Artificial Intelligence (AI) is a technology through which machines can simulate human intelligence and perform activities and task that require human intelligence such as learning, reasoning, problem solving etc.
- Facial Recognition system matches a human face from a digital image or video frame based on the database it derives its information from, typically government approved IDs. The software then compares it with selected facial features to identify the human face

Benefits

- The final solution of detection and capturing of image is at 4 FPS (Frames Per Second) that helps detecting moving passengers clearly
- The technology doesn't require passengers to undress and reduces cases of false positive (5% of total cases) compared to traditional body scanners

Implementation Complexity

- The technology implemented present challenges for clearly identifying some objects like a knife, when seen from the sides as compared to the tip
- The AI is trained using the images captured from the camera and there's a lack of proper image samples for people carrying threats. Thus, the AI is unable to discriminate images based on the normal appearance

Main Sources

- How is AI being used to improve airport security?: https://sequestim.com/how-is-ai-being-used-to-improve-airport-security/
- Aurora AI website: https://aurora-ai.com/

Use Case: Tax Fraud
Country: France
Technology: Data mining and Machine learning

Identification and prediction of tax fraud using ML

Business Problem - Context

- The French Tax Authority (DGFiP) observed fraud in tax returns as a major problem for the country's economy
- DGFiP identified **slow and inefficient** tax filing systems and processes as the root cause for fraudsters being able to identify loopholes and evade the taxes

Use Case Description

- In 2014, The French Government deployed "Fraud targeting and maximization of requests value (CFVR)" that aims **to identify mistakes and fraud in tax** returns by using various parameters developed through the data collected about taxpayers
- The CFVR is fed with data from various **government-owned and private sector databases** such as banking accounts, tax declarations, foreign company listing, etc. ,by utilizing data mining to make accurate decisions with issues pertaining to tax frauds.
- CFVR analyses data from past (utilizing rule-based algorithm from previous technology) to find patterns in the behavior of businesses or households. This is achieved by ML, as the rule-based algorithm trains CFVR by using data derived from previously used technology

Related Technologies

- Machine Learning or ML is a sub-set of Artificial Intelligence (AI), it utilizes data samples and algorithms to mimic the humanized way of learning. Based on these learning and feedbacks, the machine continuously improves its accuracy in application. Machine learning algorithms perform text mining and text analysis to classify the text contained in the mails
- Data Mining is a procedure of extracting information from huge sets of data. These information can be used in applications such as Market Analysis and Management, Corporate Analysis & Risk Management, and Fraud Detection

Benefits

- In 2019, a total of approximately **56,000** cases were identified out of which **18,000** cases led to recovery of tax income
- In the same year, the application produced a **success rate of one in three** (1 in 3 cases scrutinized based on CVFR data resulted in recovery of tax)
- A total of **EUR 785 million** of tax income was recovered in 2019

Implementation Complexity

- Improving efficacy of targeting tax frauds by working on algorithms
- Limited access to data sources; looking to add more data sources from digital platforms to cover a larger number of taxpayers (businesses and households)

AI Black Box

- Identifies fraud by analyzing data from past controls and identifying behavioral patterns of businesses and households
- The ML tools observe network analysis to identify businesses and persons related/known to the fraudsters
- The data that is fed into the algorithms comes from government-owned and private-sector databases

Main Sources

- Automation in French Tax Authority: https://algorithmwatch.org/en/france-tax-automated-dgfip/

Use Case: BREXIT
Country: UK
Technology: Machine learning, Predictive supply chain

AI backed risk detection and compliance check system

Business Problem - Context

- With BREXIT expected to lead a transformation in the UK customs, the country has been reported to process more than **400 million customs declaration** per year
- Under the current system, the Customs Handling of Import and Export Freight (CHIEF) is incapable of handling the **increased volumes** of declaration claims, largely due to **capacity constraints**
- The new customs laws will require use of **data-driven solutions** to predict and manage the flow and provide a frictionless and intuitive service

Use Case Description

- As of 2021, the revenue and customs department have collaborated with Capgemini and Early metrics, a company that specializes in rating startups and other emerging technologies, to prepare a highly data-driven system at the border with the help of artificial intelligence backed **risk detection and compliance check system**
- **Predictive supply chain technology** will be used to proactively prepare the required documents in advance, for the arrival of the goods carrying vehicles. They will then be required to pass through checkpoints and their number plates will be scanned and logged with AI-backed surveillance technology. Simultaneously, the freight on board will be scanned for counterfeit or dangerous material
- Machine learning will be incorporated with IoT to create a comprehensive **digitized customs clearance solution** wherein a set algorithm will be responsible for compliance check

Related Technologies

- Predictive supply chain technology is basically supply chain solutions that are based on artificial intelligence and machine learning. With the help of machine learning, the technology mines huge sets of historical data, finds patterns and subsequently creates a predictive a model that is then able to make autonomous decisions as part of the supply chain process.
- Use of AI in video analytics and surveillance leverages deep learning advancements and processes signals from digital video by using a unique algorithm to perform security-related tasks such as identifying impostors, tracking people and objects, and detecting behaviors

Benefits

- The system will allow consolidation of processes, systems, and coordination of government messaging
- A key advantage will be that of ease of scalability and relatively lower variable cost

Implementation Complexity

- One of the challenges is to obtain large sets of reliable data as it will require extensive collaboration between the government and industry
- Another point of concern is the skill-gap with regards to the implementation of the new technology

Main Sources

- Capgemini report on Digital Future: Digital Futures - The Future of UK Border | Capgemini

Use Case: Passenger identity verification from massive datasets
Country: USA
Technology: Machine learning

Human guided machine learning to verify passenger identity

Business Problem - Context

- The US Customs and Border Protection (CBP) agency was looking to develop and release the 'Global Travel Assessment System (GTAS)', an open-source platform that could receive and store standard air traveler information (Advanced Passenger Information (API) and Passenger Name Record (PNR)) enabling real-time risk-modeling
- The CBP wanted to use biometric data available from API/PNR to **match passengers against an existing database**. Additionally, they required discrete passenger identity matching from API and PNR data, along with historical data to build personalized profiles for passengers over time

Use Case Description

- The CBP in 2018, collaborated with Tamr, a company that provides data unification solutions using machine learning, to enhance algorithms used for **passenger identification and matching for GTAS**. They used human-guided machine learning to address CBPs requirements by:
  - Developing Tamr's machine learning platform to **resolve passenger identity from massive datasets** of passenger (API) and reservation (PNR) data;
  - 'Training' of the machine learning platform with the help of continuous input from region-specific subject matter experts in the travel domain
- The output was shown in the form of probability of match (in a decimal number form), which was used by managers to make informed decisions

Related Technologies

- **Human-guided machine learning** is a process where subject matter experts provide inputs to the system in real-time, which is then used as a baseline for future predictions and modelling. It works using a bottom-top approach by first using algorithms to conduct the heavy lifting of identifying relationships within the data, and engaging humans when necessary for training or validation
- As most of the workload is taken up by machine learning, with subject matter experts only validating the datasets, the technology is highly reliable, scalable as it doesn't solely depend on the human brain and is very quick to respond.

Benefits

- The application provides the necessary capabilities to pre-screen travelers, ensuring criminals and terrorists are identified correctly and on time
- The benefits also extend into public wherein identification of people traveling from a particular country of outbreak can be done

Implementation Complexity

- Preparation of datasets and pre-processing the inputs usually are a time-intensive exercise.
- Moreover, despite the involvement of subject matter experts, the system is vulnerable to genuine human error, leading to inaccuracies in the algorithm's output.

Main Sources

- US Department of Homeland Security: https://emerj.com/ai-case-studies/the-department-of-homeland-security-uses-ai-enhanced-entity-resolution-for-its-global-travel-assessment-system-gtas/
- Tamr website - Tamr Provides Rapid, Trusted Entity Resolution To The Department of Homeland Security - Tamr Inc.

Use Case: AI backed RADAR for customs
Country: USA
Technology: Deep learning, AI backed RADAR system

Artificial intelligence backed RADAR system applied in customs

Business Problem - Context

- The U.S. Customs and Border Protection (CBP) agency monitors everything that happens along US' borders on land, air, and sea. It hunts for suspicious individuals and activities while at the same time letting "trusted travelers" through expeditiously. With 328 ports of entry into the US and almost a million people going through them each day, the **workload is huge**
- To manage this workload efficiently for informed decision making and increased strength of security, CBP wanted to use artificial intelligence

Use Case Description

- In 2016, the CBP collaborated with Echodyne, a company that creates radar technology for defense, security and autonomous applications, to **leverage its Metamaterial Electronically Scanning Array (MESA) to automate certain process** to reduce workload.
- The radar was integrated with artificial intelligence, hardware sensors and machine-learning software**. Deep learning algorithms were fed into the system that helped in surveillance and entity detection** (identification of people or objects based on pre-fed datasets) providing warning signs, classification of potential threat and any abnormal activity in the environment

Related Technologies

- The MESA radar technology leverages electronically scanned radar technology. The active electronically scanned arrays enable surveillance with higher update rates and can double detection range as well. Variable range and angle sectors are enabled while still maintaining a 360- degree background for situational awareness
- Deep Learning, a subset of Machine Learning (ML), improves automation, performs analytical and physical tasks without human intervention. This technology is generally deployed to input and process unstructured data such as texts and images
- Use of AI in video analytics and surveillance leverages deep learning advancements and processes signals from digital video by using a unique algorithm to perform security-related tasks such as identifying impostors, tracking people and objects, and detecting behaviors

Benefits

- The deep learning algorithms allowed the AI backed radar system to offer a more stable communication based on optimal utilization of radio frequency spectrum and minimizing interference between users
- It was able to provide comprehensive surveillance capabilities with special emphasis on threat detection even under dynamically changing environments such as change in lightings, poor weather conditions, etc.

Implementation Complexity

- A key complexity that remains is the risk of malicious manipulation through system hacking along with concerns of data leak
- Sometimes the entity detection of these systems may not be as accurate as they are expected to be leading to incorrect decision making

Main Sources

- US Department of Homeland Security: https://emerj.com/ai-sector-overviews/artificial-intelligence-homeland-security/

Use Case: Voicebots for Tax-related queries
Country: Ireland
Technology: NLP, NLU, Sentiment Analysis

AI voicebots for Tax-related queries

Business Problem - Context

- The Office of the Revenue Commissioners of Ireland receives about **3 million calls** a year pertaining to tax clearance and other finance-related topics. Many of these calls are repetitive and answering the queries is a time-consuming process
- Irish Government wanted to run a pilot program to leverage artificial intelligence to effectively answer tax-related queries. The objective was to provide a **24/7 automated service**, improving the customer experience

Use Case Description

- In 2018, the Office of the Revenue Commissioners of Ireland, in collaboration with Accenture, implemented a **Virtual Digital Agent (VDA),** which is a voicebot designed to answer calls related to tax-clearance. The voicebot was designed to include 6 call categories in total
- Virtual Digital Agent uses a combination of Natural Language Programing(NLP), speech synthesis and text-to-speech to pick up the call, understand the queries and appropriately answer them.
- The voicebot incorporated **200 unique dialogue steps** and addressed 18 use cases. It developed an ability to recognize **21 caller intents**
- The voice bot includes **smart suggestions** where the conversations are based on **current status of customer** rather than 'How can I help you' greeting at the beginning of the conversation

Related Technologies

- Using Natural Language Processing (NLP), AI algorithms can pre-process user input messages into tokens that are linguistically representatives. Normalization techniques are applied to remove human mistakes or slight deviations from a certain word. Those normalized tokens can then be grouped using pattern matching techniques to get an understanding of the user's request
- Natural language understanding (NLU) comes into play to extract intents and entities from the processed text inputs. Responses databases are then queried in order to generate response candidates from which another algorithm selects the most suited one
- In voicebots, sentiment analysis is also often used to analyze the user's experience and transfer the conversation to a human if deemed necessary

Benefits

- Voicebot handled up to 50% of calls from start to finish
- About 70% of first-time applicants interacted with the voicebot
- About 75% of tax -clearance applicants managed to retrieve their access number through the voicebot

Implementation Complexity

- Implementation of secure authentication model which is suitable for AI-based NLP technologies remains a challenge
- Some cases require communication through multiple channels, which necessitates the development of omni-channel platform

AI Black Box

- The customer's call is initially connected to cloud where speech gets converted to text.
- The text is sent to another cloud where caller's intent and appropriate response are determined
- The text response is finally converted to speech

Main Sources

- Department of public expenditure and reform of Ireland website: https://www.ops.gov.ie/actions/innovating-for-our-future/innovation/emerging-technology/artificial-intelligence/revenue-chatbot/
- Conversational AI voicebot- Accenture: https://www.accenture.com/_acnmedia/PDF-96/Accenture-HPS-IRA-Creds-d5-Final.pdf#zoom=50
- Irish Revenue Commissioners website: https://www.ops.gov.ie/app/uploads/2021/03/Revenue_VDA_Dept_of_Defence_2021.pdf

Use Case: Virtual assistant for re-directing customer queries
Country: Estonia
Technology: NLP, NLU, Sentiment Analysis

Virtual assistant for re-directing customer queries

Business Problem - Context

- The citizens of Estonia have a lot of queries most of which are directed towards specific agencies of the government
- The citizens are not fully aware of the roles and responsibilities shouldered by the agencies. As a result, citizens end up **contacting the incorrect agency** for resolution of their queries
- The Estonian Government wanted to develop AI-based tools to help citizens direct their queries to the **appropriate agency**

Use Case Description

- In 2020, The Estonian Government started developing a **voice-based virtual assistant** under the **KrattAI concept,** which is part of the National AI Strategy program
- Through the voice-based interactions, the virtual assistant will be able to direct citizens to the appropriate public agency depending upon the nature of query. The virtual assistant will **use NLP to identify the query** and subsequently transfer it to the correct agency
- Citizens would also be able to execute tasks such as **filing taxes** and **passport applications** by making a voice request to the agency
- The virtual assistant is currently being tested in the **police and border control** departments

Related Technologies

- Using Natural Language Processing (NLP), AI algorithms can pre-process user input messages into tokens that are linguistically representatives. Normalization techniques are applied to remove human mistakes or slight deviations from a certain word. Those normalized tokens can then be grouped using pattern matching techniques to get an understanding of the user's request
- Natural language understanding (NLU) comes into play to extract intents and entities from the processed text inputs. Response databases are then queried to generate response candidates from which another algorithm selects the most suited one
- Sentiment analysis is also often used to analyze the user's experience and transfer the conversation to a human if deemed necessary

Benefits

- The virtual assistant will be able to provide personalized services through voice-based interactions
- Since the query is transferred to the right agency, a lot of time which was previously wasted in contacting the right agency, is now saved

Implementation Complexity

- Since the virtual assistant will handle a large number of interactions, privacy issues might arise
- Virtual assistant has to be designed to handle queries which requires two or more agencies to collaborate

AI Black Box

- Estonian government is working with Estonian block chain company Guardtime to develop measures of anti-tampering and data integrity in the algorithms used for implementation of AI.
- The block-chain system ensures transparency and integrity, and can identify hacks in just a second

Main Sources

- Voice-based virtual assistant: https://www.itu.int/en/myitu/News/2020/11/18/08/19/Estonia-public-services-Kratt-AI-governance
- AI Black Box: https://www.globalgovernmentforum.com/estonian-citizens-to-access-public-services-through-virtual-assistants/
- Implementation challenges: https://www.e-marketingassociates.com/blog/pros-and-cons-of-a-virtual-assistant

Use Case: False reporting identification
Country: Spain
Technology: Cognitive Robotics, Process Automation

Cognitive robotics, process automation to spot false crime reporting

Business Problem - Context

- With cases of **fake reporting of cases rising in Spain**, especially for low level crimes, there has been an increased need for a model that helps identify such cases with the help of technology. The fraudulent reporting of cases brings with itself significant consequences for individuals.
- Additionally, it wastes valuable police resources and is often used in combination with other fraudulent behavior

Use Case Description

- To detect false reports of cases, the Spanish national police adopted the VeriPol AI system in 2018. It was designed to be integrated into the existing Spanish National Police information system called SIDENPOL, allowing for easier use and integration into existing work practices
- VeriPol uses a **combination of Natural Language Processing (NLP) and machine learning classification algorithms** that can **estimate the probability of false police reports** with high accuracy rates
- Moreover, the system also provides insights into characteristics of false and true police cases, making identification much easier. For example, pilot studies found that false police reports are more likely to include shorter statements, focused on the objects that were stolen and lacking details

Related Technologies

- Cognitive robotics involves endowing a system software (or a robot) with higher-level cognitive functions that enable them to reason, act in a robust manner under changing, completely known, and unpredictable environments. It integrates reasoning, perception and action with a uniform theoretical and implementation framework
- Using Natural Language Processing (NLP), AI algorithms can pre-process user input messages into tokens that are linguistically representatives. Normalization techniques are applied to remove human mistakes or slight deviations from a certain word. Those normalized tokens can then be grouped using pattern matching techniques to get an understanding of the user's request

Benefits

- The system can both detect false reports early, leaving more police resources available to focus on other tasks and reports
- It will also deter people from filing fake statements in the future
- The system will help provide more insights into how people lie to police officers and the features of a false police case

Implementation Complexity

- There has been a sense of resistance towards the adoption of the technology by different police stations. This has been hypothesized using the fact that there has been no significant dip in cases of crime, despite the drop in false reporting.
- Many police officers consider its use to be limited to theoretical applications only and does not integrate well with the real-time work that the police officers do.

Main Sources

- Challenges: Challenges in adoption
- AI in Public Services: https://publications.jrc.ec.europa.eu/repository/bitstream/JRC120399/jrc120399_misuraca-ai-watch_public-services_30062020_def.pdf https:/inews.co.uk/news/science/how-artificial-intelligence-is-spotting-fake-police-reports-214425

Visit our website:

Deloitte.

Funded by
the European Union

Find out more
about the Technical
Support Instrument: