INDUSTRIAL
STRATEGY

The Use of Emerging Technologies for Regulation

Annex 1 – Case Studies

BEIS Research Paper Number 2020/041

Authors

Kate McEntaggart, Janne Fillet, Beatrice Marinoiu, Helen Quaresma, Simona de Paolis

Acknowledgements

The participation of interviewees to inform the case studies is gratefully acknowledged. The study team is also grateful to previous work carried out on data-driven and algorithmic regulation for BEIS by Dr Panos Panagiotopoulos and the Better Regulation Executive.

Contents

Introduction

This annex includes full summaries of each of the 12 case studies used as evidence for the findings identified in the main report. The 12 case studies are listed in order in Table 1 below.

Table 1: Case studies

Example	Technology function
The Intellectual Property Office (IPO) and Machine Learning (ML)	The IPO explored the use of machine learning (ML) tools to improve processes in the registration of patents and trade marks. ML has been used to help examiners identify similar trade marks and for patents, prior art more efficiently than this was usually done.
The Rural Payments Agency (RPA) and CROME	RPA used a type of supervised machine learning to develop a crop map of England (CROME) based on satellite images and image classification methods. This can then be used to better assess/allocate rural payments and avoid fraud.
A consortium of private firms and public sector bodies' use of sensor technologies for mosquito classification (the Vectrack Project, EU)	The Vectrack system uses sensors and machine learning to collect and classify mosquitos according to several criteria such as gender, provenance and risk of disease in real time.
Ofcom and the use of blockchain for a common numbering database	Ofcom has piloted the use of blockchain to support the development of a common database of landline numbers in the UK.
The Finnish Government's use of a Virtual Agent Network	The Starting Up Smoothly project explores the use of chatbots to provide an integrated approach to customer service between different organisations. It focused on the test case of linked chatbots to address the needs of foreign entrepreneurs looking to start businesses in Finland.
The Ministry of Justice (MoJ)'s development of an intelligent search tool	The MoJ developed an intelligent search tool for prison reports using natural language processing (NLP) techniques. The tool brings together reports that were previously dispersed in different locations and allows users to conduct searches that return contextually similar results in addition to exact matches.
North Tyneside Council's use of RPA	North Tyneside launched a project to apply robotic process automation (RPA) to repetitive tasks within the Council's services.
The Financial Conduct Authority (FCA)'s Data Strategy	This case study focuses on the FCA's adoption of a new Data Strategy, rather than a specific tool. The Data Strategy involves several organisational change projects: bringing in infrastructure, hiring internal data science teams, improving data management practices and promoting an innovation-friendly culture.
Natural England's use of eDNA	Natural England employed the use of environmental DNA (eDNA) for sampling Great Crested Newts, a protected species under UK and European Law.
The Environment Agency's use of webcams	The Environmental Agency employed the use of webcams and image differentiation technologies to monitor culverts and respond to flood hazards.

The Driver and Vehicle Stands Agency (DVSA)'s Intelligent Risk Rating Assessment tool	The DVSA has developed a tool that uses ML algorithms for pattern identification on MOT Testers. This provides risk ratings facilitating targeted MOT inspections based on past behaviour, allowing resources to be focused on garages with highest risk scores.
The City of Chicago and WindyGrid	WindyGrid is a platform that aggregates City data from across departments. The platform has been used as a basis to develop features for predictive analytics, which leverage this data to improve city services.

1. The Intellectual Property Office (IPO) and Machine Learning

Introduction

The following case study details the Intellectual Property Office's (IPO) exploration of the use of machine learning (ML) and artificial intelligence (AI) to help improve some of the IPO's processes. The IPO is an executive agency in the UK responsible for intellectual property rights. It handles applications and maintains registers for patents, trademarks and designs.

Best practices

The IPO case illustrates several instances of good practice described in the main report:

1. The IPO took a problem-focused approach to choosing and developing its tools, and the interviewees emphasised the importance of this approach. They considered a range of options—for the patents tool, this was a formal part of the feasibility study—including off-the-shelf and previously tested algorithms. More established technologies, such as the use of more basic algorithms, were also incorporated as part of the solution.

2. The IPO case provides a good example of learning from others. International cooperation is common amongst intellectual property offices and working with these pre-existing contacts has helped the IPO to both inform the early development of the tool and to identify future projects.

3. Infrastructure—notably cloud computing—was already in place, and this gave the IPO access to the specialised hardware needed for machine learning processes.

4. The projects were carried out by a considered mix of in-house staff and external contractors to match the skills and capacity of the IPO, and the IPO plans to build up internal capacity to better support such projects in future. External contractors both brought in essential skills and were also able to share skills and learning with the organisation.

5. The projects have made use of internal user testing: helping to engage staff in the process and ensuring that the development team understood staff needs and requirements.

The IPO has explored the use of machine learning for two main applications: one for *trade marks* and another for *patents*. The trade marks tool is at an advanced stage of development and is ready to be implemented in practice. The patents tool was not in development at the time this case study was written, the IPO having recently completed a feasibility study.[1]

[1] The results of this feasibility study have been published and can be accessed at: Cardiff University (2020) AI-assisted patent prior art searching – feasibility study. Available online at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/880212/AI-assisted-patent-prior-art-searching.pdf

This case study is based on two interviews with IPO staff members: one familiar with the trade marks case and the other familiar with the patents case.

The trade mark tool

When an application for a trade mark is submitted, it normally undergoes two types of assessment. The IPO first completes an *absolute grounds* check, where the submitted trade mark is checked against a pre-set list of attributes that would disqualify a trade mark from being registered. This includes, for example, checking whether the proposed trade mark uses the Queen's crest or vulgarities. The trade mark is then examined on *relative grounds*. This examines the application in relation to trade marks that have already been registered.

To examine an application on relative grounds, the trade mark needs to be compared against a large data set of approximately 2 million trade marks. The current approach to these examinations involves using pre-existing tags, for which there is an internationally agreed standard. For example, if a trade mark contains a picture of a lion, this would be tagged with the internationally recognised code for lions, and a trade mark examiner would compare any new application for a trade mark with a lion against other trade marks tagged with that code.

The tool that has been developed assists with this examination process, using AI to provide a ranking search result of similar trade marks for a relative grounds check. Rather than searching the tags, it matches images based on similarity, providing examiners with more directly relevant results. This AI component was part of a broader business solution that also includes non-AI components, such as an *absolute grounds* check for foul language. The IPO intends to use these components both for a tool to assist examiners and for a pre-apply tool for trade mark applicants. The pre-apply tool would provide applicants with an overview of possible conflicts between their trade mark and existing registered trade marks, and a better understanding of both relative grounds and absolute grounds.

The tool for patent prior art searches

The feasibility study for the use of AI technologies in the patent prosecution process related to patent prior art searches. Prior art searches are a complex process currently undertaken by patent examiners to identify similarities between a new patent application and existing patents, as well as other non-patent sources such as academic literature, journal publications and other public disclosures. These are currently undertaken by patent examiners specialised in different technologies. These searches can be time-consuming and current search tools can provide a low proportion of relevant results. Unlike for the trade mark case, the complex way that patents are drafted meant that there was no obvious solution to pursue from the outset. Therefore, the feasibility study was intended to determine the extent to which AI could help examiners identify the most relevant prior art more efficiently. It comprised a literature review, user research, the development of a proof-of-concept and an evaluation of the proof-of-concept. The results of the feasibility study have indicated that further research is needed before the IPO can develop an appropriate tool.

Timeline

Both projects were funded through the Regulators' Pioneer Fund.[2] The IPO applied for funding in August 2018, received funding in November 2018 and began the projects in January 2019. The first phase of the trade marks project – the proof of concept – finished in July 2019. The core AI development was 4 to 5 months long, before the launch of the full implementation, with a plan to go live internally in March 2020. The feasibility study for patent prior art searches using AI began in January 2019 and finished in December 2019.

Image recognition is a fairly well understood use-case for AI and had already been explored by some of the IPO's peers in the context of intellectual property verification and registration. Similar work had been done by the World Intellectual Property Organisation (WIPO) and by IP Australia, which the IPO reviewed. In the development phase, IPO staff visited the WIPO for discussions on the WIPO's algorithms and the IPO's project. The insights gained were used to build a business case for the development of the trade mark tool.

Relevant technologies

Cloud computing infrastructure needed to be in place for the development of the trade mark tool and broader business solution. The tool was hosted on Microsoft Azure. Graphics Processing Unit (GPU) instances on Azure were used to host the algorithm model and conduct training.

A mix of off-the-shelf and newly developed algorithms have been used. The core algorithms in use are open source algorithms. Modelling used standard computer vision architectures and pre-trained models, including R-CNN algorithms (to look for relevant areas within an image to compare against) and a VGG 19 neural network (to vectorize the images for comparison). Vectorization converts images into a format that can be more easily processed by algorithms.

In addition to the image recognition portion of the trade mark tool, some Natural Language Processing (NLP) has also been incorporated. For this, seven different word-based algorithms are in use to identify words that look or sound similar to other words (e.g. homophones as well as those similar in spelling).

The patents tool prototype also uses a mix of algorithms, including NLP and computational semantics. However, as noted in the feasibility study, there is further research needed to identify algorithms better suited to the prior art search process.

Developing the technology

Several factors appear to have contributed to the successful development of the technological solution at the IPO. These include:

- the use of a third-party contractor to provide technical expertise;

[2] The Regulators' Pioneer Fund is a competition-based intervention to finance regulator-led projects and is overseen by the Better Regulation Executive within BEIS. The Fund aims to develop a UK regulatory environment that encourages business-led investment, particularly in science and technology-driven innovation.

- an assessment of the suitability of existing tools;
- a shared understanding of needs and possibilities between the IPO and the contractor; and
- having high quality data to support ML.

In both cases, third-party contractors were used to help deliver the projects. A private consulting firm worked on the trade mark tool, while researchers at Cardiff University completed the patents feasibility study. However, the IPO has within its own staff individuals with relevant skills. In the case of the trade marks tool, in-house staff built the holistic solution and developed the front-end components.

According to the IPO, the decision to use a **third-party contractor** for the AI component of the trade marks project was done to ensure that they had access to the right expertise. There was also a sense that external consultants would have experience working on these issues across organisations and industries, and would therefore be more likely to **'think outside the box'**. The interviewee strongly recommended the use of external contractors for conducting such work, even if only used at initial development phases of the interview. For the patents feasibility study, they noted that, were this project to be implemented in the future, they would likely contract out components of the work again.

In both cases, **assessing the suitability of existing techniques** was needed. For the patents feasibility study, this involved a literature review to identify what had already been done. In the case of prior art searches, the review found no existing AI technologies that could fully automate the search process. For the trade marks case, image recognition being a well-understood application of AI, several existing algorithms were available. Some alterations were needed, however. For example, many of the existing algorithms for image recognition have been developed for photograph-style images. The images used for trade marks are generally cartoon-style images and therefore many of the existing algorithms are not directly transferable. Although the concept is similar, some additional work needed to be done to ensure that these were applicable.

Another key step in the process for both projects was to ensure there was **a good understanding of needs and possibilities from both sides**: at the IPO's and at the contractor's. One challenge mentioned by both interviewees is a misunderstanding from some within the organisation as to what AI can offer. In the trade marks case, the third-party contractor helped to address this to some extent by running a general session on AI for staff. According to the IPO, this session was well received and helped many in the organisation better understand what AI could address and what it could not. For the patents project, the interviewee also highlighted the need to give academic researchers a firm understanding of the processes currently used by examiners. They observed and interviewed IPO patent examiners in a range of different technology areas to define key user requirements before building their concept model.

Both tools use ML and therefore required **data to train algorithms**. For both patents and trade marks, the IPO owned all of the relevant data needed to train the tools. The available data sets are well-suited for machine learning applications: they are well-structured and high-quality, and because patent and trade mark data is publicly accessible, there were no legal or usage concerns. Patents data presents additional challenges due to its complexity. For example, patents can sometimes be narrow in scope (e.g. related to one part of a system), but the text within the patent will reference a broader system. This nuance is straightforward for human examiners, but it can be more challenging for algorithms to identify which parts of the text are the most important. In the case of prior art searches, this contributed to the algorithms used for the proof-of-concept suggesting some non-relevant results when ranking documents.

For the trade marks case, there were some general challenges around the large **volume of relevant data**. Although having access to sufficient, high-quality data is important for successful machine learning applications, large volumes of data can contribute to logistical challenges. This meant that sufficient time was needed for data preparation and processing. Testing and validation had been done with a smaller data set, and although the majority of the data for trade marks is high quality, there were a few instances in the larger data set where image files were missing or corrupted. This led to a few false starts during the training, where the data was loaded and the data set broke partway through. The interviewee noted that if the patents feasibility study were to be taken forward as a deployable tool, the volume of data would be an even bigger problem. The trade mark tool needs to search around 2 million trade marks and there is limited meta data; by contrast, a fully functioning patents tool would need to search over 110 million patent documents, which contain more metadata, are lengthier and more complex than the trade mark equivalents.

Implementation, monitoring and next steps

Further insights have been collected on good practices and learning points for the implementation and monitoring of the solutions, with regards to the trade marks tool, which was being rolled out at the time this case study was completed. For the patents tool, as the feasibility study had just ended, no decision had been made yet as to how the IPO would proceed.

Once the trade marks tool has been rolled out, the tool will require **monitoring** and **improvement**. For that purpose, the IPO have incorporated a feedback model, where examiners can assess the ranking of similar trade marks that the algorithm returns as they use the tool. This data will then allow the IPO to continuously train and improve the algorithm. The IPO has planned to carry out monitoring and updates in-house. To support this, there were plans to **hire junior machine learning specialists** who would be responsible for continuously updating the model and the algorithm based on user feedback. If more significant changes were needed, they would potentially bring the third-party contractor back, depending on how quickly the IPO could grow the internal capability.

At the time of this case study, the IPO was going through a significant digital transformation programme, which would likely take two to three years to complete. The first step to this transformation was **improving the infrastructure** on the input side (for example, improving electronic filing capabilities for customers). The machine learning tools have fed into this strategy, but they are not the priority. Other digital capabilities, such as ensuring customers can receive all reports electronically and access these on a central online platform, would need to be established first. However, one interviewee stressed that this did not mean that research into AI applications should be put on hold. Rather, because research and testing in this area takes time, it could be conducted in parallel to other projects.

The IPO has considered how the use of these tools could be developed further and expanded upon in the future. One possibility considered was to use algorithms to make decisions on certain applications, rather than simply assisting examiners in their decisions. This would relate mostly to **automating rejections** for trade marks on absolute grounds: for example, where foul language is detected. However, using the tool to make these types of decisions, rather than as decision-making aide, could have legal implications (e.g. the right to explanation).

Another potential future legal issue relates to the **use of tagging data**. The current system of tagging images is based on international treaties and participation in a worldwide standard. At the time of completing the case study, the algorithm would still make use of some of that tagging information, but as the image recognition process improves, this could make tagging no longer necessary to identify similar images. A move away from that tagging system would have implications for those international treaties in a regulatory regime that is highly dependent on international cooperation.

As noted in the previous section, international cooperation has facilitated the development of these tools. The IPO has learnt from the experiences of international colleagues, and they have identified opportunities for further collaboration. For example, IP Australia has a data set that uses freeform tagging for logos, which could potentially be used to make further improvements to the algorithms for trade marks. The process of receiving that data set was still at an early stage, but cooperation has been fairly straightforward. For the patent prior art search tool, the interviewee also noted that the way forward will likely involve **working collaboratively with international partners**. Collaboration is straightforward because patent examiners in different jurisdictions approach these tasks in the same way and are not in competition with one another. Once completed, the feasibility study could be discussed with peers from third countries.

The tool developed for trade marks could also be applied to some extent to patents and designs, since the latter are based largely on drawings. Much of the algorithm that has been created for trade marks could be reused. Since some patents will contain images, a similar technique could be applied there too, although this had not been extensively explored yet at the time of completing the case study.

Impact

As these tools had not yet been fully implemented at the time of completing this case study, their impact was not yet measured. The trade marks tool was to be integrated into both a pre-apply tool for applicants and a tool for examiners. The IPO expected that the pre-apply tool would help **improve the quality of applications received**. A similar initiative by IP Australia resulted in an 18% increase in successful applications; the IPO's tool may yield similar benefits. Alongside the improvement in application quality, the tool for examiners should lead to a **more efficient examination process**: the expectation was a 20% reduction in the time taken to examine an application. The IPO was to conduct analysis to evaluate the impact and were working on increasing tracking for that purpose.

The patent tool was expected to reduce the time examiners would spend searching for prior art. The interviewee noted that this could have one of two impacts. On the one hand, examiners could **process applications more rapidly**. On the other hand, examiners could use the time otherwise gained to **increase the overall quality** of their work.

Measuring improvements in quality, rather than efficiency, presented a challenge. This could be done through qualitative research with relevant stakeholders, such as patent attorneys familiar with patent processes in different jurisdictions. However, it generally takes between four to five years for a patent to be granted and then additional time before it would be challenged in court. This means that improvements to the quality of examinations would be apparent only several years after implementation.

Lessons learnt

For both cases, interviewees noted that **stakeholder management** was important to the success of the projects and that this was something that they had initially underestimated. The work done to **engage and help staff understand the algorithm**, its likely performance and what expectations should be was important to helping avoid misunderstandings and achieving buy-in. Interviewees felt this could have been usefully done earlier and more assertively.

Interviewees noted the many challenges to developing AI tools, which stem from misunderstandings of AI and what it can solve. They noted that many would see AI as "a silver bullet", however it is not always the right solution. For the trade marks case, the interviewee felt that this strong focus on the AI component meant that there was **not enough focus on the non-AI elements of the solution** under development, such as the more basic algorithms and components incorporated. As such, there was an important need to step back and consider the broader need and business problem.

Similarly, the interviewees emphasised the importance of **having a clear conception of the end-goal** when developing an AI tool. Rather than thinking about AI in a broad way, which is a common approach and can lead to many unnecessary tangents, it is important to know exactly what the algorithm should automate or augment and then identify how AI can be used to assist that.

The interviewees also noted that there is an expectation that AI will completely automate or replace jobs, but in the case of both tools developed by the IPO, these were conceived as human-in-the-loop tools from the outset. The intention was to change the way work was done and improve the efficiency and quality of outputs: to achieve these things, engaging the relevant staff in developing and understanding the process was essential.

One interviewee noted that the uncertainty around the development of algorithms meant that there was a challenge in scoping and defining the business benefits of research. Because it was difficult to know from the outset how well an algorithm would perform, it was **challenging to guarantee a business benefit** in terms of hard figures.

2. The Rural Payments Agency (RPA) and CROME

Introduction

The following case study details the Rural Payments Agency's (RPA) work to develop the Crop Map of England (CROME), which uses satellite imagery and image classification methods to develop a better understanding of crop types and land usage in England. The RPA is an agency under the Department for Environment, Food and Rural Affairs (Defra) and is responsible for the payments of rural subsidies to farmers and landowners in England.

Best practices

The RPA case illustrates several instances of good practice described in the main report:

1. Building on and improving on work that has come before, while actively considering what future projects work may lead to. The RPA's tool used elements from a previously developed tool, and the team has actively considered how they might build on this in future.

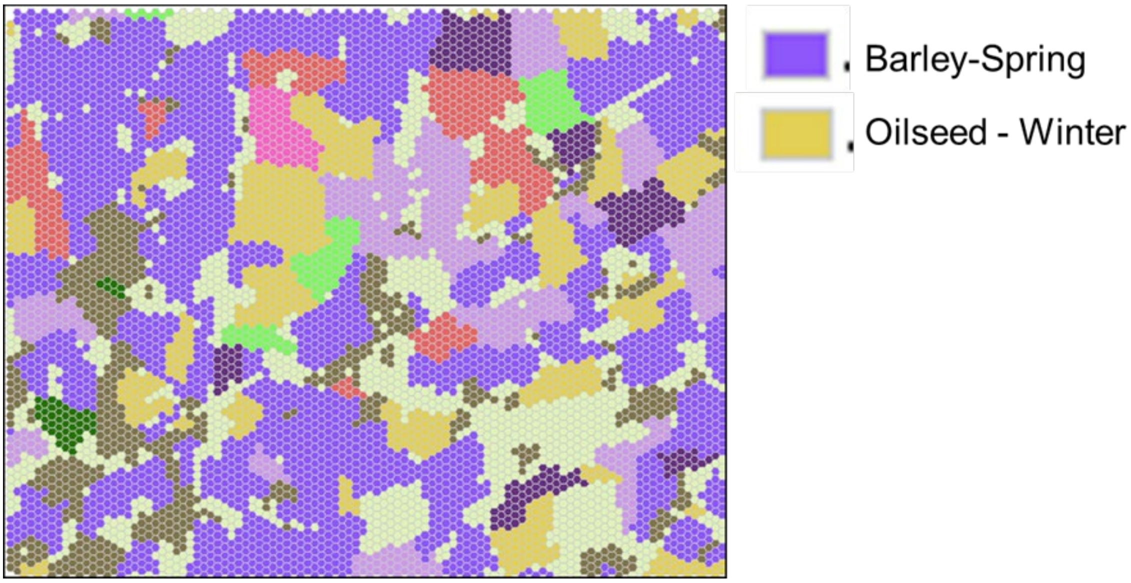
2. Identifying and developing solutions to intellectual property issues early on: considering alternate data sources. Because certain types of satellite data were not open source, the RPA needed to identify an alternative option.

3. Creating an open source tool with multiple uses, thereby increasing the tool's impact. The tool is useful not only for determining agricultural payments, but also has uses for other Defra delivery bodies, as well as private sector organisations, charities and academia.

Subsidies are currently set according to the EU's Common Agricultural Policy (CAP). The CAP includes area-based payments: the more land that is owned by a beneficiary, the higher the subsidy payment. CAP payments in the UK amount to around €4 billion per year. Controls on payment applications were previously done through regular field inspections (20%) and through remote checks using very high-resolution satellite imagery (80%). However, since 2015, the CAP has also included "greening" requirements, stipulations for payment that reflect the impact of agriculture on the environment. These include some stipulations for crop diversification to qualify for payment. Therefore, to appropriately allocate subsidy payments and conduct required controls, the RPA needed clearer oversight of the land that is managed by CAP beneficiaries and what is being grown on that land. The previous approach to satellite imagery—the Land Parcel Identification System (LPIS)—only allowed the RPA to ascertain whether land was being cultivated for agricultural purposes or non-agricultural purposes: it could not distinguish between crop types. As it was based on optical satellite imagery, images produced this way were also susceptible to cloud cover.

The development of CROME addressed the requirement to enforce the greening crop diversification rules and to provide an alternative source of crop information which would not be impacted by cloud cover. The idea of producing a crop map was brought forward as it was deemed to be both useful for the RPA and for other purposes across government. It began as a research and development project in 2015, a pilot was undertaken in 2016 and the new process produced the first full data set in early 2017. This was based on images obtained between January and August of 2016. The resulting map is a polygon vector dataset containing information on the crop types grown in the 2.6 million fields in England. The map is based on tessellated representation of the land cover based on hexagon grids classifying England into the main crop types. The map also shows grassland and non-agricultural land covers, such as water resources (rivers, lakes etc). The map is now produced annually.

Figure 1: Representation of land cover based on hexagon grids



Source: RPA interviewee

Since its creation, the crop map has been used for a number of other purposes beyond the RPA's use. This includes use by other Defra delivery bodies to control water quality management, manage flood defence areas and to monitor biodiversity. For example, a number of bird species prosper next to certain crops. The map helps to identify these crop areas and observers can monitor bird species. Defra has also used the map for statistical and monitoring purposes, to assess the types of crops planted and in which fields. It has also proved to be a useful tool for Historic England to provide archaeologists with land cover information surrounding ancient monuments.

The map is open access and free for public use. As such, it has also been accessed by private sector organisations, charities and academia.

Relevant technologies

Optical satellite imagery alone was not enough to establish a clear picture of crop diversification. The solution instead relies on using outputs from both time-series optical satellites and radar satellites. Optical satellites provide pixel-based imagery, while radar satellites use microwave pulses. Together, these two satellite techniques were able to provide high-resolution multi-spectral satellite images suitable for processing.

An automated process was then used to classify the satellite data. This was done through the use of a supervised classification algorithm using random forest machine learning techniques. Data from on-site control visits was also incorporated to train the machine learning algorithm to improve accuracy. The current CROME data set classifies land and the crops planted with 95% accuracy.

Developing the technology

The UK Space Agency provided £15,000 of financing for the research and development phase of the project, which led to the development of the algorithm methodology building the automated process. This funding came from the Space for Smarter Government Programme, which was established in 2014 to encourage the uptake of space products, data and services across UK government departments[3]. The subsequent pilot phase was funded by the RPA and cost an additional £30,000. Running costs are largely human resource costs, and amount to around 1 full-time equivalent.

The initial piloting and development of CROME was undertaken jointly with an external contractor. This was done as the specific **tools and knowledge were not available in-house**. Identifying a contractor to carry out this work was straightforward: the RPA received several responses to its call for tender and has received high quality services from its contractor.

A significant amount of **data was brought together from different sources** to produce CROME. Data on land ownership, boundaries and the presence of buildings (e.g. farmhouse, barns etc.) was taken from the LPIS, which is run by the RPA. On-site inspection data was also held by the RPA. Radar satellite data came from the EC Copernicus space programme, but the optical satellite imagery was obtained through external contractors.

One licensing issue arose linked to the use of topographic mapping data. The LPIS uses topographic data from the UK National Mapping Agency. This data entailed intellectual property rights, which meant that the CROME crop map could not use topographic field boundaries and be an open data product. To resolve this issue, the RPA opted to produce the CROME map using a hexagon-based tessellation, which is not subject to any intellectual property conditions.

The map uses mostly geographic data and it **does not incorporate any personal data**.

The sheer volume of this data presented a **challenge for storage**. At the time of completing this study, the data was stored on RPA hardware. However, the RPA noted that ideally it should be processed and stored on the cloud.

Implementation, monitoring and next steps

The CROME map has been fully operational since 2017. No specific issues have been encountered since the start of its operation.

The map requires **regular data harvesting and data processing**. The imagery is acquired from February to August (when crops are planted). When the satellite data is received by the RPA, it undergoes lengthy pre-processing and post processing before it is fed into the model, the process takes around 3 days per month.

[3] Further details available online at: https://spaceforsmartergovernment.uk/

In addition to the CROME product process, RPA has been considering ways to further improve the accuracy and functionality of the LPIS map of field boundaries and land covers products. One potential future development is the use of computer vision deep learning **techniques** for the land cover classification. This would offer an alternative to the manual photo-interpretation and digitisation undertake by analysts. At the time of completing the study, a small pilot study has been conducted with a supplier to understand the what the use of deep learning would entail. This study concluded that the use of deep learning would require significant research and development and access to specialised hardware through the cloud, in part to increase the speed of processing. The use of deep learning techniques would also require a set of training data. Bringing this forward would therefore require additional financial resources.

Impact

The impact of CROME has not been fully measured. However, it has enabled the RPA to administer the CAP scheme more cost effectively and reduced the risk of EU budget deductions (EC penalties) for England under the CAP, which are triggered if RPA does not administer the scheme correctly. If the RPA had not developed the CROME crop map and instead used inspectors to physically inspect the additional fields on an annual basis as part of the enforcement regime, the RPA would have incurred addition costs of approximately £400,000 per year.

In addition to a positive impact on the RPA's work, the map has also had an impact on the Environment Agency's work and has helped them to reduce the number of physical on-site visits required for water quality testing (this is because proximity to certain types of crop has specific impacts on water quality). The Environment Agency has been designing and piloting a model that will overlay the CROME data with other data sets held by the Environment Agency to better estimate impacts on water quality.

Since 2018, the EU has required Member States to develop such crop maps to support subsidy controls. As the RPA was one of the early developers of such a tool, the team has since been consulted by some Member States (e.g. Lithuania) for assistance and advice.

There are a variety of potential uses for the data: for example, it could also be used to make more strategic decisions about irrigation or by water companies to help plan water treatment and distribution. For this reason, it was made publicly accessible and open to download. This has been a success so far and there are indications that it is being used widely, by private sector organisations, charities and academia. Website statistics indicated that as of March 2020, over 900 people have downloaded the full data set (despite its significant size) and over 5000 people have viewed the page.

3. A consortium's use of sensor technologies for mosquito classification (the Vectrack Project)

Introduction

The following case study details the Vectrack project, which has developed the use of sensors for mosquito surveillance. Mosquito surveillance is generally carried out by public health authorities to monitor and prevent the spread of mosquito-borne disease. The Vectrack project has been led by a private firm, Irideon, in consortium with a wider team of public and private partners.

This case study has been based on interviews with two members of the project team, some written inputs from a participating public health authority and documentation supplied by the project team.

Best practices

The Vectrack case illustrates several instances of good practice described in the main report:

1. Making use of existing off-the-shelf tools: the sensor technologies used in this project had already been developed and used by the project lead in different contexts. In this project, they applied this existing technology to a new problem.

2. Considering the project's needs carefully and identifying an appropriate team accordingly. The interviewee noted that finding the right partners was one of the most important parts of this project.

3. Considering the challenges of collecting new data: the Vectrack project has relied on collecting large amounts of new data to train algorithms.

Protocols used to control mosquitos currently rely on the use of mosquito traps. These traps are used to capture and manually sample mosquito types in different locations every week. Classifications are then carried out in a laboratory, and it generally takes several days before results are available. The process to obtain information is labour-intensive and slow. This makes pre-emptive intervention challenging and increases the risk of epidemics.

The initial idea to apply sensors to this problem was developed by Irideon, a technology firm specialising in Internet of Things (IoT) sensor technologies. Irideon then set up a consortium by partnering with public health institutes in Catalonia, Spain (IRTA) and Portugal (INSA), as well as AVIA-GIS, a Belgian company that specialises in spatial risk mapping and modelling. The consortium then applied for the European Commission's innovation support programme Fast Track to Innovation and began to receive funding in November 2019. By developing sensors that transform mosquito traps into IoT sensors, Vectrack automates the previously manual process of collection and classification. This enables immediate feedback on the species, sex and age of the mosquitos: data on this can then be submitted to the public health authorities in real time.

Figure 2: A mosquito bite

Source: CDC, Public Health Image Library

In addition to the automation of mosquito surveillance, the Vectrack project is seeking to combine sensor data with satellite data to enable large-scale geo-mapping of mosquito population movements. This combined data will then feed into spatial models, and along with data on meteorological and environmental conditions, this can be used to generate large-scale risk maps.

The technology was in a development phase at the time of completing this research. Participating laboratories (IRTA-CRESA in Spain and INDA-CEVDI in Portugal) were helping to train the machine learning algorithms by collecting sensor data on different species of mosquito. The Vectrack project team was also collaborating with entities in Africa and South America to gather data on a wide range of species. Alongside the work done to train the algorithms, field trials were also ongoing in the city of Barcelona. The next step was to manufacture more sensors to deploy in other cities, so that field trials were conducted in a range of locations. Future additional trials were planned in Valencia, Lisbon, the Algarve, Madeira, Guadalupe, Tahiti, São Paulo, Burkina Faso, Mozambique and Montpellier.

Relevant technologies

The Vectrack tool is based on Senscape®, an IoT platform designed by Irideon. This platform has also been used for other IoT applications, including as a platform for security applications and smart beekeeping. These IoT mosquito sensors, attached to traditional mosquito traps, identify mosquitos' wingbeat frequencies, flight kinetics and morphology. Time-stamped sensor results are sent via 3G to the Senscape® Hub application, which uses machine learning algorithms to identify mosquito species, sex and age.

Developing the technology

The development of the Vectrack project has relied on:

- achieving sufficient funding – particularly as a project led by a private firm;
- bringing together a multi-disciplinary team; and
- providing algorithms with sufficient divergent data.

Interviewees from across the project team noted that the most significant obstacle to the development of this technology has been **funding**. Although the beneficiaries of this technology will be public health authorities, development was led by a private firm and there was a need to generate revenues to keep the project live. The project had cost around 3 million EUR at the time of completing the interviews and was expected to require a further 2 million EUR of funding before completion. Some funding had been received from the European Union. Additional funding would need to come similarly from grants or social investors. One interviewee noted that there have been barriers to receiving funding directly through public sector beneficiaries because of public procurement rules. One local public health authority involved in the project noted that their own internal funding for projects and studies was generally quite low, therefore they relied on national or international grants for any significant work to take place, as has been the case for the Vectrack project.

Interviewees also noted the importance of bringing together **a multi-disciplinary team**. The project scope required that entomologists, engineers and developers all work together: a key component to making the project work had been in ensuring that all team members understood their own roles and what is expected of them, as well as the roles and disciplines of other team members. The interviewee explained that assembling the team was facilitated by Irideon's previous experience working with multi-disciplinary teams. Choosing those partners was enabled by having a clear plan for the project: knowing from the outset what outcomes are planned and what steps and skills are required to achieve those outcomes.

Data to train algorithms has been a key condition of success. At the time of interview, work was still being done to improve the technology's accuracy in identifying mosquitos, as this was still below expectations. Specifically, more work needed to be done to provide the algorithms with diverse data, so that they could identify species and genders across a variety of environmental conditions. To train the algorithms appropriately, mosquitoes with specific characteristics and divergent morphological traits need to be reared in laboratories. However, handling mosquitoes is a high-risk activity which cannot be carried out in normal laboratories. Rather this implies working with organisations that are certified and equipped with Biosafety level 3 laboratories. Therefore, it has been essential to work with the **right partners**. Training was expected to take time. To improve the accuracy of the tool, training would also need to be done on species that are not available in Europe. This would require collaboration with partners in third countries.

Another **legal condition** was the requirement to attain a CE marking, a certification mark which would confirm that the tool complies with electrical safety laws and would not interfere with other electrical equipment. This was not difficult to achieve but did require pre-testing in a certified laboratory before deployment.

Impact

Impact cannot yet be measured for the Vectrack project as it is still under development.

The primary expected impact is allowing public health authorities to receive mosquito surveillance information in real time (rather than with a 1-2 week delay), meaning that they can **take any necessary action without delay**. This would substantially reduce the time entomologists need to spend identifying species. Because entomologists who conduct these identifications are normally knowledgeable about local species, they may take longer to recognise non-local species. The sensor-trap would rely on a wider library of species and would be able to identify these more quickly and more accurately. Current accuracy is over 95% for counting insects and distinguishing between mosquitoes and other insects and around 92% for differentiating between genus's for two of the most dangerous species (Culex and Aedes). Accuracy is around 93% for distinguishing between male and female mosquitoes and 75% for estimating the age of female mosquitoes.

In addition to improving the efficiency of previous processes, the use of sensors would also **offer new capabilities to public health authorities**. Receiving information in real-time on where mosquitoes have been trapped means that researchers could promptly collect ribonucleic acid (RNA) from the traps. Testing this RNA can determine whether mosquitoes are carrying viruses, meaning that public health authorities can better prepare for potential epidemics. Conventional methods make collecting this information much more challenging because the RNA degrades quickly. The project team has found that sensor data can be used to identify the insects' age, something very difficult to do through morphological characteristics alone. Age is an important factor when determining mosquito risk, therefore this can help to produce better risk estimates.

Having better data on mosquito populations would also make it easier **to test more sustainable alternatives for population control**. For example, the sterile technique involves releasing sterile males and relying on them attempting to breed with females, who would then lay eggs that will not hatch. Collecting population data in real-time means that researchers could more thoroughly assess the feasibility and effectiveness of such techniques. Where these techniques are found to be effective, this could help to reduce reliance on insecticides.

4. Ofcom and the use of blockchain for a common numbering database

Introduction

The following case study details Ofcom's exploration of the use of blockchain for the development of a common numbering database to manage telephone numbers. A common numbering database would be a shared database of all landline telephone numbers in the UK. Ofcom is the UK's regulator for communication services, responsible for overseeing and regulating the use of broadband, home phones and mobile services, as well as protecting consumers from nuisance and scam calls.

Best practices

The Ofcom case illustrates several instances of good practice described in the main report:

1. Taking a problem-focused approach: the goal of this project is to develop a common numbering database. Considering the requirements for a common numbering database, blockchain appeared to be a potential solution. Ofcom's work has focused on developing a proof-of-concept and testing the feasibility of this approach.

2. Considering the project's needs carefully when developing the team: Ofcom brought in external skills, including some hiring of in-house staff, but chose to lead the project internally to support building knowledge and understanding in-house.

3. Collaborating, engaging and sharing the tool with users from an early stage. Because the telecoms industry are an important partner for Ofcom in this work, they have focused on engaging them extensively and from an early stage.

At the time of the case study (May 2020), Ofcom was facing several challenges that could be addressed by the development of a common numbering database. Ofcom's approach to managing telephone numbers is to issue large blocks of 1000 or more telephone numbers to telecoms operators. Each operator then manages its individual telephone numbers (e.g. assigning them to their customers) and porting (transferring) them into and out of their control when customers switch operators.[4] Ofcom maintains a database of these allocated blocks of numbers, but this is not suitable to be updated each time individual phone numbers are ported between providers. This means that Ofcom does not have access to a complete and up-to-date record of which telecoms operators are using which phone numbers. This creates challenges for investigating and enforcing regulations intended to protect consumers from nuisance and scam calls and the development of technical measures to authenticate phone numbers when used as Caller IDs.

Ofcom also identified issues with the UK's existing porting arrangements. Ofcom considers there is a lack of consistency and transparency in existing porting mechanisms, which are largely determined by telecoms providers. The lack of automation in some of these processes

[4] Porting is the name given to the facility that enables consumers to keep their phone numbers when switching providers.

creates additional costs and makes the porting mechanism prone to mistakes, data errors and mismatches.

In addition to existing challenges with porting, transparency and oversight, Ofcom is overseeing a major technology transfer in the UK telephone networks. Telecoms providers are gradually moving their landline customers from the country's traditional telephone network – the 'public switched telephone network' (PSTN) – to newer 'internet protocol' (IP) technology. The move to IP will lead to new vulnerabilities (such as nuisance callers and criminals using software to deliberately change the Caller ID, a practice known as 'spoofing'), but also provides an opportunity to address existing challenges with new solutions and new processes.

To address both these existing issues and likely risks that will emerge from the move to IP, Ofcom has identified a need for a common numbering database. One potential solution for the development of a common numbering database is the use of distributed ledger technology, or blockchain. Blockchain technology provides a system for recording the transaction of assets (in this case, phone numbers) in multiple places at the same time, securing the data in a traceable and transparent manner. As such, Ofcom considers that the use of blockchain would have several potential benefits:

- it would support the migration of phone services to IP technology;
- it would enable telecoms operators to introduce security protocols to reduce nuisance calls; and
- it could support more efficient processes for porting numbers and routing calls when customers switch between competing providers.

The work to apply the use of blockchain to a common numbering database was funded through the Regulators' Pioneer Fund.[5] The proof-of-concept project began in October 2018 and was completed in March 2020. Ofcom's goal is to implement a common numbering database by 2022.

Relevant technologies

Ofcom's proof-of-concept database is based on blockchain technology. Blockchain technology locks the ownership and status of digital assets (in Ofcom's case, of telephone numbers) with a cryptographic key in a database platform that is distributed amongst members. For changes to be made, all members need to agree.

There is no standard database or software for blockchain. For this purpose, Ofcom has developed their own database.

Developing the technology

Ofcom's development of a blockchain solution for a common numbering database has relied on:

[5] The Regulators' Pioneer Fund is a competition-based intervention to finance regulator-led projects and is overseen by the Better Regulation Executive within BEIS. The Fund aims to develop a UK regulatory environment that encourages business-led investment, particularly in science and technology-driven innovation.

- internal leadership of the project and the development of their own software;
- a mix of internal and external developers;
- the recruitment of new in-house talent to support development; and
- building on existing engagement with the telecoms industry to ensure a user-friendly and relevant solution.

Ofcom initially considered an off-the-shelf database solution and spoke with a range of potential contractors at the outset. However, most of the solutions on the market were costly and not tailored to Ofcom's needs. External solutions would have also made Ofcom dependent on third-party software and may have led to licensing issues. Ofcom therefore decided to **develop the database and its supporting software internally**. Although the project was led by an Ofcom team, external contractors were used to support development and to address internal capacity restraints. This included the use of some external support in India.

Conducting this work internally led Ofcom to **recruit new talent** (on a temporary basis) with skills in blockchain development. Finding the right talent was a challenge: although Ofcom benefitted from London's significant technology and fintech sectors (and therefore relevant local expertise), they found it difficult to compete for that talent alongside private firms, who were able to pay higher salaries. Ofcom was eventually able to build an appropriate team, but for the same reasons, retaining talent also proved difficult. One of the web developers on the team left partway through, which caused significant delays to the planned timeline.

Ofcom noted that they may have faced fewer challenges if they had brought in an external contractor to lead the project at the outset. However, interviewees felt that deciding to use internal resources as much as possible proved to be a rewarding experience for both the staff involved and the organisation more broadly. Being involved in the various phases of the project helped the team better understand their processes in order to optimise them. Also interviewees felt that the experience of this work will likely help to prepare the team for any future work that involves designing and developing software.

The development process was continuously adaptive and iterative, in part because Ofcom was exploring a new technology and because it required collaboration with industry, largely established communications providers with whom Ofcom has existing relationships. Ultimately, a common numbering database is only useful if it is taken up by the telecoms industry, meaning that **building engagement** was important and that the solution needed to be **user friendly**, relevant and suitable for industry users. **Conducting continuous discussions with telecoms providers** and relevant stakeholders was therefore an important part of the process. Ofcom noted that holding such consultations was time consuming and that many of the stakeholders involved were unfamiliar with the technology. To help address this, Ofcom created a special repository where interested parties could download the source code and experiment with it in order to accelerate their own innovations.

Ofcom did not encounter any significant legal barriers during the development phase. One potential future issue relates to the "**right to be forgotten**"[6]: because blockchain is specifically designed to be resilient to changes within the database, enacting requests to have data removed would be challenging. Ofcom has consulted with the Information Commissioner's Office (ICO) on this and will need to develop some specific tools using techniques such as data pruning that can address this issue, should the platform (or a similar version of it) ever be

[6] The right to be forgotten refers to the right to have personal data erased from internet searches or other databases. This is included in the GDPR in the UK.

deployed in a ‘real world’ scenario. Although no other issues have arisen, Ofcom is observing the evolution of smart contracts and other use cases of blockchain technology across different markets, to identify any other impending legal issues.

Implementation, monitoring and next steps

Subject to Ofcom’s policy process and discussion with industry, any new database would not be launched until 2022. For the database to be successful, it would need to be implemented by the telecoms industry. The current intention is to develop a solution that is both useful and feasible for industry, so that industry remains engaged and adopts this voluntarily. Once the database has gone live, Ofcom will need to continue to engage regularly with telecoms providers, vendors and standards bodies to establish a credible roadmap for implementing the necessary measures.

Ofcom also noted that the implementation of the new database will not eliminate all nuisance calls (such as those from non-UK phone numbers, which would not be part of the database), and further work will be required to broaden its scope. This will be done in part through engagement with international partners to tackle issues of shared concern, as well as through further conversations with industry.

Impact

As the project is still under development, the impact of Ofcom’s use of blockchain on porting, transparency and oversight has not been measured. The expected impact is the development of a platform which could be used to support the establishment of a numbering database in the UK. Such a database would then support greater transparency and oversight of phone numbers and porting mechanisms, and an associated reduction in nuisance and spam calls. However, interviewees noted that measuring impact will be a challenge as impacts may not be clearly evident from the outset and implementation will take time. It is possible there will be further delays related to the transition to IP technology, as this will be a significant and complex project on its own.

5. The Finnish Government’s use of a Virtual Agent Network

Introduction

The following case study details *Starting up Smoothly*, a collaboration between Finland’s Immigration Service (Migri), Tax Administration (Vero) and the Patent and Registration Office (PRH). The aim of the project was to develop a test example of a virtual agent (e.g. chatbots) network that would bring together the customer services of these normally separate organisations. The example tool developed by the *Starting up Smoothly* project was intended for foreign entrepreneurs looking to start businesses in Finland.

Best practices

The Starting up Smoothly case illustrates several instances of good practice described in the main report:

1. Taking a problem-focused approach and building on other projects: the idea to develop a network of chatbots was developed when one agency was developing their own chatbot and found that focusing on their own content would only address part of the customer’s journey.

2. Using an internal pilot to demonstrate the functionality of the tool, engage staff and gather valuable feedback.

3. Upskilling staff and using subject-matter experts to help create content. The interviewee also noted here the importance of upskilling staff not only with the required technical skills, but also paying attention to other skills that might be relevant (in this case – good writing for chatbots).

4. Starting small: the purpose of this project was in part to help develop an approach to working across agencies on a customer service tool in an effective way. A niche case was chosen to narrow the scope and focus the project on specific content. The lessons learned from this are now being used to support broader Machine Learning work in the Finnish government (e.g. the AuroraAI project).

Before *Starting up Smoothly*, the approach to customer service relied on human operators answering questions from the public. These operators were only able to answer questions relevant to the services their organisation carried out. They were unable to help customers with queries that fell outside the organisation's remit.

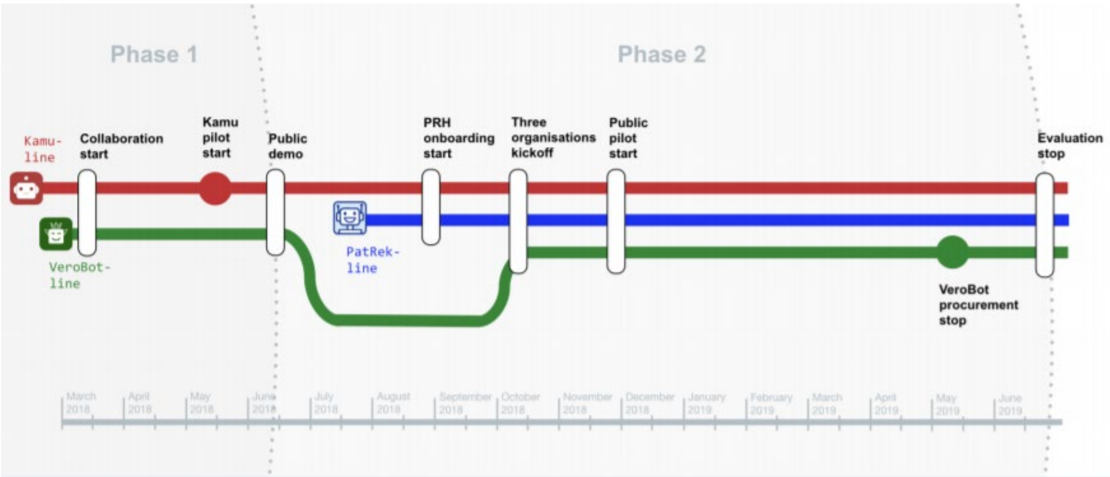
Beyond the lack of integration, the previous approach to customer service was time consuming and inefficient. Human agents would need to take the time to read, review and respond to each question. Using a virtual agent rather than a human operator was considered at first by Migri to optimise their customer service. However, during development it became apparent that likely users of Migri’s virtual agent would also need to navigate services from other departments and agencies. Creating a virtual agent for Migri would therefore only address a part of the customer’s journey.

It was from that observation that the *Starting up Smoothly* project was initiated. Although the aim of the project was to develop an operational service tool for foreign entrepreneurs, it was at first intended as a feasibility study. The main purpose was to assess the extent to which a joint customer service tool could help bring together siloed government services associated with a specific life event, without the need to make significant changes to existing operations and remits. For this feasibility study, the scope was limited to government services required by foreign entrepreneurs looking to start a business in Finland.

This case study is based on two interviews: one with a staff member of Migri, providing a policy perspective on the case, and one with the external partner, boost.ai, who provided the technology. It has also drawn on an evaluation published by the project team in 2019.[7]

The project was developed in two distinct phases (see Figure 3). The first phase began in March 2018, with an initial demonstration version of the tool based on collaboration between Migri and Vero ready in June 2018. PRH then joined the collaboration in August 2018. A public pilot based on the inputs of all three organisations was then launched in November 2018.

Figure 3: Starting up Smoothly project timeline



Source: Miessner, S. (2019) Starting up Smoothly – Experiment evaluation.

Relevant Technologies

The *Starting up Smoothly* tool is based on a Virtual Agent Network (VAN), a concept put forward by the contractor, boost.ai. Separate virtual agents (chatbots) were developed by each of the organisations involved, with the VAN allowing customers to be connected with multiple virtual agents in a single chat window. Queries would then be transferred between agents as appropriate.

The virtual agents were developed using existing conversational artificial intelligence (AI) techniques. This combines natural language processing (NLP) with natural language understanding (NLU) . The virtual agents are able to not only recognise what customers are asking for, but also to place importance on the context of customer questions.

[7] Miessner, S. (2019) Starting up Smoothly – Experiment evaluation. Available online at: https://migri.fi/documents/5202425/0/Starting+up+Smoothly+experiment+evaluation_CMYK.PDF/87688320-dfef-9246-6c24-c1ac8e436103/Starting+up+Smoothly+experiment+evaluation_CMYK.pdf

Understanding the context helps virtual agents know when to hand off a user to a digital colleague within their network or to forward a customer to human support.

The virtual agents used conversation flow data. These are based on **intents**[8]**, questions-answer pairs**[9] and **transfer possibilities**. ML models are then trained using this data, multiplying the questions-answer combinations and intents using synonyms and alternative wordings. The technology is able to detect false responses in the conversation logs and these logs are used to continually adjust and improve its performance.

Developing the technology

The Finnish government's development of Starting Up Smoothly has relied on:

- the use of an external contractor and off-the-shelf platform;
- training and up-skilling of staff in ML, the software and in copy-writing;
- inter-organisational collaboration, including through the use of a shared physical space for development; and
- internal and external pilots of the tool.

As the skills and platform to develop virtual agents were not available in-house, an **external contractor** was brought in. Before *Starting Up Smoothly* began, Migri was already in the process of developing a stand-alone virtual agent with boost.ai, a company specialising in conversational AI. Although virtual agents are a well-understood use case for AI, Migri needed an option that would be operational in Finnish, Swedish and English to address the different language needs of individuals needing immigration assistance. As such, their options for suitable contractors were limited. At the time, boost.ai was one of the few developers offering a Finnish language model.

Boost.ai offered a ready-made platform for conversational AI. Based on previous work for private sector clients, it also offered a starter pack, including a corpus of dummy questions-answer pairs and sector specific questions to quick start their client's content creation process. However, as Migri was their first public sector client, no off-the-shelf starter pack was available and content needed to be created from scratch.

Content was created by non-technical subject matter experts from within the organisations. The contractor provided an initial two-day training, **instructing staff on the basic concepts of AI, NLP, NLU and the contractor's specific technology**. The technology was designed by the contractor so that individuals without a technical background could develop and maintain content themselves with minimal training. The contractor has a dedicated education team to deliver such training.

According to the project team, this training was successful and allowed staff to create content that could then be used to train the algorithms. However, creating high-quality content was more challenging than anticipated. Firstly, finding the right **language skills** in-house to address the need for multiple chatbot languages (i.e. native speakers across the three languages most likely relevant to foreign entrepreneurs) was a challenge and posed a barrier

[8] Refers to what the customer wants to achieve
[9] Refers to the multiple ways a question can be formulated aiming at the same answer.

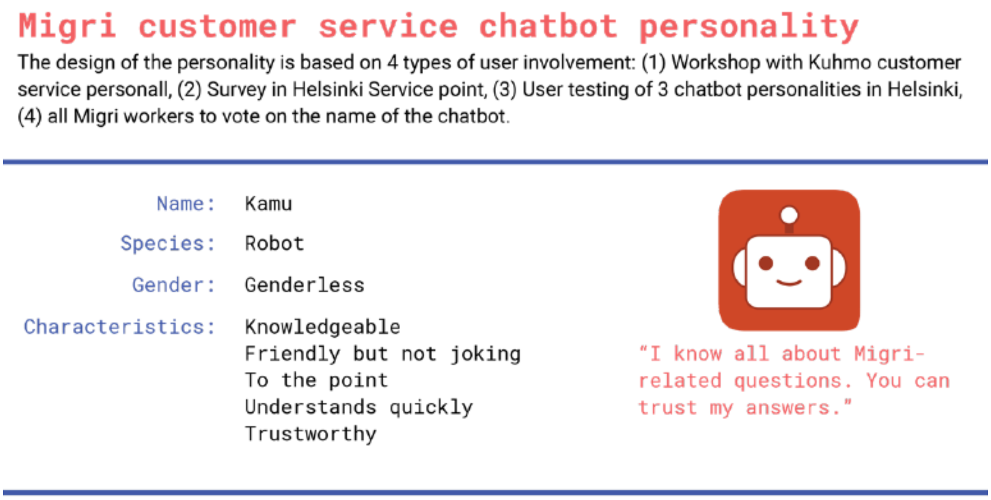
to development. Ultimately, the project team was not able to develop content in all three languages across all three agents as initially intended. Secondly, subject matter expert staff did not have experience in writing for chatbots. This type of writing has to be clear, simple and conversational, and was unfamiliar to in-house staff. Further training at a language centre was therefore required to upskill staff in **writing for chatbots**.

Each organisation developed their own separate virtual agent, with its own name and personality (see Figure 4 for an example). The VAN then helped to transfer users between agents. This design had several advantages:

- It helped to avoid customer confusion, making sure that customers understood that their requests were being handled by different organisations;
- It allowed each organisation to remain in charge of their own content;
- It would mean that once the network has been established, organisations could remain in control of their own budgets and procurement; and
- It meant that chat logs could be reviewed directly by the relevant organisation, helping to avoid data protection issues.

Although in theory, content could be developed completely independently, in practice, some **inter-organisational collaboration** was needed to agree which life events to cover, which services belong to that event and to identify opportunities for transfers. PRH, Migri and Vero worked together to develop a high-level overview of possible content. The organisations then worked individually to expand and specify the content. Common content creation guidelines were also established to ensure consistency**. A physical space to actively cooperate** with peers in other departments and visualise conversation flows facilitated this development. There were also challenges associated with collaboration, particularly when PRH joined the project at a later stage. Communication on time commitments and the division of responsibilities was not clear enough from the outset and led to unnecessary confusion.

Figure 4: Starting up Smoothly chatbot personality



Source: Miessner, S. (2019) Starting up Smoothly – Experiment evaluation.

Engaging staff in the project was identified as a key challenge, as many staff members involved in customer service were sceptical about the use of the tool. Some assumed incorrectly that the tool would be automatically extracting information from the agencies' websites and therefore would be prone to errors, others felt that the AI more generally would not be able to provide the answer quality required. Following the initial development phase, the conversational agents were **piloted internally with organisation staff**. This helped to gather initial feedback, improve understanding of what the tool could and could not do and achieve buy-in across the wider organisation. A **pilot testing the technology directly with customers** was held later on in the process. The user feedback received was then used by the team to improve the content and algorithms. According to the project team, collecting feedback from these pilots was an important step to creating a successful tool. Testing the content with customers directly helped to reveal blind spots, as some information that seemed obvious to subject matter experts and organisation staff was unclear to their intended customers. However**, identifying customers from the target audience** (entrepreneurs looking to start a business in Finland) was challenging. Significant planning and recruitment efforts were required to find the right people to test the virtual agents. The development team needed to make use of a wide range of networks and channels to identify potential testers.

Implementation, monitoring and next steps

Following the public launch of the virtual agents, work shifted from creating content towards **maintaining existing content, monitoring virtual agent performance, documenting the working process and evaluating the project**.

However, as this was intended as a feasibility study, fewer resources were allocated to operations following the public launch. Migri has continued to actively develop its virtual agent, while the other two agencies have not. While the other two virtual agents are being monitored, no further content is currently being developed.

Conversation logs are **monitored on a regular basis**. Analysing the logs, identifying errors and creating new intents and transfers strengthens data quality and improves the experience for customers. The interviewee noted that over time, errors and issues with content have decreased, and monitoring can be done less frequently.

No further developments are envisaged specific to the *Starting up Smoothly* project. However, in early 2020, the Finnish Ministry of Finance launched the **AuroraAI national artificial intelligence programme** (2020-2022). This was inspired in part by the *Starting up Smoothly* experience and aims to use AI to “bring people and services together in a better way”.[10]

Impact

From the outset, the work done under *Starting up Smoothly* was not intended to significantly impact the work of the three organisations involved, although it was meant to improve customer service for the target audience. No resources were allocated to publicising the service, due to the uncertainty of its future use. This, along with the limited scope of the service (focusing on the needs of foreign entrepreneurs) and the lack of content in all three languages has meant that actual use was low.

[10] More information available online: https://vm.fi/en/article/-/asset_publisher/kansallinen-tekoalyohjelma-auroraai-alkaa-tavoitteena-saada-ihmiset-ja-palvelut-kohtaamaan-paremmin-tekoalyn-avulla?fbclid=IwAR0I7MVtxprKBbAyuiCtinIIw-FMwpD83D4kAQMADQXm1IBG_ws_ocLSvEE

Migri's virtual agent was already under development before the collaboration began, meaning that a more comprehensive range of content was developed separately. It has been incorporated into Migri's website and has helped them to improve their customer service efficiency. Other measures were taken at the same time to improve customer service, and together with the virtual agent, this led Migri to improve from being able to respond to 15% of customer enquiries to being able to respond to 70%.

In addition to improving efficiency, the use of virtual agents and a virtual agent network is also anticipated to improve the quality of customer service. As compared to human agents, the virtual agents are always accessible, customers can take their time to reread and understand the content and are able to see the answers to all of their questions together in a single window.

Starting up Smoothly's impact as a feasibility study has been significant. Within the organisations involved, it has helped to introduce staff to the use of AI and ML and induced a broader acceptance among staff towards new technologies and ways of working, including cross-organisationally. Most importantly, the results of the project have offered a model for the creation of an automated, user-centred solution across government departments which can be developed relatively quickly while maintaining operational independence. The idea of a network of virtual agents across the Finnish government is now being brought forward by the Finnish Ministry of Finance through the AuroraAI project.

6. The Ministry of Justice's development of an intelligent search tool

Introduction

The following case study details the development of an intelligent search tool for prison inspection reports by the Ministry of Justice (MoJ). MoJ is responsible for government policy for the criminal, civil and family justice systems in England and Wales. As part of this work, MoJ uses prison inspection reports to support its policy decisions and improve its services.

Best practices

The Ministry of Justice case illustrates several instances of good practice described in the main report:

1. The previous development of a bespoke in-house analytical platform: this provided a basis upon which the team was able to build their search tool.

2. Learning from the experiences of others and monitoring new technologies: the interviewee noted that running self-organised away days with the data science team helped to develop innovative new ideas.

3. Using an early model of the tool to demonstrate usage and potential to staff. This helped to build engagement and interest.

Prison conditions are assessed by four official, independent bodies in the UK.[11] These reports are then used by MoJ staff as an evidence base to support policy. For example, a regular review of prison reports can help to identify emerging trends in prisons, such as the circulation of a new drug or a type of inmate conflict. At the time the tool was first developed, MoJ had over 250,000 sentences of unstructured text in over 500 reports detailing the inspection of prisons and other institutions. Neither the reporting nor the administration of reports (e.g. where, how and when they were published) followed a protocol and information was not centralised in one database. Both the number of reports and the lack of consistency between reports prevented staff from quickly retrieving relevant information. This meant that staff's use of these reports was irregular and limited. The lack of accessibility made it difficult to conduct any significant analysis in a way that could be used to inform prison inspections or policy decisions.

In response to this issue, MoJ's data science team initiated the development of a prototype tool that would bring this data together and allow for an "intelligent" search of data to address the lack of consistency between reports. A neural network was trained on the prison reports to learn how reports are written and how specific words are used in prison contexts. The search does not look for synonyms, but rather looks to identify similar contexts. For example, the word "spice" is a common term used in prison reports referring to a type of drug. When a user searches the database for "spice", the search tool would return reports which include the word "spice" as well as reports where other words are used in a similar context and have a similar meaning, such as "drugs".

The entire project took around 14 months in total. Development started in September 2017 and the tool was deployed in June 2018. It was finalised in December 2018.

This case study is based on publicly available information[12] and one interview to provide a policy and data science perspective on the case.

Relevant technologies

The intelligent search tool uses natural language processing (NLP) and natural language understanding (NLU) techniques. It allows for the searching of words or sentences, and retrieves both documents in which the exact word or phrase occurs, as well as likely synonyms and similar contexts. The neural network is trained on the prison reports and tracks how people use specific words in prison contexts. In order to extract the content from the reports, PDFs are scraped using Python (a programming language) to retrieve these reports.

Developing the technology

MoJ's development of its intelligent search tool relied on:

- an internal team experienced in NLP;

[11] These are: Her Majesty's Chief Inspectorate of Prisons, the Independent Monitoring Boards, Prisons and Probation Ombudsman and the Office for Standards in Education, Children's Services and Skills
[12] Information retrieved from the information hub of the UK Government, "Collection: A guide to using artificial intelligence in the public sector", available online: https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector#examples-of-artificial-intelligence-use

- a previous similar prototype and development infrastructure in place;
- building internal engagement; and
- publicly available data.

The development of the tool was done entirely in-house. The interviewee highlighted that the team's **experience with NLP** facilitated the development of the solution: MoJ have a well-established data science team with a strong interest in NLP and ML. The idea for the tool came from the data science team themselves. They had been introduced to some of the concepts during an away day and built a basic prototype for retrieving information related to Parliamentary Q&As. The interviewee explained that the team regularly holds self-organised away days focusing on new practices or data science techniques, and that these have been a helpful exercise for developing innovative new ideas. The tool for inspection reports **built on the concepts developed in this prototype**. The team had also already developed the MoJ **Analytical Platform**: a cloud-based solution with tools available for analysts and data scientists, which was used by the team to develop the intelligent search tool.

One obstacle to the development of this technology was **ensuring data quality**. The data used to train the algorithms and determine word and sentence vectors influences the quality of the tool. Even though the software was in place, scraping PDFs slowed down the development process and the lack of a protocol for writing and uploading reports caused errors in the system. Some reports had formatting issues that also proved challenging for algorithms. For example, in one report, some text had been changed to white so that it would be invisible to the human eye, but it was still extracted by the tool. Therefore, the team was unaware of the existence of the text while the algorithm was processing it. These types of errors added to the challenge of training the algorithms.

Another challenge faced by the development team was the need to **convince staff** and executives of the benefits of this solution. This was a challenge in part because organisational restructuring and staff turnover meant that the intended customers kept changing (both personnel and teams). According to the interviewee AI solutions and prototypes require significant internal buy-in order to be implemented. The team found that clearly identifying the problem and demonstrating the usage and potential of the technological solution (e.g. by having an early model of the tool to demonstrate) helped to build engagement and interest.

The interviewee noted that user experience was only considered to a limited extent at the start of development and **more user testing would have been beneficial**. The interviewee explained that not enough testing occurred in part because prison inspectors had limited time available to participate and the organisational structure meant that data science teams did not have direct access to relevant user groups. Involving users was also challenging in this case due to preconceptions among staff on what ML could deliver and the extent to which it can replicate human interpretation. For example, in early consultations, staff put forward sentiment analysis as a desirable feature, allowing users to filter "positive" and "negative" reports. However, the interviewee noted that sentiment analysis in the context of prison reports would not always provide meaningful information, as NLP is not able to detect nuance, indirect language and underlying meaning. Important messages could be missed. The interviewee noted that making contact as early as possible in the development process and focusing on what *can be done*, rather than what cannot be done, helps to build a more constructive discussion with users.

No legal issues were encountered in accessing data since prison reports are publicly available.

Implementation, monitoring and next steps

Limited resources have been invested for the development team to monitor the tool. At the time of the interview, the tool had been in use for about a year and is hosted on the MoJ Analytical Platform. When new reports are uploaded, they are scanned by the system and automatically added to the tool's library. Any new report is scraped and run through the algorithm. This ensures that the data remains up-to-date. However, according to the interviewee, the tool itself is not adequately monitored or updated. It has **not been designed to handle large amounts of data**, so there is a concern that as more reports are added, the system may no longer be able to process the data. The interviewee noted that a new solution will need to be developed if the tool is to be scaled up.

Considering the data quality challenges, the interviewee noted that **stricter data validation requirements** should have been set for new reports that are uploaded to the database. Specifically, because PDFs are difficult to scrape and MoJ's scraping software is brittle, small changes in a PDF can lead to problems in the tool. This could have been addressed by rejecting PDFs in the wrong format at the point of upload, however, according to the interviewee, there was not enough time in the development process to include this feature. The existing minimal requirements lead to lower quality data, which hinders the tool's performance.

Maintenance and monitoring has also not occurred in part because the initial plan was to develop a prototype, rather than to scale up and implement a workable tool. This stems in part from the data science team's focus on developing innovations, rather than supporting existing tools. The interviewee noted that there can be a tendency among data scientists to look toward new innovations rather than supporting existing ones.

There were no plans to add to the tool, although suggestions have been made to add probation reports to the tool's library. However, the jargon used in probation reports and the different context could confuse the algorithm and result in irrelevant hits.

Impact

Although certain metrics have been collected, including numbers of uses (over 1,000 at the time of interview), there had been no formal analysis or evaluation of impact. The interviewee explained that this was due to a common tendency amongst the data science team to move on to new projects, rather than building and improving old tools. The interviewee noted that the tool and the capabilities it has delivered went beyond initial expectations.

The tool has delivered **efficiency gains** and **improvements to service quality**. The increase in accessibility means that staff are now consulting reports more frequently than they were previously able to. The tool has also offered **new capabilities to prison inspectors and policymakers**. Easy access to report information enables data-driven decision-making regarding prison inspections and policy decisions within MoJ. Staff can now more easily identify reoccurring issues or emerging trends across prisons.

Unexpectedly, use of the tool has spread beyond its target audience. The tool was developed for prison inspectors and prison staff, but it has reached people beyond these envisioned users, including the private officers of Ministers. This suggests that the technology has **made reports more accessible** and broadened the scope of use of prison report data.

7. North Tyneside Council's use of RPA

Description of the case study

North Tyneside Council is a metropolitan district local authority located in England. It is responsible for a range of local government services in North Tyneside, including social care, waste management, housing and planning.

Best practices

The North Tyneside Council case illustrates several instances of good practice described in the main report:

1. Making use of existing off-the-shelf tools to develop bespoke solutions: the team used a third-party software provider and received training from this provider.
2. Taking a problem-focused approach: the solution was only applied in areas that met certain criteria, where RPA would likely make a difference to efficiency.
3. Upskilling existing staff to participate directly in the development process: staff who were subject-matter experts were trained in RPA and some were able to use this to directly automate simple processes themselves.

Figure 5: Location of North Tyneside Council within England

North Tyneside Council's robotic process automation (RPA) project began in 2014. The project was part of a long-term (10 year) contract between North Tyneside Council and a private firm, ENGIE, which had been won in 2012. In the context of this partnership, ENGIE had been delivering a variety of back-office services, including information and communication services (ICT) for North Tyneside Council. The RPA project was led by ENGIE and driven in part by commitments made to North Tyneside Council to improve the efficiency of the delivery of public services at the Council. RPA supported this by targeting routine and repetitive activities, such as those performed in customer service centres. The project aimed to successfully automate such activities with the help of RPA software and thus reduce staff workload.

The principle of the solution was to programme rules into the software so that it can independently input and manipulate data and communicate that to other IT systems. This enables the software to carry out existing rules-based processes with no or very little human supervision.

RPA was first implemented within the Revenues and Benefits services of the Council. Since then, the technology has been implemented across other services of the authority, such as Council Tax, Debt Recovery and Employee Services. Automation continues to be scaled to other areas, such as recruitment and payroll.

In the Benefits service, RPA was used to process online housing benefit claims. Historically, these have been submitted using a 26-page paper application. As claims are now submitted online, the software is programmed to automatically search for the necessary data in the Council's records, input data in the benefit claims forms and gather relevant evidence (such as proof of occupancy, income etc) for each new claim. The assessment of claims and the calculation of benefit entitlements are then carried out by human benefit assessors.

Within the Revenues service, RPA was used for online Council Tax Direct Debit applications. The software verifies data provided in the form by comparing it with data in the Council's databases and interacts with the customer through automated e-mails to confirm details. This enables the application process to be fully automated, without any human intervention.

RPA was also incorporated within the debt recovery process for Council Tax payments. Previously, liability orders were processed manually to allocate fees or to classify cases based on their various enforcement stages. For some cases, once a previous debt is identified, the liability order is assigned to a bailiff. For others, cross checking multiple documents is necessary in order to establish the debtor's situation. Where verification of the debtor's history is straightforward, the case is processed automatically before being assigned to a task manager. The software can also calculate the compliance fees, issue notices, and update the enforcement status and the staff then takes over the processing of the case.

In both services, before the implementation of RPA, all applications were processed by human assessors. The Council used a combination of paper forms and online forms to process the applications. The interviewee noted that online forms were often printed, signed, scanned and sent back via post or email. Before the introduction of RPA, the application processes had been recently rebuilt as an online form that could be completed directly within the Council's websites. The introduction of these online forms was a necessary prerequisite for the subsequent use of RPA.

This case study is based on one interview with the private contractor and some additional documentation.

Relevant Technologies

The databases for the Benefits and Council Tax services are managed through a database software solution that is provided by a third-party supplier.

An off-the-shelf third-party software, Blue Prism, was used for the development of RPA processes. The processes themselves were programmed by staff.

Developing the technology

North Tyneside Council's development of RPA relied on:

- use of an off-the-shelf third-party software for RPA and the receipt of training from that provider;
- involving subject-matter experts to map out processes;
- collaborating with staff to identify further processes for automation; and
- training subject-matter experts at the Council in RPA.

The implementation of RPA for the benefits claims process was used as the Proof of Concept of the project. An **off-the-shelf third-party software for RPA** (Blue Prism) was used, and Blue Prism **provided the project team with certification and training**. For some more complex processes, the project team also worked directly with the online form provider to develop RPA processes.

As a first step, a small team of **subject-matter experts (e.g. Council staff who were knowledgeable in the area) was tasked with mapping the benefits claims process**. The project team created a process flow chart identifying each of the actions required to process an application and the decision-making logic behind those actions. A team of six developers from ENGIE then programmed that process so that the software could complete it as a human would do. Software was also programmed to identify elements that are uncertain as exceptions. Consequently, whenever uncertainties occur, the software flags these for human intervention.

Programming the process took time, in part because the benefits claims process is complex and the development team needed to code all possible outcomes of the steps involved in the process.

The application of RPA to the benefits claims process was considered an effective approach to improving efficiency, so the team **moved to identify further processes for automation**. The development team led workshops with staff to identify relevant tasks. This engaged staff and highlighted an opportunity to have some subject-matter experts undergo Blue Prism's RPA training, so that they could actively participate in the automation of their own processes. **Training subject-matter experts** meant that the development of some processes, particularly more simple processes, could happen directly within those teams. For more complex processes, such as the benefits claims process, external software developers were brought in to work alongside certified members of staff.

The interviewee noted that although this was a successful approach, there were some challenges in training subject-matter experts. Firstly, developing RPA skills was not easy and required some understanding of programming and not all subject-matter experts were able to complete the training. Secondly, retaining staff that had successfully completed the training was a challenge. At the time, there was a high demand for RPA specialists and the interviewee noted that some staff who had undergone RPA training ended up leaving for higher-paid roles elsewhere.

When deciding on what processes to automate, the interviewee noted two main deciding factors:

- the length of time to automate a process. Some processes take more time to automate than others: processes where there are many possible actions and variables are time-consuming to automate. Processes that rely on external documents or information that is not standardised in terms of format or content also require additional steps and therefore take time to automate; and
- how frequently a process is used. Some processes and services are used less frequently, meaning that the potential for efficiency gains through automation are limited and automation is not a cost-effective approach. For benefits claims, the volume of applications was high (around 500 per month), meaning that there were significant efficiency gains possible.

According to the interviewee, the most suitable processes for RPA were therefore simple processes and processes that are used most frequently. For complex processes, they noted that one approach to automating these is to split the process into several separate stages of work and automate each stage separately.

Council Tax Direct Debit applications were automated next. This was an example of a relatively simple process and could be automated to the point where no human intervention was needed. The simplicity meant that the RPA process could be developed quickly and with significant input from subject-matter experts, as opposed to more complex processes which require more input from technical experts in RPA.

The quality of existing data in the systems posed a challenge for the automation of some processes. The Council holds a large database with information on residents, but sometimes this data is duplicated or incomplete. This can cause errors when the RPA software attempts to retrieve the data needed to be inputted in the forms. To address this, the project team invested resources to clean the data (e.g. removing duplicates and amending incorrect or incomplete entries) and understand these risks when they programmed the RPA tool. This helped to improve the accuracy of the automated process.

In the services where RPA was introduced, non-certified staff members were also upskilled to work with the new automated processes. For example, staff needed to be able to assess reports produced by the software in order to intervene where human intervention is required. In some instances where automation led to significant efficiency gains, staff were re-trained to shift their focus from data inputting to working on more complex tasks.

Implementation, monitoring and next steps

Following the initial successful introduction of RPA, the project team has encountered a challenge in making further use of these tools. The automated processes had been designed to follow software for housing benefits and Council Tax provided by a third-party software firm. However, at some point after North Tyneside had begun its use of RPA, the third-party software firm began to update the layout of their software regularly, so that certain features moved positions. This meant that the RPA that had been designed in reference to a previous layout would no longer function and would need to be reprogrammed, causing delays to the Council's processes. The interviewee explained that the need to constantly update the programming had become too burdensome and the use of RPA was no longer deemed cost effective for most processes.

North Tyneside Council was therefore switching instead to purchasing licences for the use of application programming interfaces (APIs) offered by the third-party software firm. These are off-the-shelf solutions that communicate between different software to automate some processes with outcomes similar to RPA. However, the interviewee explained that because these are off-the-shelf solutions, they often do not provide automation that is as tailored to the Council's needs as RPA. For example, the interviewee explained that in the case of the Direct Debit process, the API only automated half of the process that they had completely automated through RPA. In this case, they are therefore continuing to use the RPA process that had been developed.

Impact

For the initial proof-of-concept on the use of RPA for benefits services, a case study was developed by ENGIE to quantify some of the **efficiency impact** associated with automating processes. This found that the introduction of RPA led to a reduction in time to process benefits claims from 36 days in 2014 to 27 days in 2016. The number of staff needed to process benefits claims therefore went from 24 staff members to 20 staff members.

In addition to improvements in efficiency, the interviewee also noted that for many processes, the use of RPA offered **improved accuracy** as compared to human staff (e.g. avoiding misspellings or typing errors).

The interviewee also explained that RPA enabled a **more effective use of human resources**. Benefit assessors could focus their efforts on assessing claims accuracy, rather than performing routine tasks, leading to reductions in fraud and speedier claims resolution. The interviewee highlighted that the project did not prompt a reduction in staff numbers (which is often a concern when automation is brought in) but rather an increase in the work quality of employees, by enabling them to move from repetitive tasks and focus on complex assessments and customer satisfaction

8. The FCA's Data Strategy

Introduction

The following case study details the Financial Conduct Authority's (FCA) development of its Data Strategy.[13] The FCA is the conduct regulator for financial services firms and financial markets in the UK.

This case study does not detail the development of a specific tool, but rather a process of organisational change the FCA has put in place to support a more strategic approach to managing, using and exploiting data and analytics.

Best practices

The FCA's data strategy illustrates several instances of good practice described in the main report:

1. Ensuring key infrastructure and skills are in place to support future work. Much of the FCA's data strategy involves ensuring that the FCA have a strong basis on which to build future innovations.

2. Developing a bespoke in-house analytical platform for data science teams: this provides a strong basis for data science teams to experiment and innovate.

3. Collaborating with a wide range of staff early on: not only did the FCA collaborate with staff in the development of their strategy, but their strategy for building and embedding in-house data science teams within the organisation is meant to facilitate collaboration between data science teams and subject-matter experts.

The FCA published its first data strategy in 2013.[14] The 2013 strategy was focused on improving data collection from firms and ensuring that data requests were not overly burdensome. However, in the years following that data strategy, technologies and techniques that used to be niche became more widely available, and those responsible for the data strategy felt that the FCA could improve the efficiency and effectiveness of its operation by making use of these.

The FCA has also seen more regulated firms using these tools across their businesses. For example, insurance firms were using data science techniques to allocate suggestions for insurance policies to consumers. At the same time, the number of firms the FCA was responsible for regulating increased with the growth of markets and extensions of the FCA's remit (e.g. responsibilities for the regulation of consumer credit), challenging the FCA to do more with the same resources. Areas where the FCA has already begun to make use of new tools and techniques, and some key elements of the new strategy include:

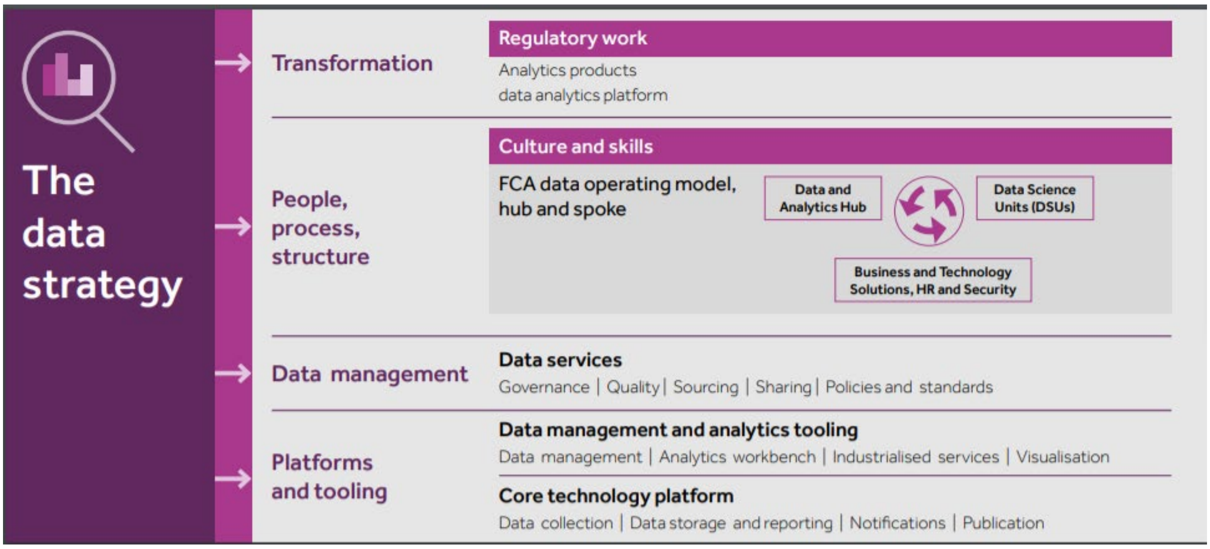
[13] Available online at: https://www.fca.org.uk/publications/corporate-documents/data-strategy
[14] Available online at: https://www.fca.org.uk/publication/corporate/fca-data-strategy.pdf

- testing and deploying new tools and techniques like web scraping, network analytics and natural language processing to transform how harm is identified and firms are regulated;
- combining new tools and techniques with business change skills to ensure tools are deployed in a sustainable manner; reviewing processes and behaviours to ensure individuals are making good use of new capabilities;
- building new flexible digital platforms to support how data is collected, stored managed and published;
- seeking to improve the flow and quality of public and commercial data the FCA collects; and
- modifying risk and control frameworks to ensure new tools and techniques are used safely.

In the longer term, bringing these data and analytics capabilities together is meant to provide an opportunity for improved approaches to regulation, including the extension of near real-time monitoring of markets and sectors. This would provide a deeper understanding of how financial services are performing and allow the FCA to react to harm more rapidly and deter misconduct.

The FCA developed its new data strategy, published in September 2019, considering these goals. Figure 6 below sets out some of the key elements of this strategy.

Figure 6: The FCA's data strategy



Source: FCA (2020) Data strategy. Available online at: https://www.fca.org.uk/publications/corporate-documents/data-strategy

Developing the strategy

The development and initial implementation of the FCA's Data Strategy has relied on:

- running workshops throughout the FCA, including senior management, to identify challenges with current activity and understand problems facing the FCA;

- active support and sponsorship from senior management;
- building infrastructure and establishing capabilities, carefully considering what tools and capabilities would be useful to the wider organisation's needs and would best exploit existing data;
- increasing and improving internal data science capacity, including the delivery of a hub-and-spoke model where data science units are embedded within the FCA's nine divisions;
- increasing data science resource with the recruitment of 70 new data scientists, analysts and engineers across the organisation; and
- promoting an innovation-friendly culture.

The data strategy team began their development of the strategy by **running workshops throughout the FCA**, including with senior management, to identify the organisation's challenges with data. This was done with the help of an external contractor. These consultations addressed issues with the FCA's data and systems, as well as questions around talent and skills, infrastructure, and whether there were opportunities for making use of data that were not being fully explored. Common challenges raised were: having the right data of the right quality at the right time; having the right systems to store, catalogue and share the data appropriately; and how data could be used to better predict where harm might occur. **Taking a holistic approach** was important to the strategy, according to interviewees, but also required an upfront investment of time (3 months) to conduct fieldwork with colleagues.

When developing the strategy, the team were **encouraged by the FCA board** to set a clear vision of what they wanted to achieve and how they could transform the organisation making use of new techniques and tools. Interviewees noted that receiving a high level of scrutiny and interest from senior management was important to developing a working strategy.

Some projects were already underway before the strategy was published in September 2019, and in some parts of the organisation, teams were already beginning to try new ways of working with data. The Data Strategy was intended to support the projects already underway, as well as establishing new ideas, projects and plans. The strategy included a delivery roadmap, broken into a series of "transition phases", each lasting around six months. This is intended to help the FCA manage the complexity of the strategy through specific blocks, helping to focus resources and allow the organisation to move forward in different areas at different speeds. The entire transformation is expected to take 3-5 years.

The initial focus of the strategy has been to **build core infrastructure and establish capabilities**. For example, one early element has been to move much of the FCA's data into a cloud-based data lake, bringing together previously siloed data into a single central location. Changes to infrastructure have also included developing and incorporating new tools. These include:

- a transition from the FCA's data collection platform, Gabriel, to a new data collection platform (RegData). This new platform was developed following feedback received from the FCA's regulated firms. It is intended to be more cost-effective, user friendly and more flexible to future changes;
- the development of a data science workbench, a custom-built cloud analytics environment that gives data scientists access to data science tools and cloud computing

power that can be used safely and securely on FCA data. This will allow data scientists to experiment with new tools and techniques on real data in a safe environment;

- bringing in Tableau™, a visualisation tool and platform that allows individuals to create dashboards of data in a more efficient and interactive way than Excel and rolling this out across the organisation. This involved providing training to 240 colleagues and establishing a centre of excellence and monthly forum to share best practice; and

- developing in-house tools that use Natural Language Processing (NLP) techniques to analyse large volumes of unstructured text data such as social media. This has been used, for example, during the Covid-19 pandemic to provide real time insight on consumer concerns.

Interviewees noted the importance of **asking the right questions before deciding a tool is the right solution.** This involves the analytics teams working closely with supervisors and specialists across the organisation to understand the issue and consider a range of potential solutions.

Interviewees noted that having the **right infrastructure and organisational structure in place was a crucial first step** to enacting the strategy's more ambitious objectives, but that **these changes require appropriate investment**. It was therefore important to have support from senior management early on to ensure this funding. This was supported by the early engagement of senior management through workshops, which helped ensure that the strategy would reflect their priorities and encouraged ownership of the strategy. Following this, there have been a number of additional sessions with leadership focused on cultural change and the required actions needed to bring this about.

Another aspect of the developing the strategy was establishing **an effective and ethical data management approach**: considering the FCA's policies for managing, governing and owning data, as well as ensuring quality. This will require reviewing data management policies to ensure appropriate governance of new tools or techniques.

In addition to technological infrastructure and changes to data management policies and approaches, another component to the data strategy has been **to increase and improve the FCA's in-house data science capacity**. This has been done by:

- employing a hub-and-spoke model. There is a central "hub" incorporating the FCA's central data services team and the advanced analytics team and "spokes" made up of nine data science units with their own managers, technical specialists and associates. A number of graduates are also allocated to each team. These data science units work within specific divisions and act as a direct and dedicated point of contact to the division who needs data expertise;

- recruitment into these new data science teams through a mixture of external and internal recruitment and a renewed focus on their graduate programme and building this out to include science, technology, engineering and math (STEM) subjects. There has also been a focus on upskilling existing data analysts and data science staff through events such as training days and workshops; and

- upskilling through a new learning and development programme and engagement with staff across the organisation (at all levels) to improve the understanding and use of data and data science. Interviewees emphasised the importance of bringing staff expertise to projects. They noted that part of that work is making sure that staff are comfortable with the use of data and data science.

Interviewees noted a number of challenges to this approach. Recruiting and retaining staff required upfront investment and therefore the support of senior management. The large size of the FCA has added challenges to engagement. There are over 4000 employees at the FCA, and across the organisation there are a wide range of experiences with data, attitudes and expectations. For this reason, the communications plan has split the audiences into those in data science units, data commissioners, practitioners and consumers so that messaging can be appropriately tailored.

Interviewees noted that in the past, they have made more use of external contractors to conduct short-term pieces of work related to data. Through working with external contractors, they have both learnt new skills and been able to attract talent in-house. Interviewees explained that bringing talent in-house and building their own capacity has several benefits. First, the FCA has **limitations around data security** (much of the data they work with contains sensitive or personal information), making it simpler to have projects carried out in-house. Secondly, in-house teams can build experience, making it easier to liaise with staff and **use data in a continual way**. Innovations do not need to be bound to specific projects or timescales, and retaining institutional knowledge means that teams can build on what has come before. Finally, keeping data science staff closer to the business makes it easier to translate business problems into suggested solutions.

The “hub” continues to engage with external bodies and academics to bring new perspectives and knowledge into the organisation.

Interviewees emphasised that **culture change** is an important component of the strategy. One aspect of this is creating a more innovation-friendly environment: allowing more time for innovation, encouraging people to try new things and allowing for failure. Another aspect is building engagement with staff at all levels and addressing any concerns about the potential impact of data science on their work. Again, interviewees highlighted the importance of senior leadership in enacting this type of change.

Impact

Because the strategy incorporates several different phases and elements, a range of impacts are expected. These include:

- increases in efficiency due to the automation of processes. This in turn should allow staff to focus on those issues and decisions that require judgement, improving quality;
- the FCA is better at predicting where harm might occur as a result of having an improved understanding of new or emerging products, market or sector risks;
- data collection is cheaper and easier for the FCA’s regulated firms; and
- the FCA can attract and retain data science talent and keeps pace with innovation in the market.

Currently, the strategy is being internally evaluated at each transition phase. Because the initial transition phases are related to establishing capabilities, this evaluation will focus on the process of how these new capabilities are delivered. As the strategy progresses toward regulatory applications, the evaluation will move to a focus on outcomes: improvements to efficiency and effectiveness at reducing harm and regulating firms and markets.

Although the strategy is currently focused on building infrastructure and new capabilities, some new tools and applications have already been enabled by the strategy. These include:

- social media monitoring: using analytics to identify issues consumers are expressing concern on and passing this information to front-line teams;
- a "financial crime heat seeker" tool: a tool that allows users to identify outliers based on specified criteria, approximately 400 data points across revenue, business model, compliance staff numbers and controls, to help to identify suspicious activity or areas of further investigation; this 'outlier' tool is currently being explored across all relevant areas;
- developing tools to profile data for quality, using techniques to look for anomalies in the data supplied by firms. For example, a data profiling tool has been created in Tableau™ employing a technique called "dynamic time warping": an algorithm that can be used to detect changes in firm behaviour over time, helping to identify potential areas for further regulatory investigation; and
- the use of web scraping to search the internet for financial promotions that could cause harm to consumers.

9. Natural England’s use of eDNA

Introduction

The following case study details Natural England’s use of environmental DNA (eDNA) to support their strategic licensing approach for great crested newts. Great crested newts (Triturus cristatus) are a protected species under European and UK law. This means that any interference with their environment is prohibited and requires a licence from Natural England. Although their population is decreasing, great crested newts are commonly found in Europe and in the UK and they are a species frequently encountered by land developers.

Best practices

The Natural England case study illustrates several instances of good practice described in the main report:

1. Conducting user testing early on: this was important to determining whether eDNA would be feasible. Extensive user testing was conducted early on, not only to gather user feedback, but also to assess the accuracy of the technique.
2. Building on projects: the use of eDNA enabled the collection of significantly more data on the presence of great crested newts than traditional sampling methods. Having this level of data meant that Natural England could combine this with other data and use species modelling to develop density maps of newts’ presence. In turn, this has helped support the development of a strategic licensing approach.
3. For external data, considering potential intellectual property issues early in the process. For the density map, dealing with IP and licence issues took significant work early on.

Figure 7: Great crested newt

Source: Natural England (2019). Protecting great crested newts. Available online at: https://naturalengland.blog.gov.uk/2019/03/11/protecting-great-crested-newts/

Assessing the presence of great crested newts on a site and attempting to relocate them can be costly for developers. In some instances, the presence of newts is only discovered after development has begun, effectively halting the project until further action is taken. This process is costly, time consuming and often poses a barrier to development plans.

The identification of great crested newts requires skilled personnel, as newts are elusive and hard to identify. Traditional methods for identifying newts include nocturnal surveys, trapping and regular checks to monitor further presence. These methods are time-consuming. Since great crested newts are found across the UK but only in around 10% of ponds, extensive sampling is required to assess their distribution. Historically, much of the traditional sampling work was undertaken by volunteers, motivated by an interest in conservation and observing wildlife. Using eDNA is a new approach to this process involving analysis of the traces of DNA left by an organism in the environment (e.g. through skin or waste products). This eDNA is procured by sampling the water or soil and then extracting the DNA through laboratory processes. This means that the animal's presence in the environment sampled can be detected even if no specimen has been caught or observed.

As the application of eDNA was being developed, beginning in 2012, there was also a move to develop a new licensing scheme for great crested newts. Rather than assessing the presence of newts and relocating them, District Level Licensing allows developers to pay into a conservation scheme, which supports the creation, maintenance and monitoring of new or existing habitats best suited to support and expand newt populations. This can be done without surveying for newts. These habitats have been identified and developed based on a map indicating the density of newt populations across England, produced using the eDNA data and species distribution modelling.

The application of eDNA techniques to survey great crested newts began around 2012 and the study elaborating the methodology was published in 2014. The development of the methodology took around two years and involved an initial scoping stage, development of laboratory assessments and field trials. This work was done in partnership with Defra, the Freshwater Habitats Trust and the Amphibian and Reptile Conservation Trust. Work to develop and implement District Level Licensing began in 2016 and implementation was ongoing at the time of writing this case study.

This case study is based on interviews with two individuals at Natural England and a review of the methodology study[15] and other documentation provided by the interviewee.

Relevant technologies

The use of eDNA relies on next generation sequencing technologies. Although DNA extraction and sequencing is not a novel process, the recent development of next generation sequencing technologies allows for much greater volumes of raw DNA and specimens to be processed.

[15] Biggs J, Ewald N, Valentini A, Gaboriaud C, Griffiths RA, Foster J, Wilkinson J, Arnett A, Williams P and Dunn F (2014). Analytical and methodological development for improved surveillance of the Great Crested Newt. Defra Project WC1067. Freshwater Habitats Trust: Oxford. Available online at: http://randd.defra.gov.uk/Document.aspx?Document=11973_WC1067_FinalReport.pdf

Developing the technology

Natural England's use of eDNA for monitoring great crested newt populations has relied on:

- the use of existing technologies and methodologies;
- a rigorous approach to testing the technology; and
- bringing together a team that could translate a proof-of-concept into a reliable, routine application.

The use of eDNA was based on **existing technologies and methodologies**. The work done by Natural England and its partners was not to develop the specific technology or tool, but rather to develop a reliable method and approach to applying this technology to licensing reform for great crested newts.

The first step was therefore to develop and test the methodology. According to interviewees, the status of the great crested newt as a European Protected Species contributed to choosing a **rigorous approach to testing the technology**. The study team sought to achieve a high level of confidence that the technology was equal to or better than methods previously used and did not present any risk of harm to newts. The initial methodological study took 2 years and directly compared the effectiveness of traditional methods with that of eDNA. Samples were collected from ponds outside of the great crested newts' known habitats to test for false positives. Professional survey teams and volunteers then conducted eDNA sampling at the same time as traditional sampling at 35 different sites. The eDNA samples were then collected by just over 80 volunteers at a further 239 sites. The study assessed the extent to which eDNA detection was affected by other factors, the extent to which eDNA collection would be practicable for volunteers (as volunteers are generally used to assist with traditional detection methods) and the extent to which eDNA may lead to false positives. This initial study concluded that eDNA offered a reliable and more efficient approach to sampling than traditional methods.

The main challenge in developing the use of eDNA for great crested newt identification was **moving from a proof-of-concept to a reliable, routine application**. This required technical expertise, expertise in innovation, experience in turning proof-of-concept projects into actual applications, and the funding to bring such skills to the project. The team had to establish eDNA technology as the methodology of choice, provide evidence that the technology is robust and ensure effective contracts were in place with trusted contractors to undertake the survey and analyse the results. Great crested newts being a high profile protected species helped secure the required funding. Initial funding (around £200,000) largely came from Defra, with some support from Natural England, the Freshwater Habitats Trust and other environmental NGOs. Testing and field trials was led by the Freshwater Habitats Trust.

Implementation, monitoring and next steps

Following the results of the study, the eDNA sampling method was implemented and is used in daily operations. Laboratory analysis is provided by external contractors. These external contractors did not necessarily provide eDNA analysis before the implementation of this method: one consequence of the development of eDNA is that it has created a niche market for eDNA laboratory analysis. Because the use of eDNA is still an emerging technology and market, Natural England designed and introduced a proficiency test for laboratories, requiring that all participating laboratories undergo this test. This helps to ensure that the results remain high quality.[16] The use of eDNA sampling has allowed Natural England and conservation organisations to collect a significant amount of data from across ponds in England. This data has in turn enabled the District Level Licensing approach, which would have been prohibitively expensive if traditional sampling methods were used. The significant improvements to the efficiency of sampling means that Natural England has been able to fund surveys across over 7000 ponds in England. This in turn generated enough data to produce density maps. The density maps have been based on eDNA data collected over three successive summers, combined with historic sampling data and data from external sources to create spatial layers and map the presence of great crested newts. Species distribution modelling was then used to produce two maps: one map indicating Risk Zones, showing what impact developments in different areas will likely have on newt populations, and another map indicating Strategic Opportunity Areas, showing where new ponds can best be built to support newt populations.

One challenge noted by the interviewee was that for the distribution modelling and maps, **external data was not always accessible open data**. General data layers on land covers or rivers may be openly available, but species data may not always be open access, incurring costs or subsequent conditions on usage. Similarly, other national datasets may need a licence for use, which can then impose restrictions on what can then be shared publicly. Complying with these licensing conditions was a challenge. As a public sector organisation, Natural England aimed to share information generated through the project under an open government licence. Working through these various licensing issues required significant work early on. The interviewee noted that they had initially underestimated the amount of work that this would require.

One interviewee noted the need to **consistently monitor the standardised protocols** used for sampling and analysis to ensure that these are kept up to date. Standardisation without monitoring could contribute to inhibiting development and innovation. District Level Licensing is intended to support this: a small component of the payments made into the scheme goes to regular modelling and mapping updates.

Impact

The impact of this innovation at Natural England has not been formally evaluated. The initial methodological study provides some indications of the likely impact of the change. It compared eDNA to traditional methods in terms of their efficiency and reliability. It found that taking a sample using traditional methods would take approximately 48 hours, while using eDNA methods would take 2 hours. Moreover, it found that eDNA methods had a **higher successful detection rate** and **required less training to undertake** as compared to traditional methods. In other words, eDNA has offered a process that is **less costly, faster and more reliable** than traditional sampling methods.

eDNA has become the standard technique for detecting the presence of great crested newts, replacing or sometimes complementing the manual techniques used before. It has enabled District Level Licensing, which is expected **to reduce burden and uncertainty on developers** and **ensure that resources spent go directly towards compensation habitat for the species**. The use of eDNA, in part, has enabled protected species licensing to be strategic. This is expected to deliver economic benefits, as well as improve environmental outcomes. As

[16] Rees, H and Gough, K (2018). Great Crested Newt eDNA Laboratory Quality Systems, Proficiency Testing and Interpretation of Results. In Practice Issue 99. CIEEM. Available at: https://cieem.net/wp-content/uploads/2020/03/InPractice99_Mar2018.pdf

the scheme is currently being implemented, the team is quantifying the benefits the scheme is helping to realise.

The density maps are open source **and publicly available**, meaning that they can be used by academics and conservation organisations to conduct further research or support other initiatives related to the conservation of great crested newts.

Natural England's experience with District Level Licensing is also **building an example for future projects and the conservation of other species**. Natural England would like to implement a landscape-scale approach to conservation and the use of tools such as eDNA can help to develop the evidence base required to support such approaches. The interviewee noted that great crested newts were a good choice for this type of approach because they are comparatively easy to sample: they are relatively common and a lot is known about the species and their requirements. For a species such as bats, building an appropriate, robust evidence base will be more challenging, but beginning with an easier case has allowed Natural England to address more general challenges and establish a framework for similar schemes.

10. The Environment Agency's use of webcams

Introduction

The following case study details the Environment Agency's use of webcams for monitoring culverts.[17] The Environment Agency is responsible for a range of monitoring and enforcement activities related to protecting the environment, including monitoring air and water quality.

Best practices

The Environment Agency case study illustrates several instances of good practice described in the main report:

1. Considering a range of options: the best solution in this case was the use of a more established technology, webcams.
2. Involving staff in the development: staff on the ground were important to the development of the use of webcams, and this helped to build up local knowledge, ownership and understanding.
3. Carefully considering the use of external contractors and mitigating against the risks of ending up with costly solutions: the Environment Agency's initial experience with procurement did not bring about the right solution. In their second attempt, they learned from this and were able to identify a more appropriate contractor.

This case study has been based on interviews with two individuals involved in innovation at the Environment Agency. Both interviewees were part of an "innovation team" and as such had an extensive experience with a range of projects. Learnings identified in this case study reference some of those broader experiences.

Figure 8: Examples of culverts

Source: Environment Agency (2010). Fluvial Design Guide. Available online at: http://evidence.environment-agency.gov.uk/FCERM/en/FluvialDesignGuide/Chapter8.aspx

[17] Culverts are the structures that allow rivers or streams to flow under roads, buildings, bridges and other similar construction

Interviewees felt that within the public sector, it is easy to default to a conservative approach to adopting new technologies: technologies that are “emerging” are perceived as costly, complicated to implement and their efficacy uncertain. In some cases, changing approach, particularly when it comes to monitoring, also means losing continuity in data. Having a long time series of data is preferable to having more accurate data in some cases, and this can be another reason for not changing methods. In spite of these concerns, the Environment Agency has been monitoring new technologies, reviewing what is happening in the private and public sector and assessing the extent to which these could be applied at the Agency, and in particular into the next iteration of a monitoring network.

The example discussed here relates to the Environment Agency’s responsibility to maintain culverts in the UK. These culverts are liable to blockages, where debris that has flowed downriver ends up trapped. This can in turn lead to flood risk. To help prevent this, culverts contain trash screens, which help to collect rubbish, logs and other debris that can block the culvert and lead to flood risk. These screens require regular cleaning and maintenance. Historically, Environment Agency staff were responsible for physically inspecting these screens and clearing them out when needed. This involved extensive travelling, and in many instances, unnecessary trips or delayed responses to issues. To avoid the need for physical inspection, the Environment Agency has implemented the use of webcams and image differencing technology.

Developing the technology

The Environment Agency’s development of the use of webcams for monitoring culverts has relied on:

- early trials conducted by the innovation team;
- grassroots adoption of the tool by staff;
- a careful identification of contractor;
- gaining support from middle management; and
- learning from others.

The Environment Agency’s use of webcams began with **trials conducted by the innovation team** transmitting simple images of the culverts. Staff could use these images to avoid travelling to the culverts. Feedback from staff on these trials was positive, but setting up the cameras and systems was complicated. There was an issue related to **privacy laws**, as they needed to ensure that cameras were not capturing images from any adjoining private land. The telemetry was also complex. Because it was perceived as a useful solution across the organisation and because of these implementation challenges, the project was then taken up centrally, rather than by the innovation team. A procurement competition was set up and the contract was awarded to a large IT firm. However, the solution developed by this firm was costly, involving significant physical infrastructure around the cameras and complex telemetry. Because of the high cost, development of the webcams stalled, and the solution proposed by the contractors was not taken further.

In the meantime, Environment Agency staff at local levels began to implement their own, cheaper versions of the technology. This was not centrally coordinated, and as such, led to a mix of different tools, processes and telemetry. After one local provider went out of business and the cameras that staff had begun to rely on were no longer functional, a project was kickstarted to bring these different local projects together centrally. A second procurement competition was set up requiring additional time and effort **to identify a contractor for the new work**. The new procurement contract went to a smaller, mid-sized company. The solution proposed was more affordable and more agile. Following this contract, there was a single supplier, a single system and a small range of cameras in use.

The interviewee explained that, although bringing this together in a centralised project was important, the grassroots nature of the systems that preceded centralisation helped to build **up local knowledge, ownership and understanding**. The interviewee cited the importance **of building engagement with "middle management**" – noting that often those at the top of the organisation and those on the ground are motivated to engage in innovation, but those who are responsible for budgets and demonstrating value will be less inclined to engage. These middle management staff are essential in bringing things from proof-of-concept to something operational. They should therefore be engaged from an early stage.

After the single centralised system was developed, the question arose of how best to monitor screens. Images would need to be checked regularly or else the system would not achieve its objective. Often when incidents are occurring, staff are busy in the field and not able to check these images. To address this problem, **image differencing technology** was applied. The innovation team had learned about this technology through an example of work done by the National Physical Laboratory for Network Rail. For Network Rail, monitoring minute differences in infrastructure is important, as it helps to forewarn of potentially more significant problems. In that respect, the tool at Network Rail fulfils the same function that the Environment Agency needed it for: an early warning system for physical disruptions that could cause major incidents at a later stage if not mitigated. The Environment Agency subsequently commissioned the National Physical Laboratory to apply the same approach to their monitoring of culverts.

Image differencing is a decision aide: staff receive an alert when something has changed. They can then check the image and determine whether it relates to a potential blockage. The sensitivity for alerts is set high, with the assumption that many of the alerts will be false positives (image differencing will also pick up unrelated changes, such as a spider on the lens). However, many of the smaller changes the system will pick up could still be significant. As the interviewee noted, the testing and implementation of the solution at Network Rail contributed to the success of the Environment Agency's experience.

Implementation, monitoring and next steps

Maintenance is included in the new procurement contract. Environment Agency staff on the ground assist with repositioning cameras where needed, but any technical issues are managed by the contractor. As this is a niche tool used by the Environment Agency and such service is only required occasionally, accessing these skills externally is more efficient. The interviewee also noted that maintenance has become easier in part because the technology available has become more advanced.

Impact

The impact of this case has not been formally evaluated, but has related mostly to increases in efficiency: staff do not need to travel to physical sites for inspection, but only for mitigation, which means their time is used more efficiently.

11. DVSA's Intelligent Risk Rating Assessment tool

Introduction

The following case study details the development of the Intelligent Risk Rating Assessment tool of the Driver and Vehicle Standards Agency (DVSA), an executive agency of the UK Department for Transport (DfT). DVSA is responsible for road safety in the UK and the compulsory annual Ministry of Transport (MOT) test for all road vehicles over a certain age.

Best practices

The DVSA case study illustrates several instances of good practice described in the main report:

1. Building on the previous establishment of other technologies and the adoption of a cloud-based system: DVSA's digitisation work established a basis on which the project team was able to build their solution.
2. User testing was conducted early on and was incorporated as part of the design process for the tool.
3. Early adopters of the tool helped to act as ambassadors, engaging other staff and increasing confidence in the solution.

DVSA relies on Vehicle Examiners (VEs) for the examination of 66,000 MOT Testers across 24,000 garages in the UK maintaining standards and eliminating fraud. Historically, inspection and enforcement activities were performed unannounced on a regular basis. Before the visit, VEs needed to complete a manual data analysis of the garage MOT test data. A risk score was assigned by a VE based on a series of questions, which could be open to interpretation. This meant that the assessments of the MOT Testers' performance on the day of the inspection could be subjective and sometimes lead to inconsistencies with those of other VEs. In addition, this process was resource intensive and did not allow for the allocation of staff resources to those garages that would most need them.

Several considerations have led DVSA to consider an alternative approach:

- the size of the market to be regulated: substantial resources were required to inspect all 66,000 MOT Testers across the 24,000 garages;
- the growing amount of data produced by MOT testers, which offered opportunities to regulate MOT testing and road safety more efficiently;
- DVSA's limited resources, which meant that a site assessment could only be completed every one to three years, a frequency deemed too low to ensure road safety and eliminate fraud; and
- the element of subjectivity in the assessment of MOTs by VEs, which could lead to inconsistencies.

DVSA wanted to increase its efficiency, improve the quality and transparency of its regulatory activities and optimise general business activities.

To achieve these objectives, DVSA invested in the development of a new risk rating model, using machine learning (ML). The model incorporates a broad range of features, including individual MOT tester performance. The model takes into account changes to testing behaviour over time, for example if a tester starts to record only high test scores, which could be an indication of fraud. In addition, it incorporates other control and risk factors such as test volumes, pass rates, and the disciplinary history of the garage.

DVSA applied a clustering model against garage test data, which grouped the MOT centres based on their testing behaviour. The information generated is then combined with the historical data about disciplinary action against garages for not adhering to correct MOT standards, which ultimately defines a risk score. The model is run every month and considers changes to testing behaviour, providing up-to-date risk scores.

The tool has been used by DVSA to make data-based decisions on field visits, targeting its resources based on risk. The output of the model is used to inform and conduct necessary inspection and enforcement operations to verify and tackle suspected fraud and inadequate standards. The solution has also increased the frequency with which risk scores are given to MOT Testers, from once every three years to once every month.

DVSA partnered with an external firm for the development of the technology. The project started in 2017. Five to six weeks were needed to study a subset of data and detect patterns to identify “good” and “bad” MOT tests, followed by a validation period after the problem and opportunity were defined. The production process took 12 months. The technology has been in use since the last quarter of 2018.

This case study is based on 2 interviews, one with a DVSA representative and another with a representative from the contractor. It is based also on the review of public documentation on the solution.

Relevant technologies

The solution relies on three technologies:

1. Amazon Web Services (AWS) cloud platform to host the service;
2. a statistical tool (clustering model) to assess the data; and
3. a risk rating algorithm using the output of the statistical tool and historical data to rate each garage.

MOT tester data is added and stored on the AWS Cloud. A clustering model is applied against garage test data, which groups MOT centres based on their testing behaviour. The information generated is then combined with the historical data about disciplinary action against garages for not adhering to correct MOT standards, which ultimately defines a risk score.

The risk rating algorithm in place focuses on individual MOT tester performance, takes into account changing testing behaviour and incorporates other control and risk factors such as test volumes, pass rates, and the disciplinary history of the garage.

Developing the technology

DVSA's development of its intelligent risk rating assessment tool has relied on:

- technologies previously adopted as part of the digital transformation of the MOT test service, such as the transition to the AWS cloud platform;
- a pre-existing relationship between DVSA and a private provider who was involved from the beginning of the process;
- user consultation and testing with VEs as subject-matter experts at three different location across the UK before the tool was rolled out;
- support from the department and the envisaged users; and
- MOT test data from garages of sufficient quality to train the algorithm used to generate the risk scores for MOT testers.

The tool built on **the adoption of a cloud-based system and other technologies previously adopted** by DVSA. DVSA and its partner for the delivery of this project had a pre-existing relationship. They had collaborated on the development of multiple digital services since 2013. The risk rating model was part of the digitalization of MOT test services which had started in 2012. This entailed incorporating various new technologies in its business operations. The mainframe-based MOT system was replaced with a cloud-based system, which proved faster and more user-friendly for people accessing the data. The data was pushed out to users through Power BI[18] dashboards allowing real time decision making. More projects followed, such as the MOT History Check, allowing the public to check a car's MOT test history among other information. The use of ML to develop the risk assessment tool built on those previous changes.

The fact that the risk rating tool followed from a suite of known and tested changes meant that it was not difficult for internal funding to be secured. The tool was part of wider investment of £23 million in improvements to the MOT process. The necessary computational and human resources were made available. The skills were found in-house, in the data science team, and through **collaboration** with the contractor. VEs, being the end-users, were also included in the development process as subject-matter experts to provide input and feedback. The interviewee noted that all parties worked on the project together as one team.

Because **user experience was taken into account from the start** through collaboration with VEs, research activities were limited and conducted on a small scale with a defined number of respondents. The tool was then tested at three different locations across the UK and scaled-up after validation.

The main challenge was **gaining the support from staff and enforcement officers**, some of whom had doubts about the extent to which the technology would be useful, whether the data produced by the tool could be trusted, and whether the business case for the change was sufficient. To address these concerns, the tool's added value to DVSA's operations was set out in non-scientific language to show how it would deliver efficiency gains and contribute to improving quality. In addition, early adopters among VEs acted as product ambassadors of the solution. This helped increase confidence among other VEs. VEs were then directly involved in the validation of the tool: the output of the predictive model was verified and checked in

[18] A business analytics service by Microsoft

cooperation with VEs to determine the extent to which it reflected reality. The validation process further clarified the utility of the tool for users.

Another challenge was **ensuring data quality**. The data used to train the algorithms and determine the risk rating influences the quality of the tool and trustworthiness of the output of the model. All data needed to support the solution was already available and **no legal barriers to access or use the data** were encountered. GDPR challenges would arise in a case where DVSA wanted to use the technology outside its operating environment (e.g. partnering with police officers). If they would have wanted to broaden the scope of users, this would require a new legal framework in order to share the data with other organisations.

Implementation, monitoring and next steps

Early integration of feedback from test users facilitated the implementation of the technology and **quick adoption among VEs**. Furthermore, VEs' experience with MOT testing data facilitated the roll out of the tool. No upskilling of users was needed because the solution was developed to support and improve their way of working rather than changing the approach: VEs benefitted from a richer, near-real-time dataset that identifies high risk garages, reducing the time that they would normally spend on pre-visit manual data analysis.

The tool has been integrated in DVSA's digital services and is being **monitored** on a weekly basis **as part of daily operations**. The modelling algorithm depends on the digital data platform that records the data from the MOT system. It uses that data to build a performance picture of the garages. The model automatically processes the data and detects patterns, which are directly available to the VEs in spreadsheets. VEs can decide which garages to spot-check based on the ranking and target the ones in the red zone, which are the garages in that pose the highest risk for underperforming or committing fraud.

The contractor ensured that DVSA would be able to maintain and improve any new solution independently. The contractor provided the necessary information and upskilling needed. This has allowed DVSA's data scientists to maintain the model and improve it if desired.

The tool has created appetite **to look at other roll outs of ML** and AI within DVSA and further investigate the potential of predictive analytics. At the time of the interview the collaboration between the contractor and DVSA continued.

Impact

The solution has brought both **efficiency gains and capacity gains**. Instead of a manual risk analysis of 24,000 garages and site visits to every garage in the UK once in one to three years, MOT testers are now being assessed on a monthly basis. DVSA enforcement officers spend 50% less time on preparing their checks. In March 2019, twice as many of the visits to high risk garages resulted in disciplinary action as compared to March 2018 (going from 33% to over 66%). Assuming the number of non-compliant garages has remained consistent, this indicates a significant improvement in the targeting of visits. Since its introduction, the model has continued to improve targeting of garages, both because greater data has been able to improve the accuracy of the tool over time and because the regular monitoring of the tool has helped to improve the model.

12. The City of Chicago and WindyGrid

Introduction

This case study details the WindyGrid platform, an open-source, predictive analytics platform developed and implemented by the City of Chicago to allow policymakers to explore and analyse the city's data in an aggregated manner.

Best practices

The City of Chicago case study illustrates several instances of good practice described in the main report:

1. Building on existing infrastructure and skills: the City of Chicago had already worked to consolidate and centralise IT departments and data management systems and had built a highly skilled in-house team.

2. Building on other projects: the WindyGrid project began as a straightforward exercise in bringing data together. Analytics projects were then developed using that data, and more tools and features were added over time.

3. Learning from the experiences of others: the project team shared the code used for WindyGrid, allowing other cities to develop similar tools and to collaborate to improve such tools.

4. Taking a problem-focused approach: although data analytics proved helpful for many City departments, it did not always offer the best solution. In such cases, departments were encouraged to look at other resources or approaches.

The project began in 2012, as the City of Chicago was preparing to host a summit of the North Atlantic Treaty Organization (NATO). To ensure preparedness and optimise public services, officials identified the need for a tool that could combine live data coming from multiple sources (e.g. surveillance footage, crime data, public transport data) into a single interactive map of the city. Such a tool would help public authorities have better oversight of incidents across the city and thereby improve emergency response.

A basic version of the WindyGrid platform was developed in response. This served as a basis for adding further layers of data and ultimately, analytical capabilities. WindyGrid is a single graphical interface that integrates operational data from across city departments (including live and historical data) on a single dashboard. The different types of data are integrated into a simple interactive screen, allowing city authorities and public safety agencies to make targeted queries and receive automatic updates and alerts.

The data processed through WindyGrid is diverse. It ranges from geospatially tagged emergency operations data to public tweets, video feeds from surveillance cameras, weather, traffic and pedestrian patterns and city bus location data. WindyGrid has been designed to enable users to query data by type, time or distance from a given location. Users can also choose between different ways of displaying data, for example by creating a heat map to show concentrations of results in the areas they are interested in.

The development of WindyGrid also led to a version of the tool called "OpenGrid" (OpenGrid.io), intended for the general public. The map brings together publicly available datasets in a single platform, including, for example: the location and subject of non-emergency service requests, food safety inspection results, potholes, street closures and filming locations.

Both WindyGrid and OpenGrid are currently in use and their underlying software and technical aspects are publicly available as open-source code.

This case study is based on one interview with a member of the development team and a review of documentation.

Relevant technologies

WindyGrid is a Geographic Information System (GIS) that pulls together the city's spatial data, allowing staff to access both historic and real-time data through a single platform. It was designed using open source software tools. The data ETLs (extract, transform, load)—the systems that take the data held by different Departments and sources and bring them into WindyGrid—were built in-house, as was the front-end of the tool.

The predictive analytics features have been built over time and in specific response to the needs of City staff.

Developing the technology

The City of Chicago's data analytics project has relied on:

- a well-established "open data" approach;
- in-house technical capabilities;
- a first tool – WindyGrid – developed in very limited time and fully in-house; and
- generous funding for an innovation programme, received from philanthropic grants and the City's own budget.

The initial version of WindyGrid was developed in a short time in the context of the 2012 NATO Summit. The team had approximately 6 weeks to develop the initial platform. The limited time and budget available **and lack of available solutions on the market** at the time led the City of Chicago to develop the project entirely in-house. The interviewee noted that the market for relevant solutions has evolved significantly since 2012, and in the present context, it would be more effective to use external support and off-the-shelf tools.

According to the interviewee, the in-house team was able to accomplish development in a relatively short time frame because the City had recently **hired talent with technological skillsets** and work had already been done **to consolidate and centralise** the City of Chicago's **IT departments and data management systems**. The interviewee noted that both of these conditions were not usual for city governments at that time.

The initial platform developed provided a basis for the team to add further capabilities. According to the interviewee, the City's leadership were supportive of making better use of the City's data, and this helped the development team receive the funding needed to continue work on WindyGrid. The City of Chicago also received around $10 million in funding for a wider data innovation project through philanthropic grants, from which the WindyGrid project benefitted.

In subsequent years, more datasets and analytical capabilities were incorporated in the platform in response to the priorities of the City's departments. Once a new use case was identified, the IT department would assess the analytics potential and create a roadmap for the development of new features. They would then work to develop a relevant model and algorithms using the platform as a basis to address those needs. The development of WindyGrid therefore relied on engagement and collaboration of multiple departments within the City. Examples of use cases include:

- capturing data on pothole complaints on Twitter for the Streets Department;
- capturing information on food poisoning from Twitter to help guide restaurant inspections;
- bringing together data from the Transit Department, Sanitation Department and Traffic Services to more efficiently route parades; and
- pulling together data from a range of sources, including weather patterns, waste collection and vacant buildings to identify likely rat infestations around a week earlier than it would be reported by residents.

There were **no legal issues related to the use of data** for WindyGrid. The City of Chicago had traditionally employed an "open data" philosophy and the local legislation allowed for the publication of operational data, as long as it did not involve personal or sensitive data or if the data was anonymised where necessary. No personal or sensitive data was included in WindyGrid. The development team avoided introducing specific types of sensitive data such as health data or data regarding children in order to avoid conflicts with federal or local legislation. Throughout the development process the general approach was to limit the project as much as possible to internally available data. When data came from external sources, safeguards were put in place through data-sharing agreements.

However, there were a few shortcomings related to the **quality of data**. As the data used was already held and owned by the city and had been collected for different purposes, the quality of the data varied. Most of the datasets were well-structured, but others were very poor and the development team had to look for external surrogate data in order to obtain their desired insights. Existing **bias within the data** was another issue. For example, crime data may be a reflection of policing practices as well as of crime risk. The interviewee noted that a good understanding of how data was generated was essential to understanding whether and how that data could be appropriately used. The project team addressed this issue by piloting all the models they were building and considering potential biases.

The interviewee explained that one challenge in bringing forward the analytical components of WindyGrid was convincing management and staff that data analytics could be a reliable and valuable tool for different use cases, in spite of minor data imperfections. Once the team started applying data analytics to specific problems and using them to develop solutions, there was a growth in interest from across departments. Positive use cases, such as those mentioned above, helped to advertise the potential of the WindyGrid tool. Staff also received training on the platform to encourage the use of the tool. The interviewee explained that in some cases, there was a need to convince staff that data analytics was not always the most appropriate solution for their problem and some departments had difficulties understanding which of their problems can be solved with predictive analytics and which would need a different type of resource or approach. For example, the interviewee explained that there was initial interest in using data analytics to address the City's backlog of tree trimming issues. In this instance, data analytics could not offer a solution, as the backlog could only be addressed by investing more resources in tree trimming.

Implementation, monitoring and next steps

Besides helping city managers analyse data and engage in predictive problem-solving, the project also aimed to support other cities that lack the resources to build the capacity themselves, by sharing the platform with them.

The City of Chicago built through WindyGrid an open-source data infrastructure and a set of algorithms that other cities can re-use with few software development costs. This means that other cities could import the architecture and the predictive algorithms developed by the City of Chicago and adjust it to accommodate their own operational requests or related public services datasets. The project aims to give other cities a roadmap to develop their own predictive analytics projects.

The interviewee noted that several components of the WindyGrid project have been replicated by other cities, which was a gratifying aspect for the project's team. Moreover, the interviewee noticed that those cities have been sharing with each other the code they have been developing building on WindyGrid.

Impact

Each of the numerous use cases of WindyGrid can be assessed in terms of their specific impact. However, the project did not benefit from a full-scale evaluation of its overall impact for the City of Chicago. The IT department gradually evaluated and assessed the new functionalities that were added to the platform in order to scale them as part of the city's annual budget process and to show return on investment for each new feature.

Consequently, the impact of the WindyGrid platform varies across the different areas and services in which it is being used in the decision-making process. Several other factors also influence the different levels of success in using WindyGrid. For example, some departments are more prepared to embrace data analytics tools if their staff are familiar with data management.

Some departments could more easily assess quantitative impacts than others. For example, for some departments such as sanitation, the average cost for garbage disposal can be easily calculated, a complex urban challenge such as homelessness prevention is more difficult to quantify and evaluate.

In the absence of a formal impact evaluation of the project, anecdotal evidence suggests that bringing different data streams and datasets together and enabling data analytics has had positive impacts across a wide range of public services.

One benefit brought by the WindyGrid tool was that it has enabled the City's services to move away from a reactive mode of intervention to a proactive mode. For example, the city found that it could anticipate future rat outbreaks with a good level of accuracy by combining information on water-main breaks and citizens' complaints involving garbage. This helped identify certain high-risk areas within the city. As a result, City authorities could pre-emptively address the rodent problems before they became visible. It was noted that WindyGrid can predict rodent activity spikes 7 days in advance.[19]

Algorithms have been used for other purposes to a similar effect. By sifting through key words, location data and geocoded tweets, Chicago's police were able to identify an alert on WindyGrid about an abandoned backpack in a public square and to intervene eight minutes earlier than they would have if they had relied only on an emergency call to the police. This helped the police deploy bomb squads in a timely manner.

Another example of WindyGrid's impact is in improving the cost-effectiveness of managing parades. By providing centralised live information, WindyGrid can be simultaneously consulted by various departments working to provide sanitation, traffic control or security for the city's parade. This facilitates coordination across departments, provides for a faster response rate in case of incidents and an overall significantly lower cost of the parade for the City of Chicago. In one example, the City estimated that WindyGrid cut down the direct cost of a parade to the city by 35%. Moreover, the tool helps the city to quickly return to normal once a parade has ended, leading to likely cost savings and benefits for private companies (such as couriers and delivery services).

[19] Chicago Digital, (2014) "Chicago's SmartData Platform: Pioneering Open Source Municipal Analytics" available at https://digital.cityofchicago.org/index.php/chicagos-smartdata-platform-pioneering-open-source-municipal-analytics/?doing_wp_cron=1591467125.3721129894256591796875

Glossary

This glossary defines some of the more technical terms used in the case studies.

Algorithms – are a set of rules, formula, or a finite sequence that defines and instructs computations.

Application Programming Interface (API) – is an intermediary tool that aids the communication of data between different systems. It allows the transition and translation of requests and responses across different systems.[20] The API takes requests, communicates what to do to the system, and then brings the response back to the users. APIs facilitate the interaction between applications, data, and devices.[21]

Artificial Intelligence – covers technologies that enable machines/computers to do tasks that typically require human intelligence. A system incorporating artificial intelligence will analyse its environment and take actions with some degree of autonomy in order to achieve the goals that were programmed in their code.[22]

Blockchain – is a shared database filled with entries that must be confirmed by all users of the database and encrypted. Blockchain is a type of distributed ledger database formed by a sequence of "blocks" that are added to the chain of transaction records. A distributed ledger technology does not necessarily have this feature.

Brittle (Software Brittleness) – refers to the increasing fragility of some older software which are seemingly operational but fail when presented with unusual or altered data.

Chatbot – a software designed to substitute human verbal interaction with automated processes. This is achieved using artificial intelligence and natural language processing systems to simulate the software's conversational skills and understanding of natural language inputs[23].

Clustering – is a technique to organise large amounts of data, setting it into significant groups in which patterns are detected, i.e. "clusters". The model maximises distance *between* clusters, while minimizing distance *within* the same cluster between similar data points. The model attempts to detect natural statistical separation within a dataset. The similarity of data samples is defined by the distance between vectors. Clustering works well on large datasets with multiple features. Output is optimised through adjusting input parameters.

Cloud – is a term used to describe a network of servers that can be accessed remotely regardless of one's location. Clouds can also be defined as online pools of resources which can be used to store and manage data, or to support the running of different applications or services. They replace the local storage space of a personal computer with an online-

[20] Gazarov, Petr (2019) "What is API? In English, please" https://www.freecodecamp.org/news/what-is-an-api-in-english-please-b880a3214a82/
[21] MuleSoft (015) "What is an API?" https://www.mulesoft.com/resources/api/what-is-an-api[22] High-Level Expert Group on Artificial Intelligence, (2019), *A definition of AI: Main capabilities and scientific disciplines,* European Commission, Brussels.
[22] High-Level Expert Group on Artificial Intelligence, (2019), *A definition of AI: Main capabilities and scientific disciplines,* European Commission, Brussels.
[23] The Ultimate Guide to Chatbots – Drift.com Available at: https://www.drift.com/learn/chatbot/#pillar-start

accessible storage space, which can be retrieved by multiple devices with the necessary credentials.[24]

Cloud-based data lake – is a cloud-based repository that allows storage of both structured and unstructured data.

Cloud Computing – refers to on-demand availability of IT services (applications, storage or programmes) through cloud access, and thus renting rather than buying hardware.

Computational Semantics – is a branch in computer science that researches how to automate the process of semantics and human communication. One of the central questions of this discipline is how to represent meaning in ways understandable by computers [25]. Computational semantics aims to perform automated meaning analysis of natural language, and it has a prominent role in the natural language processing (NLP) applications [26].

Data analytics – refers to the techniques used to analyse raw data to recognise patterns and identify behaviours[27]. Data analytics entails the process of extraction, gathering and categorisation of data for inference. The techniques involved in these processes vary depending on the desired end results and structure of the data.

Deep learning – is an area of machine learning. In machine learning, the algorithm is given a set of relevant data to analyse and learn from it to generate predictions or to group data objects. In deep learning, the algorithm is given raw data and it is programmed to partially solve feature extraction issues, particularly relevant for unstructured data such as images, speech and text. Deep learning models typically require a substantial amount of data to generate accurate predictions. Although all statistical machine learning improves with higher amounts of data, deep learning data requirements are more substantial. For that reason, smaller datasets work better with classical machine learning methods rather than with deep learning.

Distributed ledger technology – is a technology that enables the creation of a de-centralised database. Instead of storing data in a single place, the data is divided and entrusted to multiple users and it creates a specific protocol. This generates a secure environment in which any change must be replicated by every user in order to occur. In this way, fraud can be avoided and there is a high level of transparency surrounding the transactions/exchanges that occur in such a system.

Dynamic Time Warping (DTW) – is an algorithm technique used in time series analysis to compare similarities in different arrays that do not synch up perfectly. Dynamic Time Warping is applicable in a wide range of domains, such as speech recognition and financial markets data.

Environmental DNA (eDNA) – is a sampling approach that extracts the genetic material directly from an environmental sample (soil, sediment, water, etc.) without the need to interact directly with the biological source. This process is enabled by the analysis of the environmental

[24] Microsoft, 2018, *What is the cloud*, Available at https://azure.microsoft.com/en-us/overview/what-is-the-cloud/]
[25] Blackburn, P. and Bos, J., 2005. Representation and inference for natural language. *A first course in computational semantics. CSLI*.
[26] Bos (2011) "A Survey of Computational Semantics: Representation, Inference and Knowledge in Wide-Coverage Text Understanding" *Language and Linguistics Compass*, 5/6. Available at: https://www.let.rug.nl/bos/pubs/Bos2011Compass.pdf
[27] Frankenfield, Jane (2019) "Data Analytics", Investopedia. Available at: https://www.investopedia.com/terms/d/data-analytics.asp

traces left by the organisms interacting with the environment. This, paired with DNA sequencing technologies, allows for a non-invasive and effective approach to biodiversity monitoring.[28]

Geographic information system (GIS) – is a tool to understand geographical patterns and relationships. It involves bringing together location-based data and displaying it relative to Earth's surface. The GIS representation can be in the form of layers, representing divergent kinds of data in a single map, such as urban buildings and natural elements of the environment.[29]

Graphics Processing Unit (GPU) instance – is a single-chip processor that alters the devices' memory in order to accelerate the creation of images and thus boost the image performance. This feature is particularly relevant in applications that require a rapid succession and feedback of graphics (such as videos).

Internet of things (IoT) – is a technology connecting ordinary smart devices or objects incorporating sensors to a computer network and to the internet. This allows devices to communicate with each other and to collect and exchange data in real time. By collecting data in this manner, large volumes of data can be extracted from objects in a short period of time and with little effort.

Intelligent Virtual Agent – are animated chatbots that are generally implemented to aid customer service on online platforms. These agents can be supplemented with AI-powered conversational tool to improve their service and efficiency[30].

Machine learning (ML) – is a subset of artificial intelligence. It refers to technology that enables machines to learn from the data that is inputted inside the machine and acquire skills to perform some tasks with partial human involvement.

Mainframe-based systems – refers to large computing systems, used primarily by big businesses and large organisations that allow for the rapid processing of substantial amounts of data.

Natural Language Processing (NLP) – is a technology that facilitates communication between computers and humans, working as a translator between machine code and human language. For example, NLP makes it possible for computers to extract information from natural texts, , or support human interpretation by classifying text.

Natural Language Understanding (NLU) – is a subarea of NLP that relates to the comprehension and understanding part of natural language by the machine. Similarly to NLP, NLU accepts text in a natural language as input and extracts information that can be used in downstream tasks. NLU refers to the core NLP tasks: it allows the system to interpret the meaning of the natural language, extracting useful information from the text[31].

Neural network (neural net) – is a class of machine learning algorithms that can learn to produce target outputs from specific inputs by analysing training examples. Neural networks

---

[28] Philip Francis Thomsen, Eske Willerslev, 2015, *Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity,* Biological Conservation, Volume 183, Pages 4-18, ISSN 0006-3207, Available at https://doi.org/10.1016/j.biocon.2014.11.019

[29] National Geographic, Resource library encyclopaedic entry. Available at: https://www.nationalgeographic.org/encyclopedia/geographic-information-system-gis/

[30] Chavez, David (2019) "A Buyer's Guide to Intelligent Virtual Agents" Avaya.com Available at: https://www.avaya.com/2019/10/buyers_guide_virtual_agents/

[31] https://expertsystem.com/natural-language-understanding-different-nlp/

can be trained in both a supervised and unsupervised manner. Neural network models differ greatly with respect to their architecture, but all use a sequence of linear transformations interspersed with nonlinearities. There are several types of building blocks that are common: fully connected layers, attention layers, convolutional layers. These building blocks are combined using nonlinear transformations and are arranged hierarchically or in a recurrent way.

Open Source – is a type of licence that is used for disseminating software code. This licence allows for the free and open distribution of the source code behind a software. Software that is derived from the original open source software must then also be freely distributed.

Polygon vector data – are a subset of vector data that are comprised of the types of geospatial data used in spatial representations. Polygon vector data (divergent from point vector data and line vector data) are constituted by three or more connected vertices, thus forming closed polygon shapes. They are used to describe areas.[32] Vector data allow the representation of discrete geographical locations into defined spatial objects.

Predictive analytics (or predictive modelling) – is a process that uses data analysis and modelling to predict future events or behaviours based on past performance or patterns. Generally, this technique involves assigning a probability or risk score to selected events depending on their likelihood of occurrence.

Random forest – is a type of supervised machine learning algorithm. In technical language, processes are being described as ‘decision trees’. The "forest" is a group of decision trees. Random forests are multiple decision trees that are merged together to draw a more accurate and stable prediction from the data it analyses.

R-CNN algorithms – are neural network algorithms for object detection, designed to create and classify feature vectors for image recognition. The output of this process allow machine to recognise the objects contained in an image.

Robotic Process Automation (RPA) – is a technology that allows the user to configure a software (also commonly referred to as “bots” or “robots”) for it to carry out a specific process in an automated manner. The “bot” can be used to mimic the actions that a human would take to complete a specific task. For example, RPA can be used to search for information in a specific system and then input it into a form. This replaces the manual inputting normally done by a human. The technology can be built on various processes and can also overlay on other software applications.

Scraping – is a process of data extraction through which a computer gathers and computes data from human-readable interfaces. The best-known application is web scraping, whereby the content of websites is translated into large amounts of data for computational use.

Source code – is a text, or code, written in programming language that provides a set of commands to be executed by a machine or programme.

Species Distribution Modelling – is a quantitative spatial technique related to quantitative biogeography. It uses tools that model and predict the environmental suitability for species, combining data on species occurrence and environmental conditions.

[32] Dempsey (2017) available https://www.gislounge.com/geodatabases-explored-vector-and-raster-data/

VGG 19 neural network – is a trained convolutional neural network used for image classification. The name derives from the team of developers – Visual Geometry Group (VGG) – and the number of layers in the network.

Word/Sentence Vectors – are numerical vectors that represent the meaning of a word or sentence. They are vectorised representations of words and/or sentences that capture certain lexical, syntactic and semantic properties of a word/sentence. This is more advanced than traditional NLP, as it allows machines to capture the meaning and association of a word or sentence rather the mere dialectical absence/presence of a word or sentence. Typically, these representations are learnt by a neural network from text data in a supervised or an unsupervised manner.

This publication is available from: www.gov.uk/beis

If you need a version of this document in a more accessible format, please email enquiries@beis.gov.uk. Please tell us what format you need. It will help us if you say what assistive technology you use.