



مركز محمد بن راشد
للابتكار الحكومي
MOHAMMED BIN RASHID CENTRE
FOR GOVERNMENT INNOVATION



ابتكر
IBTEKR



Data Science Report for Designing Innovative Services

January 2023

Table of Content

Introduction	04
UAE Introduces Region’s First Platform to Analyse Natural and Human-Made Phenomena	06
United Arab Emirates	
Using Machine Learning & Data for Military Recruitment	10
United States of America	
The London Datastore	14
United kingdom	
Public Private Partnership in Dublin	20
Ireland	
Using Data Science to Improve Workplace Safety Inspections	24
Chile	
Using AI for Recruitment	28
Sweden	
Police uses AI to Spot False Police Reports	32
Spain	
Kansas City: Using Predictive Analytics for Road Maintenance	36
United States of America	
The AuroraAI Programme	40
Finland	

INTRODUCTION

Analytics is an important and rapidly developing business functionality that is used by increasing numbers of organizations around the world. Reflecting this trend, government agencies and departments are now taking the opportunity to leverage contextual real-time information into their decision-making processes. Through this report, we present to you, examples of data analytics use cases through which countries around the world are improving public service delivery while saving costs and increasing efficiencies.

In the past, when reliable data were limited and data collection and analysis required time and money, evidence-based decision making was not easily implemented. Due to the advances of a new environment where real-time data are generated exponentially through, among others, social media, web searching, and the Internet of Things, public organizations are now able to collect data at an unparalleled speed. This has helped them in developing a deeper understanding of public issues and in some cases even predicting future events, and hence preempting excessive damages. For example, in Kansas, the authorities use data to predict impending road damage and pothole creations. This helps them in taking measures before potholes develop and hence saves them the high cost of pothole repairs. Similarly, in Chile, the government uses predictive analytics to improve workplace safety inspections to avoid unfortunate events such as the 2010 cave-in at the San José copper-gold mine that caused several fatalities.

In this report, you will also find examples of governments using data to make hiring decisions and hence improving their performance. Some of these come from Sweden and the United States, where the governments are using analytics to make recruitment decisions by using algorithms built through past data on successful candidates, which allows them to make better recruitment decisions to prevent higher turnovers and improve the performance of their organizations.

The use of data analytics can truly enhance public service delivery and improve the satisfaction levels, and the well being of citizens. For example, countries like Finland are using data to provide seamless services across various departments to their citizens upon certain important life events or moments. So, for example, a person who is about to go through a certain life-event – such as changing jobs – is automatically offered public, private and third sector services to proactively support the transition from one job to another without a period of unemployment.

The positive effects of using data in the public sector are nearly endless. The use of data analytics not only allows the government to pinpoint areas that need attention, but it also gives them that information in real time. It allows governments to make faster decisions, and it allows them to monitor those decisions and quickly enact changes if necessary. The examples provided in this report can help you in developing a deeper understanding of the uses of data analytics and we hope that they encourage you to increasingly employ data in your work and realize its full benefits.



United Arab Emirates

UAE Introduces Region's First Platform to Analyse Natural and Human-Made Phenomena

The Application Development and Analysis Department at the Mohammed Bin Rashid Space Centre (MBRSC) has launched a scientific platform that provides analytical research on natural and man-made phenomena and their impact on environmental resources in the UAE. The first of its kind in the region, the platform focuses on studying the environment in depth, using remote sensing systems, image processing, geographic information system (GIS), and AI. It serves as a database for new studies to enable users and scientists to benefit from studies and research papers offered by MBRSC.

The platform includes analytical and research studies presented regularly to various government entities. These navigate a wide array of topics, from natural events and their impact on environmental resources, such as water, air quality and vegetation, to changes to the country's coasts, to phenomena impacting marine life, like red tide.





The MBRSC adopts the scientific method in producing these reports through four main steps. The first two steps are to identify the phenomenon using MBRSC’s cutting-edge devices and technologies and develop a hypothesis based on the experience of a specialised team. The hypothesis is then tested in line with international standards to determine its accuracy. It undergoes long experiments with different types of variables, and findings are presented in the form of charts and graphs for easier analysis. Then comes the conclusion phase, where data are connected into information of scientific value. It is the last stage where a complete analytical study is produced to help decision makers make informed decisions.

The launch of the platform highlights the UAE’s efforts to drive smart transformation and integration of innovative tech solutions, in light of the progressive thinking adopted by MBRSC to align with the country’s vision.

The platform offers all analytical studies in an interactive manner. It allows users to identify environmental changes in a specific area during a specific period of time – with the help of interactive maps – and compare these changes to a different period to come up with a scientific interpretation of the variables. This interactive feature also enables tracking the characteristics of a particular phenomenon. The user can obtain accurate figures on change rates and determine whether the changes are positive or negative, which significantly helps in measuring the impact of relevant public policies.

The platform was developed using ArcGIS, a leading GIS software used by all UAE government entities concerned with urban development and monitoring natural and man-made environmental events, among others.

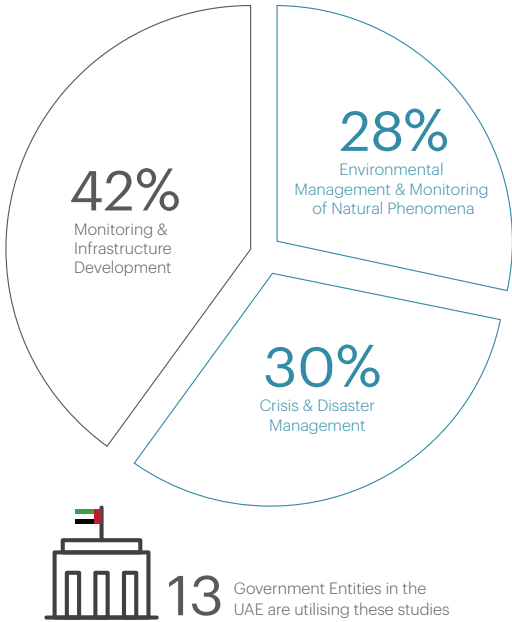
Access to these studies is not limited to government entities, but this scientific tool can benefit other institutions that are not concerned with environmental phenomena, such as real estate developers and educational bodies. Such institutions can tap into space technology at the MBRSC’s Application Development and Analysis Department, as well as its offering of integrated solutions that meet their various requirements.

The platform also makes its studies and images available to various global bodies, most importantly international disaster relief and rescue organisations, which can use these images to enhance their performance for the benefit of humanity. Currently, the MBRSC is working to grow its global partnerships and platform’s user base, and to link the platform to international organisations, which benefit from 20% of its research and studies.

Work on this project began in March 2020, when three UAE female engineers worked hand in hand to launch the platform despite COVID-19 working conditions. Supervised by MBRSC, the team managed to conquer a series of challenges, including the conversion of all data generated by their analytical studies into simplified scientific readings that could be shared on an interactive digital platform. Another challenge was the comparison between environmental events at different times, a highly important feature for those concerned with studying environmental phenomena over long periods of time.

The pandemic-imposed remote work arrangements required the team to intensify efforts to complete the project as efficiently as possible. The engineers also faced the huge responsibility of extracting scientific data upon which decisions of government entities would rely, not to mention other obstacles they were able to overcome to achieve maximum scientific benefit.

The platform has made numerous achievements within a short span of time and a relatively small budget of approximately AED10,000. The platform currently includes 36 interactive studies across the fields mentioned earlier. While 42% of the studies are focused on project monitoring and infrastructure development, 30% are on crisis and disaster management and 28% are on environmental management and monitoring of natural phenomena. As many as 13 government entities in the UAE are utilising these studies.





United States of America

Using Machine Learning & Data for Military Recruitment

With less than 1 percent of young Americans both willing and able to join the military, even fewer of whom can pass special operations selection, each military service in the US is looking at creative ways to use the skills and attributes would-be operators can bring to the table. That includes prioritizing technology experience for some services, while in others, using data and artificial intelligence to pinpoint who might be the best fit.



Each service has its own specific outreach and recruiting practices, but the recruiting process for all services has many commonalities and works on two levels. The national level focuses on outreach to develop an awareness of opportunities and to attract prospective recruits through service websites (i.e., goarmy.com, navy.com, airforce.com); advertisements on television, on radio, and in print; educational tours; chat rooms; direct mail; call centers (1-800 numbers); and mobile tours and events (such as sporting activities or air shows).

At the local level, individual recruiters conduct high school, cadet, community college, and other community programs or events; conduct other local outreach through personal contacts, emails, telephone prospecting, public service announcements, and other face-to-face encounters; assist walk-ins; and conduct interviews and discussions with prospective recruits attracted by national-level efforts. Interested prospects who appear to be qualified are taken through appropriate aptitude and physical tests. If they pass these tests, they can then meet with a job counselor to discuss available jobs for which they qualify, review the enlistment options and training sequences for those jobs, and sign an enlistment contract. The current process is time consuming. For example, the Air Force reported that it takes it 110 contacts with different individuals to get a single recruit under contract and into basic military training. Moreover, the services face challenges in meeting recruiting goals and face resource shortages in manpower, dollars, or both. Further, despite all the costs and effort involved, attrition levels are high. Hence, officials keep searching for efficiencies and early success indicators that go beyond selection criteria. For example, the Marine Corps has studied non-cognitive testing, which measures qualities such as impulse control and "grit," as a way to discover officer candidates who have what it takes to make it through the training pipeline and on to a successful career.

Since late 2019, officials with U.S. Special Operations Command are experimenting with artificial intelligence and machine learning to determine whether it's possible to highlight the qualities that distinguish an ideal operator. Officials began by collecting and assessing a wide swath of data regarding candidates for Marine Corps Forces Special Operations Command. The data being collected covers as wide a range as possible, excluding only information that is protected from a privacy standpoint. The focus is increasingly moving away from an assessment and selection program that's based on performance and more toward one based on attributes.

By first identifying the unique traits that seem to make up the most ideally suited operator, the same practice could be employed on incoming classes — sifting through Marines that don't possess the best traits for an operator and identifying those that are particularly well suited for the rigors of the special operations community. The idea wouldn't necessarily be to use artificial intelligence to disqualify potential candidates, but rather as a special operations recruitment tool. Once the algorithm develops a mature model of what the ideal traits for a special operations are, it could feasibly scour Marine Corps personnel records for the best possible candidates and flag them for recruiter engagement. By identifying the Marines that are most likely to succeed, AI could feasibly lower attrition rates without lowering the command's standards, thanks to seeking out the best possible candidates before the first day of training even begins.

Marine officials aim to have a minimal viable product in six months that they could begin using to use machine learning to identify who would be a good Marine Raider.

In the last few years, there have been many other examples of the US military using data for various aspects of recruiting. For example, the US Navy generates monthly production reports, creates trend and risk analysis, discerns relationships among data points, identifies gender differences, and develops recruiting predictive tools using a variety of statistical methods such as aptitude predictors and recruiter performance analytics. Production reports are being used both as descriptive summaries of historical data and, increasingly, for predictive and prescriptive reporting. The Navy is using analytics to inform much of its recruiting decision-making. Examples include manpower allocation, resource allocation, and program and process improvements.

As another example of use of data for recruiting, US Army Cadet Command (USACC) uses predictive analysis to determine what recruiting campaigns are required. The model is tailored to each school and market, historical trends, and performance. The Army uses machine learning, cadet demographics, and surveys to examine past performance and to inform recruiting campaigns. It wants to extend this analysis to look at particular majors, such as STEM and nursing.



United kingdom

The London Datastore

Created in January 2010, the London Data Store is a data-portal pioneer. At its very core, it is a free and open data-sharing platform where anyone can access data relating to the capital. Whether you're a citizen, business owner, researcher or developer, the site provides over 700 datasets to help you understand the city and develop solutions to London's problems. The Datastore was one of the first of its kind developed by a major city anywhere in the world, making freely available large amounts of data about the capital.





Today the platform has around 60,000 users each month and is home to more than 6,000 datasets — up from around 500 when it was first launched in 2010. Initially the Datastore was used primarily as a tool to improve transparency and accountability. Information on what the Greater London Authority (GLA) spends and data that underpins mayoral strategies were published to allow the public to scrutinize the decisions and activity of the Mayor and the GLA.

However, very quickly it became apparent that opening up the city’s data could enable a whole range of positive outcomes beyond simply greater civic transparency. Today policymakers at City Hall are increasingly seeing data as a means to address some of the city’s challenges through maps and apps open to the public.

- ✓ London Rents Map, which 85,000 Londoners used last year to help them find an affordable home
- ✓ Schools Atlas, which allows for school place planning as well as being a tool for parents selecting schools for their children
- ✓ Cultural Infrastructure Map, which helps people enjoy and preserve music venues, studios and community halls in their neighborhood’s
- ✓ A range of air quality mapping using data from a network of sensors showing Londoners pollution levels in their local areas, as well as prioritizing new electric bus routes.
- ✓ Minimizing disruption from roadworks: The organizations that manage the city’s energy, transport and water infrastructure use the datastore extensively. Mapping Application to share data with each other has enabled them to better target investment and minimize disruption by coordinating street works.

London’s population is estimated to grow to nearly 11 million by 2050

and data, such as that available via the Datastore, will become even more important as the City Hall works with its partners to positively impact Londoners’ lives. At their heart, all of these initiatives are about making data accessible to the people who can use it to make a positive difference to our city.

While the Datastore has already achieved a lot with open data, this is just the tip of the iceberg when it comes to all of the data that is collected, used and shared in London. With this in mind the London Data Store is exploring how the Datastore can support the sharing of data that isn’t necessarily suitable to be published for public consumption. Additionally, the type of data that the Datastore is being used to share is changing. In 2018 the Datastore gained the capability to share ‘secure’ data (restricted and/or licensed data including permission, privacy, publication and distribution; as well as data that is presently held privately).

With the adoption of sensors there are now live feeds of data being shared through the Datastore, a change from the usual static tables of data in a CSV or excel format. Data from sensors is a departure because it is dynamic and different in nature from the monthly or quarterly published statistical data which was the Datastore’s mainstay until then.

An example of an initiative sharing private data arose when data from utilities and private developers was brought together to create the Infrastructure Mapping App, which allows coordination of roadworks to minimize disruption and congestion.

Although the intention was to cater for data analysts, a number of applications have been developed to allow data to be visualized and used by a non-specialist audience and today users, of which there are about 60,000 each month, are split about 50/50 between the two groups.

In parallel, Transport for London (TfL) developed the Unified API, which opened a number of datasets for buses and trams. This led to the creation of around 600 apps used by 42% of Londoners, directly supported more than 500 jobs and generated annual economic benefits and savings of up to £130 million annually for travelers, London and TfL itself.

With new use cases such as these, the Datastore can no longer be considered simply a repository of open data. Therefore in 2019 the London Datastore undertook a discovery exercise with help from the Open Data Institute (ODI) to help redefine the Datastore’s evolving mission.

Over the last three months of 2019 the Datastore listened to hundreds of users in London’s data community — including both users and non-users of the Datastore coming from local authorities, academia, the private sector, civil society, think tanks, media organizations, the NHS and other public sector bodies — to explore their needs when it came to accessing data to do their jobs.

Users indicted that the platform does a good job of making London's data more accessible — although there is room for improvement, particularly around search and navigation as well some of the functionality around sharing secure data. However, many of the difficulties that people faced cover a range of cultural and process barriers. Many of London's institutions have recognized for several years the value in opening up and collaborating with data. However, this common understanding has not always translated into action.

The discovery process identified the two largest barriers to this was a lack of clarity on what standards should be adhered to (i.e. what “good” looks like) and a lack of leadership to pull people around a common challenge. Hence beyond making some technical improvements to the Datastore platform to ensure it is fit for purpose as a central registry of London's data, the discovery findings have opened up a wider conversation around what guidance and leadership is needed to ensure a positive impact with the city's data.



Ireland

Public-Private Partnership in Dublin

Founded by the four Dublin Local Authorities, Smart Dublin's goal is to future-proof the Dublin region by trialing and scaling innovative solutions to a wide range of local challenges. It brings together technology providers, academia and citizens to transform public services and enhance quality of life. The program's coverage extends over a range of topics including government, economy, people, mobility, living and the environment. One of Smart Dublin's remit is to identify challenges that local authorities face and engage with the private sector to test new technologies to solve them. One challenge was getting accurate spending data, precise information about spending patterns in the city which the city aims to solve through this public-private partnership.





In November 2017, Mastercard and Dublin City Council (DCC) announced a three-year innovation partnership to support Dublin's city planning and future development efforts as part of its Smart Dublin programme. Under this project, Mastercard develops its SpendingPulse reports to be released as part of DCC's Quarterly Economic Monitor. Mastercard SpendingPulse™ aggregates credit card transactions from the proprietary Mastercard payments networks across companies, public or private to generate sector level values.

Given the breadth of coverage, SpendingPulse™ can also be used to assess total market trends, regional sales, and industry benchmarking. The Mastercard SpendingPulse™ DataFeed contains sales volume information for eight countries, each with at least seven years of monthly history. Within each country, a variety of sectors, subsectors, and segments are provided. In addition, where possible, Mastercard breaks out eCommerce sales from traditional brick-and-mortar sales.

SpendingPulse™ is built upon a proprietary methodology that spans three key pillars: the sources of transactions, point of sale location, and sector classification. A combination of aggregate and anonymous sales activity from the Mastercard payments network estimates for other forms of payment (i.e. cash or check) and supplemental data from third-party sources are used to create indexed sales values. To calculate country and city specific sector values, both the merchant and the location of the transaction must be identified. Mastercard relies on the point of sale information to map a transaction to a given geography.

The aim of the three-year agreement between Dublin and Mastercard is to use the data from Dublin to improve city intelligence and performance. This unique partnership with Mastercard has allowed Dublin to develop up-to-date insights on its economic performance at a time when the United Kingdom was facing a number of external threats such as Brexit. The project is a great example of private sector data being used for smarter purposes and supporting future city planning.

The reports on spending in Dublin prepared by Mastercard include insights on retail spending and tourism patterns drawn from anonymized and aggregated transaction data as well as other means of payments such as cash and cheques. This helps the city officials understand the spending patterns of Dubliners and tourists, thus allowing the city authorities to compare the capital's performance to the whole of Ireland. Over the course of the three-year agreement there is also a commitment to drive innovation opportunities in transit, planning and tourism and to engage with Dublin's thriving startup ecosystem.

As part of the partnership with Dublin City Council, Mastercard will work to engage with Dublin startups to come up with new innovations to promote a cashless society while also developing new innovations in using the city's data.

Another collaboration between Mastercard and Dublin is through the City Possible program lead by Mastercard since 2019. City Possible brings together different cities to understand what local authorities want to know, and connects them with private companies who can respond and build products in turn. City Possible is a new public-private partnership model focused on meeting the needs of people in cities.

For the past few years Smart Dublin has worked with Mastercard and this network of cities. One tool that has emerged out of the City Possible initiative is called City Insights. The tool contains anonymized spending data paired with geospatial locations within a city. Smart Dublin, along with other cities in the network, have been giving feedback on City Insights in order to refine it into a product.

The City Insights tool helps local authorities assess how planned and unplanned events impact the local economy. Through this tool, the city government has access to information on how much people spend in different neighborhoods on different days. For example, it can get a better picture of how specific events such as St Patrick's Day or the upcoming Euro 2020 matches in Dublin affect the retail sector. The data can potentially be used in the future to show the success or failure of policy decisions such as the creation of strategic development zones (SDZs).

Mastercard hopes that City Possible members will take advantage of the City Insights initiative to help mitigate potential troublesome issues with the city, like natural disasters. On the other hand, it can also be used to bolster economic growth and identify opportunities for local investment.



Chile

Using Data Science to Improve Workplace Safety Inspections

ILO 176 – a convention that protects people working in mining - has been in existence for more than a decade and has been adopted by 25 countries. It upholds the rights of workers to refuse work they consider unsafe, to leave a mine they consider dangerous and to elect their own health and safety representatives. It seeks to protect the rights of workers if they complain about their working conditions.



Every year, thousands of Chileans are killed or injured in work-related accidents. This was recently brought to light during the 2010 Copiapó mining accident where due to a cave-in at the San José copper-gold mine in the Atacama Desert near Copiapó, 33 men were trapped 700 meters (2,300 ft).

For Chile, one study reported a rate of 15.4/100,000 worker fatalities in 2003. According to the findings of another study, Chile's mortality rate from work-related injuries, including commuting injuries, was 7.2/100 000 workers for 2014, and 7.3/100 000 workers for 2015. Moreover, the analysis of worker fatality from occupational injuries in Chile in 2014 and 2015 reveals that more workers die from incidents on the job than during commuting.

Chile's labor ministry, Dirección del Trabajo (DT), was tasked with increasing workplace safety through inspections and enforcement. But DT's inspections were largely reactive: complaints would come in and then an inspection was undertaken and completed, often after an injury or death.

In this context, the benefits from preventative inspections are considerable, as these inspections can help find safety issues before bad things happen. Understanding the merits of such measures, DT started moving to preventative inspections. They hired data scientists, built some models, and even ran field trials. But their efforts ran into many challenges, ranging from having data on what labor facilities even exist to identifying new facilities that might be at risk of violating workplace safety. Tax records report the location of the company's headquarters, but the government may not know where other corporate facilities exist. DT's inspectors wasted a lot of time going to non-existent facilities (the facility had closed or moved). Inspectors were assigned to regions and sometimes to industries and violation types. The inspection list is generated once a month, but DT doesn't collect all of the labels until after another list is produced. Only a small portion of facilities is ever inspected —DT does not know if a violation exists unless inspectors go there.

DT employs a limited number of inspectors and more often than not they inspect facilities with no hazards. Their capacity allowed them to inspect about 20% of the companies located in Chile. Their hope was to identify and inspect the most likely offenders and also create a culture of anticipating inspection. By accomplishing these goals, DT could ensure workplaces in Chile became more compliant with labor laws leading to a safer and better environment for workers.

In 2018, DT partnered with Data Science for Social Good (DSSG), a University of Chicago based program, to help address these problems. DSSG worked with DT to build more accurate and useful machine learning models to help DT perform preventative inspections. DT provided DSSG with data from the last 10 years of inspections including employee numbers and extensive company and facility specifications. DSSG also obtained copper prices from the Commodity Exchange and other macroeconomic data from the Organization for Economic Co-operation and Development.

DSSG used this data to build a model to assess whether a facility would be found to have a violation if inspected. In addition, it also built a "similarity" model, assessing how similar a facility is to those previously inspected. Fewer than half of facilities had been inspected in the past, and they differ in significant ways from the rest, not the least of which is that they haven't had safety complaints. By plotting risk score against similarity, DT can now easily compare a facility's risk and a score is assigned based on the level of risk of being noncompliant to workplace safety rules. Using these scores, DT can determine which facilities are at a risk of being noncompliant and hence inspect them physically.

Bias detection: violation and inspection model

This project was handed over to DSaPP (Center for Data Science and Public Policy at the University of Chicago) for improving the machine learning models and designing and conducting field trials. The trials used both a list of facilities produced by the model and a random sampling stratified by region and industry to:

- Test the accuracy of the existing model and its predicted risk scores
- Calculate a base rate and identify existing biases in the inspection process

The DSSG expect their work to affect a reduction in inspection bias by efficiently using results from the inspection model and incorporating geographical and industry constraints in predicting risk. Future work to ensure safety in Chile's workplaces includes:

- Developing a user-friendly interface integrating the model with DT's existing systems
- Providing DT with motivation and tools to develop further data Science-supported projects

The project also serves as a capacity-building opportunity for DT, with the goal of incorporating data Science methods across Chilean public services.



Sweden

Using AI for Recruitment

In 2018, Upplands-Bro municipality had a foreign-born population of just over 27%, putting it significantly above the Swedish national average. But the unemployment rate among foreign-born individuals is three times that of the Swedish-born population, suggesting foreigners experience more difficulty finding work.



Companies around the world are looking to make their hiring processes bias-free. Research shows diversity brings many advantages to a business, including increased profitability and creativity, new perspectives, ideas and experiences, and better problem-solving abilities. This can help create more effective and resilient organizations.

The Upplands-Bro municipality, located outside Stockholm, has set digitalization as a priority for the community and is currently rolling out a variety of future technologies like artificial intelligence. The municipality took the next steps towards modernization by pushing their digital agenda for unbiased recruitment in May 2019.

They are the first clients to work with the Swedish recruitment and staffing agency, TNG's, newly launched social interview robot, Tengai Unbiased.



TNG has worked towards the development of an unbiased recruitment process over the past 15 years and is now taking the next step to disrupt the job interview process by preventing unconscious bias affecting hiring decisions through this innovation.

Tengai is the first physical AI interview robot in the world. She is unbiased by design as she doesn't see age, gender, clothing, backgrounds, and looks. She comes with inclusion and diversity software and performs blind interviews and then evaluates and measure soft-skills and personality traits. Questions are always asked in the same way, in the same order, making the interview fair for each candidate. And Tengai gives human feedback, such as nodding or smiling, to encourage candidates to elaborate on their answers. Recruiters and managers are then given a transcript of the interview and decide whether to put the candidate forward to the next round of interviews, based on the answers alone. Because Tengai doesn't engage in small talk before or after the interview, no assumptions are made based on any information gleaned.

To build Tengai, TNG partnered with a Stockholm-based social robotics and conversational AI startup Furhat Robotics. Furhat created a platform for social robots that can build conversations with humans by speaking, listening and showing emotions. The platform includes the three-dimensional bust with a projection of a human-like face as well as the operating system with intelligent natural language processing tools to understand and participate in a conversation. Furhat didn't build the platform for a specific use case, but rather offers developer tools for third parties to create their own applications. TNG designed Tengai on top of Furhat's social robotics platform. It began by creating an HR application, including question trees and social skills, to conduct an unbiased competence-based interview process. TNG also trained the robot using machine learning tools and data collected from multiple test interviews conducted by a diverse pool.

Most AI systems make decisions by looking at historical data, which only perpetuates existing biases. To avoid such pitfalls, TNG made it a point to not train Tengai on historical data. Instead, it was trained on [competency-based] behavioral anchors for every specific competency that TNG was trying to measure. These behavioral anchors demonstrated a certain competency, depending on the job role. The process looks something like this: The hiring manager first identifies the most essential competencies needed for the role and then designs a few questions—typically to pose a situation that would require certain behaviors that demonstrate the competencies associated with that job. For example, if the role has a customer service focus, Tengai might say: "Tell about a time when you were not effective in meeting a customer's needs and what steps you took to correct the situation."

To ensure that no human biases creep in, TNG turned to a mix of recruiters and volunteers, with diverse races, religions and gender, to train Tengai in the initial stages. In addition, they ensured the robot's algorithm did not look for any data on age, gender, ethnicity, or religion during the interview.

Besides removing human bias, Tengai also helps job seekers get more candid during the interview. The majority of the candidates interviewed by Tengai felt reported that they could answer Tengai's questions more honestly and sincerely than with a human recruiter. Tengai seemed to make them more comfortable, and they opened up and talked about difficult issues. Further, interviewees don't need to worry about the interviewer's body language or second-guess what the interviewer wants to hear.

Tengai is redefining what an objective assessment of the interview process looks like by pushing recruitment tech to a new level and assisting recruiters and managers to make data-driven hiring decisions, instead of decisions based on gut feeling.



Spain

Police uses AI to Spot False Police Reports

In many parts of the world the filing of false police statement is a crime that carries serious consequences, such as jail terms and heavy fines. False statements not only contaminate police databases and damage the outcomes of criminal investigations, but also waste significant amounts of public resources that could be dedicated to pursuing other crimes. Yet, false reports are very common, especially in the reporting of low level crimes such as robbery. Until recently, the only strategy for catching false reports was asking seasoned police officers to review suspicious reports, but this approach was not always effective.



Veripol

To spot false statements in a timely manner, in Spain, the police is using a computer tool called VeriPol. VeriPol, created by scientists in Wales and Spain, uses a combination of automatic text analysis and advanced machine learning techniques, to successfully identify false robbery reports with over 80 per cent accuracy. The tool has now been rolled out across all of Spain to help support police officers and indicate where further investigations are necessary.

VeriPol is specific to reports of robbery and can recognise patterns that are more common with false claims, such as the types of items reported stolen, finer details of incidents and descriptions of a perpetrator. Its research team, which included computer Science experts from Cardiff University and Charles III University of Madrid, believe the tool could save the police both time and effort by complementing traditional investigative techniques, whilst also deterring people from filing fake statements in the first place.

VeriPol is partly based on a process known as natural language processing – a branch of artificial intelligence that helps computers understand, interpret and manipulate human language. For example, the computer tool uses algorithms to identify and quantify various features in text, such as adjectives, acronyms, verbs, nouns, punctuation marks and numbers and figures. Historical police reports that were known to be false have been fed through VeriPol so that it could code each one and begin to 'learn' the specific patterns.

1,122 Robbery reports

VeriPol was trained on a total of 1,122 robbery reports the national police had closed—meaning either the thief had been convicted or the complainant had confessed to fabricating the crime. It then tested how accurately the algorithm classified a sample of 659 reports as true or false, compared with two human experts. VeriPol outperformed the police by 15 and 20 percent, respectively. The results, published in June 2018 in Knowledge-Based Systems, have also helped researchers understand how people lie to the police. Fabricated reports, for example, tend to describe a specific modus operandi (the attacker usually wears a helmet or attacks from behind). They also use shorter sentences and lack information about the actual incident.

83% cases closed

VeriPol was put to task on a real-life pilot study in the urban areas of Murcia and Malaga in Spain in June 2017. In just one week, 25 cases of false robbery reports were detected in Murcia, resulting in the cases being closed, and a further 39 were detected and closed in Malaga. In comparison, over the course of eight years between 2008 and 2016, the average number of false reports detected and cases closed by police officers in the month of June was 3.33 for Murcia and 12.14 for Malaga. After VeriPol had assigned a high probability of falsehood to the reports and the claimants were further interrogated, around 83 per cent of cases were subsequently closed.

Police officers across Spain are now using VeriPol and integrating it into their working practices. Ultimately, the police hope that by showing that automatic detection is possible people will be deterred from lying to the police in the first instance.

Researchers elsewhere, who were not involved in the research, think that VeriPol's success could be replicated in other countries—particularly where police departments are short-staffed. They hope that VeriPol would also be used to spot other often-staged crimes, such as home burglary or car theft.



United States of America

Kansas City: Using Predictive Analytics for Road Maintenance

Potholes are caused by numerous factors including, moisture damage, freeze-thaw damage, pavement base failure, heavy traffic loads etc. Potholes pose a liability to both vehicle and pedestrian traffic. Potholes can cause flat tires, bent rims, and vehicle alignment issues and, in some cases, auto accidents may occur. Recent statistics have indicated that potholes cost American drivers over \$6 billion dollars per year.



Typical cost to repair potholes is approximately \$35 to \$50 per pothole. There may be an initial mobilization cost of about \$100 to \$150 to bring trucks and crew out to the repair site. Size of potholes as well as materials used can also affect the repair cost.

An impact from a deep pothole can equal the impact of a 35-mph accident. Damage to vehicles attributable to potholes cost an average of over \$300 per vehicle per year and over \$1,000 during the lifetime of a vehicle. Moreover, potholes cause countless injuries each year to cyclists and pedestrians. In fact, they are the leading pedestrian tripping hazard for property and business owners



One Vehicle



300\$

Per Year



1000\$

Over The Entire Life of Vehicle

In Kansas City, Missouri, in 2017, the local government used Amazon Web Services, new street-level connected devices, and data analytics tools from the Chicago-based software company Xaqt, to prevent potholes before they started — rather than engage in full-scale street repairs after potholes had occurred.

The city’s road maintenance backlog had grown to about \$40 million in 2018.

The local government then introduced ballot initiatives that could inject an additional \$8 to \$10 million into roadway maintenance. The officials wanted to make the most of these funds. The platform from Xaqt helped the city to get a better handle on the \$3 million to \$6 million the city spent on road maintenance annually in two ways — by predicting where potholes were likely to occur and by prioritizing items on its backlog.

Pothole repairs last just six to 12 months, while long-term roadway repairs can last a decade or longer, which means if the city can know where potholes will occur, its limited resources will go further toward preventative maintenance. The technology does more than predict the number of potholes — it also pinpoints the exact pothole location. Decisions are made via Xaqt’s algorithm, which uses machine learning to interpret the data.

Furthermore, the program engaged not just the city staff at the mayoral level, but at the operator level — the individuals who actually interfaced with human beings at city streets level.

Kansas City first began installing smart streetlights that can automatically dim — and also track pedestrian and vehicle traffic — in 2015. Since then, the city has built a suite of tools to better understand how well things are running. A network of 178 smart streetlights can now be found primarily along the city’s Main Street.

Xaqt didn’t have a pothole prediction tool ready to sell the city. The project started with a request from the city’s public works director for a tool that would provide “demonstrative and immediate” improvement to street repairs.

According to Xaqt, they began by talking to public works to better understand what causes road damage, such as the expansion and contraction of asphalt in tough conditions such as heavy bus use, water main breaks or extreme temperature changes. The company and the city looked at historical data and each generated prediction of where the trouble areas were, and they compared notes. Xaqt discovered that even when their algorithm wasn’t right, they were highly effective at identifying the segments of road that were in high need of repair, which then allowed the city to prioritize their resources around road inspections.

The city’s pothole prediction tool is a recent example of how the local government is plying its data against its pressing problems and the administration’s priorities. With its new tools, the city predicts it can also identify vacant structures with 85 percent higher accuracy. Other examples include: an online map that shows live and historical data on parking availability, traffic flow and street car locations downtown and a digital inclusion map that combines the city’s internet availability using data from the Federal Communications Commission and the U.S. Census Bureau to show how internet access aligns with poverty.

Like many government organizations, Kansas City is also managing an aging workforce. Therefore, the city is trying to capture all the knowledge of city staff while it can and using this data to build into AI and machine learning applications for better city management.

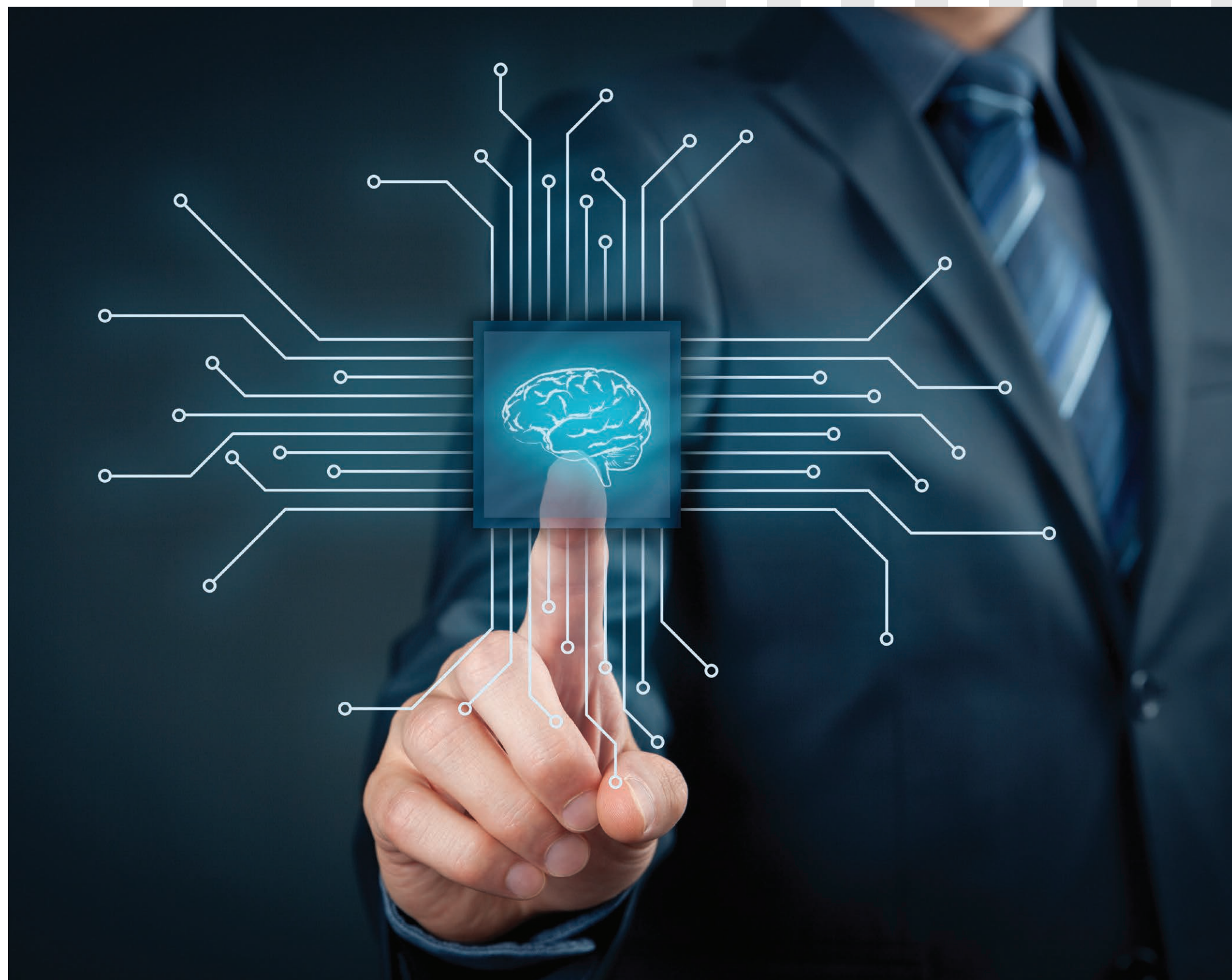


Finland

The AuroraAI Programme

On 6 February 2020, the Ministry of Finance launched the AuroraAI programme. Continuing until the end of 2022, the programme lays the foundation for using artificial intelligence to bring services and people together in a better way.

The AuroraAI programme creates the prerequisites for smoother service use in different situations and life-events. In other words, it reduces the amount of time a citizen spends going from one government office to another to get more information and benefits that he may be entitled to claim. The programme will also improve the functional and technical possibilities for cooperation between public administration and other sectors in addition to strengthening the ethical use of artificial intelligence as part of public services provision and operating models.



The core idea of the AuroraAI programme involves proactively offering services to people according to their own life-events. AuroraAI helps service users and service providers find each other. The model is based on people's needs, whereby artificial intelligence helps citizens and companies to utilise services in a timely and ethically sustainable manner. This will provide people with access to a new way of taking care of their overall well-being and, at the same time, will promote service providers' ability to form customer-oriented and dynamic service chains in collaboration with other operators and to manage their activities based on up-to-date information.

life-event – such as changing jobs – will automatically be offered public, private and third sector services to proactively support the transition from one job to another without a separate period of unemployment. At Aurora's heart is reinforcement learning, whereby the software identifies the combinations of services – from public and private providers – that prove most popular with particular user groups over time. These services get prioritized and promoted, while the less popular combinations are laid by the digital wayside.

The model will be supported by development of the AuroraAI network, which will enable the technical prerequisites for interaction and interoperability between services functioning on different platforms. The AuroraAI network serves as an interactive link between services provided by public administration organisations and the services of other sectors. The AuroraAI network will be available to people and organisations for use in selected life-events by the end of 2022.

The AuroraAI programme is based on open, cross-sectoral networking. In order to facilitate the work, an AuroraAI cooperation group will be set up as well as implementation support teams around different themes. The organisations participating in the programme will jointly improve technology and digitalization capabilities in the public sector and develop cooperation between the public and private sectors in accordance with the government programme.

Aurora would be the first AI assistant built by and for a country's public sector, complementing the slew of private AI assistants like Amazon's Alexa, Apple's Siri, and the Google Assistant. The Aurora plan has some precedent in Finland's public services, as the Finnish Immigration Service already uses a robotic assistant to answer calls in a multitude of languages and direct callers automatically to different organizations for particular questions and requests.

The term set by the Ministry of Finance for the preliminary study on the Aurora national artificial intelligence programme ran from 15 September 2018 to 28 February 2019. The preliminary study was carried out as extensive, open networking between the public, private and third sectors and, at the same time, was one of the suggestions made in the report AI Finland. The preliminary study identified the kinds of changes activities based on human-centricity and life event-thinking signify, for example, for the provision of services and for management. In addition, during the preliminary study, the first test version of the AuroraAI network of smart services and applications was created, as was the Development and implementation plan for AuroraAI 2019–2023.

As a rule, the AuroraAI products will be made available for public, private and third-sector operators. The programme aims to create conditions and processes that will enable the development and maintenance of this AI-based service model and the AuroraAI network after the end of the programme.

The AuroraAI programme is based on Prime Minister Sanna Marin's government programme, according to which Finland wants to be recognized as a technology development pioneer that enhances and implements the opportunities created by digitalization and technology advances across administrative and sectoral boundaries. As derived from the government programme, the aim of the AuroraAI programme is to use the AuroraAI network to facilitate smoother functioning of everyday life and business in a data-secure and ethically sustainable manner.

Aurora's trial/pilot period began in September 2018, focusing on three specific scenarios: moving to a new place of study, taking courses to improve employment opportunities, and supporting children and parents in changing family relationships. The project will be further developed, with even more life events falling under its remit between 2020 and 2022.

